ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

DISTRIBUTED LEARNING POLICIES FOR POWER ALLOCATION GAMES

Panayotis MERTIKOPOULOS
Elena V. BELMEGA
Aris L. MOUSTAKAS
Samson LASAULCE

March 2011

Cahier n° 2011-07

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.enseignement.polytechnique.fr/economie/
mailto:chantal.poujouly@polytechnique.edu

Distributed Learning Policies for Power Allocation Games

Panayotis Mertikopoulos, Elena V. Belmega, Aris L. Moustakas and Samson Lasaulce

Abstract

We analyze the problem of distributed power allocation for orthogonal channels by considering a non-cooperative game whose strategy space corresponds to the users' distribution of transmission power over the network's channels. When the channels are static, we find that this game admits an exact potential function and this allows us to show that it has a unique equilibrium almost surely. Furthermore, using the game's potential property, we derive a modified version of the replicator dynamics of evolutionary game theory which applies to this continuous game, and we show that if the network's users employ a distributed learning scheme based on these dynamics, then they converge to equilibrium exponentially quickly. On the other hand, a major challenge occurs if the channels do not remain static but fluctuate stochastically over time, following a stationary ergodic process. In that case, the associated ergodic game still admits a unique equilibrium, but the learning analysis becomes much more complicated because the replicator dynamics are no longer deterministic. Nonetheless, by employing results from the theory of stochastic approximation, we show that users still converge to the game's unique equilibrium.

Our analysis hinges on a game-theoretical result which is of independent interest: in finite player games which admit a (possibly nonlinear) convex potential function, the replicator dynamics (suitably modified to account for nonlinear payoffs) converge to an ε-neighborhood of an equilibrium at time of order $\mathcal{O}(\log(1/\varepsilon))$.

Index Terms

Non-cooperative games; Nash equilibrium; potential games; power allocation; replicator dynamics.

I. Introduction

As a result of the decentralized nature of future and emergent wireless networks, game theory has become an important tool to describe and analyze distributed resource allocation problems in which wireless nodes cannot be assumed to adhere to centrally controlled protocols. The main focus of these considerations has been to devise policies and algorithms that the network's nodes can use to optimize their resources (power, bandwidth, etc.) on their own, so, following [1], the main questions that arise are *a*) whether there exist "equilibrial" policies which

P. Mertikopoulos is with the Economics Department of the École Polytechnique, Paris, France; E. V. Belmega with the Alcatel-Lucent Chair on Flexible Radio in Supélec, Gif-sur-Yvette, France; A. L. Moustakas with the Physics Department of the National and Kapodistrian University of Athens, Greece; and S. Lasaulce with the Laboratoire des Signaux et Systèmes, Supélec, Gif-sur-Yvette, France.

Part of this work was submitted in ValueTools '11: 6th International Conference on Performance Evaluation Methodologies and Tools.

are stable against unilateral deviations; *b*) whether these (Nash) equilibria are unique; and *c*) whether they can be reached in short enough time by distributed learning algorithms which require only local information.

Accordingly, an important paradigm which has attracted significant interest in the literature concerns the allocation of transmission power over orthogonal communication channels – see [1–3] for a survey. From the point of view of centralized wireless systems, this has been a relatively well-studied subject, especially with respect to optimal power allocation schemes which allow users to reach the boundary of the rate region assuming full channel knowledge and centralized control [4, 5]. On the other hand, more recent examinations focus on reaching socially stable power allocations [6–10], because, even if the globally optimal, capacity-achieving power allocation is known, it might be unstable under unilateral deviations by selfish users (and thus useless in a decentralized setting).

In this paper, we consider the problem of uplink communication in multi-user networks composed of possibly several receivers that operate on distinct, non-interfering channels, and we focus on giving definitive answers to points (a)–(c) above, analyzing the the equilibrial structure of the problem and its convergence aspects. Despite its apparent simplicity, this parallel multiple access channel (PMAC) model has several relevant applications such as, for instance, in 802.11-based wireless local area networks (WLANs) which include several non-overlapping channels over which the users need to independently optimize their uplink transmission powers [11] (a special case of the need to have simultaneously connected users that upload information over several overlapping heterogeneous wireless networks [12, 13]). Furthermore, the PMAC model can be applied to distributed (or partially distributed) soft handoffs in cellular systems [14], distributed power allocation in digital subscriber lines (DSLs) [7], and, finally, it may also form the basis of throughput-maximizing power control in multi-carrier code division multiple access (CDMA) systems [8].

Our analysis will focus on the simple single-user decoding (SUD) scheme in which the transmitted signal of each user is decoded separately by the receiver(s) by treating the incoming signal of other users as additive (Gaussian) noise. The main reason why we are using SUD instead of successive interference cancellation (SIC) is that the former is known to have lower decoding complexity and signalling overhead than the latter – a consequence of SUD not having to broadcast the decoding order to the transmitters [15]. In addition however, there is also the important limitation that the equilibrial structure of the corresponding PMAC model using SIC instead of SUD is not well-understood even for static channels (let alone for stochastically fluctuating ones).

In a series of related papers [9, 10], Scutari et al. studied non-cooperative power allocation games in the context of static Gaussian interference channels (ICs). There, the existence of a Nash equilibrium (in the "pure" sense of Rosen [16]) is a consequence of the convexity properties of the users' achievable rate functions and follows directly from Theorem 1 in [16]. In fact, under suitable (but somewhat stringent) conditions on the channel matrices, this equilibrium solution is shown to be unique, and, echoing the results of [17], it was demonstrated that, under the same conditions, iterative water-filling algorithms converge to the game's equilibrium.

At a formal level, the static PMAC setting is a special case of the IC framework of [9, 10], but, as we shall see, the derived sufficient conditions of [9, 10] typically fail in the static PMAC case, leaving the issues of uniqueness and convergence wide open. Therefore, although the (global) capacity region of this channel is fully determined

[2, 4, 5], the distributed version of this channel remains unresolved. To account for this, the authors of [14] recently introduced a power allocation game in which users choose how to allocate their available power among the receivers and showed that the game admits an exact potential function in the sense of Monderer and Shapley [18]. As in the more general multiple-input multiple-output (MIMO) multiple access channel (MAC) case [19], the game's potential function is convex, so Nash equilibria correspond to the (compact and convex) minimum set of the potential. However, the game's potential is, in general, not *strictly* convex, so one would expect that uniqueness of Nash equilibria fails along with strict convexity and the sufficient conditions of [9, 10]. Rather surprisingly, we find that this is not the case: even though these conditions do not hold in the static PMAC context, the Nash equilibrium of the game is almost surely unique (Theorem 6).

As far as convergence to equilibrium is concerned in the static channel case, the authors of [6] considered a setting similar to our own (incorporating a linear pricing model, but limited to a single channel) and exhibited various power control algorithms which converge to equilibrium under certain "mild-interference" conditions. Furthermore, one of the main results of [7] was to show that if the transmitters know the local channel state and the overall inteference-plus-noise covariance matrix, then, assuming that the interference level among users is not too large, iterative water-filling converges to the equilibrium set of the game. The authors of [17] enhanced this result by removing this condition for a suitably modified iterative water-filling algorithm, but asynchronous/sequential iterative water-filling is much harder to analyze because the sufficient conditions of [9, 10] fail in the PMAC model.

Instead of focusing on water-filling algorithms, we present a simpler learning scheme based on the replicator dynamics of evolutionary game theory [20] which involves the same (or often less) information from the side of the players, and which does not require them to solve a nonlinear water-filling problem. Dynamics of this type have been studied extensively in finite Nash games (finitely many actions with payoffs which are multilinear with respect to the users' strategies) and continuous population games [21–23], but, in *nonlinear* games (such as the one we have here), their properties are not as well understood. The reason for this difficulty is that nonlinear payoff functions cannot be written as multilinear combinations of "coordinate-spfecific" utilities (representing pure strategies in finite games or species/phenotypes in evolutionary ones), so it is not clear what payoffs to use in order to update the components of the users' power allocation profile. Nevertheless, using the potential property of the game, we were able to identify the correct modified version of the users' payoffs which allows the replicator dynamics to converge to the game's (a.s.) unique equilibrium *unconditionally* and at an exponential rate: *users reach an ε-neighborhood of the game's (unique) equilibrium in time which is at most of order* $\mathcal{O}(\log(1/\varepsilon))$ (Theorem 8).

On the other hand, a major challenge occurs if the channels are block-fading, i.e., they remain constant over a packet transmission and have undergone a stochastic variation by the beginning of the next transmission block. In this fading regime, the static game equilibrium and the convergence properties of the replicator dynamics lose much of their relevance because, if users base their updating on the instantaneous channel gain coefficients, learning becomes a *stochastic* process. Since we are considering a relatively fast power allocation scheme which tries to adapt as soon as possible to the variability of the observed channel, the rate of these stochastic fluctuations can be assumed to be of similar magnitude to the power adaptation rate. As a result, by using the theory of stochastic approximation [24] (as opposed to the framework of stochastic differential equations where the adaptation rate is an order of magnitude lower than that of the underlying noise process), we show that the stochastically perturbed replicator dynamics converge to the (unique) equilibrium of an averaged game whose payoff functions correspond to the users' achievable ergodic rates (Theorem 10).

The behavior of the replicator dynamics is similar when the channel is changing faster than the power updating rate and the throughput of each user between updates is exactly their ergodic rate. This case falls within the analysis of power allocation games in fading MIMO channels [25], where the authors proved existence and uniqueness of equilibrium using Rosen's theorems on concave games [16] (but did not conduct any dynamical analysis).

Our own convergence analysis relies heavily on a novel game-theoretic result which is of independent interest: *in games which admit a (strictly) convex potential function, the replicator dynamics converge to (the game's unique) equilibrium at an exponential rate* (Theorem 13). In this way, and in addition to its applicational significance, the PMAC model becomes a simple and illustrative case study which exemplifies the power of the replicator dynamics when combined with the potential game property. Indeed, despite its apparent simplicity compared to the full IC model, much of the analysis of [9, 10] does not apply; instead, it is the potential properties of the PMAC game that guarantee equilibrium uniqueness, and the (properly modified) replicator dynamics allow users to equilibrate exponentially quickly. In fact, since equilibria are located at the minimum set of the potential, this result essentially means that the replicator dynamics associated to a convex objective function converge to an ε-neighborhood of its minimum set in time which is of order $\mathcal{O}(\log(1/\varepsilon))$.

Notational Conventions

We will use bold uppercase letters to denote matrices and "†" to denote their Hermitian transpose; also, for any set denoted by a script letter, the corresponding unscripted character will denote its cardinality, as in $S \equiv \operatorname{card}(\mathcal{S})$.

Now, if $\mathcal{S} = \{s_\alpha\}_{\alpha=1}^S$ is a finite set, we will denote by $\mathbb{R}^\mathcal{S}$ the (real) vector space spanned by $\mathcal{S}$. Then, given the natural identification between $\mathcal{S}$ and the canonical basis $\{e_\alpha\}_{\alpha=1}^S$ of $\mathbb{R}^\mathcal{S}$, we will use α to refer interchangeably to either s_α or e_α, depending on the context. Similarly, we will also identify the set $\Delta(\mathcal{S})$ of probability measures on $\mathcal{S}$ with the standard $(S-1)$-dimensional simplex of $\mathbb{R}^\mathcal{S}$: $\Delta(\mathcal{S}) \equiv \{x \in \mathbb{R}^\mathcal{S} : \sum_\alpha x_\alpha = 1 \text{ and } x_\alpha \geq 0\}$.

Finally, as far as players and their strategies are concerned, we will follow the original convention of [26] and employ Latin indices for players $(k, \ell, \dots)$, while reserving Greek ones $(\alpha, \beta, \dots)$ for their ("pure") strategies. Also, when we want to sum over the strategy indices $\alpha \in \mathcal{A}_k$ of a particular player k, we will use the shorthand notation $\sum_\alpha^k \equiv \sum_{\alpha\in\mathcal{A}_k}$; otherwise, the summation sign $\sum_\alpha$ will run over all indices $\alpha \in \mathcal{A} \equiv \bigcup_k \mathcal{A}_k$ by convention.

II. System Model

Following [14], the basic setup of our model is as follows: we consider a finite set $\mathcal{K} = \{1, \dots, K\}$ of wireless single-antenna transmitters (the *players* of the game) who wish to transmit to a group of single-antenna receivers (which can also be considered as a single receiver). Each of these receivers operates on a given channel in the finite set $\mathcal{A} = \{1, \ldots, A\}$ and these channels are assumed to be orthogonal (in the frequency domain typically). Each user may then transmit over a subset $\mathcal{A}_k \subseteq \mathcal{A}$ of these channels, and, to keep things reasonable, we will be assuming that $A_k \equiv \operatorname{card}(\mathcal{A}_k) \geq 2$ for all users $k \in \mathcal{K}$ (so users have at least *some* choice on how to transmit their data).

In particular, if $x_{k\alpha} \in \mathbb{C}$ is the user's transmitted message on channel $\alpha \in \mathcal{A}_k$ and $h_{k\alpha}$ denotes the respective channel coefficient (a zero-mean complex Gaussian variable), then the received signal on channel α will be:

$$y_\alpha = \sum\nolimits_k h_{k\alpha} x_{k\alpha} + z_\alpha, \tag{1}$$

where $z_\alpha \sim \mathcal{CN}(0, \sigma_\alpha^2)$ represents the thermal noise process on channel $\alpha \in \mathcal{A}$. Accordingly, user $k \in \mathcal{K}$ can split his transmitting power among the channels $\alpha \in \mathcal{A}_k$ subject to the constraint:

$$\sum\nolimits_\alpha^k p_{k\alpha} \leq P_k, \tag{2}$$

where $p_{k\alpha} = \mathbf{E}\left[|x_{k\alpha}|^2\right]$ represents the power with which user k transmits on channel α and P_k is the user's maximum power. As a result, the *power allocation* of the k-th user will be given by the point $p_k = \sum_\alpha^k p_{k\alpha} e_{k\alpha} \in \mathbb{R}^{\mathcal{A}_k}$, and, analogously, the *power profile* which collectively reflects all of the users' power allocations will be represented by $p = (p_1, \ldots, p_K) \in \prod_k \mathbb{R}^{\mathcal{A}_k} \cong \mathbb{R}^{\mathcal{Q}}$, $\mathcal{Q}$ being the *disjoint union* (categorical coproduct) $\mathcal{Q} \equiv \coprod_k \mathcal{A}_k = \{(k, \alpha) : \alpha \in \mathcal{A}_k\}$ (not to be confused with their regular union which amounts to the entire channel set $\mathcal{A} \equiv \bigcup_k \mathcal{A}_k$).

The performance metric that we will examine for the users of this wireless network is their achievable transmission rates. Obviously, the transmission rate for user k in channel $\alpha \in \mathcal{A}_k$ depends on the corresponding signal to interference-plus-noise ratio (SINR) which is given by:

$$\mathsf{sinr}_{k\alpha}(p) = \frac{g_{k\alpha} p_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell \neq k} g_{\ell\alpha} p_{\ell\alpha}}, \tag{3}$$

where $g_{k\alpha} = |h_{k\alpha}|^2 \geq 0$ denotes the channel gain coefficient of user k with respect to $\alpha \in \mathcal{A}_k$.

As mentioned in the introduction, we will be assuming that receivers employ the single-user decoding (SUD) scheme due to its lower complexity compared to successive interference cancellation (SIC) techniques. On that account, the users' achievable rates will depend only on their power allocation policies through their SINR, but the exact functional form of this dependence hinges on the time-variability of the channel gain coefficients $g_{k\alpha}$. We will thus consider two distinct cases with regards to this issue: *i)* one for *static* channels with fixed $g_{k\alpha}$, and *ii)* fast (or shadow) *fading* channels where the channel coefficients $g_{k\alpha}$ evolve over time as a stochastic process.

Remark. We should stress here that the channel gains $g_{k\alpha}$ are the only stochastic parameters in our model. So, unless explicitly mentioned otherwise, any probabilistic statement we make in this paper will refer to the probability law of the random variables $g_{k\alpha}$; similarly, $\mathbf{E}[\cdot]$ will denote expectation over the channel gains: $\mathbf{E} \equiv \mathbf{E}_g$.

A. Static Channels

If the communication channel remains static and the receiver employs the SUD scheme, the spectral efficiency of user k in the power profile p will be given by [6, 14]:

$$u_k(p) = \sum_{\alpha \in \mathcal{A}} b_\alpha \log\left(1 + \mathsf{sinr}_{k\alpha}(p)\right) = \sum_{\alpha \in \mathcal{A}} b_\alpha \log\left(1 + \frac{g_{k\alpha} p_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell \neq k} g_{\ell\alpha} p_{\ell\alpha}}\right), \tag{4}$$

where $b_\alpha = B_\alpha/B > 0$ is a normalized version of the bandwidth B_α of the channel $\alpha \in \mathcal{A}_k$, rescaled to unity by the total bandwidth factor $B = \sum_\alpha B_\alpha$. As for the channel gains $g_{k\alpha}$, we assume that they are drawn from a continuous probability distribution on $[0, \infty)$ at the outset of the game, and that they then remain fixed for the duration of the transmission – see also the relevant assumptions in [9, 14]. For notational convenience, we will also employ the convention that $g_{k\alpha} = 0$ for all channels $\alpha \notin \mathcal{A}_k$.

Now, as intuition would suggest (and as was shown rigorously in [14]), when the users' utility is based solely on their spectral efficiency (4), it is clearly to their best interest to transmit at the highest possible total power, i.e., satisfying (2) as an equality.[1] In this way, we obtain the normal form game $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{K}, \{\Delta_k\}, \{u_k\})$, where:

1) The set of *players* of $\mathfrak{G}$ corresponds to the transmitters $\mathcal{K} = \{1, \ldots, K\}$.
2) The *strategy space* of player k is the (scaled) simplex $\Delta_k \equiv P_k\,\Delta(\mathcal{A}_k) = \left\{p_k \in \mathbb{R}^{\mathcal{A}_k} : p_{k\alpha} \geq 0 \text{ and } \sum_\alpha p_{k\alpha} = P_k\right\}$; as is customary, we will denote the game's space of strategy profiles $p = (p_1, \ldots, p_K)$ by $\Delta \equiv \prod_k \Delta_k$. Up to a scaling factor, this amounts to interpreting a power allocation policy as a probability distribution, which is not a common approach; this will in fact lead us to the learning technique proposed further.
3) The players' *payoffs* (or *utilities*) are given by the spectral efficiencies $u_k : \Delta \to \mathbb{R}$ of (4).

Of course, the game $\mathfrak{G}$ defined in this way does not follow the original form of [26] because *a*) the players are not mixing probabilities over a finite set of possible actions, and *b*) even though the players' strategy spaces happen to be simplices, their payoffs are not multilinear.[2] On the other hand, since Δ is a convex polytope and the utilities u_k of the users are concave functions of their power allocations p_k, we immediately see that the game $\mathfrak{G}$ is *concave* in the sense of Rosen [16]. Moreover, it was shown in [14] that $\mathfrak{G}$ is actually an *exact potential* game,[3] i.e., that it admits a (global) potential function $\Phi : \Delta \to \mathbb{R}$ such that:

$$u_k(p_{-k}; p'_k) - u_k(p_{-k}; p_k) = \Phi(p_{-k}; p_k) - \Phi(p_{-k}; p'_k), \tag{5}$$

for all players $k \in \mathcal{K}$, and for all power allocations $p_k, p'_k \in \Delta_k$ of user k and $p_{-k} \in \Delta_{-k} \equiv \prod_{\ell \neq k} \Delta_\ell$ of k's opponents.[4] In fact, the authors of [14] provided the following explicit form for the potential function Φ:

$$\Phi(p) = -\sum\nolimits_\alpha b_\alpha \log\left(\sigma_\alpha^2 + \sum\nolimits_k g_{k\alpha} p_{k\alpha}\right), \tag{6}$$

which, incidentally, is (minus) the system-achievable sum-rate when implenting SIC techniques.

B. Stochastically Fluctuating Channels

When the channels can no longer assumed to be static, we need to explicitly model their variations. This will become relevant in two contexts. First, when the channel is assumed to be block-fading, i.e., constant over each

[1]There are scenarios where the power constraint (2) is not saturated, such as those with pricing or energy-efficient metrics [6], but we will not deal with this issue here (see Section V though).

[2]In other words, the interpretation of power allocation policies as mixed strategies does not extend to the game's payoff functions.

[3]In the finite player sense of Monderer and Shapley [18], and not in the continuous sense of [23].

[4]The change of signs in (5) from [18] is deliberate so as to conform with physics, where one looks at the *minima* of the potential function.

block of transmission but time-varying across different transmissions. Second, when the channel is assumed to be fast-fading, i.e., varying even within the same transmission block. Clearly, in both cases, we will need to take into account the changes of the other users' power allocation policies as well, an issue which we will address when discussing the stochastic approximation later.

In the fast-fading channel model, the channel gains $h_{k\alpha} \equiv h_{k\alpha}(t)$, $k \in \mathcal{K}$, $\alpha \in \mathcal{A}_k$, are i.i.d. zero-mean complex Gaussian random variables. [5] In that case, assuming that users saturate their power constraints for each channel use, the relevant choice of payoff functions is the ergodic transmission rates of [27]:

$$\overline{u}_k(p) = \sum_{\alpha\in\mathcal{A}} b_\alpha \, \mathbf{E}_g \left[\log\left(1 + \frac{g_{k\alpha} p_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell\neq k} g_{\ell\alpha} p_{\ell\alpha}}\right)\right], \tag{7}$$

with b_α and σ_α being as in (4). Therefore, in exactly the same way as above, we obtain the *ergodic game* $\overline{\mathfrak{G}} \equiv (\mathcal{K}, \{\Delta_k\}, \{\overline{u}_k\})$, which has the same strategic structure as its static counterpart $\mathfrak{G}$ but payoff functions given by the ergodic rates (7).

Similarly to the static case, it is easy to show that $\overline{\mathfrak{G}}$ admits the following exact potential:

$$\overline{\Phi}(p) = -\sum\nolimits_\alpha b_\alpha \, \mathbf{E}_g \left[\log\left(\sigma_\alpha^2 + \sum\nolimits_k g_{k\alpha} p_{k\alpha}\right)\right]. \tag{8}$$

However, the exact form of the ergodic potential $\overline{\Phi}$ will now depend on the law of the channel coefficients $g_{k\alpha}$. In Gaussian fast-fading channels (Rayleigh fading), we have $g_{k\alpha} = |h_{k\alpha}|^2$, where $h_{k\alpha} \sim \mathcal{CN}(0, \sqrt{\gamma_{k\alpha}})$, $\gamma_{k\alpha} \geq 0$, so the $g_{k\alpha}$ are χ^2-distributed. In that case, the integral calculations of [28, eq. (11)] readily yield the following explicit form for the potential $\overline{\Phi}$ (and, by taking the appropriate differences, for the ergodic payoffs (7) as well):

Proposition 1. *In Gaussian fast-fading channels with $h_{k\alpha} \sim \mathcal{CN}(0, \sqrt{\gamma_{k\alpha}})$, the ergodic potential $\overline{\Phi}$ is given by:*

$$\overline{\Phi}(p) = -\sum\nolimits_{\alpha\in\mathcal{A}} b_\alpha \sum\nolimits_{k\in\mathcal{K}} \zeta(r_{k\alpha}^{-1}) \prod\nolimits_{\ell\neq k} \frac{r_{k\alpha}}{r_{k\alpha} - r_{\ell\alpha}} - \sum\nolimits_\alpha b_\alpha \log(\sigma_\alpha^2), \tag{9}$$

where $r_{k\alpha} = \gamma_{k\alpha} p_{k\alpha}/\sigma_\alpha^2$ and $\zeta(x) \equiv \int_0^\infty (x+t)^{-1} e^{-t}\,dt = -e^x \operatorname{Ei}(-x)$, with Ei *denoting the exponential integral function* $\operatorname{Ei}(x) = -\int_{-x}^\infty t^{-1} e^{-t}\,dt$.

This explicit expression for the potential function $\overline{\Phi}$ will be very important in the analysis of Section IV-B where it allows us to calculate the game's equilibrium (the minimum of $\overline{\Phi}$ and the convergence rate (23) of the replicator dynamics. Furthermore, we note for posterity that the authors of [19] showed that $\overline{\Phi}$ is strictly convex, so its minimum set consists of a single point. On the other hand, the static potential Φ is convex, but not necessarily *strictly* so:[6] indeed, any two power profiles $p, p' \in \Delta$ such that $\sum_k g_{k\alpha} p_{k\alpha} = \sum_k g_{k\alpha} p'_{k\alpha}$ for all $\alpha \in \mathcal{A}$ will also have $\Phi(p) = \Phi(p')$. This simple observation will be of crucial importance in determining the Nash set of the game, which will be the subject of the next section.

[5] Of course, in addition to this model, other temporal variation may be considered (such as slow fading due to shadowing), but for the sake of simplicity we will focus on the fast-fading situation.

[6] This is precisely the subtle (but crucial) mistake that underlies the equilibrium uniqueness argumentation of [14].

III. Equilibrium Analysis

Since we have a finite number of players (the transmitters $k \in \mathcal{K}$), the notion of Nash equilibrium takes the form of stability in the face of unilateral deviations. More specifically:

Definition 2. We will say that the power profile $q \in \Delta$ is at *Nash equilibrium* in the game $\mathfrak{G}$ (resp. $\overline{\mathfrak{G}}$) when

$$u_k(q) \geq u_k(q_{-k}; q'_k), \quad (\text{resp. } \overline{u}_k(q) \geq \overline{u}_k(q_{-k}; q'_k)) \tag{10}$$

for all $k \in \mathcal{K}$ and for every deviation $q'_k \in \Delta_k$ of player k. In particular, if q satisfies the strict version of the inequalities (10), then it will be called a *strict equilibrium*.

Note that a (pure) Nash equilibrium necessarily exists in $\mathfrak{G}$ and $\overline{\mathfrak{G}}$ since they are both concave games in the sense of Rosen [16]. The main issue here is therefore to determine whether it is unique or not and, for this purpose, we will exploit the games' potential property.

A. Static Channels

As is standard in convex potential games [29], to calculate the set of Nash equilibria Δ^* of the game, we only need to look at the (necessarily convex) minimum set of its potential function. Obviously, if the potential is strictly convex, then Δ^* will consist of a single equilibrium point, but strict convexity typically fails for the static potential Φ, so we will pause here to introduce the concept of *degeneracy*. To that end, let $T_p\Delta$ denote the tangent space of Δ at p. Since Δ is an affine polytope embedded in $\mathbb{R}^{\Omega} \cong \prod_k \mathbb{R}^{\mathcal{A}_k}$, it is easy to see that for every interior point $p \in \text{Int}(\Delta)$, $T_p\Delta$ will be isomorphic to the $(Q-K)$-dimensional subspace which is "parallel" to the polytope Δ:[7]

$$T_p\Delta \cong Z = \left\{z \in \mathbb{R}^{\Omega} : \textstyle\sum_{\alpha}^{k} z_{k\alpha} = 0 \text{ for all } k \in \mathcal{K}\right\}. \tag{11}$$

However, as we just noted, some of these $Q-K$ directions will be *degenerate*, in the sense that the potential Φ remains constant as we move along them. Specifically, the set of (almost surely independent) constraints

$$\textstyle\sum_k g_{k\alpha} z_{k\alpha} = 0, \quad \alpha \in \mathcal{A}, \tag{12}$$

cuts itself a subspace W of R^{Ω} whose intersection with Z corresponds to the total of $A+K$ constraints:

$$a) \quad \sum\nolimits_{\alpha} z_{k\alpha} = 0, \qquad k \in \mathcal{K}; \tag{13a}$$

$$b) \quad \sum\nolimits_{k} g_{k\alpha} z_{k\alpha} = 0, \qquad \alpha \in \mathcal{A}. \tag{13b}$$

Therefore, if $Q = \sum_k A_k \leq A+K$, the system of (almost surely independent) tangency and degeneracy constraints (equations (13a) and (13b) respectively) will only admit the trivial solution $z=0$;[8] otherwise, there will be tangent directions $z \in Z \setminus \{0\}$ along which the potential Φ stays constant. To keep track of all this, we have:

[7] Recall here our notational conventions: since $\Omega \equiv \coprod_k \mathcal{A}_k$, then $Q \equiv \text{card}(\Omega) = \sum_k A_k$ (where $A_k = \text{card}(\mathcal{A}_k)$).

[8] These conditions are remarkably similar to the MIMO rank condition $\text{rank}(\mathbf{H}^\dagger \mathbf{H}) = \sum_{k=1}^K n_{t,k} \leq n_r + K$ of [15]. Alternatively, the irreducibility condition observed in [30] for wireline networks is actually the *inverse* one.

Definition 3. The subspace $W \leq \mathbb{R}^{\mathcal{Q}}$ defined by the constraints (13b) will be called the space of *degenerate* (or *redundant*) directions of the game $\mathfrak{G}$. Moreover, we define the *degeneracy index* of $\mathfrak{G}$ to be:

$$\operatorname{ind}(\mathfrak{G}) \equiv \dim(W \cap Z), \tag{14}$$

where Z is the tangent space determined by the tangency constraints (13a).

An easy way to visualize this concept is to represent the game $\mathfrak{G}$ as a bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ whose vertex set $\mathcal{V}$ consists of the transmitters $k \in \mathcal{K}$ plus the receivers $\alpha \in \mathcal{A}$, and a receiver $k \in \mathcal{K}$ is linked to a receiver $\alpha \in \mathcal{A}$ by an edge in $\mathcal{E}$ if and only if $\alpha \in \mathcal{A}_k$ (that is, $\mathcal{V} = \mathcal{K} + \mathcal{A}$ and $\mathcal{E} = \mathcal{Q}$). Our previous discussion then yields:

Proposition 4. *With probability* 1*, the game* $\mathfrak{G}$ *is degenerate if and only if the associated bipartite graph* $\mathcal{G} = (\mathcal{K} + \mathcal{A}, \mathcal{Q})$ *has more edges than vertices (i.e., there are more links than wireless devices). In that case:*

$$\operatorname{ind}(\mathfrak{G}) = \operatorname{card}(\mathcal{E}) - \operatorname{card}(\mathcal{V}) = Q - A - K = -\chi(\mathcal{G}), \tag{15}$$

where $\chi(\mathcal{G})$ *is the Euler characteristic of the graph* $\mathcal{G}$ *viewed as a* 1*-dimensional simplicial complex.*

We thus see that typical uplink scenarios where several channels are shared by a relatively large number of wireless users will exhibit degeneracy (a.s.) because the number of links exceeds the number of receivers plus transmitters. On account of this difficulty, a promising way to determine whether the MAC game admits a *unique* equilibrium is to take advantage of the plethora of sufficient conditions that have been established in the literature for this purpose [9, 10, 31]. These conditions invariably depend on the spectral radius and positive-definiteness of certain matrices whose entries are determined by the channel gain coefficient ratios $S_{k\ell}(\alpha) \equiv g_{k\alpha}/g_{\ell\alpha}$ (for instance, one such condition is that the spectral radius $\rho\,(\mathbf{S}(\alpha))$ be less than 1 for all channels $\alpha \in \mathcal{A}$). Unfortunately however, conditions of this sort typically fail in the parallel MAC s etting [32] because the spectral radii of these matrices are greater than or equal to 1, so the water-filling operator of [9] is not necessarily a contraction.

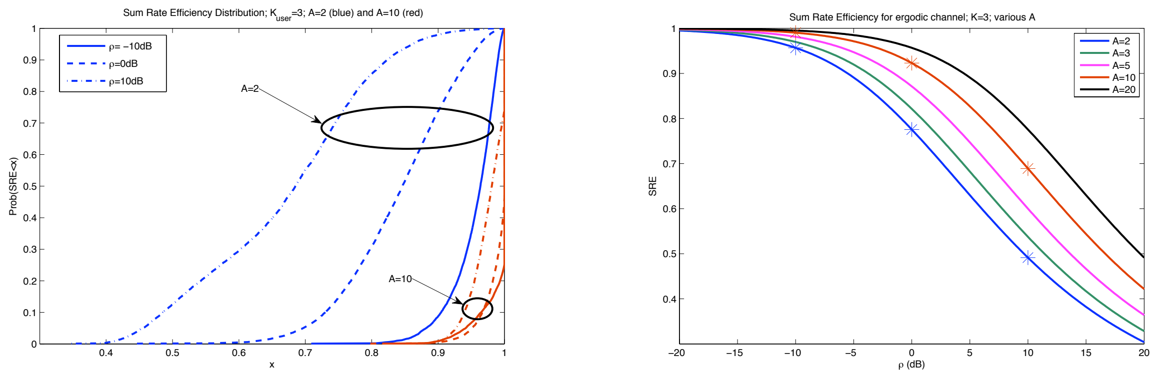
Hence, despite their intrinsic theoretical value, the sufficient conditions that have been established in the literature are quite problematic in the PMAC framework because they are met only in some very special cases (which can be attacked directly anyway). Consequently, in order to address the issue of uniqueness, we will adopt the approach of [32], where power profiles are represented by suitably constructed graphs:

Definition 5. Given a power profile $p = (p_1, \ldots, p_K) \in \Delta$, construct a graph $\mathcal{G}_p = (\mathcal{V}_p, \mathcal{E}_p)$ as follows:

1) The vertices $\mathcal{V}_p$ coincide with the network's receivers: $\mathcal{V}_p = \mathcal{A}$.
2) For every $k \in \mathcal{K}$, pick a receiver $\alpha \in \mathcal{A}_k$ to which k assigns positive power $p_{k\alpha} > 0$ and join α (called the *hub* of user k in $\mathcal{G}_p$) to every other receiver $\beta \in \operatorname{supp}(p_k) \setminus \{\alpha\}$ by an edge.

Every graph $\mathcal{G}_p$ constructed in this way will be said to *represent* the power profile p.

Bearing this construction in mind, the analysis of [32] shows that any graph $\mathcal{G}_q$ which represents an equilibrial



(a) Distribution of the sum-rate efficiency for the static game.

(b) The sum-rate efficiency for the ergodic game as a function of SNR.

Fig. 1. In Fig. 1(a) we analyze the sum-rate efficiency (SRE) for static channels. In particular, we plot the CDF of the SRE for static channels h_{ka} drawn out of a i.i.d. complex Gaussian distribution with variance ρ for all channels a and users k. We see that for $A = 10 > K$ the SRE attains the maximum value 1 with a finite probability which decreases with ρ. This can be attributed to the high probability to have a Nash equilibrium at a vertex of Δ in which case the sum rate and the total achievable rate are equal. In Fig. 1(b) we plot the SRE for ergodic channels as a function of SNR. For convenience, we represent with stars the values of ρ for which we have calculated the SRE distribution in Fig. 1(a).

power profile $q \in \Delta^*$ will have to be acyclic (a.s.).[9] We can then use this acyclicity property to conclude that any equilibrial profile must lie in the interior of an at most $(A-1)$-dimensional face of Δ (a.s.), and a dimension-counting argument finally yields:

Theorem 6. *The static channel game $\mathfrak{G}$ has a unique Nash equilibrium (almost surely).*

B. Stochastically Fluctuating Channels

On the other hand, by exploiting the averaging effect which is present in ergodic rates, it can be shown that the ergodic potential $\overline{\Phi}$ *is* strictly convex [19]. This directly leads us to the following result:

Theorem 7. *The ergodic game $\overline{\mathfrak{G}}$ admits a unique Nash equilibrium.*

C. Analysis of Sum-Rate Efficiency

It is instructive here to discuss the efficiency of the game's (unique) equilibrium, in both the static and ergodic regime. Clearly, the globally optimal solution of the problem is achieved when the sum rate is maximized, but this point is not necessarily reached. We may therefore analyze the efficiency of the Nash equilibrium by calculating the so-called sum-rate efficiency (SRE), a quantity which is closely related to the game's price of anarchy:

$$\mathrm{SRE} = \frac{\sum_k u_k(q)}{C_{sum}}, \quad (\text{resp. } \overline{\mathrm{SRE}} = \frac{\sum_k \overline{u}_k(q)}{\overline{C}_{sum}} \text{ for the ergodic game}) \tag{16}$$

[9] Compared to [32], users in our case have their own distinct channel sets $\mathcal{A}_k$. This complicates the logistics somewhat, but the essence of the arguments of [32] remains unaltered.

where the numerator is the sum of the achievable rates of all users at equilibrium, while the denominator $C_{sum}, \overline{C}_{sum}$, represents the maximum achievable aggregate sum-rate. Interestingly, both quantities are maximized with the same power allocation, because the potential function is equal to (minus) the achievable sum-rate of the network.

In Fig. 1(a) we plot the SRE for randomly drawn static channels. We see that while it can deviate significantly from its maximum value (unity) for $A < K$, in the opposite case, the SRE is typically close to its maximum value – and, in fact, equal to it with positive probability. We may attribute this to the fact that for $K < A$ there is a finite (and relatively large) probability that the systems equilibrium is at a vertex of Δ: this implies that each user is transmitting on a single channel, hence explaining why the system reaches sum-rate optimality. What is remarkable is the fact that using a sub-optimal decoding scheme such as SUD in a distributed scenario we can reach the centralized system's sum-capacity when the NE is on a vertex. Numerical simulations show that this occurs with probability slightly higher than $1/2$; in that case, more complicated SIC techniques yield no performance benefits over the simpler SUD scheme.

On the other hand, in Fig. 1(b), we plot the SRE for the case of the ergodic channel. There, we see that while the SRE is nearly optimal for small SNR, it deviates strongly from its maximum value for larger SNR values.

IV. Convergence to Equilibrium

Even though Theorems 6 and 7 guarantee that there is a unique power allocation which is stable under selfish unilateral deviations, it is not at all clear whether users will be able to calculate it in decentralized environments where only partial/local information is available at the terminal (e.g., as in distributed or partially distributed cognitive radio networks). Consequently, our goal in this section will be to present a simple distributed learning scheme which allows users to converge to the unique equilibrium of the static game $\mathfrak{G}$, and also to determine the speed of this convergence.

From the point of view of learning in games [21], this question has attracted considerable interest and two of the most well-studied paradigms are best-response (BR) algorithms and reinforcement learning (RL). On the one hand, BR dynamics comprise a class of updating rules where players are assumed to monitor their opponents' power allocation policies and choose the "best response" (with respect to their individual payoffs) in the next round of updating. Unfortunately, in addition to this perfect monitoring requirement, BR schemes also require users to be able to calculate their best responses, so their implementation in large decentralized networks poses a formidable challenge. A partial solution to this problem for static channels is provided by the well-known iterative water-filling algorithm and its asynchronous/sequential variants [7, 9, 10, 17] where the perfect monitoring condition is relaxed and users only need to know their local channel and overall noise-plus-interference covariance matrix. In that case however *a*) users cannot update their powers simultaneously; *b*) they must still solve a nonlinear fixed point problem at each update period; and *c*) equilibrium convergence is conditional on the interference level being small enough (except in the case of [17]). In fact, the conditions of [9, 10] do not hold at all in the PMAC case [30], while the interference threshold of [7] below which users converge to equilibrium becomes zero in the large K regime where one would encounter decentralized heterogeneous networks [33].

On the other hand, RL algorithms (such as fictitious play and its various incarnations [21]) rely on the players knowing the payoffs that they receive – or would have received by playing a different strategy. Thanks to this "fictitious" information (which is often hard to come by), these algorithms enjoy some very strong convergence properties in potential games: for instance, stochastic fictitious play converges to equilibrium in finite potential games [34]. However, such learning algorithms have been designed for discrete action sets, so it is very hard to adapt them to games with continuous action spaces such as the ones we are studying here.

On account of these limitations, we will adopt instead a hybrid approach: by assuming that users have similar information as in water-filling algorithms (which is easy to obtain), we will present a simple learning scheme based on the replicator dynamics of evolutionary game theory [20–22] where players update the frequency with which they employ an action proportionally to the difference between its payoff and their average one. This scheme's reinforcement aspect is intimately related to logistic fictitious play [21, 35] and suffers from the same drawback, namely that it typically applies only to *finite* action sets. However, by exploiting the simplicial structure of our model and its potential properties, we derive a replicator equation which works for continuous action spaces and which allows users to converge to equilibrium *unconditionally* and at an *exponential rate* (Theorems 8 and 10).

A. Static Channels

Since the replicator equation is an algorithm for discrete action spaces, one would hope that the discrete payoff differences that go in the dynamics could be captured by the channel-specific rates:

$$u_{k\alpha}(p) = b_\alpha \log\left(1 + \mathsf{sinr}_{k\alpha}(p)\right) = b_\alpha \log\left(1 + \frac{g_{k\alpha} p_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell\neq k} g_{\ell\alpha} p_{\ell\alpha}}\right), \tag{17}$$

which lead to the replicator equation:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t)\left(u_{k\alpha}(p(t)) - P_k^{-1} \sum\nolimits_\beta^k p_{k\beta}(t) u_{k\beta}(p(t))\right), \tag{18}$$

where the averaging in the second term of (18) is mandated by the fact that $p(t)$ must remain in Δ for all $t \geq 0$.

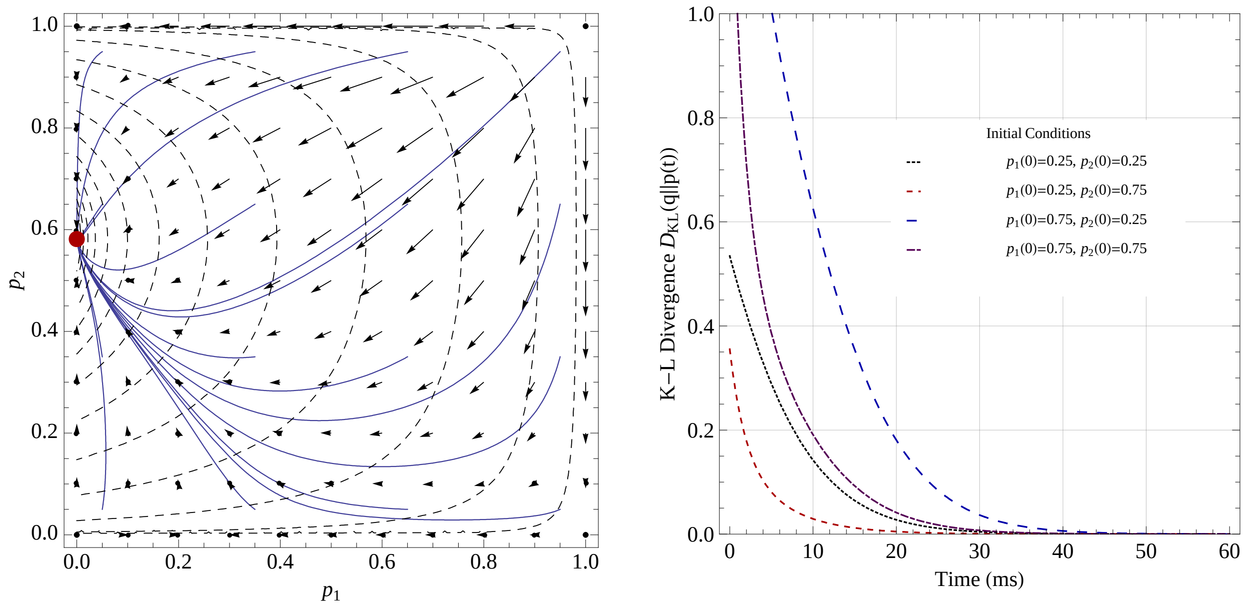
Unfortunately, despite its apparent similarity to the replicator dynamics of evolutionary game theory [22], the total rate u_k of (4) cannot be expressed as a linear combination of the channel-specific rates $u_{k\alpha}$, so (18) is not particularly well-behaved with respect to the game $\mathfrak{G}$ – for instance, Nash equilibria are not stationary. Thus, given that players invariably want to increase their rewards, we will instead consider the directional derivatives:

$$v_{k\alpha}(p) \equiv \frac{\partial u_k}{\partial p_{k\alpha}} = \frac{b_\alpha g_{k\alpha}}{\sigma_\alpha^2 + \sum_\ell g_{\ell\alpha} p_{\ell\alpha}}. \tag{19}$$

Clearly, if user k transmits with positive power to channel $\alpha \in \mathcal{A}_k$, he will be able to calculate the value of the corresponding gradient $v_{k\alpha}(p)$ simply by knowing $g_{k\alpha}$ and his SINR in channel α. On that account, any learning scheme based on the $v_{k\alpha}$'s will be inherently distributed and local in nature (and simpler than solving a water-filling problem to boot), so we will consider the replicator dynamics corresponding to the *marginal payoffs* $v_{k\alpha}$:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t)(v_{k\alpha}(p(t)) - v_k(p(t))), \tag{20}$$

with v_k denoting the user average $v_k(p) = P_k^{-1} \sum_\beta^k p_{k\beta} v_{k\beta}(p)$.



(a) A typical phase portrait exhibiting convergence to equilibrium

(b) The K-L divergence w.r.t. equilibrium as a function of time

Fig. 2. Convergence to the (unique) Nash equilibrium for a game with 2 users sharing 2 static channels A and B. The contours in Fig. 2(b) represent the level sets of the K-L divergence with respect to the game's equilibrium (red dot), and the variables p_1 and p_2 represent the fraction of the power allocated to channel A by each user (gain and noise parameters were chosen arbitrarily). In Fig. 2(b) we plot the evolution of the K-L relative entropy w.r.t. the game's equilibrium over time (with sampling period $\delta = 1$ ms).

It is worth comparing (20) to power updating schemes that have appeared in the past. Clearly, the schemes discussed in [7, 9, 10, 17] are quite different due to their waterfilling character. Closer in spirit are the power control algorithms developed in the 90's aiming to minimize transmitted power [36–38]. In these cases, the instantaneous SINR was compared to a target value and the power was iteratively updated proportionally to the difference. Nevertheless, the differences with those algorithms are quite significant, not only in terms of the actual updating, but also with respect to their convergence properties.

Remark 1. It is also important to note here that the marginal payoffs $v_{k\alpha}$ are similar but *not* equal to the SINR (3) of user k at a given channel $\alpha \in \mathcal{A}_k$ since *a*) the numerator of (19) includes the bandwidth fraction $b_\alpha = B_\alpha/B$, and *b*) the denominator of (19) gives the *total* power of the received signal on a given channel, including the user's own signal $g_{k\alpha}p_{k\alpha}$. The $v_{k\alpha}$ are also related but do not coincide with the popular metric of SINR per unit power of [3, 8]. All these metrics can be calculated based on the same feedback and, at a first glance, they might appear more natural than using $v_{k\alpha}$; be that as it may, we shall see that our (perhaps unconventional) choice leads to very strong convergence properties for (20).

A strong indication that (20) is a step in the right direction is that the rest points of (20) are characterized by the (waterfilling) property that for every pair of nodes $\alpha, \beta \in \operatorname{supp}(p)$ to which user k allocates positive power, we will also have $v_{k\alpha}(p) = v_{k\beta}(p)$. Hence, comparing this to the KKT constrained minimization conditions obtained in

[14], we immediately see that the Nash equilibria of $\mathfrak{G}$ are stationary in the replicator equation (20). This result is well-known in finite games with multilinear payoffs [21] and in continuous population games [23], but the converse does not hold: for instance, every vertex of Δ is stationary in (20) without being a Nash equilibrium., so stationarity in (20) does not imply (Nash) equilibrium.

Nevertheless, the game's (unique) Nash equilibrium is the *only* attracting state of the dynamics:

Theorem 8. *Let $q \in \Delta$ be the (a.s.) unique equilibrium of $\mathfrak{G}$. Then, every solution orbit of the replicator dynamics (20) which begins at finite Kullback-Leibler divergence with respect to q will converge to it. Moreover, there exists a positive constant $c > 0$ such that:*

$$D_{KL}\left(q \,\|\, (p(t)\right) \leq h_0\, e^{-ct} \text{ for all } t \geq 0. \tag{21}$$

where D_{KL} denotes the Kullback-Leibler divergence and $h_0 \equiv D_{\mathrm{KL}}\left(q \,\|\, p(0)\right)$.

In other words, replicator solution orbits converge to an ε-neighborhood of a Nash equilibrium in time which is at most of order $\mathcal{O}(\log(1/\varepsilon))$.

Remark 1. Recall that the *Kullback-Leibler divergence* (or *relative entropy*) is [22]:

$$D_{\mathrm{KL}}(q \,\|\, p) = \sum\nolimits_k D_{\mathrm{KL}}(q_k \,\|\, p_k) = \sum\nolimits_k \sum\nolimits_\alpha^k q_{k\alpha} \log\left(q_{k\alpha}/p_{k\alpha}\right). \tag{22}$$

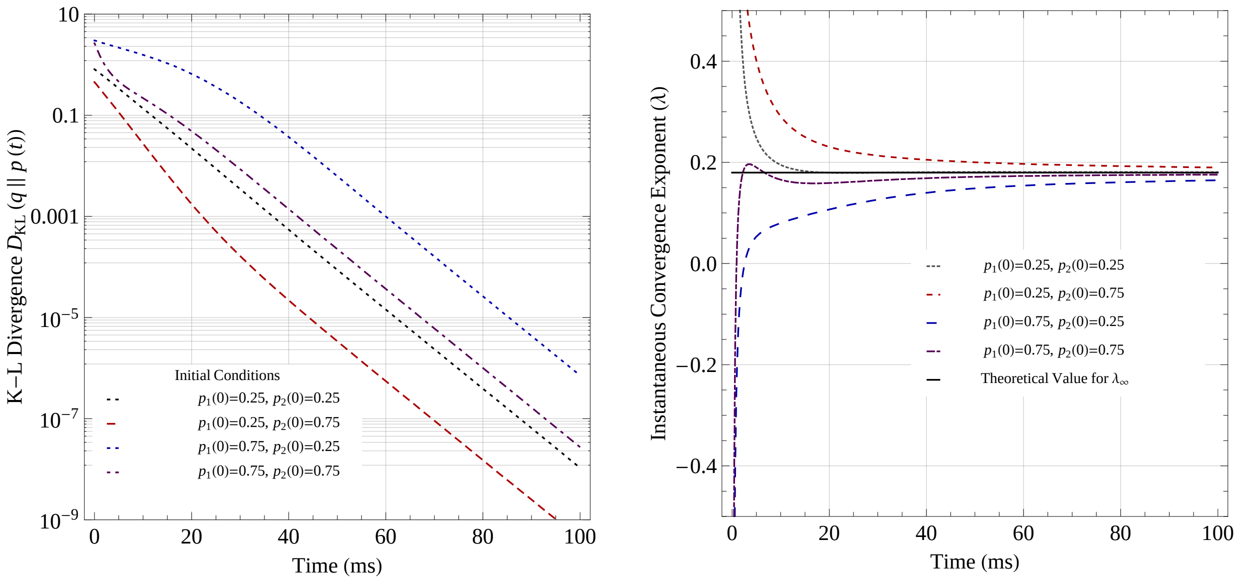
Clearly, $D_{\mathrm{KL}}(q_k \,\|\, p_k)$ is finite if and only if p_k allocates positive power $p_{k\alpha} > 0$ to all channels $\alpha \in \mathrm{supp}(q)$ which are present in q_k; more succinctly, the domain of the relative entropy function $H_q \equiv D_{\mathrm{KL}}(q \,\|\, \cdot)$ consists of all power allocations which are absolutely continuous w.r.t. q. Thus, if users start with a uniform power allocation (a natural choice if they do not know anything about the network), they are guaranteed to converge to equilibrium.

Remark 2. From an evolutionary perspective, the analysis of [23] does allow for nonlinearities in the phenotypes' utility functions, but the species' average utility (measured as the number of offspring in the unit of time) is still a convex combination of these phenotype-specific payoffs. The resulting replicator equation is formally identical to (20), and, by exploiting the game's potential, Sandholm shows that any ω-limit point of a replicator trajectory must be Nash. Theorem 8 (see also Theorem 13) extends these results by showing that (almost) every replicator orbit converges to Nash equilibrium and by showing that the speed of this convergence is exponential.

Remark 3. The reader will note that we have not specified the norm under which "ε-neigborhoods" are taken in Theorem 8. This is intentional: since all norms in finite-dimensional vector spaces are equivalent, the order of the convergence time will remain the same in all norms. This is also why we chose to state our results in terms of the K-L divergence D_{KL} (instead of a more conventional choice like the L^2 norm): although D_{KL} is not a metric in itself, it has all the relevant attributes and, in many ways, it is the most natural distance measure on Δ.

Remark 4. Obviously, the exponent c of (21) is of crucial importance because it bounds the convergence speed of the replicator dynamics with respect to the K-L divergence. More specifically, if we set:

$$\lambda(t) \equiv -\frac{1}{t} \log \frac{D_{\mathrm{KL}}(q \,\|\, p(t))}{D_{\mathrm{KL}}(q \,\|\, p(0))}, \tag{23}$$



(a) Exponential decay of the K-L divergence from equilibrium

(b) The asymptotic convergence exponent of the replicator dynamics

Fig. 3. The speed of convergence to the game's equilibrium in a game with 2 users and 2 static channels (notation and parameters as in Fig. 2). In Fig. 3(a) we see that the K-L divergence of the users' power allocation with respect to the game's equilibrium decays exponentially quickly, while, in Fig. 3(b), we see that the instantaneous convergence exponent $\lambda(t) = -t^{-1} \log D_{KL}(q \,\|\, p(t))$ is asymptotically equal to the predicted value λ_∞ of Theorem 8.

then Theorem 8 simply states that the "instantaneous" convergence exponent $\lambda(t)$ is bounded below by c.

Consequently, it is pretty important to be able to assess the value of c that goes into (21). The full expression can be found in Appendix A, but one case is important enough to mention here. Indeed, if q is a strict equilibrium, say $q = \sum_k P_k e_{k,\alpha_k}$, then we have:

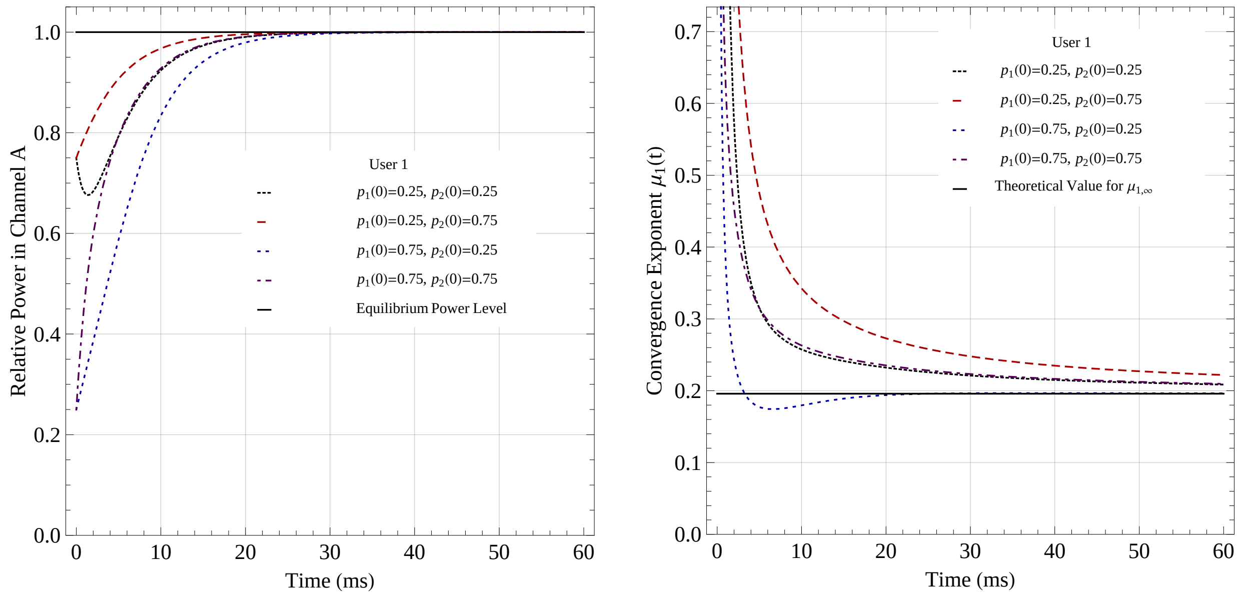
$$c = \min_k \left\{ P_k/h_0 \left(1 - e^{-h_0/P_k}\right) \Delta v_k \right\}, \tag{24}$$

where $\Delta v_k = \min\{v_{k,\alpha_k}(q) - v_{k\beta}(q) : \beta \in \mathcal{A}_k \setminus \{\alpha_k\}\}$. We thus see that the exponent of (21) exhibits a very natural behavior: it increases with the payoff differences Δv_k (which measure the "attracting strength" of q), and decreases with the initial relative entropy h_0 (which describes how far the initial condition $p(0)$ is from q).

From this expression, and given that $\lim_{\xi\to 0}(1 - e^{-\xi})/\xi = 1$, we see that the asymptotic convergence exponent $\lambda_\infty = \liminf_{t\to\infty} \lambda(t)$ is bounded from below (and away from zero) by the quantity:

$$m = \min_k \Delta v_k = \min_k \min_{\beta\neq\alpha_k} \left\{ v_{k,\alpha_k}(q) - v_{k\beta}(q) \right\} > 0. \tag{25}$$

In fact, a closer inspection of the calculations of Appendix A reveals that this expression is *sharp*, in the sense that $\lambda_\infty = m$ (Fig. 3). We thus see that (21) and (24) not only bound the convergence speed of the replicator dynamics



(a) A user's power allocation converges to its Nash value

(b) Plot of the instantaneous convergence exponent $\mu_1(t)$

Fig. 4. Numerical analysis of a PMAC game with 2 users sharing 2 static channels A and B, and admitting a strict equilibrium (our notation is as in Figs. 2 and 3, and, to reduce clutter, we focus only on user 1). In Fig. 4(a) we see that the fraction of power that user 1 employs on channel A converges to its Nash value very quickly; furthermore, in Fig. 4(b), we see that in the limit $t \to \infty$, the instantaneous convergence exponent $\mu_1(t) \equiv -t^{-1} \log(1 - p_1(t))$ approaches the value predicted by Proposition 9.

(20), but they also provide an estimate which becomes exact in the local limit $h_0 \to 0$:[10]

Proposition 9. *If $q = \sum_k P_k e_{k,\alpha_k}$ is a strict equilibrium of $\mathfrak{G}$ and $\Delta v_k = \min_{\beta \neq \alpha_k}\{v_{k,\alpha_k}(q) - v_{k\beta}(q)\}$, then, as $t \to \infty$:*

$$p_{k,\alpha_k}(t) \sim P_k\left(1 - e^{-\Delta v_k t}\right). \tag{26}$$

Similar results can also be obtained for non-strict equilibria where at least one user is water-filling his power between the receivers (Fig. 3(b)), but the relevant expressions are a bit more cumbersome, so we prefer to omit them – see Appendix A for details.

Remark 5. The authors of [39] show that iterated water-filling has an asymptotic exponential rate of convergence as a consequence of the water-filling operator being a contraction. However, the sufficient conditions which guarantee the contraction property fail in the PMAC case [32], so there is virtually no overlap of our results.

Remark 6 (The Discrete-time Replicator Dynamics). Needless to say, to properly implement the differential equation

[10]Since Theorem 8 shows that replicator trajectories converge to equilibrium, we could also obtain the *asymptotic* convergence exponent λ_∞ by linearizing (20) near q and looking at the eigenvalues of the linearized system. Since linearization corresponds to taking the limit $h_0 \to 0$, we only mention here (without proof) that these eigenvalues are precisely the payoff differences $v_{k\beta}(q) - v_{k,\alpha_k}(q)$, $\beta \in \mathcal{A}_k \setminus \{\alpha_k\}$ that go into (25).

(20) as a learning algorithm, one should instead consider the discrete time variant:

$$p_{k\alpha}(n+1) = p_{k\alpha}(n) + \delta\, p_{k\alpha}(n)\left(v_{k\alpha}(p(n)) - v_k(p(n))\right), \tag{27}$$

where δ is the discrete time step. Clearly, the continuous dynamics (20) are recovered in the limit $\delta \to 0$, so the only challenge in this discretization is to show that $p(t)$ remains in Δ for all t. To that end, summing (27) over all $\alpha \in \mathcal{A}_k$ immediately gives $\sum_{\alpha}^{k} p_{k\alpha}(n) = P_k$ (assuming that this holds for $t = 0$), while, if we set $v = \max_{k,\alpha} \max_{p\in\Delta} |v_{k\alpha}(p)|$ and choose $\delta \leq (2v)^{-1}$, an easy induction argument also shows that $p_{k\alpha}(n) \geq 0$ (again, assuming that this holds for all $k \in \mathcal{K}$, $\alpha \in \mathcal{A}_k$ when $n = 0$).

B. *Stochastically Fluctuating Channels*

The key assumption of the results of the previous section was that the channel gain coefficients $g_{k\alpha}$ remain static for the whole duration of the transmission. Instead, if the channel gains evolve over time following a block-fading model, then the replicator equation (20) becomes a non-deterministic dynamical system because the coefficients $g_{k\alpha}$ that go into the gradients $v_{k\alpha}$ are now continuous stochastic processes that evolve over time.

In general, there are two directions one can take in order to incorporate the effects of stochastic fluctuations in a dynamical system. The first involves the derivation of a stochastic differential equation (SDE) for the dynamic variables (here the users' power profiles; see e.g., [32, 35] for related results). This approach roughly corresponds to assuming that the fluctuations evolve at a much quicker pace than the updating process, or, in our case, that over an (infinitesimal) update period of length δ, the channel fluctuations have undergone a random walk of size $\sqrt{\delta} \gg \delta$. In contrast, when the fluctuations occur at a comparable rate as the power updating, their variations over the same interval are of order δ, so the proper way to proceed here is to employ the theory of stochastic approximation [24, 40].

To that end, we will begin by rewriting the discrete form (27) of the replicator dynamics as:

$$p_{k\alpha}(n+1) = p_{k\alpha}(n) + \delta(n)\, p_{k\alpha}(n)\left[v_{k\alpha}(p(n), g(n)) - v_k(p(n), g(n))\right], \tag{28}$$

where $g(n) = \{g_{k\alpha}(n)\}$ are the channel gain coefficients at the n-th iteration of the replicator algorithm, $\delta(n)$ is a time-dependent *learning parameter* to be discussed later, and the $v_{k\alpha}$'s are defined as in (19) – note that they now depend on time via the channel coefficients g. In particular, our assumptions for the channel gain coefficients $g_{k\alpha} = |h_{k\alpha}|^2$ is that the $h_{k\alpha}$'s are complex Gaussian i.i.d. variables.

Then, if we set $\bar{v}_{k\alpha} = \mathbf{E}[v_{k\alpha}]$ and $\eta_{k\alpha} = v_{k\alpha} - \bar{v}_{k\alpha}$ (where $\mathbf{E}[\cdot]$ denotes expectation with respect to the law of the channel gains g), we immediately obtain:

$$p_{k\alpha}(n+1) = p_{k\alpha}(n) + \delta(n)\, p_{k\alpha}(n)\left[\left(\bar{v}_{k\alpha}(p(n)) - \bar{v}_k(p(n))\right) + \left(\eta_{k\alpha}(p(n), g(n)) - \eta_k(p(n), g(n))\right)\right], \tag{29}$$

where the $\bar{v}_{k\alpha}$ are now deterministic functions and the $\eta_{k\alpha}$ are zero-mean stochastic processes. The first thing to note here is that, assuming the parameters $\delta(n)$ are chosen small enough, this discrete-time stochastic dynamical system respects the simplicial structure of Δ (that is, $p(n) \in \Delta$ for all $n \geq 0$). To see this, we may restrict ourselves without loss of generality to a game with one user and two choices, A and B (the general argument is similar). In that case, if we drop the user index for simplicity and let $p_A(n) \equiv p_{1,A}(n)$ denote the power that the user sends to channel A at the n-th iteration of the dynamics, we need to show that we can choose $\delta(n)$ small enough so that $0 \leq p(n) \leq 1$ for all $n \geq 0$ and $g_{A,B} \geq 0$ drawn *randomly* at each step.[11] So, if we assume inductively that this is true for some $n \geq 0$, a little algebra yields:

$$p_A(n+1) - p_A(n) = \delta(n)\, p_A(n) \left[\frac{g_A(n)(1-p_A(n))}{\sigma_A^2 + g_A(n) p_A(n)} - \frac{g_B(n)(1-p_A(n))}{\sigma_B^2 + g_B(n)(1-p_A(n))} \right], \tag{30}$$

and by the symmetry of this last equation under the transformation $p \rightsquigarrow 1-p$, it suffices to show that $p(n+1) \geq 0$. However, since the first term of the LHS of (30) is positive and the second is *uniformly* bounded (by 0 from below and 1 from above), we see that choosing $\delta(n) \leq$ yields $p_A(n+1) \geq 0$, while the complementary inequality $p_A(n+1) \leq 1$ follows in exactly the same fashion.

Now, going back to (29), a trivial application of the dominated convergence theorem allows us to interchange differentiation and integration, so we obtain:

$$\overline{v}_{k\alpha}(p) = \mathbf{E}\left[\frac{\partial u_k}{\partial p_{k\alpha}}\right] = \frac{\partial}{\partial p_{k\alpha}} \mathbf{E}[u_k] = \frac{\partial \overline{u}_k}{\partial p_{k\alpha}}. \tag{31}$$

In other words, we see that the mean-field payoffs of (29) are simply the gradients of the ergodic rates (7). Then, if we remove the zero-mean noise processes $\eta_{k\alpha}$ from (29), the general theory of [24, 40] tells us that the replicator equation (29) is just a stochastic approximation of the *mean-field ordinary differential equation*:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t)\left[\overline{v}_{k\alpha}(p(t)) - \overline{v}_k(p(t))\right], \tag{32}$$

so, one would hope that the asymptotic properties of (29) are approximate versions of the properties of (32).

With regards to the latter, recall that $\overline{\Phi}$ is strictly convex [25], so the analysis of Appendix A readily yields:

Theorem 10. *The mean field dynamics (32) converge to the unique Nash equilibrium of the ergodic game $\overline{\mathfrak{G}}$; in particular, every interior solution orbit of (32) converges to an ε-neighborhood of the game's equilibrium in time which is at most of order $\mathcal{O}(\log(1/\varepsilon))$.*

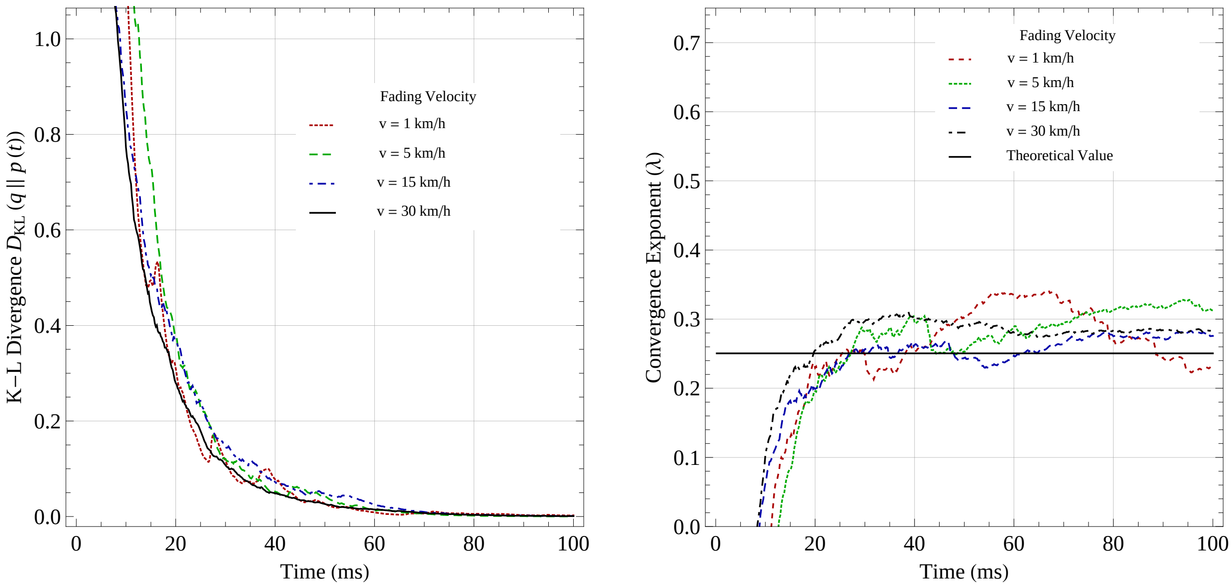
The implications of this theorem are similar to those of (8), so it is more important to concentrate on whether these asymptotic properties apply to the *stochastic* dynamical system (28). To make this question precise, let $\tau(n) = \sum_{j=1}^n \delta(n)$ and define the *interpolated processes* $\hat{p}_{k\alpha} : \mathbb{R}_+ \to \mathbb{R}$ by taking the values $\hat{p}_{k\alpha}(\tau(n)) = p_{k\alpha}(n)$ from (28) and interpolating linearly between these samples for $t \in [\tau(n), \tau(n+1))$:

$$\hat{p}_{k\alpha}(t) = p_{k\alpha}(n) + \frac{t - \tau(n)}{\tau(n+1) - \tau(n)}\left(p_{k\alpha}(n+1) - p_{k\alpha}(n)\right). \tag{33}$$

Then, if $\Theta : \Delta \times \mathbb{R} \to \Delta$ is the flow generated by the mean-field equation (32) on Δ,[12] we will say that:

[11] This randomness is precisely what causes the problem: if the process $g(n)$ contains an unbounded subsequence (which it almost surely does), the marginal payoffs $v_{A,B}$ are no longer uniformly bounded on Δ. By comparison, in the static case there are no such issues.

[12] In other words, $\Theta(p,t) \equiv \Theta^p(t)$ describes the solution trajectory of (32) which starts at p.



(a) Convergence to equilibrium

(b) The rate of convergence

Fig. 5. Convergence of the fading replicator dynamics to the unique Nash equilibrium of the ergodic game $\overline{\mathfrak{G}}$. To test the asymptotic properties of the replicator dynamics even beyond the regime of Theorem 12, we considered a 2×2 game and a constant sampling period of $\delta = 1ms$. The channel gain coefficients $g_{k\alpha}$ follow Rayleigh fading $g_{k\alpha} = |h_{k\alpha}|^2$ with the $h_{k\alpha}$ now being temporally correlated, following the simple to implement temporal correlation model $h(n+1) = rh(n) + \sqrt{1-r^2}\, z(n+1)$, where $z \sim \mathcal{CN}(0,1)$ and $r = \exp(-f\delta \cdot v/c)$ is the temporal correlation coefficient – we simulated user velocities $v = 1, 5, 15, 30$ km/h and a carrier frequency of $f = 2$ GHz. We see that the system equilibrates rapidly (the K-L divergence $D_{\mathrm{KL}}(q \,\|\, p(t))$ vanishes – Fig. 5(a)), while the stochastic convergence exponent $\lambda(t) = -\frac{1}{t}\log D_{\mathrm{KL}}(q \,\|\, p(t))$ settles down remarkably close to the predicted value λ_∞ for the ergodic potential (Fig. 5(b)).

Definition 11. The interpolated process $\hat{p} = \{\hat{p}_{k\alpha}(t)\}$ is an *asymptotic pseudo-trajectory* to the flow Θ if:

$$\lim_{t\to\infty} \sup_{0\le s\le T} \left\| \Theta(\hat{p}(t), s) - \hat{p}(t+s) \right\| = 0 \text{ for all } T > 0, \tag{34}$$

i.e., if $\hat{p}$ follows the flow Θ but makes an asymptotically vanishing amount of jumps every T units of time.

In light of this definition,and given that the stochastic process $p(n)$ is uniformly bounded (recall that $p(n) \in \Delta$ for all $n \ge 0$), an easy adaptation of standard results from the theory of stochastic approximation [24, Chapter 2, Theorem 2 and Corollary 4] shows that if the learning parameters $\delta(n)$ vanish rapidly enough as $n \to \infty$, then the noise processes $\eta_{k\alpha}$ will be dominated by the predictable (mean field) part of (29). More specifically, we obtain:

Theorem 12. *Assume that the learning parameters $\delta(n)$ of the stochastic dynamics (28) are ℓ^2-summable but not ℓ^1-summable (i.e., $\sum_{n=1}^\infty \delta(n) = \infty$ but $\sum_{n=1}^\infty \delta^2(n) < \infty$). Then, the replicator learning algorithm (28) for stochastically fluctuating channels converges (a.s) to the unique Nash equilibrium of the ergodic rates game $\overline{\mathfrak{G}}$.*

Remark 1. The most usual choice for the learning parameters δ is $\delta(n) = 1/n$. This variable rate can then be taken to mean that the wireless users update their power allocation policies every $1/n$ units of time, in which case the actual time which has elapsed after n iterations is $\mathcal{O}(\log n)$.[13] So, even though the number of iterations needed might turn out to be pretty high, the actual elapsed time will be significantly lower.

Remark 2. In case users employ a constant step size $\delta(n) \equiv \delta$ and no learning-induced discounting, the replicator dynamics evolve faster but convergence to equilibrium will be in the distribution sense [24]. This is an important (and largely unsolved) problem in the theory of stochastic approximation, so we will not address its theoretical aspects here. Instead, by resorting to numerical simulations (Fig. 5), we see that the replicator dynamics *do* converge to equilibrium in distribution (and almost surely if the equilibrium is strict – Fig. 5(a)). Moreover, this convergence is approximately exponential, with a convergence exponent remarkably close to the one predicted by the analogue of (25) with the ergodic payoffs $\bar{v}_{k\alpha}$ substituted in place of the static ones (Fig. 5(b)).

V. Conclusions and Future Directions

In this paper, we studied the distributed power allocation problem for orthogonal uplink channels (representing for example multiple wireless access points that operate in orthogonal frequency bands) by introducing a game which admits a convex potential function. In both cases of static and fading channels, the associated game admits a unique Nash equilibrium (Theorems 6 and 7) and we showed that a simple distributed learning scheme based on the replicator dynamics converges to it from (almost) any initial condition (Theorems 8 and 12). In fact, by proving a general result for convex potential games, we showed that the speed of this convergence is exponential: users converge to an ε-neighborhood of an equilibrium in time which is at most of order $\mathcal{O}(\log(1/\varepsilon))$ (Theorem 13).

There is a number of important extensions of this work which showcase the strength of the replicator dynamics and their asymptotic properties in potential games. First off, instead of the users' achievable rates, one could consider energy-efficient metrics where users do not saturate their power constraint (2) – e.g., as in the pricing model of [6] where a high price of transmission power might restrain users from transmitting at maximum power. In that case, the users' action spaces are simplices of one dimension higher (accounting for their variable total power), and much of the analysic carries through, though space and time considerations do not allow us to elaborate on this point.

An other important extension concerns non-orthogonal channel models such as the MIMO MAC problem. In that case, the game's strategy space is no longer a product of simplices, but the entire space of positive-definite precoding matrices with constrained trace, so the replicator equation (20) no longer applies. Nonetheless, one can still write down a suitably modified *matrix-valued* replicator equation which allows users to converge to equilibrium, but, again, given space constraints, we postpone this direction for the future.

Finally, we should also stress here that the key property of the replicator equation (20) is that it converges to the minimum of the game's potential. We may thus take a "dual" approach and employ the replicator dynamics as a learning algorithm which allows users to converge to the minimum of an arbitrary objective function Ψ by replacing the marginal payoffs $v_{k\alpha}$ with the gradients $\frac{\partial \Psi}{\partial p_{k\alpha}}$. So, if we focus for example on the users' aggregate

[13] On the other hand, since there are practical limitations to how fast users can update their powers, an alternative interpretation is that wireless users update their policies at every tick of a counter and they discount their update by $1/n$ at each step.

rate $\Psi(p) = \sum_k u_k(p)$, the only real catch is that the gradients $\psi_{k\alpha} \equiv \frac{\partial \Psi}{\partial p_{k\alpha}}$ cannot be calculated based on purely local information at the transmitter. However, some straightforward analysis shows that if the receiver broadcasts a specific real number for each channel,[14] then the values of the gradients $\psi_{k\alpha}$ *can* be calculated using this broadcast and the same information needed for the usual replicator dynamics (channel coefficients and SINRs). Consequently, for the relatively small cost of a broadcast message per channel, the users of the network will be able to converge exponentially quickly to the power allocation profile which maximizes their aggregate rate.

Appendix A
Convergence Speed of the Replicator Dynamics

Given a point $q \in \Delta$, we will say that a function $F : \Delta \to \mathbb{R}$ is *q-stable* when:

$$f(\theta) \equiv F(q + \theta z) \text{ is convex and increasing for all } z \in T_q^c\Delta \text{ and for all } \theta > 0 \text{ such that } q + \theta z \in \Delta. \tag{35}$$

In the above, $T_q^c\Delta$ denotes the *solid tangent cone* to Δ at q, that is:

$$T_q^c\Delta \equiv \left\{ z \in T_q\Delta : z_{k\alpha} \geq 0 \text{ for all } \alpha \in \mathcal{A}_k \text{ with } q_{k\alpha} = 0 \right\}. \tag{36}$$

In other words, $T_q^c\Delta$ consists of the tangent vectors $z \in T_q\Delta$ which point towards the interior of Δ – obviously, $T_q^c\Delta = T_q\Delta$ for all $q \in \mathrm{Int}(\Delta)$. Consequently, a (not necessarily convex) function $F : \Delta \to \mathbb{R}$ will be q-stable[15] for some $q \in \Delta$ when F increases convexly along all rays in Δ that emanate from q.

Clearly, strictly convex functions are q-stable with respect to their (unique) global minimum, and weakly convex functions with a unique minimum q at the boundary of Δ are also q-stable. In particular, both the static and ergodic potential functions Φ, $\overline{\Phi}$ are q-stable (see also [32]), so Theorems 8 and 10 are a special case of the following:

Theorem 13. *Assume that the game $\mathfrak{Q} \equiv \mathfrak{Q}(\mathcal{K}, \{\Delta_k\}, \{\phi_k\})$ admits a q-stable potential function F in the sense of (5). Then, the replicator dynamics (20) for the marginal payoffs $\phi_{k\alpha} = \frac{\partial \phi_k}{\partial p_{k\alpha}}$ converge to q for any initial condition that starts at finite K-L divergence $h_0 \equiv D_{\mathrm{KL}}(q \,\|\, p(0))$ from q. Moreover, for any such initial condition, there exists a positive constant $c > 0$ such that:*

$$D_{\mathrm{KL}}(q \,\|\, p(t)) \leq h_0\, e^{-ct} \quad \textit{for all } t \geq 0. \tag{37}$$

Our main goal in this appendix will be to prove Theorem 13 and, in so doing, to prove Theorems 8 and 10 as well. This will be done in a series of lemmas revolving around the so-called "evolutionary index" [30]:

$$L_q(p) = -\sum\nolimits_{k,\alpha} (p_{k\alpha} - q_{k\alpha})\, \phi_{k\alpha}(p). \tag{38}$$

Since $\phi_{k\alpha} = -\frac{\partial F}{\partial p_{k\alpha}}$, the evolutionary index can be interpreted as the directional derivative of F along the tangent direction $q - p \in T_q\Delta$.[16] More importantly however, an easy calculation reveals that L_q is just (the negative of) the

[14] This channel-specific number is just $\chi_\alpha \equiv \sum_k \left[\left(1 + \mathsf{sinr}_{k\alpha}^{-1}(p)\right) \left(\sigma_\alpha^2 + \sum_{\ell \neq k} g_{\ell\alpha} p_{\ell\alpha}\right) \right]^{-1}$.

[15] The reason for this terminology is that q will be globally evolutionarily stable for any game with potential F – see also [41].

[16] This also explains the terminology "evolutionary index": q is evolutionarily stable if and only if L_q is positive in a neighborhood of p.

time derivative of the relative entropy H_q with respect to the replicator dynamics (20):

$$\dot{H}_q(p) = \sum\nolimits_{k,\alpha} \frac{\partial H_q}{\partial p_{k\alpha}} \dot{p}_{k\alpha} = -\sum\nolimits_{k,\alpha} q_{k\alpha} \left(\phi_{k\alpha}(p) - P_k^{-1} \sum\nolimits_\beta^k p_{k\beta}\phi_{k\beta}(p)\right) = \sum\nolimits_{k,\alpha} (p_{k\alpha} - q_{k\alpha})\, \phi_{k\alpha}(p). = -L_q(p) \tag{39}$$

The first thing that we will show is that $L_q(p) > 0$ for all $p \in \Delta \setminus \{q\}$; this will imply that H_q is Lyapunov for the replicator dynamics (20), and the convergence part of Theorem 13 will follow from Lyapunov's theorem. Indeed, if we fix some $z \in T_q^c\Delta$, note that (38) may be rewritten as:

$$f'(\theta) = \sum\nolimits_{k,\alpha} \left.\frac{\partial F}{\partial p_{k\alpha}}\right|_{q+\theta z} z_{k\alpha} = \theta^{-1} L_q(q + \theta z), \tag{40}$$

for all sufficiently small $\theta > 0$ such that $q + \theta z \in \Delta$. Then, on account of (35), we obtain:

$$L_q(p) = \theta f'(\theta) \geq f(\theta) - f(0) = F(p) - F(q). \tag{41}$$

This last estimate shows that $L_q(p) > 0$ for all $p \neq q$, thus proving the convergence part of Theorem 13.

Now, the basic idea to prove the convergence time estimate (37) will be to establish an inequality of the form $L_q(p(t)) \geq cH_q(p(t))$ for some positive constant $c > 0$; with $\dot{H}_q(p) = -L_q(p)$, (37) will then follow from Grönwall's lemma. However, to establish such an inequality, we will need to refine the estimate (41) by quite a bit, and also to obtain an equally fine growth estimate for $H_q(p)$.

The basic idea in refining (41) will be to show that L_q grows (at least) linearly along the directions which are not supported in q, and (at least) quadratically along those which *are* supported in q. More precisely, let $V_q = \{x \in \mathbb{R}^\Omega : x_{k\alpha} = 0 \text{ if } q_{k\alpha} = 0\}$ be the subspace of directions of $\mathbb{R}^\Omega$ which are supported in q, and let $V_q^\perp$ be its orthogonal complement in the usual Euclidean (L^2) product of $\mathbb{R}^\Omega$. Then, by decomposing $z \in \mathbb{R}^\Omega$ as $z = z_{/\!/} + z_\perp$ with $z_{/\!/} \in V_q$ and $z_\perp \in V_q^\perp$, we may introduce two very useful seminorms $\|\cdot\|_{/\!/}$ and $|\cdot|_\perp$via the equations:

$$\|z\|_{/\!/}^2 \equiv \|z_{/\!/}\|_2^2 = \sum\nolimits_{k,\alpha}^{/\!/} z_{k\alpha}^2, \tag{42a}$$

$$|z|_\perp \equiv \|z_\perp\|_1 = \sum\nolimits_{k,\alpha}^{\perp} |z_{k\alpha}|, \tag{42b}$$

where the indices 1 and 2 refer to the L^1 and L^2 norms of $\mathbb{R}^\Omega$ respectively, and the notation $\sum_{k,\alpha}^{/\!/}$, $\sum_{k,\alpha}^{\perp}$ is a shorthand for summing over the directions that are in V_q and $V_q^\perp$ (i.e., $\sum_{k,\alpha}^{/\!/} \equiv \sum_{k,\alpha: q_{k\alpha}>0}$ and similarly for $\sum_{k,\alpha}^{\perp}$). We then get:

Lemma 14. *Let F be q-stable. Then:*

$$L_q(p) \geq F(p) - F(q) \geq m\, |p - q|_\perp + \tfrac{1}{2} r\, \|p - q\|_{/\!/}^2, \tag{43}$$

where $m = \min_k\{\phi_{k\alpha}(q) - \phi_{k\mu}(q) : q_{k\mu} = 0,\ q_{k\alpha} > 0\}$, and r is the minimum of the Rayleigh quotient $R_{q+z}(z) \equiv \langle z, \mathbf{M}(q + z)z\rangle/\|z\|^2$ for the Hessian $\mathbf{M}(p) = \frac{\partial^2 F}{\partial p_{k\alpha}\partial p_{\ell\beta}}$ of F, restricted over all $z \in T_q^c\Delta$ such that $q + z \in \Delta$.

Proof: Note first that q is the unique minimum of F, so the Karush-Kuhn-Tucker (KKT) conditions give:

$$\phi_{k\alpha}(q) = -\left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q = -\lambda_k \quad \text{for all } \alpha \in \mathcal{A}_k \text{ such that } q_{k\alpha} > 0, \tag{44a}$$

$$\phi_{k\mu}(q) = -\left.\frac{\partial F}{\partial p_{k\mu}}\right|_q < -\lambda_k \quad \text{for all } \alpha \in \mathcal{A}_k \text{ such that } q_{k\mu} = 0. \tag{44b}$$

Then, a first order Taylor estimate with Lagrange remainder readily yields:

$$f(\theta) = f(0) + f'(0)\theta + \tfrac{1}{2}f''(\xi)\theta^2 \tag{45}$$

for some $\xi \in (0,\theta)$, so (43) will follow once we properly estimate the linear and quadratic terms of (45).

As far as the linear term of (45) is concerned, we have:

$$\begin{aligned} f'(0) &= \sum\nolimits_{k,\alpha} z_{k\alpha} \left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q = \sum\nolimits_{k,\alpha}^{/\!/} z_{k\alpha} \left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q + \sum\nolimits_{k,\alpha}^{\perp} z_{k\alpha} \left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q \\ &= \sum\nolimits_k \left(\lambda_k \sum\nolimits_{\alpha}^{/\!/} z_{k\alpha} + \sum\nolimits_{\alpha}^{\perp} z_{k\alpha} \left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q\right) = \sum\nolimits_{k,\alpha}^{\perp} z_{k\alpha} \left(\left.\frac{\partial F}{\partial p_{k\alpha}}\right|_q - \lambda_k\right) \geq m\,|z|_{\perp}, \end{aligned} \tag{46}$$

where λ_k is the Lagrange multiplier of (44), the equality in the second line follows from the fact that $\sum_\alpha z_{k\alpha} = 0$ (which implies that $\sum_\alpha^{\perp} z_{k\alpha} = -\sum_\alpha^{/\!/} z_{k\alpha}$), and the last inequality is an immediate consequence of the definition of m.

Similarly, for any $\xi \in (0,\theta)$ and $z \in T_q^c\Delta$, we easily get:

$$f''(\xi) = \sum\nolimits_{k,\ell} \sum\nolimits_{\alpha,\beta} \left.\frac{\partial^2 F}{\partial p_{k\alpha}\partial p_{\ell\beta}}\right|_{q+\xi z} z_{k\alpha} z_{\ell\beta} = \left\langle z, \mathbf{M}(q+\xi z) z\right\rangle = R_{q+\xi z}(\xi z)\,\|z\|^2, \tag{47}$$

where $R_p(w) = \langle w, \mathbf{M}(p) w\rangle$, $p \in \Delta$, $w \in T_p\Delta$, denotes the Rayleigh quotient of the Hessian $\mathbf{M}$ of F. Hence, if r is the minimum of $R_{q+w}(w)$ over the compact set $B_q = \{w \in T_q^c\Delta : q + w \in \Delta\}$, we will also have $f''(\xi) \geq r\|z\|^2$, and (43) follows by simply plugging the estimates (46) and (47) back into (45) and noting that $\|z\| \geq \|z\|_{/\!/}$. ■

In light of Lemma 14, it would clearly suffice to establish an inequality of the form $H_q(p) \leq b_\perp|p-q|_\perp + b_{/\!/}\|p-q\|_{/\!/}^2$ for some positive constants $b_\perp, b_{/\!/} > 0$. Unfortunately, such an inequality cannot hold *globally*, because $H_q(p)$ blows up whenever q is not absolutely continuous with respect to p. Still, we do get a *local* inequality of this type:

Lemma 15. *Let $q \in \Delta$. Then, for every (nonzero) $z \in T_q^c\Delta$ and for every $a > 1$, the equation*

$$H_q(q+\theta z) = a\,|z|_\perp \theta + \tfrac{1}{2}a \sum\nolimits_{k,\beta}^{/\!/} z_{k\beta}^2/q_{k\beta}\,\theta^2, \tag{48}$$

admits a unique positive root $\theta_a \equiv \theta_a(z)$. Consequently:

$$H_q(q+\theta z) \leq a\,|z|_\perp \theta + \tfrac{1}{2}a \sum\nolimits_{k,\beta}^{/\!/} z_{k\beta}^2/q_{k\beta}\,\theta^2 \text{ for all } \theta \leq \theta_a(z). \tag{49}$$

Proof: The proof of this lemma is a tedious (and, at times, frustrating) exercise in analysis. To simplify it as much as possible, let $h(\theta) \equiv H_q(q+\theta z)$ denote the LHS of (48), and denote the RHS of the same equation by $ag(\theta)$. Then, if we set $w(\theta) = h(\theta) - ag(\theta)$, we immediately obtain $w(0) = 0$, $w'(0) = |z|_\perp(1-a) \leq 0$, and $w''(0) = \sum_{k,\beta}^{/\!/} z_{k\beta}^2/q_{k\beta}(1-a) < 0$.[17] As a result, $w(\theta)$ will initially dip down to negative values, but since w blows up to positive infinity as $q + \theta z$ approaches a face of Δ which does not support q absolutely, it follows that the equation (48) admits at least one positive root.

To show that this root is actually unique, we will need to study higher derivatives of w. As a first step in this direction, it is not hard to show that $w^{(4)}(\theta) = h^{(4)}(\theta) > 0$, which in turn implies that $w''(\theta)$ is strictly convex.

[17]Note that the seminorm $|z|_\perp$ vanishes if $z \in V_q$, but $\|z\|_{/\!/}$ vanishes on $T_q^c\Delta$ if and only if $z = 0$.

However, since $w''(0) < 0$ and 0 is already a root of $w(\theta)$, the mean value theorem ensures that $w(\theta)$ can have *at most* two positive roots.

So, assume that $w(\theta)$ has two distinct positive roots ξ_1, ξ_2 with $0 < \xi_1 < \xi_2$. Then, by the mean value theorem, the existence of two positive roots ξ_1 and ξ_2 implies that there exist ξ_1', ξ_2' with $0 < \xi_1' < \xi_1 < \xi_2' < \xi_2$ such that $w'(\xi_1') = 0 = w'(\xi_2')$ (recall that $w(0) = 0$). Furthermore, since $w(\theta)$ vanishes at 0, is negative near $\theta = 0$, and blows up for larger θ, it follows that w changes sign *exactly* one time (otherwise we would have three distinct positive roots). As a result, at least one of the positive roots ξ_1, ξ_2 must be a local extremum of w. For simplicity, assume that $w'(\xi_2) = 0$ (implying that w is positive near ξ_2), in which case we get three distinct points (ξ_1', ξ_2' and ξ_2) where w' vanishes. Then, a second application of the mean value theorem similarly implies the existence of ξ_1'', ξ_2'' with $0 < \xi_1'' < \xi_2''$ such that $w''(\xi_1'') = w''(\xi_2'') = 0$. However, $w''(\theta)$ is strictly convex and has $w''(0) < 0$, so it cannot have two distinct positive roots, a contradiction. The inequality (49) then follows from the negativity of $w(\theta)$ in the interval $(0, \theta_a(z))$. ■

This lemma is the first (crucial) step in the proof of Theorem 8. The next one is the following:

Lemma 16. *Let $p(t)$ be a solution trajectory of the replicator dynamics which starts with finite relative entropy $h_0 = H_q(p(0))$ w.r.t. the Nash equilibrium q of $\mathfrak{G}$. Then, there exists $b > 1$ such that:*

$$H_q(p(t)) \leq b\,|p(t) - q|_\perp + \frac{b}{2q_0}\|p(t) - q\|_{/\!/}^2, \quad \textit{where } q_0 = \min_{k,\alpha}\{q_{k\alpha} : q_{k\alpha} > 0\}. \tag{50}$$

Before we delve into the actual proof of this lemma, note that for every $p \in \Delta$, there exists a unique $z \in T_q^c\Delta$ such that $q + \theta z \in \Delta$ if and only if $\theta \leq 1$ – in essence, this is the tangent vector which is parallel to $p - q$ and which connects q to the face (or the boundary of the face) whose interior contains p. In this "projective" image, the collection of all such vectors will be called *the projective sphere* around q, and will be denoted by S_q:

$$S_q = \{z \in T_q^c\Delta : q + \theta z \in \Delta \text{ iff } 0 \leq \theta \leq 1\}. \tag{51}$$

In its turn, this gives rise to the (uniquely defined) *projective representation* of any $p \in \Delta$ as:

$$p = q + \theta z, z \in S_q \text{ and } 0 \leq \theta \leq 1. \tag{52}$$

As it turns out, the ratio of lengths $\theta = |p - q|/|z|$ will be particularly useful in streamlining some of the arguments that follow, so we will call it the *projective distance* of p from q.

Proof of Lemma 16: To begin with the actual proof, fix some $a > 1$. Then, by Lemma 15, we know that the equation (48) admits a unique positive root $\theta_a(z)$, so let $h_a(z) = H_q(q + \theta_a(z)z)$ and set $h_a = \max\{h_a(z) : z \in S_q\}$, S_q being the projective sphere around q. Furthermore, set $h_c = \max\{h_0, h_a\}$, let $\theta_c(z)$ be the unique positive root of the equation $H_q(q + \theta_c(z)z) = h_c$, and, similarly, define $b(z)$ via the equation:

$$h_c = b(z)g(\theta_c(z)), \tag{53}$$

with $g(\theta) = |z|_\perp\theta + \frac{1}{2}\sum_{k,\beta}^{/\!/} z_{k\beta}^2/q_{k\beta}\theta^2$ as in the proof of Lemma 15. It is then easy to see that $b(z) \geq a$ since, otherwise, we would have the contradictory inequality: $h_c = b(z)g(\theta_c(z)) < ag(\theta_c(z)) < h(\theta_c(z)) = h_c$ (note that $\theta_c(z) > \theta_a(z)$ so the inequality (49) has to be reversed).

This argument was necessary to show that $b(z) > 1$, which allows us to apply Lemma 16 a second time and get:

$$H_q(q+\theta z) \leq b(z)\left(|z|_\perp \theta + \tfrac{1}{2}\sum\nolimits_{k,\beta}^{/\!/} z_{k\beta}^2/q_{k\beta}\theta^2\right), \tag{54}$$

for all $\theta \leq \theta_c(z)$. We now claim that if $p(t) = q + \theta(t)z(t)$ is the projective decomposition of $p(t)$ with respect to q, then $\theta(t) \leq \theta_c(z(t))$. Indeed, should this ever fail, we would have:

$$H_q(p(t)) > b(z(t))g(\theta(t)) > b(z(t))g(\theta_c(t)) = h_c \geq h_0, \tag{55}$$

which contradicts the fact that H_q is Lyapunov for the replicator dynamics (20) – cf. (39).

Thus, with $\theta(t) \leq \theta_c(z(t))$ for all $t \geq 0$, we get:

$$H_q(p(t)) \leq b(z(t))\left(|z(t)|_\perp \theta(t) + \tfrac{1}{2}\sum\nolimits_{k,\beta}^{/\!/} z_{k\beta}^2(t)/q_{k\beta}\theta^2(t)\right), \tag{56}$$

and (50) follows by taking $b = \max\{b(z) : z \in S_q\}$. ■

After all these preliminary steps, the proof of Theorem 13 is relatively straightforward:

Proof of Theorem 13: With notation as in Lemmas 14 and 16, let $c = \min\{m/b, rq_0/b\}$. We then get:

$$L_q(p(t)) \geq m\,|p(t) - q|_\perp + \tfrac{1}{2}r\,\|p(t) - q\|_{/\!/}^2 \geq cH_q(p(t)), \tag{57}$$

and Grönwall's lemma yields (finally!): $H_q(p(t)) \leq h_0 e^{-ct}$. Since the KKT inequalities (44b) are strict for any direction $e_{k\alpha}$ of $\mathbb{R}^\Omega$ which is not supported in q (i.e., $q_{k\alpha} = 0$), we will have $m > 0$ and, consequently, $c > 0$ as well. ■

All that remains is to calculate explicitly the specific form that c takes when q is strict. In that case, given that the intersection of V_q with the tangent cone $T_q^c\Delta$ is trivial, the quadratic part of the estimate (43) can safely be ignored and some easy reshuffling yields the (simplified) inequality:

$$L_q(p) \geq \frac{1}{2}\sum\nolimits_k \|p_k - q_k\|_1 \Delta\phi_k, \tag{58}$$

where $\Delta\phi_k = \min_{\mu\neq\alpha_k}\{\phi_{k,\alpha_k}(q) - \phi_{k\mu}(q)\} > 0$ and $\|\cdot\|_1$ denotes the L^1 norm of $\mathbb{R}^{\mathcal{A}_k}$.

As is to be expected, the estimate (50) is similarly simplified as well, because we do not have to carry out the full analysis of Lemma 16. Instead, note that the projective decomposition of any $p_k \in \Delta_k \setminus \{q_k\}$ will be $p_k = q_k + \theta_k z_k$ where the direction indicator $z_k \in T_{q_k}^c\Delta_k$ has $z_{k,\alpha_k} = -P_k$. So, with $p_{k,\alpha_k} = P_k(1-\theta_k)$, we readily obtain:

$$H_q(p) = -\sum\nolimits_k P_k \log(1-\theta_k). \tag{59}$$

Now, let θ_k^* be defined by the equation $h_0 = H_{q_k}(q_k + \theta_k z_k)$, i.e., $\theta_k^* = 1 - \exp(-h_0/P_k)$, implying that $-P_k\log(1-\theta_k) \leq h_0\theta_k/\theta_k^*$ if and only if $0 \leq \theta_k \leq \theta_k^*$ (because of convexity). Just as in Lemma 16, we then claim that:

$$H_q(p(t)) = -\sum\nolimits_k P_k \log(1-\theta_k(t)) \leq h_0 \sum\nolimits_k \theta_k(t)/\theta_k^*, \tag{60}$$

where $\theta_k(t)$ is defined via the projective decomposition $p_k(t) = q_k + \theta_k(t)z_k(t)$. To see that this is so, it suffices to show that $\theta_k(t) \leq \theta_k^*$ for all $t \geq 0$. However, if $\theta_k(t) > \theta_k^*$ for some $t \geq 0$, then we would also have $H_{q_k}(p_k(t)) > h_0$, and, hence, $H_q(p(t)) > H_q(p(0))$ as well, a contradiction – recall that $H_q(p(t))$ is decreasing by (39). Thus, combining (58) and (60), we only need pick c such that $P_k\,\Delta\phi_k \geq ch_0/\theta_k^*$, and the sharpest choice is:

$$c = \min\nolimits_k \left\{P_k/h_0\left(1 - e^{-h_0/P_k}\right)\Delta\phi_k\right\}. \tag{61}$$

References

[1] S. Lasaulce, M. Debbah, and E. Altman, "Methodologies for analyzing equilibria in wireless games," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 41–52, September 2009.

[2] G. A. Gupta and S. Toumpis, "Power allocation over parallel gaussian multiple access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3274–3283, July 2006.

[3] F. Meshkati, H. V. Poor, and S. C. Schwartz, "Energy-efficient resource allocation in wireless networks: an overview of game-theoretic approaches," *IEEE Signal Process. Mag.*, vol. Special Issue on Resource-Constrained Signal Processing, Communications and Networking, May 2007.

[4] R. S. Cheng and S. Verdu, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[5] D. N. Tse and S. V. Hanly, "Multiaccess fading channels Part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 11, pp. 2796–2815, November 1998.

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

[7] 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, May 2002.

[8] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1115–1129, 2006.

[9] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: a unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, September 2008.

[10] ——, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[11] S. Chieochan, E. Hossain, and J. Diamond, "Channel assignment structures for infrastructure-based 802.11 WLANs: a survey," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 1, p. 124, first quarter 2010.

[12] S. Lasaulce, A. Suarez, R. de Lacerda, and M. Debbah, "Using cross-system diversity in heterogeneous networks: throughput optimization," *Elsevier Journal of Performance Evaluation*, vol. 65, no. 11, pp. 907–921, December 2008.

[13] P. Mertikopoulos and A. L. Moustakas, "Correlated anarchy in overlapping wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1160–1169, September 2008.

[14] S. M. Perlaza, E. V. Belmega, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks," in *ValueTools '09: Proceedings of the 4th International Conference on Performance Evaluation Methodologies and Tools*, 2009.

[15] E. V. Belmega, "On resource allocation games in MIMO wireless channels," Ph.D. dissertation, Université Paris Sud 11, 2010.

[16] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[17] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, 2004.

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

[19] E. V. Belmega, S. Lasaulce, M. Debbah, and A. Hjørunges, "Distributed power allocation policies in MIMO channels," in *EUSIPCO '10: Proceedings of the 2010 European Signal Processing Conference*, 2010.

[20] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1-2, pp. 145–156, 1978.

[21] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, ser. MIT Press Series on Economic Learning and

Social Evolution. The MIT Press, 1998, vol. 2.

[22] J. W. Weibull, *Evolutionary Game Theory*. Cambridge, MA: MIT Press, 1995.

[23] W. H. Sandholm, "Potential games with continuous player sets," *Journal of Economic Theory*, vol. 97, pp. 81–108, 2001.

[24] V. S. Borkar, *Stochastic approximation*. Cambridge University Press and Hindustan Book Agency, 2008.

[25] E. V. Belmega, S. Lasaulce, M. Debbah, M. Jungers, and J. Dumont, "Power allocation games in wireless networks of multi-antenna terminals," *Springer Telecommunications Systems Journal*, pp. 1–14, May 2010. [Online]. Available: http://dx.doi.org/10.1007/s11235-010-9305-3

[26] J. F. Nash, "Non-cooperative games," *The Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, September 1951.

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

[28] A. L. Moustakas and S. H. Simon, "Optimizing Multiple-Input Single-Output (MISO) communication systems with general Gaussian channels: nontrivial covariance and nonzero mean," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2770–2780, October 2003.

[29] A. Neyman, "Correlated equilibrium and potential games," *International Journal of Game Theory*, vol. 26, pp. 223–227, 1997. [Online]. Available: http://dx.doi.org/10.1007/BF01295851

[30] P. Mertikopoulos, "Stochastic perturbations in game theory and applications to networks," Ph.D. dissertation, University of Athens, 2010. [Online]. Available: http://sites.google.com/site/mertikop/files/Thesis.pdf

[31] Z.-Q. Luo and J.-S. Pang, "Analysis of iterative waterfllining algorithms for multi-user power control in digital subscriber lines," *EURASIP J. Appl. Signal Process. [Online]*, May 2006.

[32] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Dynamic power allocation games in parallel multiple access channels," in *ValueTools '11: Proceedings of the 6th International Conference on Performance Evaluation Methodologies and Tools*, submitted. [Online]. Available: http://arxiv.org/abs/1011.5610

[33] J. Huang and Z. Han, *Cognitive Radio Networks: Architectures, Protocols, and Standards*. Auerbach Publications, CRC Press, 2010, ch. Game theory for spectrum sharing.

[34] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *Journal of Economic Theory*, vol. 68, pp. 256–265, 1996.

[35] P. Mertikopoulos and A. L. Moustakas, "The emergence of rational behavior in the presence of stochastic perturbations," *The Annals of Applied Probability*, vol. 20, no. 4, pp. 1359–1388, 2010.

[36] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, November 1993.

[37] N. Bambos, S. Chen, and G. Pottie, "Radio link admission algorithms for wireless networks with power control and active link quality protection," UCLA School of Engineering and Applied Science, UCLA-ENG-94-25, 1994.

[38] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 1341–1347, September 1995.

[39] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband non-cooperative systems based on game theory—part II: algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1250–1267, March 2008.

[40] H. J. Kushner and G. G. Yin, *Stochastic approximation algorithms and applications*. New York, NY: Springer-Verlag, 1997.

[41] W. H. Sandholm and J. Hofbauer, "Stable games and their dynamics," *Journal of Economic Theory*, vol. 144, pp. 1710–1725, 2009.