

Distributed Power Allocation with SINR Constraints Using Trial and Error Learning

Luca Rose, Samir Medina Perlaza, Mérouane Debbah, Christophe J. Le Martret

▶ **To cite this version:**

Luca Rose, Samir Medina Perlaza, Mérouane Debbah, Christophe J. Le Martret. Distributed Power Allocation with SINR Constraints Using Trial and Error Learning. IEEE WCNC 2012, Apr 2012, Paris, France. pp.1835 - 1840, 10.1109/WCNC.2012.6214083 . hal-00770611

HAL Id: hal-00770611

https://centralesupelec.hal.science/hal-00770611

Submitted on 8 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed Power Allocation with SINR Constraints Using Trial and Error Learning

Luca Rose
Thales Communications, France.
luca.rose@thalesgroup.com

Samir M. Perlaza, Mérouane Debbah
Alcatel - Lucent Chair in Flexible Radio
Supelec, France.
(samir.perlaza, merouane.debbah)@supelec.fr

Christophe J. Le Martret
Thales Communications, France.
christophe.le_martret@thalesgroup.com

***Abstract*—In this paper, we address the problem of global transmit power minimization in a self-configuring network where radio devices are subject to operate at a minimum signal to interference plus noise ratio (SINR) level. We model the network as a parallel Gaussian interference channel and we introduce a fully decentralized algorithm (based on trial and error) able to statistically achieve a configuration where the performance demands are met. Contrary to existing solutions, our algorithm requires only local information and can learn stable and efficient working points by using only one bit feedback. We model the network under two different game theoretical frameworks: normal form and satisfaction form. We show that the converging points correspond to equilibrium points, namely Nash and satisfaction equilibrium. Similarly, we provide sufficient conditions for the algorithm to converge in both formulations. Moreover, we provide analytical results to estimate the algorithm's performance, as a function of the network parameters. Finally, numerical results are provided to validate our theoretical conclusions.**

***Keywords*: Learning, power control, trial and error, Nash equilibrium, spectrum sharing.**

I. Introduction

In this paper, we consider a network where several transmitter-receiver pairs communicate through a common bandwidth divided into several orthogonal sub-bands, thus, subject to mutual interference. All devices must guarantee certain quality of service (QoS), expressed in terms of signal to interference plus noise ratio (SINR). The behaviour of the devices is designed for achieving a stable network operating point (equilibrium) where the maximum number of communicating pairs satisfy their QoS with the minimum global power consumption. Network operating point must be achieved by the radio devices in a fully decentralized way by selecting their power allocation policy, i.e., selecting a sub-band and a power level for each transmission. In this scenario, all communications take place in absence of a centralized controller and neither cooperation nor exchange of information between different pairs are considered. For instance, this scenario may model the case of tactical radios and *ad hoc* networks, where, in order to set power and channel, current solutions require a certain level of cooperation with exchange of information between the devices or a manual setting.

The closest works to ours are [1], [2] and [3]. In [1], variational inequality theory is used to design a centralized power control algorithm for this scenario, in [2] the authors show that, if the assumption of S-modularity holds for the corresponding game, then best response dynamics [4] converges to a generalized Nash equilibrium (GNE); in [3], the authors provide, under the assumption of low interference, a sufficient condition for the convergence of the iterative water-filling algorithm to a GNE. It is worth noting that the works in [1], [2] and [3] assume a compact and convex set of actions, i.e, the possible power allocation (PA) vectors may take any value in the corresponding simplex. Conversely, in our work, we consider a finite action set by quantizing the possible available powers into a certain amount of levels. Basically, this is because in practice power levels must be expressed in a finite amount of bits. Moreover, several authors have pointed out that better global performance (e.g. spectral efficiency) is achieved when the set of PA vectors is substantially reduced [5], [6], [7], [8]. Indeed, this effect has been reported as a Braess kind paradox [9]. In this paper, our contribution is twofold: first, we present a fully decentralized learning algorithm able to keep the SINR level above a certain threshold a high proportion of the time, by means of only *one bit* feedback and relying only on local information [4]; second, we analytically study the convergence properties and the convergence point, which is shown to be an efficient Nash equilibrium (NE) in terms of global performance. The paper is organized as follows. In Sec. II, we describe the wireless scenario and we formalize the problem; in Sec. III, we model the system as a game in normal form and in satisfaction form; in Sec. IV we present the trial and error algorithm as introduced in [10]; in Sec. V, we present a formal analytical study of the convergence properties (i.e., expected number of iterations to reach the NE and the satisfaction equilibrium (SE)), as well as, the expected fraction of time the system is at the NE and at the SE; in Sec. VI we validate our analysis through numerical simulations; the paper is concluded in Sec. VII.

II. System model

Let us consider the system described in Fig. 1. Here, a set $\mathcal{K} \triangleq \{1, ..., K\}$ of transmitter-receiver pairs share a set $\mathcal{C} \triangleq \left\{b^{(1)}, ..., b^{(C)}\right\}$ of orthogonal sub-bands. Transmitter k is allowed to transmit over one sub-band at a time at a given power level. We denote by $p_k \in \mathcal{P}$, with $\mathcal{P} \triangleq \{0, ..., P_{MAX}\}$, $|\mathcal{P}| = Q$, and $b_k \in \mathcal{C}$, the power level and the frequency sub-band chosen by transmitter k respectively. We denote by

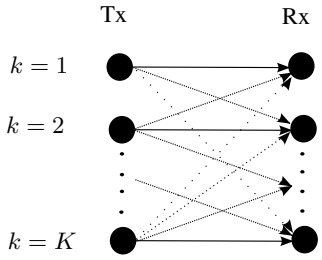


Fig. 1. System model

$\boldsymbol{p} = (p_1, p_2, ..., p_K)$ the network power allocation vector, by $\boldsymbol{b} = (b_1, b_2, ..., b_K)$ the spectrum occupation vector and by $\boldsymbol{a} = (a_1, a_2, ..., a_K)$ a network configuration vector, where $a_k = (p_k, b_k)$. To communicate, pairs have to achieve a sufficient SINR level, i.e., $SINR_k > \Gamma$ where we denote by Γ the minimum SINR threshold allowing transmission. Receivers treat interference as Gaussian noise, thus:

$$SINR_k(\boldsymbol{a}) = \frac{p_k g_{k,k}^{(b_k)}}{\sigma^2 + \sum_{l \in \mathcal{K} \setminus k} p_l g_{k,l}^{(b_l)} \mathbb{1}_{\{b_l = b_k\}}}, \tag{1}$$

where $g_{k,l}^{(b)}$ represents the channel power gain between transmitter l and receiver k over sub-band b, σ^2 is the power of the thermal noise assumed constant over the whole spectrum and $\mathbb{1}_{\{\}}$ represents the indicator function. In our scenario, we assume block-fading channels, i.e., channel realizations are time invariant for the whole transmission. Our objective is the satisfaction of the SINR constraints for the largest possible set of pairs by using the lowest global energy consumption. Formally, we want the network configuration vector $\boldsymbol{a}^*$ to be a solution of the following optimization problem

$$\begin{cases} \min_{\boldsymbol{p} \in \mathcal{P}^K} \sum_{k=1}^{K} p_k(n) \\ \text{s.t. } SINR_k(\boldsymbol{a}) > \Gamma \qquad \forall k \in \mathcal{K}^* \end{cases}, \tag{2}$$

where we denote by $\mathcal{K}^* \subseteq \mathcal{K}$ the largest subset of links able to simultaneously achieve a sufficient SINR level. Generally, to achieve this goal a central controller knowing all the network's parameters is required. In the following sections, we propose a decentralized algorithm demanding no information on the network which will steer the system to a solution of (2).

III. Game Formulation

In this section, we model the scenario presented in Sec 1 in two different formulations: a normal-form game and a satisfaction-form [11] game.

A. Normal form formulation

We model the network described above by the game in normal form

$$\mathcal{G} = \left(\mathcal{K}, \mathcal{A}, \{u_k\}_{k \in \mathcal{K}}\right). \tag{3}$$

Here, $\mathcal{K}$ represents the set of players, $\mathcal{A}$ is the joint set of actions, that is, $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times ... \times \mathcal{A}_K$ where $\mathcal{A}_k = \mathcal{C} \times \mathcal{P}$ and we introduce the utility function $u_k : \mathcal{A} \to \mathbb{R}$ defined by:

$$u_k(\boldsymbol{a}) = \frac{1}{1+\beta}\left(\frac{P_{MAX} - p_k}{P_{MAX}} + \beta \mathbb{1}_{\{SINR_k(\boldsymbol{a}) > \Gamma\}}\right), \tag{4}$$

where β is a design parameter discussed in Sec V. This function has been designed to be monotonically decreasing with the power consumption, and monotonically increasing with the number of players who achieve the minimum SINR. In the following, we show that, with this utility function, the NE of the game $\mathcal{G}$ can solve the problem stated in (2). Moreover, note that to evaluate (4) each transmitter only requires local information, since $\mathbb{1}_{\{SINR_k(\boldsymbol{a}) > \Gamma\}}$ can easily be fed back by the receiver with 1 bit.

Definition 1: (Interdependent game). $\mathcal{G}$ is said to be interdependent if for every non-empty subset $\mathcal{K}^+ \subset K$ and every action profile $\boldsymbol{a} = (\boldsymbol{a}_{\mathcal{K}^+}, \boldsymbol{a}_{-\mathcal{K}^+})$[1] such that $\boldsymbol{a}_{\mathcal{K}^+}$ is the action profile of all players in $\mathcal{K}^+$, it holds that:

$$\exists i \notin \mathcal{K}^+, \exists \boldsymbol{a}'_{\mathcal{K}^+} \neq \boldsymbol{a}_{\mathcal{K}^+} : u_i(\boldsymbol{a}'_{\mathcal{K}^+}, \boldsymbol{a}_{-\mathcal{K}^+}) \neq u_i(\boldsymbol{a}_{\mathcal{K}^+}, \boldsymbol{a}_{-\mathcal{K}^+}) \tag{5}$$

In the following, we assume that game $\mathcal{G}$ is interdependent. This is a reasonable assumption, since, physically, this means that no link is isolated from the others. The solution concept used under this formulation is the Nash equilibrium, which we define as follows:

Definition 2: (Nash equilibrium in pure strategies). An action profile $\boldsymbol{a}^* \in \mathcal{A}$ is a NE of game $\mathcal{G}$ if $\forall\, k \in \mathcal{K}$ and $\forall a'_k \in \mathcal{A}_k$

$$u_k(a_k^*, a_{-k}^*) \geq u_k(a'_k, a_{-k}^*). \tag{6}$$

To measure the efficiency of each NE, we introduce the *social welfare* function, defined by the sum of all individual utilities: $W(\boldsymbol{a}) = \sum_{k=1}^{K} u_k(\boldsymbol{a})$.

B. Satisfaction form

The satisfaction form is a game theoretical formulation modelling scenarios where players are exclusively interested in the satisfaction of their individual QoS constraints. Let us define the game as

$$\mathcal{G}' = \left(\mathcal{K}, \mathcal{A}, \{f_k\}_{k \in \mathcal{K}}\right), \tag{7}$$

where $\mathcal{K}$, $\mathcal{A}$ follow the previous definitions and the satisfaction correspondence $f_k : \mathcal{A}_{-k} \to \mathbb{R}$ is defined by

$$f_k(\boldsymbol{a}_{-k}) = \left(a_k \in \mathcal{A}_k : SINR_k(a_k, \boldsymbol{a}_{-k},) \geq \Gamma\right). \tag{8}$$

The solution concept used under this formulation is the satisfaction equilibrium (SE) defined as:

Definition 3: (Satisfaction equilibrium). A satisfaction equilibrium of game $\mathcal{G}'$ is an action profile $\boldsymbol{a}' \in \mathcal{A}$ such that $\forall k \in \mathcal{K}$,

$$a'_k \in f_k\left(\boldsymbol{a}'_{-k}\right). \tag{9}$$

Moreover, we measure the effort of player k due to the use of a particular action $\boldsymbol{a}_k$ by using the effort function [11] $\Phi_k : \mathcal{A}_k \to [0, 1]$. We can, then, define an efficient satisfaction equilibrium (ESE) as:

Definition 4: (Efficient satisfaction equilibrium). A satisfaction equilibrium $\boldsymbol{a}'$ is said to be efficient, if $\forall k \in \mathcal{K}$

$$a'_k \in \arg \min_{a_k \in f_k\left(\boldsymbol{a}'_{-k}\right)} \Phi_k(a_k) \tag{10}$$

[1]Here, $a_{-\mathcal{K}^+}$ refers to the action profile of all the players that are not in $\mathcal{K}^+$

In brief, an ESE is an action profile where all players are satisfied and no player may decrease its individual effort by unilateral deviation. Since our optimization problem is to minimize the overall transmit power, we identify the effort by the function: $\Phi_k(a_k) = p_k$.

IV. Algorithm Description

In this section, we briefly describe the trial and error (TE) algorithm introduced in [10], [12]. Later, we characterize the degrees of freedom of the system to fit in our scenario. In TE learning, each player k locally implements a state machine, at each iteration n, a state is defined by the triplet:

$$Z_k(n) = \{m_k(n), \bar{a}_k(n), \bar{u}_k(n)\}, \tag{11}$$

where $m_k(n) \in \{C, C+, C-, D\}$ represents a "mood", i.e., a characteristic that defines the machine reaction to its experience of the environment, $\bar{a}_k \in \mathcal{A}$ and $\bar{u}_k \in [0,1]$ represents a *benchmark* action and *benchmark* utility, respectively. There are four possible moods: content (C), watchful ($C-$), hopeful ($C+$), discontent (D). In the following, we characterize the behaviour of each player in every possible mood.

- **Content**

If at stage n player k is content, it uses the benchmarked action $\bar{a}_k(n)$ with probability $(1-\epsilon)$ and experiments a new action $a'_k(n)$ with probability ϵ. If at stage n the player decided to experiment, at stage $(n+1)$ it evaluates the utility $u'_k(n+1)$ associated with $a'_k(n)$ as follows: if $u'_k(n+1) < \bar{u}_k(n)$ then $Z_k(n+1) = Z_k(n)$, otherwise if $u'_k(n+1) > \bar{u}_k(n)$, then, with probability $\epsilon^{G(u'_k(n+1)-\bar{u}_k(n))}$, a new action and utility benchmark are set out, i.e., $\bar{u}_k(n+1) = u'_k(n+1)$ and $\bar{a}_k(n+1) = a'_k(n)$, respectively. Here, $G(\cdot)$ must be such that:

$$0 < G(\Delta u) < \frac{1}{2}, \tag{12}$$

we opt for a linear formulation: $G(\Delta u) = -0.2\Delta u + 0.2$.

- **Hopeful-Watchful**

If player k achieves an increment or a decrement in its utility without having experimented at the previous stage, then the mood become hopeful or watchful, according to the following rule: (i) if $u'_k(n+1) > \bar{u}_k(n)$ then, $m_k(n+1) = C+$, $\bar{a}_k(n+1) = \bar{a}_k(n)$ and $\bar{u}_k(n+1) = \bar{u}_k(n)$; (ii) if $m_k(n+1) = C-$, then $\bar{a}_k(n+1) = \bar{a}_k(n)$ and $\bar{u}_k(n+1) = \bar{u}_k(n)$. If player k observes an improvement also at the next stage (i.e., $u'_k(n+2) > \bar{u}_k(n+1)$), then the mood switches to content and the benchmark utility is updated with the new one: $m_k(n+2) = C$ and $\bar{u}_k(n+2) = u'_k(n+1)$. On the contrary, if a loss is observed also at the next stage (i.e., $u'_k(n+2) < \bar{u}_k(n+1)$), then the mood switches to discontent $m_k(n+2) = D$.

- **Discontent**

If player k is discontent, it experiments a new action ($a'_k(n)$) at each step n. We refer to this behaviour as *noisy* search. When the corresponding utility $u'_k(n+1)$ is observed, with probability $p = \epsilon^{F(u'_k(n+1))}$ the mood turns to content $m_k(n+1) = C$, a new action and utility benchmark are set up, $\bar{u}_k(n+1) = u'_k(n+1)$ and $\bar{a}_k(n+1) = a'_k(n+1)$, while, with probability $(1-p)$ it continues the noisy search. Note that function F must be such that

$$0 < F(u) < \frac{1}{2K}, \tag{13}$$

we opt for a linear formulation: $F(u) = -\frac{0.2}{K}u + \frac{0.2}{K}$.

A. Algorithm properties

Hereunder, we restate Theorem 1 in [10] and Theorem 1 in [12] using our notation.

Theorem 1: Let $\mathcal{G}$ have at least one pure Nash equilibrium and let ϵ be small enough. Then, a pure Nash equilibrium is played at least $(1-\delta)$ of the time.

This theorem introduces a different notion of convergence. Generally [4], we say that an algorithm converges when it approaches a certain solution as $n \to \infty$ while, here, it means that this solution is played with a high probability an high proportion of the total time.

Theorem 2: Let $\mathcal{G}$ have at least one pure Nash equilibrium and let each player employ TE, then a Nash equilibrium that maximizes the sum utility among all equilibrium states is played a large proportion of the time.

Note that, generally, different equilibria are associated with different social welfare values. Learning algorithms available in the literature [4], [13], do not always take into consideration the problem of equilibrium selection, which is a central issue when aiming at global performance.

V. Main results

A. Working point properties

In this section, we present our results based on the previous analysis. Proofs are omitted due to space constraints. Based on the game theoretical formulation in Sec. III and the algorithm properties in Sec. IV we state the following:

Theorem 3: Let $\mathcal{N} \neq \emptyset$ be the set of NE of $\mathcal{G}$, let $\beta > K$ and let us denote by K_l the number of players satisfied at the l-th NE. Then, TE converges to the NE where K_l is maximized.

This theorem states that, if $\beta > K$, then TE converges to a state where the largest possible number of players are satisfied *and* are at the NE. Here, β represents the interest a network designer has in satisfying the largest set of players over the minimization of the network power consumption. The next two theorems allow us to link this result with the original global design problem expressed in (2).

Theorem 4: Let (i) $\mathcal{A}^\dagger \neq \emptyset$ be the set of solutions of (2) with $\mathcal{K}^* = \mathcal{K}$, (ii) $\mathcal{N} \neq \emptyset$ be the set of NE of $\mathcal{G}$ and $\mathcal{N} \cap \mathcal{A}^\dagger \neq \emptyset$ and let $\beta > K$. Then, TE learning converges to an action profile $\boldsymbol{a}^*$ such that $\boldsymbol{a}^* \in \mathcal{N} \cap \mathcal{A}^\dagger$ and is an ESE.

This theorem links together the concept of ESE of game $\mathcal{G}'$, the NE of game $\mathcal{G}$ and the solutions of (2). Indeed, when the assumptions are met, the TE algorithm will reach a network state where: (i) all players are satisfied, (ii) the network power consumption is minimized. Note that, generally, it is possible for (2) to have a solution that is not a NE of $\mathcal{G}$.

Theorem 5: Let (2) have no solution for $\mathcal{K}^* = \mathcal{K}$ and fix $\beta > K$. Let K^* be the largest number of players that can be simultaneously satisfied and let $\mathcal{K}^*_m$ be the m-th set, such that $|\mathcal{K}^*_m| = K^*$, where (2) has a solution; let also be $\mathcal{A}^*_m$ the corresponding set of solutions.

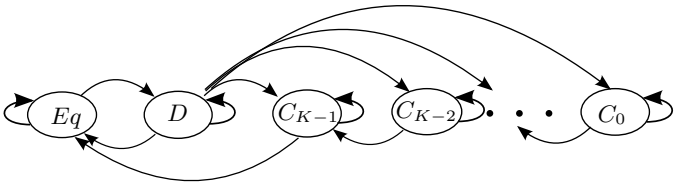


Fig. 2. Markov chain describing the TE algorithm in the network. The state Eq represents a state where all players are in equilibrium (i.e., SE or NE). C_{K-k} represents a state where $K-k$ players are using a *correct* action (i.e., an action that is satisfying or is optimal w.r.t. the others). D represents a state where one player is *discontent*.

Let us define $\mathcal{A}^* = \bigcup_m \mathcal{A}_m^*$ and $\mathcal{N} \neq \emptyset$ the set of NE. Then, TE learning converges to an action profile $\boldsymbol{a}^*$ such that $\boldsymbol{a}^* \in \mathcal{N} \cap \mathcal{A}^*$.

The previous theorem states that, when some players cannot satisfy their SINR condition, the TE algorithm selects the subset $\mathcal{K}_m^*$ among all possible $\mathcal{K}^*$ such that: (i) the highest number of players are satisfied, (ii) the network power consumption is minimized (with the unsatisfied players employing 0 power).

Corollary 1: Let $\forall k$ and $\forall b$ be $g_{k,k}^{(b)} \geq \frac{\Gamma\sigma^2}{P_{MAX}}$, let $C \geq K$ and fix $\beta > K$. Then, TE converges to a solution of (2)*.*

Basically, this corollary means that, if transmitters and receivers are satisfiable on each channel (high SNR regime), then TE converges to an optimal working point.

B. Convergence analysis

TE algorithm defines a discrete time Markov chain (DTMC) on the set of the states. Studying the behaviour of the algorithm on the complete chain is an intractable problem due to the number of states, transitions and parameters. In the following, we provide an approximated DTMC that allows us to estimate: (a) the expected converging time at the NE and at the SE, (b) the expected fraction of time the system is at the NE and at the SE. Under the light of the description made in Sec. IV, we state the following: (i) the fraction of time spent in the *watchful* or *hopeful* states is negligible compared to the one in *discontent* or *content* one; (ii) at any time, the probability of having more than one player *discontent* is negligible. In the following, we assume $C > K$ and a simplified channel model, defined as

$$\begin{cases} g_{k,k}^{(c)} = 1 & \forall k, \forall c \\ g_{j,k}^{(c)} = \frac{1}{2} & \forall k, \forall j \neq k, \forall c \end{cases}. \tag{14}$$

In Sec. VI we will show that these results are good approximations also under less restrictive conditions. The resulting DTMC for studying TE behaviour is represented in Fig. 2. When interested in convergence time and occupancy frequency of the NE, state Eq represents the NE, and C_{K-k} a state where $K-k$ players are using an individually optimal action and D a state where one player is *discontent*. The transition probabilities we evaluate are listed hereunder, the detailed description is omitted due to space constraints.

$$P(NE, D) = \frac{K(K-1)^2\epsilon^2}{C^2}\left(\frac{Q-1}{Q}\right)^2 \tag{15}$$

$$P(D, NE) = \frac{(C-K+1))}{CQ} \tag{16}$$

$$P(D, C_{K-k}) = \frac{(C-K+k)}{C^k}\frac{(K-1)!}{(K-k)!} \tag{17}$$

$$P\left(C_{K-k}, C_{K-k-1}\right) = (K-k)\frac{C-k}{CQ}\epsilon^{1+G(\Delta u)}. \tag{18}$$

The analysis of this DTMC allows us to state the following theorems:

Theorem 6: The expected number of iterations needed before reaching the NE for the first time $\bar{T}_{NE}$ is bounded as follows:

$$\begin{aligned} \bar{T}_{NE} &\leq \frac{CQ}{\epsilon^{(1+G(\Delta u))}(C-K)}\left(1 + \log\left(\frac{K(C-K+1)}{C+1}\right)\right) \\ \bar{T}_{NE} &\geq \frac{CQ}{\epsilon^{(1+G(\Delta u))}(C-K)}\left(\gamma + \log\left(\frac{K(C-K)}{C}\right)\right); \end{aligned}$$

where, $\gamma \simeq 0.577$ is the Euler-Mascheroni constant.

Note that, the time demanded to converge is directly proportional to the degree of freedom (i.e., $|\mathcal{A}_k| = CQ$) and inversely to the experimentation probability ϵ. Nonetheless, as we shall see, choosing a large ϵ increases the instability of the NE and, consequentially, the network performance.

Theorem 7: The expected fraction of time the system is at a NE $(1-\delta)$ is:

$$(1-\delta) = \frac{1}{1 + P(NE, D)T_{BNE}}, \tag{19}$$

where

$$\begin{aligned} T_{BNE} &\leq \sum_{k=1}^{K} P(D, C_{K-k})T_{CNE}(k) + \frac{P(D, NE)}{\left(1 - P(D, D)\right)^2} \\ T_{CNE}(k) &= \frac{CQ}{\epsilon^{1+G(\Delta u)}(C-K)}\left(\gamma + \log\left(\frac{K(C-k+1)}{C+1}\right)\right) \\ P(D, D) &= 1 - P(D, NE) - \sum_{k=1}^{K} P(D, C_{K-k}). \end{aligned}$$

Here, $(1-\delta)$ depends on $\frac{1}{\epsilon^2}$ as in (15). This means that, the larger the ϵ the shorter the time the system is at a NE. To evaluate convergence time and occupancy frequency of the SE we, again, make use of Fig. 2. In this case, state Eq represents the SE, C_{K-k} is a state where $K-k$ players are satisfied and D a state where one user is *discontent*. The corresponding transition probabilities are listed hereunder.

$$P(SE, D) = \frac{K(K-1)^2\epsilon^2}{C^2}\left(\frac{Q-1}{Q}\right)^2 \tag{20}$$

$$P(D, SE) = \frac{(C-K+1)}{C} \tag{21}$$

$$P(D, C_{K-k}) = \frac{1}{C^k}\frac{(K-1)!}{(K-k)!}(C-K+k) \tag{22}$$

$$P\left(C_{K-k}, C_{K-k-1}\right) = (K-k)Q_S\frac{(C-k)\epsilon^{1+G(\Delta u)}}{CQ}. \tag{23}$$

Given the model in (14) and $C > K$, the term $Q_S < Q$ represents the number of power quantization levels that a player can employ to successfully achieve $SINR_k > \Gamma$ on any free channel. We can, then, state the following theorems:

Theorem 8: The expected number of iterations needed before reaching the SE for the first time $\bar{T}_{SE}$ is bounded as follows:

$$\bar{T}_{SE} \leq \frac{CQ/Q_S}{\epsilon^{(1+G(\Delta u)}\,(C-K)}\left(1+\log\left(\frac{K\,(C-K+1)}{C+1}\right)\right)$$
$$\bar{T}_{SE} \geq \frac{CQ/Q_S}{\epsilon^{(1+G(\Delta u)}\,(C-K)}\left(\gamma+\log\left(\frac{K\,(C-K)}{C}\right)\right).$$

Under assumption (14), being satisfied is a weaker condition than being at the NE, thus, it results that $T_{NE} \geq T_{SE}$. Predictably, larger P_{MAX} and lower Γ increasing Q_S, are able to improve the converging speed.

Theorem 9: The expected fraction of time the system is at a SE F_{SE} is:

$$F_{SE} = \frac{1}{1+P(SE,D)T_{BSE}} \tag{24}$$

where

$$T_{BSE} \leq \sum_{k=1}^{K} P(D,C_{K-k})T_{CSE}(k) + \frac{P(D,SE)}{(1-P(D,D))^2}$$
$$T_{CSE}(k) = \frac{CQ}{\epsilon\,(C-K)\,Q_S}\left(\gamma+\log\left(\frac{K\,(C-k+1)}{C+1}\right)\right)$$
$$P(D,D) = 1-P(D,SE)-\sum_{k=1}^{K} P(D,C_{K-k}).$$

VI. Simulation Results

The purpose of this section is threefold. First, we run simulations to numerically validate the DTMCs introduced in Sec. V, second, we validate the results on more general channel models, then, we evaluate the performance of the algorithm in terms of satisfaction and power employed. The first two experiments have been run for two different sets of parameters. The first set is composed by: $K = 3$, $C = 4$, $\epsilon = 0.02$ and $6 \leq Q \leq 10$. The second set is composed by $K = 4$, $C = 5$, $\epsilon = 0.02$ and $6 \leq Q \leq 10$. In our first experiment, we run 10^7 iterations to estimate $(1-\delta)$ under two different channel models: the simple channels expressed in (14) and a Rayleigh channel. The results are summarized in Figure 3. As we can see, the analysis, brought on particular channel model, proves to be sufficiently precise also under more general formulations. In our second experiment, we estimated the converging time and compared with the analytical results in Figure 4. As we can see, increasing the action set dimension, i.e., increasing C or Q, brings slower convergence rate since the algorithm requires more time to explore all the possibilities. Note that, here, convergence time means the time needed by the system to work at the NE for the first time. The third experiment's parameter set is composed as follows: $K = 4$, $C = 5$, $\epsilon = 0.02$ $Q = 8$ with the simplified channel model as in (14). Here, we have run 10^3 tests, each one composed by 6000 iteration of TE. The results are showed in Fig 5, where the upper curve represents the fraction of players satisfied, while the lower curve represents the ratio between the average power employed by the network and the optimal power that should be employed to satisfy all the players. In average, in accordance with Figure 3, the system reaches an optimal equilibrium (all players satisfied and minimum amount of power employed) after around 2200 iterations. Note that, even though, for some specific scenarios, this number may be too high, the configurations selected by the algorithm before the convergence are just slightly inefficient. Indeed, a configuration where all the players are able to satisfy their SINR constraints is averagely reached after 600 iterations. Moreover, before this, we observe that only a fraction of satisfiable players is satisfied, in spite of the amount of power used.

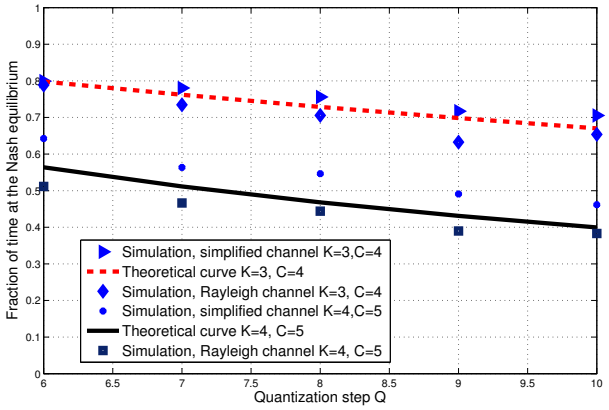


Fig. 3. Fraction of time the system is at the NE. Comparison between theoretical line and simulation results for two set of data and different channels: Rayleigh and simple one as in (14).

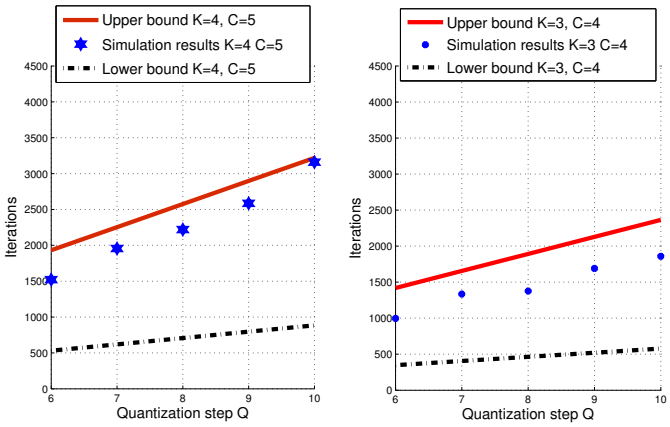


Fig. 4. Expected time for the system to reach the Nash equilibrium. On the left, results for a system with $K = 4$ players and $C = 5$ channel. On the right, results for a system with $K = 3$ players and $C = 4$ channels.

VII. Conclusion

In this paper, we have studied a power control problem in a self configuring decentralized network. We presented a new decentralized algorithm able to steer the network into a working point where the maximum number of transmitter-receiver pairs achieves a sufficient SINR while minimizing the network power consumption. The algorithm does not assume

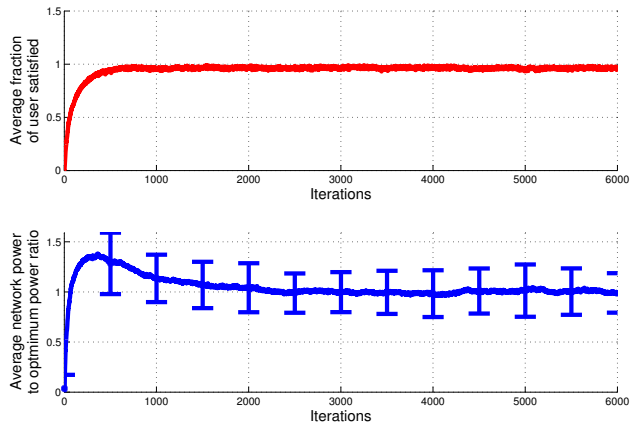


Fig. 5. Average fraction of players satisfied versus iterations, and ratio between the average power employed and the optimal level of power that should be employed to satisfy the whole network. Available number of channel $C = 5$, number of players $K = 4$, quantization level $Q = 8$.

any prior knowledge of the network and can learn an efficient equilibrium with only *one* bit of feedback. By assuming a particular channel realization, we have analytically estimated the expected performance of the algorithm through a Markov chain description of the algorithm behaviour. Finally we have shown through Monte-Carlo simulations that the analysis is approximatively correct also for general channel models.

VIII. Acknowledgement

This research work was carried out in the framework of the CORASMA EDA Project B-0781-IAP4-GC.

References

[1] H. Iiduka, "Fixed point optimization algorithm and its application to power control in CDMA data networks," *Mathematical Programming*, Oct. 2010.

[2] E. Altman and Z. Altman, "S-modular games and power control in wireless networkse," *IEEE Trans. Automat. Contr.s*, vol. 48, no. 5, pp. 839–842, May 2003.

[3] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. on Info. Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.

[4] L. Rose, S. Lasaulce, S. M. Perlaza, and M. Debbah, "Learning equilibria with partial information in decentralized wireless networks," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 136 –142, Aug. 2011.

[5] O. Popescu and C. Rose, "Water filling may not good neighbors make," in *In Proceedings 2003 IEEE Global Telecommunications Conference - GLOBECOM 03*, 2003, pp. 1766–1770.

[6] L. Rose, S. M. Perlaza, and M. Debbah, "On the Nash equilibria in decentralized parallel interference channels," in *IEEE Workshop on Game Theory and Resource Allocation for 4G*, Kyoto, Japan, Jun. 2011.

[7] E. Altman, V. Kamble, and H. Kameda, "A Braess type paradox in power control over interference channels," *1st International ICST Workshop on Physics Inspired Paradigms for Wireless Communications and Network*, May 2008.

[8] S. M. Perlaza, M. Debbah, S. Lasaulce, and H. Bogucka, "On the benefits of bandwidth limiting in decentralized vector multiple access channels," *in Proc. 4th Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Comm. (CROWNCOM)*, May 2009.

[9] D. Braess, A. Nagurney, and T. Wakolbinger, "On a paradox of traffic planning," *Transportation Science*, vol. 39, pp. 446–450, November 2005.

[10] H. P. Young, "Learning by trial and error," Tech. Rep., 2008.

[11] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Quality-of-service provisioning in decentralized networks: A satisfaction equilibrium approach," *IEEE Journal of Selected Topics in Signal Processing*, Feb. 2012.

[12] B. S. Pradelski and H. P. Young, "Efficiency and equilibrium in trial and error learning," Tech. Rep., 2010.

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