RESEARCH   **Open Access**

Optimal resource allocation in fetmocell networks based on Markov modeling of interferers' activity

Stefania Sardellitti[*], Alessandro Carfagna and Sergio Barbarossa

Abstract

Femtocell networks offer a series of advantages with respect to conventional cellular networks. However, a potential massive deployment of femto-access points (FAPs) poses a big challenge in terms of interference management, which requires proper radio resource allocation techniques. In this article, we propose alternative optimal power/bit allocation strategies over a time-frequency frame based on a statistical modeling of the interference activity. Given the lack of knowledge of the interference activity, we assume a Bayesian approach that provides the optimal allocation, conditioned to periodic spectrum sensing, and estimation of the interference activity statistical parameters. We consider first a single FAP accessing the radio channel in the presence of a dynamical interference environment. Then, we extend the formulation to a multi-FAP scenario, where nearby FAP's react to the strategies of the other FAP's, still within a dynamical interference scenario. The multi-user case is first approached using a strategic non-cooperative game formulation. Then, we propose a coordination game based on the introduction of a pricing mechanism that exploits the backhaul link to enable the exchange of parameters (prices) among FAP's.

Introduction

Femtocell networks are composed of cells having a coverage radius in the order of tens of meters, providing enhanced indoor coverage through the use of femto-access points (FAPs) or home-enhanced node B (HeNB), in the long-term evolution (LTE) terminology [1,2]. A typical scenario is sketched in Figure 1, where we can notice the wireless links among femto user equipments (FUE), macro user equipments (MUE), macro base stations (MBSs) and FAPs. More specifically, the wireless links are classified as useful or interfering depending on whether they refer, respectively, to the useful link between a transmitter and its intended receiver or to other receivers falling within its coverage area. Being installed in residential areas, e.g., home, offices, etc., the FAP's are typically interconnected with each other through a wired link, usually an ADSL subscriber line which allows the access to a broadband Internet network, as depicted in Figure 1. One of the ideas proposed in this article is to exploit the backhaul to set up a local coordination

among nearby FAPs to improve the efficiency of the radio resource management (RRM), without the presence of a centralized control.

Femtocells are becoming more and more attractive due to their benefits to both cellular operators and subscribers. On the one hand, operators see femtocells as a way to improve indoor coverage and to off-load wireless traffic from the macro cellular network to the wired network, thus releasing wireless channels to additional mobile users. On the other hand, subscribers see femtocells as a way to get higher quality services, either higher data throughput or better voice quality, thanks to a better indoor coverage, and seamless connectivity.

Following the current evolution of cellular standardization process, in this study we assume an LTE framework and we focus on the downlink channel, which assumes an OFDMA strategy. In this context, femtocell networks offer advantages with respect to Wi-Fi, as they avoid vertical hand-off and offer better QoS.

In view of a potential massive deployment of FAP's, a special attention has to be devoted to RRM. In fact, different from MBSs, FAPs are typically installed by the subscribers and maintained without global planning, with

*Correspondence: stefania.sardellitti@uniroma1.it
Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy

Figure 1 Femtocell network scenario.
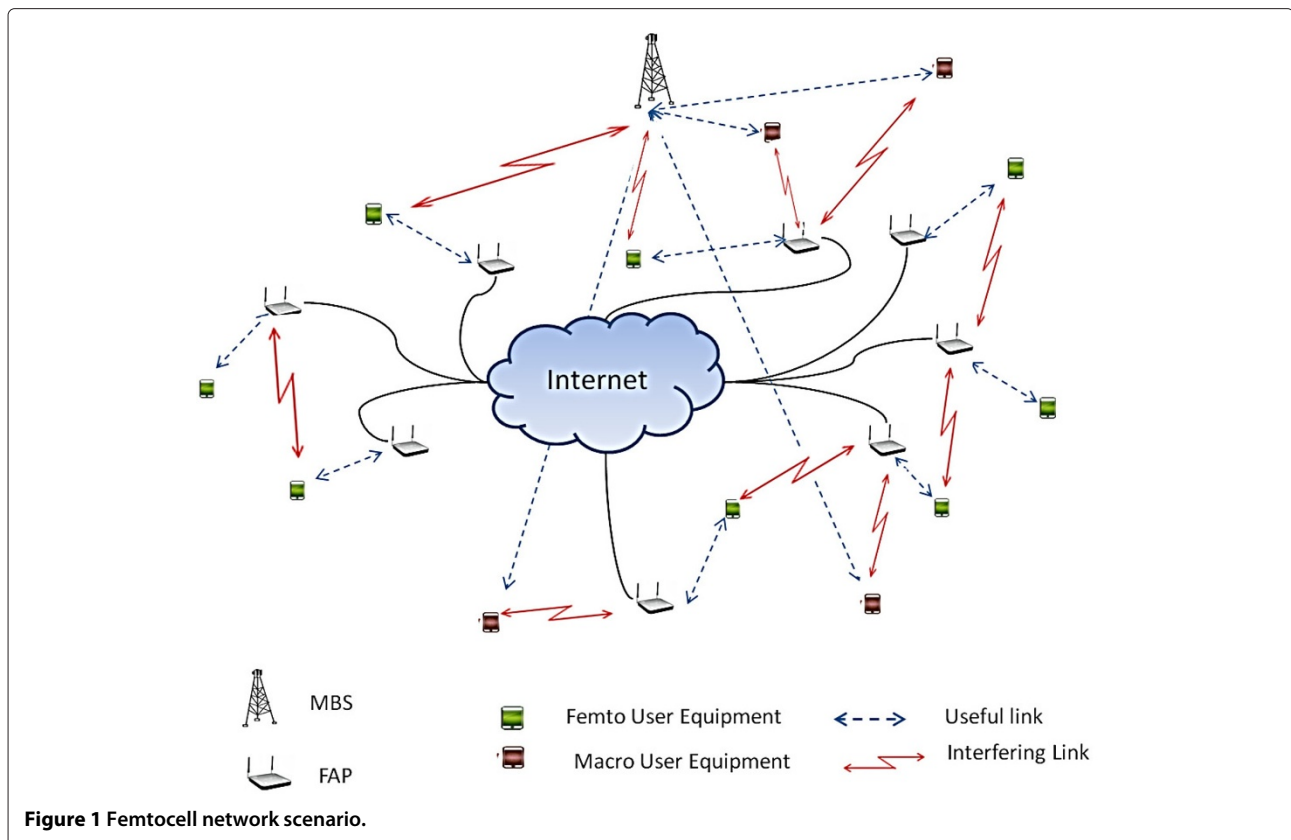
no special consideration about traffic demands or interference with other cells, either femto or macro cells. Hence, a dense deployment of FAPs might induce an intolerable interference from FUE's to MUE's or to other FUE's. Interference management is then arguably one of the major challenges to be faced in femtocell networks.

The goal of this study is to propose an algorithm for optimizing power/bit allocation over a joint time–frequency domain, incorporating a statistical model of the macro-users activity. Since the interference is unknown, the proposed algorithm follows a Bayesian approach, which allocates power/bits over successive time/frequency slots depending on a preliminary sensing and estimation of the parameters of the interference model. We assume a Markov modeling for simplicity, but the approach can be generalized to more sophisticated models, like e.g., [3,4]. More specifically, in this study the interference over different frequency subchannels is modeled as a set of statistically independent homogeneous discrete-time Markov chains (DTMCs). We consider a single-user allocation first, where a single FAP finds the optimal resource allocation according to two alternative strategies: (i) maximize the *expected* rate, conditioned to the result of the sensing and estimation phase, under a transmit power constraint; (ii) minimize the transmit power under the *expected* rate constraint.

Opportunistic spectrum access (OSA) in multicarrier networks where the channel occupancy follows a Markovian evolution has already been studied in the framework of cognitive radio (CR) in [5,6], for example. Chen et al. [5] develop an optimal OSA scheme aimed at optimizing spectrum sensing and access policies *jointly*. They assumed that the secondary transmitter receives error-free ACK signals from the secondary's receiver, whenever the transmission is successful, and this information is used to track the state of the primary channels. Interestingly enough, Chen et al. [5] establish a separation principle that decouples the design of spectrum sensor and access policy. A similar context is studied in [6,7], where the authors combine learning and dynamic spectrum access. Both Chen et al. [5] and Unnikrishnan and Veeravalli [6] consider an objective function that depends only on the available cognitive bandwidth and puts a constraint on the collision probability with the primary users. Anandkumar et al. [8] and Liu and Zhao [9] formulate the multi-user OSA problem as a decentralized multi-armed bandit problem [10]. In such a framework, each user *learns* the channel availability statistics and designs a channel access rule in order to maximize the transmission throughput (or equivalently minimize the system *regret*, defined as the loss in secondary throughput due to learning errors and collisions under distributed access).

In [9], which is an extension of the single-user policy proposed in [10] to the multi-user case, Liu and Zhao propose a family of distributed learning and access policies known as time-division fair share. For these policies, they prove the minimum growth rate of the system regret, which is shown to behave logarithmically with respect to the number of time slots. Moreover, Liu and Zhao [9] distinguish the case of known number of secondary users from the case in which this number is unknown but estimated at each user through feedback. An alternative scheme for distributed resource allocation between CRs incorporating aggregated interference control is analyzed in [11,12], where the authors propose a form of real-time multi-agent reinforcement learning, known as decentralized Q-learning [13], to manage the aggregated interference. The objective function to be minimized is an expected discounted cumulative cost related to the difference between the effective signal-to-noise plus interference ratio (SINR) and a target SINR, which has to be guaranteed to the primary system. This SINR is measured at some control points located in the protection contour of the primary network and it is fed back to the secondary base stations that adjust their transmit power consequently. One of the interesting aspects of such an approach is that it is model-free and does not require the knowledge of the transition probabilities of the underlying Markov process. Finally, Geirhofer et al. [14] propose an interference-aware resource allocation for OFDMA systems, based on the sensing and prediction of the ad hoc users from the infrastructure users.

Different from the previous studies, in this article we propose a Bayesian radio access method enabling (possibly multiple) FAP's to allocate power/bits over a time–frequency grid based on the current belief on the interference level, as obtained from previous sensing. Since the interference cannot be known in advance, we use a Bayesian approach and formulate the utility function as the expected value of the utility conditioned to previous measurements. The goal is to relax the requirement on sensing time and allocate resources over a certain number of future time slots, depending on the interference model and on our prediction capability.

The article is organized as follows. We consider first the radio access of a single FAP and we maximize the expected rate, averaged over the interference activity model, under a transmit power constraint. In this case, the solution can be found in closed form and it represents a sort of generalized water-filling algorithm, with water level depending on the interference activity probabilities. Then, we illustrate an alternative approach consisting in the minimization of the transmit power, subject to a constraint on the minimum average femto-user rate. Then, we generalize the proposed approaches to the multi-FAP scenario, where we analyze the interaction among FAPs using a game-theoretic approach. In particular, we consider first a purely competitive game, where each FAP adopts a purely selfish strategy. Since the competitive game might lead to inefficient Nash equilibria, we also propose a coordinated game where, thanks to the exchange of a few parameters through the backhaul link, the FAP's coordinate their action to improve upon inefficient Nash equilibria and maximize the sum-rate or minimize the sum-power.

Single-user Bayesian adaptive allocation

Femtocell networks are fully compliant with cellular standards. Given the current evolution of 3G systems, in this article we are concerned with an LTE system and the goal is to allocate power over a time–frequency grid adaptively, as a function of the current occupancy. This implies that the channel and interference power must be sensed at the beginning of each frame. Given the low mobility of indoor users, the channel can be assumed to be nearly constant over the frame. However, the interference from macro-users may vary along the frame depending on the macro-user activity. A correct power allocation across time and frequency would require a non-causal knowledge of the interference, which is of course unavailable. To circumvent this inconvenience, we propose a time–frequency resource allocation based on a Markov modeling of the interference activity over each frequency subchannel. More specifically, we assume that the activity of the macro users over the frequency subchannels is modeled as a set of statistically independent homogeneous DTMCs. The parameters of the statistical model are assumed constant within the frame, but they may vary over successive frames. Each FUE estimates the interference power and the transition probabilities of the interference activity over each subchannel and feeds this information back to the associated FAP, which computes the optimal power allocation over a time–frequency frame, following a Bayesian approach.

In this section, we assume that the interferers do not react to the strategy of the FAP of interest. In the subsequent sections, we will extend the study to the case where the other FAPs react to the choice of nearby, interfering FAPs, thus generating an iterative process, whose stability properties will properly be studied.

Markovian interference model

The interference activity over each frequency subchannel is modeled as a DTMC. We use the random binary variable $S_{k,m}$ to indicate the macro activity over the kth subchannel, at time m: $S_{k,m} = 1$ if the subchannel is busy, with interference power $\sigma_I^2(k, m)$, while $S_{k,m} = 0$, if the subchannel is idle. We consider different orders for the DTMC so that we can test the effect of the order on the performance of the proposed strategy. As an example, for

the order $L = 1, 2, 3$, we introduce the following transition probabilities.

$$p_{jl}^k = \Pr(S_{k,m} = l \mid S_{k,m-1} = j) \quad \text{for} \quad L = 1 \tag{1}$$

$$p_{ijl}^k = \Pr(S_{k,m} = l \mid S_{k,m-1} = j, S_{k,m-2} = i) \quad \text{for} \quad L = 2 \tag{2}$$

$$p_{rijl}^k = \Pr(S_{k,m} = l \mid S_{k,m-1} = j, S_{k,m-2} = i, S_{k,m-3} = r)$$

$$\text{for} \quad L = 3 \tag{3}$$

where k is the subchannel index and (j), (i, j), (r, i, j) are, respectively, the states for $L = 1, 2, 3$, i.e., all the binary sequence in $\{0, 1\}^L$. The probability of being in the state $i \in \{0, 1\}$ at time m over the kth subchannel is denoted with $\pi_i^{(k,m)} = \Pr(S_{k,m} = i)$. In the case of a first-order DTMC, starting from an initial time slot $m_0 = 1$, the probability $\pi_i^{(k,m)}$ can be obtained recursively as

$$\begin{pmatrix} \pi_0^{(k,m)} \\ \pi_1^{(k,m)} \end{pmatrix} = \begin{pmatrix} \omega_1^{(k)} & 1 - \mu_1^{(k)} \\ 1 - \omega_1^{(k)} & \mu_1^{(k)} \end{pmatrix} \begin{pmatrix} \pi_0^{(k,m-1)} \\ \pi_1^{(k,m-1)} \end{pmatrix} \tag{4}$$

$$m = 2, 3, \ldots,$$

where $\omega_1^{(k)} := p_{00}^{(k)}$, $\mu_1^{(k)} := p_{11}^{(k)}$, whereas the initial state $(\pi_0^{(k,1)}, \pi_1^{(k,1)})$ is obtained by observing the channel state at the time slot of index m_0. Equation (4) can be written in compact matrix form as

$$\mathbf{\Pi}^{(k,m)} = \mathbf{P}_1^{(k)} \mathbf{\Pi}^{(k,m-1)} \tag{5}$$

where $\mathbf{\Pi}^{(k,m)} = [\pi_0^{(k,m)}, \pi_1^{(k,m)}]^T$ and the entries $p_{jl}^{(k)}$ of the transition matrix $\mathbf{P}_1^{(k)}$ are given in (4). Let $\beta_{k,m} = \Pr(S_{k,m} = 0)$ and $\gamma_{k,m} = \Pr(S_{k,m} = 1)$ the probabilities that the channel k at time m is, respectively, idle and busy. Then, we can iteratively calculate them at time m from Equation (4) as $\beta_{k,m} = \pi_0^{(k,m)}$ and $\gamma_{k,m} = \pi_1^{(k,m)}$. The generalization to higher orders is straightforward and the formulas are reported in Appendix 1, for convenience. In Appendix 1, we also report the formulas used to estimate the transition probabilities from the observations.

Maximum expected rate optimization

Having introduced the interference model, our goal now is to find the bit/power allocation over an OFDM frame composed of N subcarriers and M consecutive time slots, in order to maximize the expected rate, taking into account the macro-users activity. The assumptions underlying the proposed approach are (1) the channels are affected by multipath, with time-invariant coefficients within each frame, supposed to be known at the transmitter side; (2) the activity of the interferers over each subchannel is modeled as a homogeneous DTMC of order L

and the transition probabilities are estimated by using the ML estimators discussed in Appendix 1; (3) the activities of the interferers over different channels are statistically independent of each other; (4) the power allocation of the interferers are independent of the power allocation of the FAP of interest.

The last assumption is made to distinguish this situation from the case where the interferers are themselves sensing the channel (interference) and adapting their strategy consequently. In this second case, each adaptive transmitter reacts to the strategies of the other, thus inducing an iterative process that must properly be studied. The first scenario, which is the subject of this section, is appropriate when the interferer is an MBS, for example. The second scenario is more appropriate to model the situation where there are a few nearby FAP's attempting to access the radio channel at the same time. The analysis of this scenario will be carried out in the next section by resorting to game-theoretic tools.

In the case where there is only one adaptive device, the FUE is supposed to measure the interference power from the macro network over each subchannel over a number of time slots that depend on the order of the Markov chain as well as on the accuracy of the estimation.

Based on the channel sensing, up to a current time slot m_0, our goal is to find out the optimal power allocation over a set of M successive time slots $m = m_0, m_0 + 1, \ldots, m_0 + M - 1$. Since the interference in the slots successive to the current one is not known, we follow a Bayesian approach. More specifically, our first optimization criterion is the maximization of the expected rate, conditioned to the current estimation of the interference power profile and of the Markov chain parameters, over each frequency subchannel. In formulas, our objective function is

$$\bar{r} = \frac{1}{M} \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^{N} \mathbb{E}_{S_{k,m}}\{r(S_{k,m})\} \tag{6}$$

where

$$r(S_{k,m}) = \begin{cases} \log\left(1 + \frac{p_{k,m}|H_k|^2}{\sigma_n^2(k)}\right) & \text{if } S_{k,m} = 0 \\ \log\left(1 + \frac{p_{k,m}|H_k|^2}{\sigma_n^2(k) + \sigma_I^2(k,m)}\right) & \text{if } S_{k,m} = 1 \end{cases}, \tag{7}$$

where $\sigma_n^2(k)$ denotes the variance of noise and H_k is the FAP channel transfer coefficient over the kth subchannel. The average rate is then

$$\bar{r} = \frac{1}{M} \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^{N} \left[\beta_{k,m} \log\left(1 + p_{k,m} a_n(k)\right) + \gamma_{k,m}\right.$$

$$\left. \times \log\left(1 + p_{k,m} a_I(k,m)\right)\right] \tag{8}$$

where $a_n(k) := |H_k|^2/\sigma_n^2(k)$ and $a_I(k,m) := |H_k|^2/(\sigma_n^2(k) + \sigma_I^2(k,m))$. Since the transition probabilities are not *a priori* known, they are estimated from the observations, using Equations (42), (44) or (45), depending on the most appropriate Markov order L. Knowing the transition probabilities, the occupancy probabilities $\beta_{k,m}$ and $\gamma_{k,m}$ at any time m, conditioned to the observation of the channel state at the first L time slots can easily be derived by using Equations (4), (39), (40). Then, denoting with $\boldsymbol{p}$ the (time–frequency) NM-dimensional power allocation vector, the max-rate optimization problem is formulated as follows

$$\max_{\boldsymbol{p}} \ \bar{r}(\boldsymbol{p})$$
$$\text{s.t.} \ \frac{1}{M}\sum_{m=m_0}^{m_0+M-1}\sum_{k=1}^{N} p_{k,m} \leq P_T \qquad (9)$$
$$0 \leq p_{k,m} \leq p^{\max}(k) \quad \forall k,m$$

where the upper limit $p^{\max}(k)$, $k = 1,\ldots,N$ represents a mask constraint useful to limit the transmit power over some prescribed channels, for example, the channels occupied by the MBSs. This is a convex problem, as $\bar{r}(\boldsymbol{p})$ is a concave function of $\boldsymbol{p}$ and the constraint set is convex. The optimum power vector $\boldsymbol{p}^*$ can be expressed in closed form by imposing the KKT conditions (see Appendix 2 for further details). The optimal power over the kth frequency subchannel, at time m, is

$$p_{k,m}^* = \left[\frac{-\tilde{b}_{k,m}+\sqrt{\tilde{b}_{k,m}^2 - 4\tilde{a}_{k,m}\tilde{c}_{k,m}}}{2\tilde{a}_{k,m}}\right]_0^{p^{\max}(k)} \qquad (10)$$

where $[x]_a^b = a$ if $x \leq a$, $[x]_a^b = b$ if $x \geq b$ and $[x]_a^b = x$ if $a < x < b$. The coefficients $\tilde{b}_{k,m}, \tilde{c}_{k,m}, \tilde{d}_{k,m}$ are related to $a_n(k)$ and $a_I(k,m)$ as follows

$$\tilde{a}_{k,m} = \lambda\, a_n(k) a_I(k,m)$$
$$\tilde{b}_{k,m} = \lambda\,[a_n(k) + a_I(k,m)] - a_n(k) a_I(k,m)$$
$$\tilde{c}_{k,m} = \lambda - a_n(k)\beta_{k,m} - a_I(k,m)\gamma_{k,m},$$

where λ is the Lagrange multiplier. Since the optimal powers $p_{k,m}^*$ are functions of λ, we can find this multiplier numerically as the solution of the power constraint $\sum_{m=m_0}^{m_0+M-1}\sum_{k=1}^{N} p_{k,m}^* = MP_T$. Expression (10) is a generalization of the well-known water-filling solution. Indeed, it can be shown that, by taking the limit for the transition probabilities going to 1 or 0, i.e., by turning the Markov chain into the degenerate case of a deterministic signal, Equation (10) converges to the water-filling solution.

Min-power optimization strategy

Since one of the most critical issues in femtocells is interference management, an alternative optimization procedure consists in minimizing the FAP transmit power,

under the constraint of guaranteeing the required rate over the link between the FAP and the associated FUE. This strategy was proposed, for example, in [15] assuming a static interference. Here, we generalize that approach to the case where the interference is dynamic and its activity evolves as a Markov chain, as described in the previous section. The objective now is to minimize the average transmit power across the N subchannels and over M consecutive time slots. Denoting with m_0 the index of the time slot where the interference power profile is measured, the goal is to allocate power over a set of consecutive slots, starting from m_0, i.e., for $m = m_0,\ldots,m_0 + M - 1$, under the constraint that the expected rate, conditioned to the observation on the initial L time slots, i.e., for $m = m_0 - L + 1,\ldots,m_0$, does not have to be smaller than a given value R_0.

The optimization problem can then be formulated as

$$\min_{\boldsymbol{p}} \quad \frac{1}{M}\sum_{k=1}^{N}\sum_{m=m_0}^{m_0+M-1} p_{k,m}$$
$$\text{subject to } \bar{R}(\boldsymbol{p}) \geq R_0 \qquad (11)$$
$$0 \leq p_{k,m} \leq p^{\max}(k), \quad \forall k = 1,\ldots,N,$$
$$m = 1,\ldots,M$$

where the expected rate is computed as in (8).

The minimization problem in (11) is a convex optimization problem, since the objective function is a linear (then convex) function of $\boldsymbol{p}$ and the set is convex. The solution can be written in closed form by exploiting the KKT conditions and following the same steps as in Appendix 2, the result is

$$p_{k,m}^* = \left[\frac{-b(k,m)+\sqrt{b(k,m)^2 - 4a(k,m)c(k,m)}}{2a(k,m)}\right]_0^{p^{\max}(k)}$$
$$\forall k = 1,\ldots,N, \quad m = 1,\ldots,M \qquad (12)$$

with

$$a(k,m) = a_n(k)\, a_I(k,m)$$
$$b(k,m) = a_n(k) + a_I(k,m) - \lambda a_n(k)\, a_I(k,m) \qquad (13)$$
$$c(k,m) = 1 - \lambda\left[a_n(k)\,\beta_{k,m} + a_I(k,m)\,\gamma_{k,m}\right]$$

where the Lagrange multiplier λ is found numerically in order to satisfy the rate constraint $\bar{R}(\boldsymbol{p}^*) = R_0$.

One important difference between the min-power and the max-rate problems is that, in the min-power case, the feasible set could be empty. If this happens, it means that the rate requirement is too high to be accommodated. Hence, either the rate requirement is lowered until the feasible set becomes non-empty, or the user is not admitted. This protocol is handled by the call admission control.

Multi-FAP case: maximum expected rate game

In a scenario containing multiple nearby FAP's implementing the radio access according to the adaptive strategy described in the previous section, each FAP may react to the power allocation of nearby FAP's, by changing its own power allocation and so on. This interaction induces an iterative mechanism whose convergence properties have to be carefully studied. The problem can be studied using the theoretical tools of game theory, which is well suited for this kind of multi-objective decision problem. In particular, given the existence of a wired backhaul connecting the FAP's, we will consider a purely competitive game, where each FAP (player) seeks to optimize its own utility function, irrespective of the other FAP's performance, and a coordination game, where nearby FAP's exchange some parameters to improve performance with respect to the purely competitive case.

Denoting again with the binary variable $S_{k,m}$ the macro-interference activity over the channel k at time m and with $\Pr(S_{k,m} = 0) = \beta_{k,m}$ and $\Pr(S_{k,m} = 1) = \gamma_{k,m}$ the probabilities that channel k is idle or busy, at time m, the expected rate of the qth FAP is[a]

$$\bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) = \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^{N} \Big[\beta_{k,m} \log \Big(1 + p_{k,m}^q a_n^q(k,m) \Big) \\ + \gamma_{k,m} \log \Big(1 + p_{k,m}^q a_I^q(k,m) \Big) \Big],$$

$$(14)$$

with $q = 1, \ldots, Q$, and

$$a_n^q(k,m) = \frac{|H_k^{qq}|^2}{\sigma_{n,q}^2(k) + \sum_{r \in \mathcal{N}_q} p_{k,m}^r |H_k^{rq}|^2}, a_I^q(k,m)$$

$$= \frac{|H_k^{qq}|^2}{\sigma_{n,q}^2(k) + \sum_{r \in \mathcal{N}_q} p_{k,m}^r |H_k^{rq}|^2 + \sigma_{I^q}^2(k,m)}$$

$$(15)$$

where $\mathcal{N}_q$ is the set of neighbors of the qth FAP and H_k^{rq} is the channel transfer function of the kth subchannel between the rth transmitter and the qth receiver. The probabilities $\beta_{k,m}$ and $\gamma_{k,m}$ evolve in time according to a Markov chain of order L. We assume, as in the previous section that the allocation over M consecutive time slots is carried out on the basis of the observation of a number of initial time slots equal to the order of the Markov chain, i.e., L.

It is worth noticing that the major difference between the expected rate in (14) with respect to (8) is that now the interference is composed of two contributions: the dynamic interference of the macro-users, whose activities evolve as Markov chains but whose power profile, when on, is fixed, and the interference from the other FAP's, whose activity is always on, but whose power profile, described by the vectors $p_{k,m}^r$ evolve as a response to the choices of the other FAP's.

Denoting by $\Omega = \{1 \ldots Q\}$ the set of Q players, with P_q the maximum transmit power over a frame and with $p_q^{\max}(k)$ the mask constraint over each subcarrier, the problem can be cast as a game, i.e.,

$$\mathcal{G}_1 : \quad \begin{array}{c} \max\limits_{\boldsymbol{p}_q} \quad \bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) \\ \text{subject to } \boldsymbol{p}_q \in \mathcal{P}_q \end{array} \quad \forall q \in \Omega \qquad (16)$$

where the feasible set of FAP q is

$$\mathcal{P}_q = \Bigg\{ \boldsymbol{p}_q \in \mathbb{R}^{NM \times 1} : \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^{N} p_{k,m}^q \leq P_q, 0 \leq p_{k,m}^q \\ \leq p_q^{\max}(k), \forall k \in \{1, \ldots, N\}, m \in \{m_0, \ldots, m_0+M-1\} \Bigg\}.$$

$$(17)$$

Since the objective function in (16) is strictly concave in $\boldsymbol{p}_q \in \mathcal{P}_q$, for any given $\boldsymbol{p}_{-q}$, and the feasible set $\mathcal{P}_q$ is compact and convex, game $\mathcal{G}_1$ admits a non-empty solution set for any set of channels and transmit power constraints of the users. In [16], we reformulated this game as a *Variational Inequality* [17] and we applied the iterative gradient projection algorithm to solve it, deriving sufficient conditions for its convergence to a Nash Equilibrium (NE).

Game $\mathcal{G}_1$ may possess multiple equilibria, which may not be Pareto-efficient.[b] To improve upon the performance of the NE of purely competitive game $\mathcal{G}_1$, we can modify the utility function of each user in order to induce the players to incorporate a social utility function, rather than being purely selfish. For example, in [18,19] it has been proposed to modify the utility function of each player so as to maximize the sum of all users' rates. In principle, this change should require a centralized solution. Nevertheless, Huang et al. [18] showed that the solution of the sum-rate game can be still achieved in decentralized form, provided that the players exchange some parameters, the so-called *prices*. These parameters induce a penalty on each player utility proportional to the rate decrease that each player strategy induces on the other players. Introducing pricing mechanisms in femtocell networks is possible, thanks to the existence of the backhaul link, which allows the exchange of prices among FAP's. Furthermore, we will show next that every FAP needs to exchange pricing coefficients only with its neighbors, thus keeping the amount of extra signaling limited.

Generalizing the approach proposed in [18] to our Bayesian formulation, we introduce the *price* coefficient:

$$\pi_{k,m}^r := -\frac{\partial \bar{R}_r(\boldsymbol{p})}{\partial I_{k,m}^r(\boldsymbol{p}_{-r})} \tag{18}$$

with $I_{k,m}^r(\boldsymbol{p}_{-r}) := \sum_{i \in \mathcal{N}_r} p_{k,m}^i |H_k^{ir}|^2$. These coefficients are proportional to the marginal decrease of user r's *expected* rate resulting from an increase of the qth node's transmit power, as $\frac{\partial R_r(\boldsymbol{p})}{\partial p_{k,m}^q} = -\pi_{k,m}^r \frac{\partial I_{k,m}^r}{\partial p_{k,m}^q} = -\pi_{k,m}^r |H_k^{qr}|^2$. The incorporation of the pricing mechanism leads to the new game[c]:

$$\mathcal{G}_2: \max_{\boldsymbol{p}_q} \bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) - \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^N p_{k,m}^q \left(\sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 \right)$$
$$\text{s.t.} \quad \boldsymbol{p}_q \in \mathcal{P}_q. \tag{19}$$

Each local problem is convex, hence the KKT conditions lead to power coefficients $p_{k,m}^q$ of each FAP that, within the interval $[0, p_q^{\max}(k))$, must satisfy the following equation

$$\tilde{a}^q(k,m)\left(p_{k,m}^q\right)^2 + \tilde{b}^q(k,m)p_{k,m}^q + \tilde{c}^q(k,m) = 0 \tag{20}$$

where, denoting with ν_q the Lagrangian multiplier, we have set

$$\tilde{a}^q(k,m) = \left(\nu_q + \sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 \right) a_n^q(k,m)a_I^q(k,m)$$

$$\tilde{b}^q(k,m) = \left(\nu_q + \sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 \right) (a_n^q(k,m) + a_I^q(k,m))$$
$$- a_n^q(k,m)a_I^q(k,m)$$

$$\tilde{c}^q(k,m) = \nu_q + \sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 - \left(a_n^q(k,m)\beta_{k,m} \right.$$
$$\left. + a_I^q(k,m)\gamma_{k,m} \right) \tag{21}$$

We can verify that, $\forall \nu_q > 0$, we have $\tilde{b}^q(k,m)^2 - 4\tilde{a}^q(k,m)\tilde{c}^q(k,m) \geq 0$, and the only solution is

$$\tilde{p}_{k,m}^b = \frac{-\tilde{b}^q(k,m) + \sqrt{\tilde{b}^q(k,m)^2 - 4\tilde{a}^q(k,m)\tilde{c}^q(k,m)}}{2\tilde{a}^q(k,m)}. \tag{22}$$

More specifically, we get

$$p_{k,m}^q = \begin{cases} 0 & \text{if } \nu_q + \sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 \geq a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m} \\ \tilde{p}_{k,m}^b & \text{if } \nu_q + \sum_{r \in \mathcal{N}_q} \pi_{k,m}^r |H_k^{qr}|^2 < a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m} \end{cases} \tag{23}$$

and the optimal power allocation vector is $p_{k,m}^{q*} = \left[\tilde{p}_{k,m}^b \right]_0^{p_q^{\max}(k)}$ where the multiplier ν_q is chosen in order

to satisfy the constraint $\sum_{k=1}^N \sum_{m=m_0}^{m_0+M-1} \left[\tilde{p}_{k,m}^b \right]_0^{p_q^{\max}(k)} = P_q$. The previous solution assumes, for each player, that the powers used by the other players are given.

In practice, the game evolves with each FAP reacting to the choices of the other FAPs. It is then fundamental to prove the convergence of this iterative mechanism. In the following, we present a version of the so-called Modified Asynchronous Distributed Pricing algorithm (MADP) proposed in [19], adapted to our formulation. To this purpose, it is useful to rewrite (19) introducing a unique index h so that the entries of the power vector $\boldsymbol{p}_q$ are p_h^q for $h = 1, \ldots, NM$. Then, defining the quantities

$$\text{SNIR}_h^{\beta^q} := \frac{p_h^q |H_h^{qq}|^2}{\sigma_{n,q}^2(h) + \sum_{r \in \mathcal{N}_q} p_h^r |H_h^{rq}|^2},$$

$$\text{SNIR}_h^{\gamma^q} := \frac{p_h^q |H_h^{qq}|^2}{\sigma_{n,q}^2(h) + \sum_{r \in \mathcal{N}_q} p_h^r |H_h^{rq}|^2 + \sigma_{I_q}^2(h)} \tag{24}$$

we can derive the qth user best response as

$$p_h^{q*} = \frac{2c_h^q}{\sum_{r \in \mathcal{N}_q} \pi_h^r |H_h^{qr}|^2 + \nu_q - \eta_h^q} - p_h^q \quad \forall h = 1, \ldots, MN, \tag{25}$$

where $c_h^q = \frac{\beta_h^q \text{SNIR}_h^{\beta^q}}{1 + \text{SNIR}_h^{\beta^q}} + \frac{\gamma_h^q \text{SNIR}_h^{\gamma^q}}{1 + \text{SNIR}_h^{\gamma^q}}$ and ν_q and η_h^q are the Lagrangian multipliers. Given this setting, the modified MADP algorithm is illustrated below.

Algorithm 1 MADP algorithm
Each FAP performs its allocation over M consecutive time slots from m_0 to $m_0 + M - 1$.

Before performing its allocation each FAP has to observe q_s samples from $m_0 - q_s + 1$ to m_0, in order to estimate the transition probabilities of the underlying Markov chain.

S.0: Each FAP q chooses an initial power profile in the set $\mathcal{P}_q$ and set $n = 0$;
S.1: Each FAP computes its interference prices $\pi_h^q(n)|H_k^{iq}|^2$ for $h = 1, \ldots, MN$ and sends them to its neighbors with index $i \in \mathcal{N}_q$;
S.2: At each time n, each FAP updates its power profile so as to maximize its utility function $\bar{R}_q$, given the other FAP's power profiles $\boldsymbol{p}_{-q}$ and price vectors according to
$p_h^q(n+1) = p_h^q(n) + \alpha_q(n)\left(p_h^{q*} - p_h^q(n)\right)$ for $h = 1, \ldots, MN$, where p_h^{q*} is given by (25);
S.3: Set $n = n + 1$, go to step S.1 and repeat until convergence is reached.

Following similar arguments as [19], we proved in Appendix 3 that there exists a small enough step size values $\alpha_q(n)$ for which the MADP algorithm converges monotonically to a fixed point.

Multi-FAP case: min-power game

As with the max-rate game, let us consider now the generalization of the min-power algorithm to the multi-FAP case. The utility of each player is now the total transmit power over the N subchannels and the M time slots

$$u_q(\boldsymbol{p}_q) = \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} p_{k,m}^q \tag{26}$$

and the constraint is that the expected rate for each FAP, conditioned to the initial observations, be not smaller than a given value R_q^0. The feasible set is now

$$\widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q}) = \Big\{\boldsymbol{p}_q \in \mathbb{R}^{NM \times 1} : \bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) \geq R_q^0, 0 \leq p_{k,m}^q$$
$$\leq p_q^{\max}(k), \forall k = 1, \dots, N, m = m_0, \dots, m_0$$
$$+ M - 1\Big\} \tag{27}$$

and the game is

$$\widetilde{\mathcal{G}}_1 = \{\Omega, \{\widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q})\}_{q\in\Omega}, \{u_q(\boldsymbol{p}_q)\}_{q\in\Omega}\}. \tag{28}$$

The optimal strategy for each player amounts to solving the following optimization problem

$$(\widetilde{P}_1) \quad \begin{aligned} \min_{\boldsymbol{p}_q} & \quad u_q(\boldsymbol{p}_q) \\ \text{subject to } & \quad \boldsymbol{p}_q \in \widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q}) \end{aligned} \quad \forall q \in \Omega. \tag{29}$$

The minimization problem in (29) for each player q, given the strategies of the others, is a convex optimization problem, since the objective function is a linear (then convex) function of $\boldsymbol{p}_q$ and the set $\widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q})$, given the power vector $\boldsymbol{p}_{-q}$ of the other players, is a convex set. Imposing the KKT conditions, as in the single FAP case, the solution can be expressed in closed form as $\boldsymbol{p}_q^* = \boldsymbol{g}(\boldsymbol{p}_{-q})$ whose entries are (see Appendix 4)

$$p_{k,m}^{q*} = \left[\boldsymbol{g}\left(\boldsymbol{p}_{-q}\right)\right]_{k,m}$$
$$= \left[\frac{-b^q(k,m) + \sqrt{b^q(k,m)^2 - 4a^q(k,m)c^q(k,m)}}{2a^q(k,m)}\right]_0^{p_q^{\max}(k)}$$
$$\times \forall k = 1, \dots, N, \ m = 1, \dots, M \tag{30}$$

with

$$a^q(k,m) = a_n^q(k,m)\, a_I^q(k,m)$$

$$b^q(k,m) = a_n^q(k,m) + a_I^q(k,m) - \lambda_q a_n^q(k,m)\, a_I^q(k,m)$$

$$c^q(k,m) = 1 - \lambda_q \left[a_n^q(k,m)\, \beta_{k,m} + a_I^q(k,m)\, \gamma_{k,m}\right] \tag{31}$$

where the Lagrange multiplier λ_q is chosen in order to satisfy the rate constraint $\bar{R}_q(\boldsymbol{p}_q^*, \boldsymbol{p}_{-q}) = R_q^0$. However, now the overall feasible set is not jointly convex with respect to the power vectors of all the users, i.e., it is not convex in $\boldsymbol{p} = (\boldsymbol{p}_q)_{q=1}^Q$. This makes the study of this game much harder than the standard NE problem. Nevertheless, game $\widetilde{\mathcal{G}}_1$ is a Generalized Potential Game (GPG) [20], with a potential Φ equal to the sum power. In such a case, the existence of a NE of the potential game can be proved directly by the existence of a maximum of the potential function Φ on the set $\widetilde{\mathcal{X}}$ of the game. To exploit the theory of GPG, we must prove that game $\widetilde{\mathcal{G}}_1$ admits a non-empty feasible set. The proof of this result is given in Appendix 5, containing the sufficient conditions under which the feasible set of the game $\widetilde{\mathcal{G}}_1$, i.e.,

$$\widetilde{\mathcal{X}} = \Big\{\boldsymbol{p} \in \mathbb{R}^{NMQ} : \bar{R}_q(\boldsymbol{p}) \geq R_q^0, 0 \leq p_{k,m}^q \leq p_q^{\max}(k),$$
$$\times \forall k, m, \forall q \in \Omega\Big\} \tag{32}$$

is compact and non-empty. Hence, game $\widetilde{\mathcal{G}}_1 = \{\Omega, \{\widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q})\}_{q\in\Omega}, \{u_q(\boldsymbol{p}_q)\}_{q\in\Omega}\}$ is a GPG with potential function $\Phi(\boldsymbol{p})$ the sum of the objective functions of all players, i.e., $\Phi(\boldsymbol{p}) = \sum_{q=1}^Q u_q(\boldsymbol{p}_q)$.

Nevertheless, being a potential game does not guarantee the equilibrium to be efficient. Hence, as with the max-rate game, efficiency can be improved by introducing pricing. The introduction of pricing leads to a modified game which can be cast as

$$\widetilde{\mathcal{G}}_2 : \begin{aligned} \min_{\boldsymbol{p}_q} & \quad u_q(\boldsymbol{p}_q) + \sum_{m=m_0}^{m_0+M-1} \sum_{k=1}^N \left(\sum_{s\in\mathcal{N}_q} \lambda_s \pi_{k,m}^s |H_k^{qs}|^2\right) \\ & \quad \times p_{k,m}^q \forall q \in \Omega \\ \text{subject to } & \quad \boldsymbol{p}_q \in \widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q}) \end{aligned} \tag{33}$$

where λ_s is the Lagrangian multiplier of user s relative to the rate constraint. The pricing coefficients $\pi_{k,m}^s$ are defined in the same way as in the max-rate case.

Numerical results

In this section, we present some numerical results in order to assess the performance of the algorithms proposed in the previous sections. Let us start with the single-FAP case. In all the simulation results, we have considered Rayleigh fading frequency-selective channels where the number of resolvable paths is 4, each one with unit variance. The number of subcarriers N is set to 12, as in LTE Primary Resource Block (PRB). In Figure 2 we show the average rate per OFDM symbol as a function of the allocated time slots obtained in the max-rate problem. The
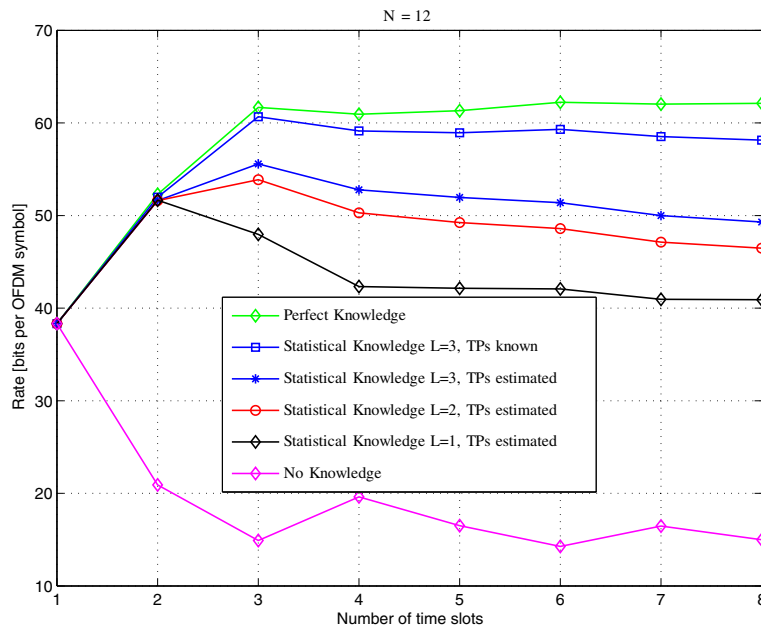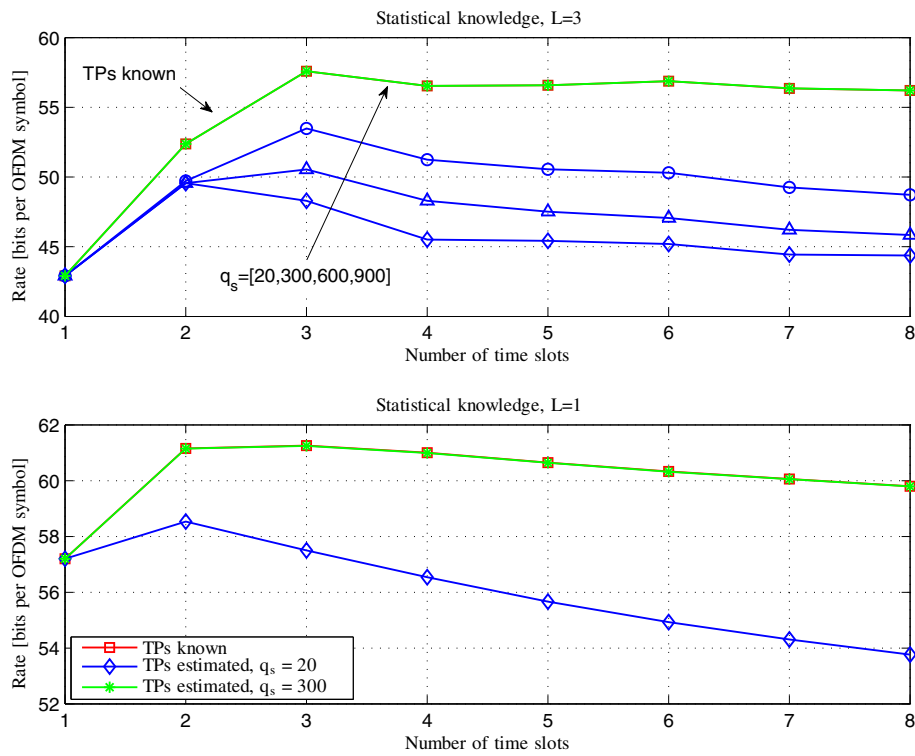
Figure 2 Single FAP max-rate: rate versus number of time slots.

simulation results have been averaged over 100 independent channel and Markov chain realizations. The different curves indicate the rate obtained by assuming different kinds of knowledge of the interference: the green curve assumes perfect knowledge of the future evolution of the macro-user activity and it is used as a benchmark case; the pink curve assumes that the interference power level over each subchannel is equal to the value observed in the first time slot; all other curves refer to the proposed algorithm, where we observe the channels in the first slot and allocate power using our proposed method. The different curves refer to different Markov orders (from $L = 1$ to $L = 3$). We also compare the case where the transition probabilities are perfectly known and the case where the probabilities are estimated from the data. The interesting behavior is that, as the order increases, our approach is able to approach the ideal case where the interference activity is non-causally known. The price to be paid is the loss of performance resulting from the estimation of the Markov parameters. Further developments could incorporate some kind of reinforcement learning to be able to allocate resources without necessarily estimating the transition probabilities, as proposed in [12], for example.

In Figure 3, we report the optimal rate of our algorithm versus the number of time slots, for different Markov orders, $L = 3$ in the upper subplot and $L = 1$ in the lower. We have considered different numbers of samples q_s used to estimate the transition probabilities. Of course, the higher is the Markov order, the greater is the number of parameters to estimate. In fact, we can notice that, for the same number of samples q_s used for the estimate,

the performance loss with respect to the ideal case of perfect knowledge of the macro-user activity (curve with red squared markers) is much higher for $L = 3$ than for $L = 1$. The aim of Figure 4 is to show the cumulative rate versus the number of slots m_0 used for the recursive estimation of the transition probabilities assuming $M = 3$ and by modeling the macro-user activity as a first-order Markov chain. We observe that, after less than 50 time slots, the performance gets very close to the asymptotic case. Finally, in Figure 5, we show the performance of the min-power allocation strategy. The utility function is the SNR [dB] at the FUE receiver obtained for different number of allocated time slots. As in the max-rate case, it is evident the advantage of increasing the order of statistical knowledge (from $L = 1$ to $L = 3$) and it can be observed a gain of about 4 dB with respect to the case where no knowledge about the macro-user activity has been assumed.

Finally, we provide some numerical examples to assess the performance of the proposed approaches (max-rate and min-power) in the multiuser case. The reference scenario is composed of one MBS and $Q = 10$ FAPs randomly distributed over a square area. The MBS activity is modeled as a third-order Markov chain and the results have been averaged over 50 independent Markov chain realizations. In Figure 6, we have reported the users' sum-rate versus the iteration index for the maximum expected rate game $\mathcal{G}_1$ in order to test the convergence of the algorithm. It can be observed that it converges in a few iterations. In Figures 7 and 8, we depict the FAPs' sum-rate versus the number of allocated time slots. In

Figure 3 Single FAP max-rate: rate versus number of time slots for different number of samples used for TPs estimation.

particular, Figure 7 refers to the purely competitive maximum expected rate game $\mathcal{G}_1$, while Figure 8 refers to the modified pricing game $\mathcal{G}_2$. In both cases, we assumed the same maximum transmit power per FAP. The three different curves in each figure indicate the sum-rate obtained by assuming perfect (non-causal) knowledge of the macro-user activity, no knowledge at all (thus assuming the interference to remain equal to the values observed in

the first slots of each frame), or only knowledge of the Markov parameters. Both figures show that acquiring a statistical knowledge (estimation) of the interference activity parameters (Markov transition rates) yields a performance advantage over the case with no information and brings the performance close to the ideal case of perfect non-causal knowledge of the interference activity. Of course, as time evolves, there is a mismatch between
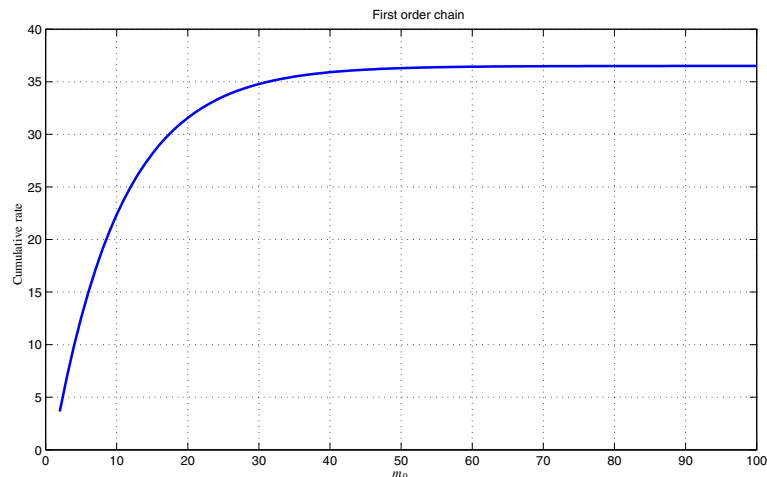
Figure 4 Single FAP max-rate: cumulative rate versus number of time slots used for the estimation of the transition probabilities.
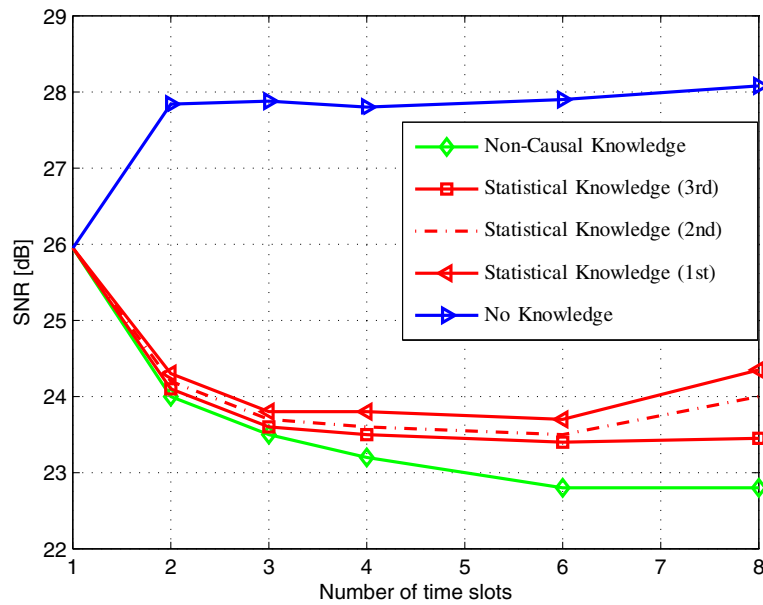
**Figure 5 Single FAP min-power: SNR versus number of time slots.**

what is predicted and the real interference so that the performance improvement tends to decrease in time. Furthermore, comparing Figures 7 and 8, it is evident that the gain achieved with the introduction of pricing.

Considering the same scenario, in Figures 9 and 10 we report the simulation results corresponding to our proposed minimum power games, $\widetilde{\mathcal{G}}_1$ and $\widetilde{\mathcal{G}}_2$. Figure 9 refers to the min-power game with no pricing, while Figure 10

refers to the game including pricing. The curves show the average SNR per FUE as a function of the number of allocated time slots. The expected target rate R_q^0 in both cases is set to 3 bps for each user. From both Figures 9 and 10 we can verify that, also in this case, the simple statistical knowledge of the transition rates yields performance close to the ideal case where the interference activity is non-causally known. Observe that the curve referring to
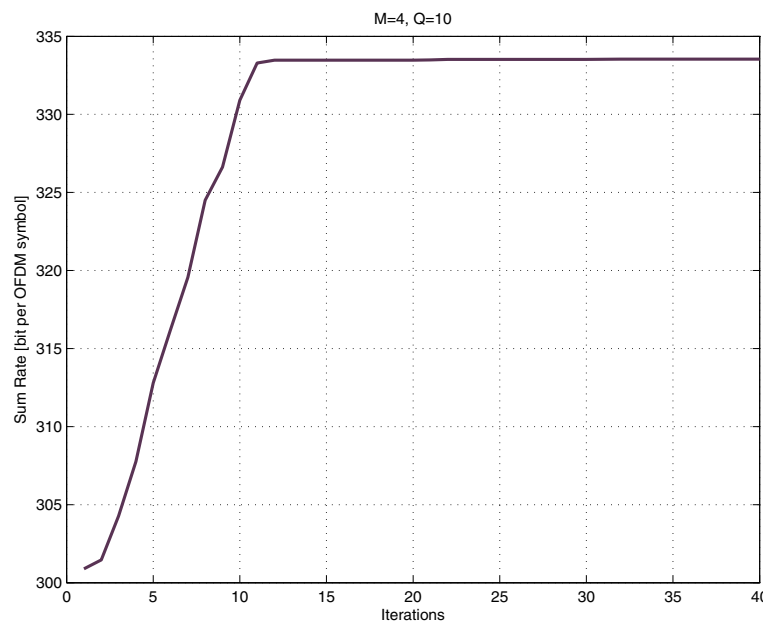

Figure 6 Multi-FAP max-rate: convergence of the game.
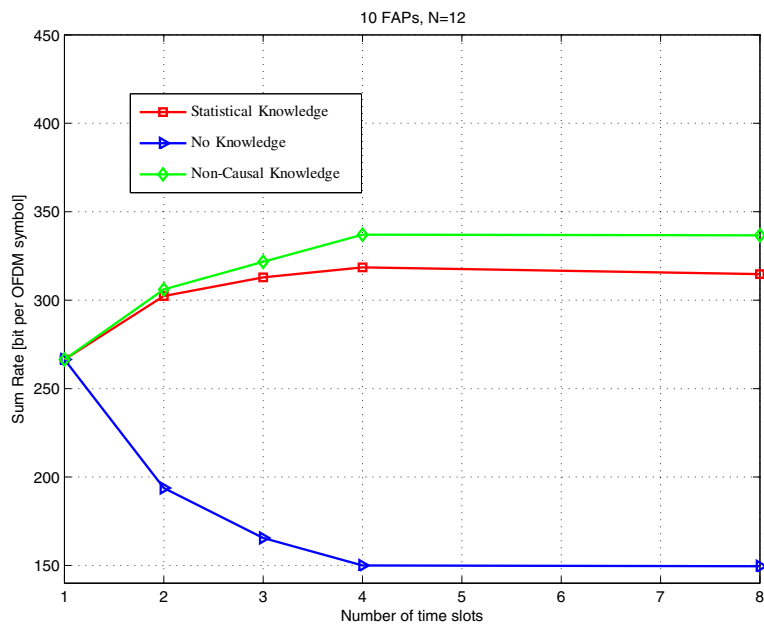
Figure 7 Multi-FAP max-rate: sum-rate versus number of time slots for the maximum expected rate game without pricing.

the non-causal knowledge tends to have zero slope asymptotically and the statistical knowledge curve presents a performance gain which tends to decrease as time evolves due to the mismatch between what is predicted and the real interference. In both Figures 9 and 10, the advantage of the statistical approach with respect to the case where there is no knowledge is considerable and by comparing
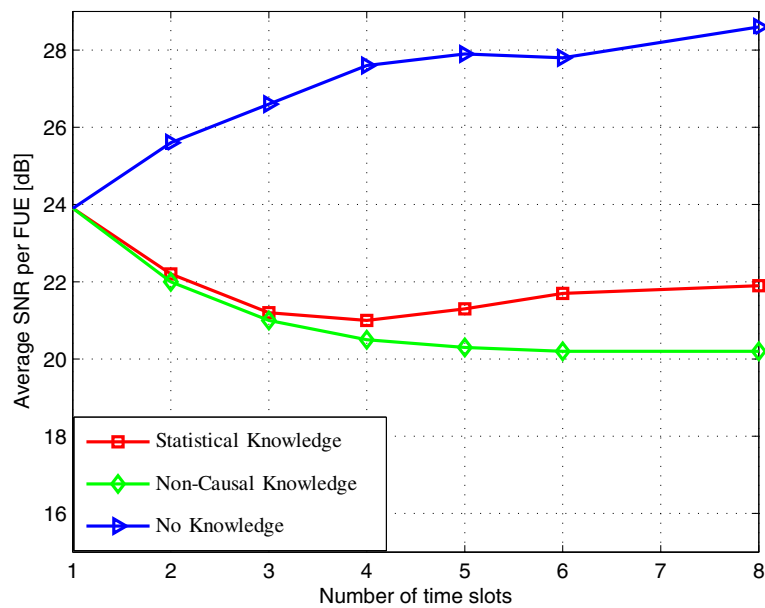
the two figures it is evident the performance gain due to the introduction of a pricing mechanism.

Conclusion

In conclusion, in this article we have shown how the estimation of the interference statistical parameters can be
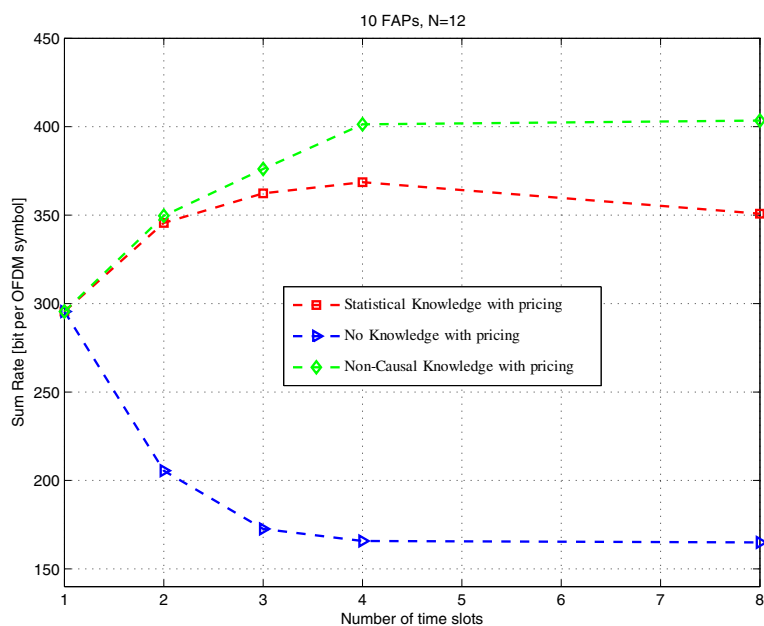
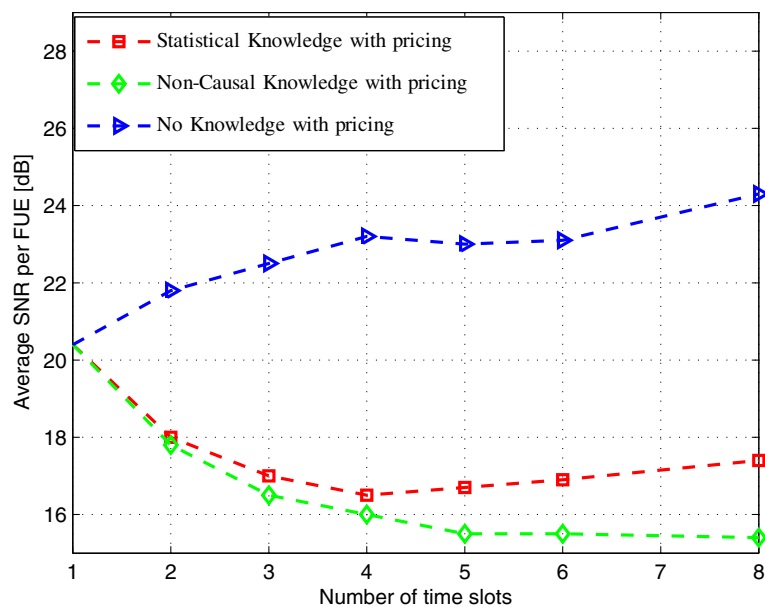**Figure 8 Multi-FAP max-rate: sum-rate versus number of time slots for the max-rate game with pricing.**

Figure 9 Multi-FAP min-power: average SNR versus number of time slots for the min-power game without pricing.

beneficial to improve the performance of a power allocation technique, provided that the statistical model fits the real data. In this study, we assume that the transition probabilities are estimated from the data. An interesting future direction consists in incorporating methods which do not really require such an estimation, but acquire the proper behavior through reinforcement learning. The interesting part of our method is that, for a given estimation of the transition probabilities, the power allocation across the

set of time slots/frequency channels is found in closed form. This is indeed useful to save convergence time with respect to gradient-based techniques.

The other important contribution of this article is the decentralized approach for resource allocation based on game theory, in the case where the interference is dynamically varying. Our Bayesian formulation of the game provides interesting results in such an uncertain environment. Finally, the introduction of coordination among

Figure 10 Multi-FAP min-power: average SNR versus number of time slots for the min power game with pricing.

FAP's based on the exchange of a few parameters (prices) through the backhaul link has been shown to provide significant performance gains with respect to the purely competitive game.

Further developments should incorporate a robust approach for the situation where the interference statistical model is not known or the statistical parameters are time-varying. One more critical aspect is the availability of the backhaul link for the exchange of prices. Since such a link is affected by random delays, it may be useful to incorporate robust mechanisms to cope with the situation where the price does not arrive within a maximum tolerable delay.

Appendix 1

Extending the Markov model used in (5) to higher-order chains, we can write the time evolution of the occupancy probabilities conditioned to the observations of the channel state at the first L time slots as

$$\mathbf{\Pi}^{(k,m-L+1,\ldots,m)} = \mathbf{P}_L^{(k)} \mathbf{\Pi}^{(k,m-L,m-L+1,\ldots,m-1)} \quad (34)$$

for $m = L+1, \ldots, \infty$ and given the initial state $\mathbf{\Pi}^{(k,1,\ldots,L)}$. More specifically, the entries of the 2^L-dimensional vector $\mathbf{\Pi}^{(k,m-L+1,\ldots,m)}$ are defined as

$$\pi_{ij}^{(k,m-1,m)} = \Pr(S_{k,m-1} = i, S_{k,m} = j) \quad \text{for} \quad L = 2,$$
$$\forall\, (i,j) \in \{0,1\}^2 \quad (35)$$

$$\pi_{rij}^{(k,m-2,m-1,m)} = \Pr(S_{k,m-2} = r, S_{k,m-1} = i, S_{k,m} = j)$$

$$\text{for} \quad L = 3, \quad \forall\, (r,i,j) \in \{0,1\}^3 \quad (36)$$

while the transition matrix $\mathbf{P}_L^{(k)}$ are expressed, respectively, for $L = 2$ as

$$\mathbf{P}_2^{(k)} = \begin{pmatrix} \omega_2^{(k)} & 0 & \theta_2^{(k)} & 0 \\ 1-\omega_2^{(k)} & 0 & 1-\theta_2^{(k)} & 0 \\ 0 & v_2^{(k)} & 0 & 1-\mu_2^{(k)} \\ 0 & 1-v_2^{(k)} & 0 & \mu_2^{(k)} \end{pmatrix} \quad (37)$$

with $\omega_2^{(k)} = p_{000}^{(k)}$, $\theta_2^{(k)} = p_{100}^{(k)}$, $v_2^{(k)} = p_{010}^{(k)}$, $\mu_2^{(k)} = p_{111}^{(k)}$ and for $L = 3$ as

$$\mathbf{P}_3^{(k)} =$$
$$\begin{pmatrix} \omega_3^{(k)} & 0 & 0 & 0 & \lambda_3^{(k)} & 0 & 0 & 0 \\ 1-\omega_3^{(k)} & 0 & 0 & 0 & 1-\lambda_3^{(k)} & 0 & 0 & 0 \\ 0 & v_3^{(k)} & 0 & 0 & 0 & \psi_3^{(k)} & 0 & 0 \\ 0 & 1-v_3^{(k)} & 0 & 0 & 0 & 1-\psi_3^{(k)} & 0 & 0 \\ 0 & 0 & \eta_3^{(k)} & 0 & 0 & 0 & \theta_3^{(k)} & 0 \\ 0 & 0 & 1-\eta_3^{(k)} & 0 & 0 & 0 & 1-\theta_3^{(k)} & 0 \\ 0 & 0 & 0 & 1-\gamma_3^{(k)} & 0 & 0 & 0 & 1-\mu_3^{(k)} \\ 0 & 0 & 0 & \gamma_3^{(k)} & 0 & 0 & 0 & \mu_3^{(k)} \end{pmatrix}$$
$$(38)$$

where $\omega_3^{(k)} = p_{0000}^{(k)}$, $\lambda_3^{(k)} = p_{1000}^{(k)}$, $v_3^{(k)} = p_{0010}^{(k)}$, $\psi_3^{(k)} = p_{1010}^{(k)}$, $\eta_3^{(k)} = p_{0100}^{(k)}$, $\theta_3^{(k)} = p_{1100}^{(k)}$, $\gamma_3^{(k)} = p_{0111}^{(k)}$, $\mu_3^{(k)} = p_{1111}^{(k)}$. Hence, assuming for simplicity of notation $\pi_{ij}^{(k,m-2,m-1)} = \pi_{ij}^{(k)}$ and $\pi_{rij}^{(k,m-3,m-2,m-1)} = \pi_{rij}^{(k)}$, the probabilities that the kth subchannel is idle (busy) at time m, i.e., $\beta_{k,m}$ ($\gamma_{k,m}$) will be at time m for $L = 2, 3$, respectively,

$$\beta_{k,m} = \omega_2^{(k)}\pi_{00}^{(k)} + \theta_2^{(k)}\pi_{10}^{(k)} + v_2^{(k)}\pi_{01}^{(k)} + (1-\mu_2^{(k)})\pi_{11}^{(k)}$$

$$\gamma_{k,m} = (1-\omega_2^{(k)})\pi_{00}^{(k)} + (1-\theta_2^{(k)})\pi_{10}^{(k)} + (1-v_2^{(k)})\pi_{01}^{(k)} \quad (39)$$
$$+ \mu_2^{(k)}\pi_{11}^{(k)}$$

and

$$\beta_{k,m} = \omega_3^{(k)}\pi_{000}^{(k)} + \lambda_3^{(k)}\pi_{100}^{(k)} + v_3^{(k)}\pi_{001}^{(k)} + \psi_3^{(k)}\pi_{101}^{(k)}$$
$$+ \eta_3^{(k)}\pi_{010}^{(k)} + \theta_3^{(k)}\pi_{110}^{(k)} + \left(1-\gamma_3^{(k)}\right)\pi_{011}^{(k)}$$
$$+ \left(1-\mu_3^{(k)}\right)\pi_{111}^{(k)}$$

$$\gamma_{k,m} = \left(1-\omega_3^{(k)}\right)\pi_{000}^{(k)} + \left(1-\lambda_3^{(k)}\right)\pi_{100}^{(k)} + \left(1-v_3^{(k)}\right)\pi_{001}^{(k)}$$
$$+ \left(1-\psi_3^{(k)}\right)\pi_{101}^{(k)} + \left(1-\eta_3^{(k)}\right)\pi_{010}^{(k)} + \left(1-\theta_3^{(k)}\right)\pi_{110}^{(k)}$$
$$+ \gamma_3^{(k)}\pi_{011}^{(k)} + \mu_3^{(k)}\pi_{111}^{(k)}.$$
$$(40)$$

The transition probabilities of a Markov chain of arbitrary order can be estimated from the observed data using the maximum likelihood strategy, as suggested for example in [21,22]. To simplify the description of the estimator we focus on a first-order Markov chain, but the extension to higher orders is straightforward. Let us assume that a set of m states, namely, $s^m \equiv s_{k,1}, \ldots, s_{k,m}$, are observed. The probability of observing a specific sequence of states is

$$\Pr\left(S^m = s^m\right) = \prod_{l=2}^{m} \Pr\left(S_{k,l} = s_{k,l}|S_{k,l-1} = s_{k,l-1}\right)$$
$$\times \Pr\left(S_{k,1} = s_{k,1}\right). \quad (41)$$

Let $n_{ij}(l)$ denote the number of times the state i at time $l-1$ switches to state j at time l. It has been proved in [21] that for a stationary Markov chain, the set $n_{ij} = \sum_{l=2}^{m} n_{ij}(l)$ forms a set of sufficient statistics. Furthermore, the maximum likelihood estimator based on the set s^m is

$$\hat{p}_{ij}(m) = \frac{\sum_{l=2}^{m} n_{ij}(l)}{\sum_{l=2}^{m} [n_{i0}(l) + n_{i1}(l)]} . \quad (42)$$

Introducing the counter $N_{ij}(m) := \sum_{l=2}^{m} n_{ij}(l)$, (42) can be rewritten in a recursive form as

$$\hat{p}_{ij}(m) = \frac{N_{ij}(m-1) + n_{ij}(m)}{N_{i0}(m-1) + N_{i1}(m-1) + n_{i0}(m) + n_{i1}(m)} .$$

(43)

The extension of the estimator to higher-order Markov chains is straightforward. As shown in [21], the estimators for a second- and third-order Markov chain are, respectively,

$$\hat{p}_{ijk}(m) = \frac{\sum_{l=3}^{m} n_{ijk}(l)}{\sum_{l=3}^{m} n_{ij}(l)}$$

(44)

and

$$\hat{p}_{ijkr}(m) = \frac{\sum_{l=4}^{m} n_{ijkr}(l)}{\sum_{l=4}^{m} n_{ijk}(l)} .$$

(45)

Appendix 2
In order to solve the problem (9) let us consider the following Lagrangian function

$$\mathcal{L}(\boldsymbol{p}) = \frac{1}{M} \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} \left[\beta_{k,m} \log_2 \left(1 + p_{k,m} a_n(k) \right) \right.$$

$$\left. + \gamma_{k,m} \log_2 \left(1 + p_{k,m} a_I(k, m) \right) \right] - \lambda$$

$$\times \left(\frac{1}{M} \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} p_{k,m} - P_T \right) + \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1}$$

$$\times \mu_{k,m} p_{k,m} - \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} \nu_{k,m} \left(p_{k,m} - p^{\max}(k) \right)$$

(46)

where $\lambda, \mu_{k,m}, \nu_{k,m}$ are the Lagrangian multipliers and the KKT conditions can be written as

$$\begin{cases} \nabla \mathcal{L} = 0 \\ 0 \leq \mu_{k,m} \perp p_{k,m} \geq 0 \qquad \forall\, k, m \\ 0 \leq \nu_{k,m} \perp (p_{k,m} - p^{\max}(k)) \leq 0 \qquad \forall\, k, m \\ 0 \leq \lambda \perp \left(\frac{1}{M} \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} p_{k,m} - P_T \right) \leq 0 \end{cases}$$

(47)

Observe that if $p_{k,m} < p^{\max}(k)$ then $\nu_{k,m} = 0$ and this system can be reduced to

$$\begin{cases} \frac{1}{M} \left[\lambda - \frac{a_n(k)\beta_{k,m} + a_I(k,m)(\gamma_{k,m} + a_n(k)p_{k,m})}{(1+a_n(k)p_{k,m})(1+a_I(k,m)p_{k,m})} \right] p_{k,m} = 0 \\ \lambda \geq \frac{a_n(k)\beta_{k,m} + a_I(k,m)(\gamma_{k,m} + a_n(k)p_{k,m})}{(1+a_n(k)p_{k,m})(1+a_I(k,m)p_{k,m})} \\ 0 \leq \lambda \perp \left(\frac{1}{M} \sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} p_{k,m} - P_T \right) \leq 0 \\ p_{k,m} \geq 0 \end{cases}$$

(48)

By exploiting the following inequality whose validity can easily be proved

$$\frac{a_n(k)\beta_{k,m} + a_I(k, m)(\gamma_{k,m} + a_n(k)p_{k,m})}{(1 + a_n(k)p_{k,m})(1 + a_I(k, m)p_{k,m})}$$

$$\leq \frac{a_n(k)\beta_{k,m} + a_I(k, m)(\gamma_{k,m} + a_n(k)p_{k,m})}{(1 + a_n(k)p_{k,m})(1 + a_I(k, m)p_{k,m})} \Big|_{p_{k,m}=0}$$

$$= a_n(k)\beta_{k,m} + a_I(k, m)\gamma_{k,m},$$

(49)

we can deduce that if $\lambda < a_n(k)\beta_{k,m} + a_I(k, m)\gamma_{k,m}$ the second inequality in (48) can hold only if $p_{k,m} > 0$ so that from the first equation in (48) it results

$$\lambda = \frac{a_n(k)\beta_{k,m} + a_I(k, m) \left(\gamma_{k,m} + a_n(k)p_{k,m} \right)}{\left(1 + a_n(k)p_{k,m} \right) \left(1 + a_I(k, m)p_{k,m} \right)} .$$

(50)

On the other hand if $\lambda \geq a_n(k)\beta_{k,m} + a_I(k, m)\gamma_{k,m}$, then $p_{k,m} > 0$ is never verified since it would imply

$$\lambda \geq a_n(k)\beta_{k,m} + a_I(k, m)\gamma_{k,m}$$

$$> \frac{a_n(k)\beta_{k,m} + a_I(k, m) \left(\gamma_{k,m} + a_n(k)p_{k,m} \right)}{\left(1 + a_n(k)p_{k,m} \right) \left(1 + a_I(k, m)p_{k,m} \right)}$$

(51)

which violates the complementary conditions. As a consequence for $\lambda \geq a_n(k)\beta_{k,m} + a_I(k, m)\gamma_{k,m}$ we have $p_{k,m} = 0$. Let us now solve Equation (50), i.e.,

$$\tilde{a}_{k,m} p_{k,m}^2 + \tilde{b}_{k,m} p_{k,m} + \tilde{c}_{k,m} = 0$$

(52)

then defining $\tilde{a}_{k,m} = \lambda\, a_n(k)a_I(k, m)$, $\tilde{b}_{k,m} = \lambda[a_n(k) + a_I(k, m)] - a_n(k)a_I(k, m)$ and $\tilde{c}_{k,m} = \lambda - a_n(k)\beta_{k,m} - a_I(k, m)\gamma_{k,m}$

$$p_{k,m} = \frac{-\tilde{b}_{k,m} \pm \sqrt{\tilde{b}_{k,m}^2 - 4\tilde{a}_{k,m}\tilde{c}_{k,m}}}{2\tilde{a}_{k,m}} .$$

(53)

Let us now make some useful observations:

1. The solutions in (53) are always real. In fact the term under the squared root is always positive $\forall\, \lambda \geq 0$ since it results

$$\tilde{b}_{k,m}^2 - 4\tilde{a}_{k,m}\tilde{c}_{k,m} = \lambda^2 \left(a_n(k) - a_I(k,m)\right)^2 + 2\lambda a_n(k)$$
$$\times a_I(k,m)\left(a_n(k) - a_I(k,m)\right)$$
$$\times \left(1 - 2\gamma_{k,m}\right) + a_n(k)^2 a_I(k,m)^2 \tag{54}$$

where $a_n(k) - a_I(k,m) > 0$ and the minimum, achieved for $\gamma_{k,m} = 1$, is given by

$$\left[\lambda\left(a_n(k) - a_I(k,m)\right) - a_n(k)a_I(k,m)\right]^2 \geq 0, \quad \forall \lambda \geq 0 ; \tag{55}$$

2. The solution

$$p_{k,m}^a = \frac{-\tilde{b}_{k,m} - \sqrt{\tilde{b}_{k,m}^2 - 4\tilde{a}_{k,m}\tilde{c}_{k,m}}}{2\tilde{a}_{k,m}} \tag{56}$$

is always negative since the inequality

$$\sqrt{\tilde{b}_{k,m}^2 - 4\lambda a_n(k)a_I(k,m)\left(\lambda - a_n(k)\beta_{k,m} - a_I(k,m)\gamma_{k,m}\right)}$$
$$> -\tilde{b}_{k,m} \tag{57}$$

is verified for all $\lambda > 0$;

3. The sign of the solution

$$p_{k,m}^b = \frac{-\tilde{b}_{k,m} + \sqrt{\tilde{b}_{k,m}^2 - 4\tilde{a}_{k,m}\tilde{c}_{k,m}}}{2\tilde{a}_{k,m}} \tag{58}$$

can be studied by considering the following inequality

$$\sqrt{\tilde{b}_{k,m}^2 - 4\lambda a_n(k)a_I(k,m)\left(\lambda - a_n(k)\beta_{k,m} - a_I(k,m)\gamma_{k,m}\right)}$$
$$> \tilde{b}_{k,m}. \tag{59}$$

In particular[d] it results

$$p_{k,m}^b = \begin{cases} > 0 & \text{for } \lambda < a_n(k)\beta_{k,m} + a_I(k,m)\gamma_{k,m} \\ \leq 0 & \text{for } \lambda \geq a_n(k)\beta_{k,m} + a_I(k,m)\gamma_{k,m} \end{cases}. \tag{60}$$

According to the above considerations the solution for $0 < p_{k,m} < p^{\max}(k)$ and $\lambda < a_n(k)\beta_{k,m} + a_I(k,m)\gamma_{k,m}$ is $p_{k,m} = p_{k,m}^b$, hence we can write

$$p_{k,m} = \begin{cases} 0 & \text{if } \lambda \geq a_n(k)\beta_{k,m} + a_I(k,m)\gamma_{k,m} \\ p_{k,m}^b & \text{if } \lambda < a_n(k)\beta_{k,m} + a_I(k,m)\gamma_{k,m} \end{cases} \tag{61}$$

so that the optimal solution can be written as $p_{k,m}^* = \left[p_{k,m}^b\right]_0^{p^{\max}(k)}$ with $\sum_{k=1}^{N} \sum_{m=m_0}^{m_0+M-1} \left[p_{k,m}^b\right]_0^{p^{\max}(k)} = P_T M.$

Appendix 3
Convergence Analysis of MADP Algorithm

Proceeding as in [19], in order to prove the convergence of the algorithm it is sufficient to show that

(a) With a proper choice of the step $\alpha_q(n)$, MADP converges to a fixed point;
(b) This point is a solution of the KKT conditions of the modified game in (19) and then it is also a solution point of the optimization problem

$$\begin{aligned} \max_{\boldsymbol{p}} \quad & \bar{R}(\boldsymbol{p}) \\ \text{s.t.} \quad & \boldsymbol{p} \in \mathcal{P} \end{aligned} \tag{62}$$

where the FAPs' sum rate is $\bar{R}(\boldsymbol{p}) = \sum_{q=1}^{Q} \bar{R}_q(\boldsymbol{p})$ and $\mathcal{P}$ is the cartesian product of the sets $\mathcal{P}_q$.

Let us denote with $U_1(n) = \bar{R}(\boldsymbol{p}(n))$ the sum utility reached at the step n of the MADP algorithm. Then for each user q we must prove that there exists a sequence $\alpha_q(n) > 0$ so that $U_1(n)$ is monotonically increasing and convergent, i.e., $U_1(n+1) \geq U_1(n)$ $\forall n$ and $U_1(n) \to U_1^*$ as $n \to \infty$. As discussed in [19], we only need to show that $U_1(n)$ is monotonically increasing, i.e., it suffices to consider a given iteration n in which user q is selected to update its power profile, and show that $U_1(\boldsymbol{p}_q(n+1)) \geq U_1(\boldsymbol{p}_q(n))$, where the total utility U_1 is now regarded as a function of $\boldsymbol{p}_q$ because only the power profile of user q is updated. Hence, our goal is to prove that $U_1(\boldsymbol{p}_q(n+1)) \geq U_1(\boldsymbol{p}_q(n))$. To do this we will use the descent lemma to bound $U_1(\boldsymbol{p}_q(n+1))$. Descent lemma [19] says that if a function $F : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and its gradient is Lipschitz continuous with Lipschitz constant equal to K then, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$F(\boldsymbol{x} + \boldsymbol{y}) \leq F(\boldsymbol{x}) + \boldsymbol{y}^T \nabla F(\boldsymbol{x}) + \frac{K}{2}\|\boldsymbol{y}\|_2^2. \tag{63}$$

One sufficient condition for Lipschitz continuity is that the l_2-norm of the Hessian matrix of $F(\boldsymbol{x})$ is bounded, in which case this bound can be used for the Lipschitz constant. It can easily be shown that it is true for $U_1(\boldsymbol{p}_q)$. Specifically, there exists a constant $B_{U_1}^q$ which upper bounds the l_2-norm of the Hessian matrix of $U_1(\boldsymbol{p}_q)$ independent of others' power profiles.

Applying the Descent lemma to $-U_1(\boldsymbol{p}_q)$, we get

$$U_1(\boldsymbol{p}_q(n+1)) \geq U_1(\boldsymbol{p}_q(n)) + \left[\boldsymbol{p}_q(n+1) - \boldsymbol{p}_q(n)\right]^T$$
$$\times \nabla_{\boldsymbol{p}_q} U_1(\boldsymbol{p}_q(n)) - \frac{B_{U_1}^q}{2}$$
$$\times \left\|\boldsymbol{p}_q(n+1) - \boldsymbol{p}_q(n)\right\|_2^2. \tag{64}$$

Hence to prove that $U_1(\boldsymbol{p}_q(n+1)) \geq U_1(\boldsymbol{p}_q(n))$, it suffices to show that

$$\left[\boldsymbol{p}_q(n+1) - \boldsymbol{p}_q(n)\right]^T \nabla_{\boldsymbol{p}_q} U_1(\boldsymbol{p}_q)(n) \geq \frac{B_{U_1}^q}{2} \left\| \boldsymbol{p}_q(n+1) - \boldsymbol{p}_q(n) \right\|_2^2. \tag{65}$$

Using the power updating rule

$$p_h^q(n+1) = p_h^q(n) + \alpha_q(n)\left(p_h^{q*} - p_h^q(n)\right) \tag{66}$$

with the best response of user q defined in (25), the inequality in (65) can be written as

$$\left[\boldsymbol{p}_q^* - \boldsymbol{p}_q(n)\right]^T \nabla_{\boldsymbol{p}_q} U_1\left(\boldsymbol{p}_q(n)\right) \geq \alpha_q(n)\frac{B_{U_1}^q}{2}\left\|\boldsymbol{p}_q^* - \boldsymbol{p}_q(n)\right\|_2^2. \tag{67}$$

Observe that

$$\frac{\partial U_1\left(\boldsymbol{p}_q\right)}{\partial p_h^q}\bigg|_{\boldsymbol{p}_q=\boldsymbol{p}_q(n)} = \frac{\beta_h^q \mathrm{SNIR}_h^{\beta q}}{1 + \mathrm{SNIR}_h^{\beta q}} \cdot \frac{1}{p_h^q(n)} + \frac{\gamma_h^q \mathrm{SNIR}_h^{\gamma q}}{1 + \mathrm{SNIR}_h^{\gamma q}}$$
$$\times \frac{1}{p_h^q(n)} - \sum_{r\in\mathcal{N}_q} \pi_h^r(n)|H_h^{qr}|^2, \tag{68}$$

then exploiting the result in (25), we can write the left-hand side (LHS) of (67) as

$$\mathrm{LHS} = \sum_{q=1}^{Q} \frac{\sum_{r\in\mathcal{N}_q} \pi_h^r(n)\left|H_h^{qr}\right|^2}{2p_h^q(n)}\left(p_h^{q*} - p_h^q(n)\right)^2$$
$$+ \sum_{q=1}^{Q} \frac{c_h^q\left(\eta_h^q - \nu^q\right)\left(p_h^{q*} - p_h^q(n)\right)}{p_h^q(n)\left(\sum_{r\in\mathcal{N}_q}\pi_h^r(n)\left|H_h^{qr}\right|^2 + \nu^q - \eta_h^q\right)}. \tag{69}$$

Now from (69), with the same steps as in [19], to ensure that

$$\mathrm{LHS} \geq \alpha_q(n)\frac{B_{U_1}^q}{2}\left\|\boldsymbol{p}_q^* - \boldsymbol{p}_q(n)\right\|_2^2, \tag{70}$$

we can choose the step $\alpha_q(n)$ as

$$\alpha_q(n) \leq \min\left\{\frac{2A_q^n}{B_{U_1}^q}, 1\right\} \tag{71}$$

where the coefficient A_q^n is defined as

$$A_q^n = \min_h\left\{\frac{\sum_{r\in\mathcal{N}_q}\pi_h^r(n)\left|H_h^{qr}\right|^2}{p_h^q(n)\,c_h^q}\right\}. \tag{72}$$

Finally, in order to prove the point (b), let U_1^* a fixed point of the algorithm such that $U_1(n) = U_1^*$ for some index n. Then since this is a fixed point, it follows that $p_h^q(n) = p_h^{q*}$, $\forall h, q$. It can then be seen that for all q, p_h^{q*} must be an optimal solution to the problem (19), given the other users current power profiles and interference price vectors. Hence, $\boldsymbol{p}(n)$ will satisfy also the KKT conditions of the problem (62).

Appendix 4

In order to find the optimal solutions of the convex problem $(\widetilde{P}_1)$ by studying the Lagrange dual problem, some additional constraint qualification conditions must hold, beyond convexity, to ensure strong duality [23]. One simple constraint qualification is Slater's condition, i.e., we must verify that some strictly feasible point exists. We can prove that the set $\widetilde{\mathcal{F}}_q(\boldsymbol{p}_{-q})$ for each user q fixed the strategies of the others is nonempty. For simplicity in this proof we assume w.l.o.g. $m_0 = 1$. More specifically, the constraint $R_q(\boldsymbol{p}) > R_q^0$ can be written as

$$\sum_{k=1}^{\mathcal{V}_N^q}\sum_{m=1}^{\mathcal{V}_M^q}\left[\beta_{k,m}\log\left(1 + p_{k,m}^q a_n^q(k,m)\right) + \gamma_{k,m}\right.$$
$$\left.\times \log\left(1 + p_{k,m}^q a_I^q(k,m)\right)\right] > R_q^0 \tag{73}$$

where we have denoted with $\mathcal{V}_N^q \subseteq \{1,\ldots,N\}$ and $\mathcal{V}_M^q \subseteq \{1,\ldots,M\}$ the subsets, respectively, of subcarriers and time slots that the player q is using during the game. Since $a_n^q(k,m) > a_I^q(k,m)$, to verify (73), it is sufficient to prove that

$$\log\left(1 + p_{k,m}^q a_I^q(k,m)\right) > R_q^0, \quad \forall k \in \mathcal{V}_N^q, \; m \in \mathcal{V}_M^q, \; q \in \Omega \tag{74}$$

and clearly it exists always a set of positive values $p_{k,m}^q, p_q^{\max}(k) \in \mathbb{R}_+$ such that

$$p_q^{\max}(k) > p_{k,m}^q \geq (e^{R_q^0} - 1)\frac{1}{a_I^q(k,m)} \qquad \forall k \in \mathcal{V}_N^q,$$
$$m \in \mathcal{V}_M^q, \; q \in \Omega. \tag{75}$$

Let us consider, for $k = 1, \ldots, N$, $m = 1, \ldots, M$, the KKT conditions of the optimization problem $(\widetilde{P}_1)$:

$$1 - \lambda_q \left[\frac{\beta_{k,m} a_n^q(k,m)}{1 + p_{k,m}^q a_n^q(k,m)} + \frac{\gamma_{k,m} a_I^q(k,m)}{1 + p_{k,m}^q a_I^q(k,m)} \right] - \mu_{k,m}^q + \alpha_{k,m}^q = 0$$

$$0 \leq \lambda_q \perp \bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) - R_q^0 \geq 0$$

$$0 \leq p_{k,m}^q \perp \mu_{k,m}^q \geq 0$$

$$0 \leq \alpha_{k,m}^q \perp p_q^{\max}(k) - p_{k,m}^q \geq 0$$

(76)

Observe that if $p_q^{\max}(k) - p_{k,m}^q > 0$, then $\alpha_{k,m}^q = 0$ so that, by eliminating in (76) the multiplier $\mu_{k,m}^q$, we obtain

$$0 \leq \left[1 - \lambda_q \left(\frac{\beta_{k,m} a_n^q(k,m)}{1 + p_{k,m}^q a_n^q(k,m)} + \frac{\gamma_{k,m} a_I^q(k,m)}{1 + p_{k,m}^q a_I^q(k,m)} \right) \right] \perp p_{k,m}^q \geq 0$$

$$0 \leq \lambda_q \perp \bar{R}_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) - R_q^0 \geq 0$$

(77)

where $\lambda_q > 0$ otherwise complementarity yields $p_{k,m}^q = 0$, $\forall k = 1, \ldots, N, m = 1, \ldots, M$, and the rate constraint is contradicted. Then, the optimum power vector must satisfy the following equation

$$a^q(k,m)(p_{k,m}^q)^2 + b^q(k,m)p_{k,m}^q + c^q(k,m) = 0 \quad (78)$$

having set

$$a^q(k,m) = a_n^q(k,m)\, a_I^q(k,m)$$
$$b^q(k,m) = a_n^q(k,m) + a_I^q(k,m) - \lambda_q a_n^q(k,m)a_I^q(k,m).$$
$$c^q(k,m) = 1 - \lambda_q \left[a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m} \right]$$

(79)

The solutions of (78) are

$$p_{k,m}^q = \begin{cases} p_{k,m}^a = \frac{-b^q(k,m) - \sqrt{b^q(k,m)^2 - 4a^q(k,m)c^q(k,m)}}{2a^q(k,m)} \\ p_{k,m}^b = \frac{-b^q(k,m) + \sqrt{b^q(k,m)^2 - 4a^q(k,m)c^q(k,m)}}{2a^q(k,m)} \end{cases}.$$

(80)

It can be proved that $\forall \lambda_q > 0$, it results $p_{k,m}^a \leq 0$, $b^q(k,m)^2 - 4a^q(k,m)c^q(k,m) \geq 0$, and

$$p_{k,m}^b = \begin{cases} > 0 \text{ for } \lambda_q > \frac{1}{a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m}} \\ \leq 0 \text{ for } \lambda_q \leq \frac{1}{a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m}} \end{cases}. \quad (81)$$

According to the above considerations the solution is $p_{k,m}^q = p_{k,m}^b$ for $0 < p_{k,m}^q < p_q^{\max}(k)$ and $\lambda_q > \frac{1}{a_n^q(k,m)\beta_{k,m} + a_I^q(k,m)\gamma_{k,m}}$ so that we can write the optimal power allocation vector as

$$p_{k,m}^{q*} = \left[p_{k,m}^b \right]_0^{p_q^{\max}(k)}$$

(82)

where the multiplier λ_q is chosen in order to satisfy the constraint $\bar{R}_q(\boldsymbol{p}_q^*, \boldsymbol{p}_{-q}) = R_q^0$.

Appendix 5

Proof that the feasible set of the game $\widetilde{\mathcal{G}}_1$ is compact and non-empty so that it can be cast as a GPG.

Let us start by the following definition of GPG given in [20]:

Definition 1. *A Generalized Nash Equilibrium Problem is a GPG if*

(a) *There exists a non-empty, closed set $\widetilde{\mathcal{X}} \subseteq \mathbb{R}^n$ such that*

$$\mathcal{X}_q(\boldsymbol{x}_{-q}) = \{\boldsymbol{x}_q \in D_q : (\boldsymbol{x}_q, \boldsymbol{x}_{-q}) \in \widetilde{\mathcal{X}}\} \quad \forall q = 1, \ldots, Q$$

(83)

where $D_q \subseteq \mathbb{R}^{n_q e}$ are non-empty, closed sets such that $\prod_{q=1}^Q D_q \bigcap \widetilde{\mathcal{X}} \neq \emptyset$;

(b) *There exists a continuous function, $\Phi(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$, named potential function, such that $\forall q \in \Omega, \forall \boldsymbol{x}_{-q}$ and for all $\boldsymbol{y}_q, \boldsymbol{z}_q \in \mathcal{X}_q(\boldsymbol{x}_{-q})$*

$$u_q\left(\boldsymbol{y}_q, \boldsymbol{x}_{-q}\right) - u_q\left(\boldsymbol{z}_q, \boldsymbol{x}_{-q}\right) > 0$$

(84)

implies

$$\Phi\left(\boldsymbol{y}_q, \boldsymbol{x}_{-q}\right) - \Phi\left(\boldsymbol{z}_q, \boldsymbol{x}_{-q}\right) \geq u_q\left(\boldsymbol{y}_q, \boldsymbol{x}_{-q}\right) - u_q\left(\boldsymbol{z}_q, \boldsymbol{x}_{-q}\right)$$

(85)

where u_q is the qth player payoff function.

According to Definition 1, we have to check the validity of the conditions (a) and (b) for the game $\widetilde{\mathcal{G}}_1$. As regard the condition (a), let us consider the feasible set of the game $\widetilde{\mathcal{G}}_1$, i.e.,

$$\widetilde{\mathcal{X}} = \{\boldsymbol{p} \in \mathbb{R}^{NMQ \times 1} : \bar{R}_q(\boldsymbol{p}) \geq R_q^0, 0 \leq p_{k,m}^q \leq p_q^{\max}(k),$$

$$\forall k \in \{1, \ldots, N\}, \forall m \in \{1, \ldots, M\}, \forall q \in \Omega\},$$

(86)

where we have assumed w.l.o.g. $m_0 = 1$. Then we have to prove the following lemma.

Lemma 1. *The feasible set $\widetilde{\mathcal{X}}$ of the game $\widetilde{\mathcal{G}}_1$ is a non-empty, closed and bounded (then compact) subset of $\mathbb{R}^{NMQ \times 1}$ if the matrices $\boldsymbol{A}_k$ defined in (91) are P-matrices, for all $k = 1, \ldots, N$, $m = 1, \ldots, M$. Sufficient conditions for which this happens are*

$$\sum_{r \in \mathcal{N}_q} \frac{|H_k^{rq}|^2}{|H_k^{qq}|^2} < \frac{1}{e^{R_q^0} - 1} \quad \forall q \in \Omega, \quad \forall k = 1, \ldots, N.$$

(87)

Proof. Let us start by considering the constraints $\bar{R}_q(\boldsymbol{p}) \geq R_q^0$ that, by considering only the subcarriers and the time slots that are effectively occupied, can be written as

$$\sum_{k=1}^{\mathcal{V}_N^q} \sum_{m=1}^{\mathcal{V}_M^q} \Big[\beta_{k,m} \log \Big(1 + p_{k,m}^q a_n^q(k,m) \Big) + \gamma_{k,m}$$
$$\times \log \Big(1 + p_{k,m}^q a_I^q(k,m) \Big) \Big] \geq R_q^0, \quad \forall q \in \Omega \tag{88}$$

where $\mathcal{V}_N^q \subseteq \{1,\ldots,N\}$, $\mathcal{V}_M^q \subseteq \{1,\ldots,M\}$ are the subsets, respectively, of subcarriers and time slots, that the player q is using during the game. We can note that $a_n^q(k,m) > a_I^q(k,m)$ then (88) is surely valid if we prove that

$$\log \left(1 + \frac{p_{k,m}^q |H_k^{qq}|^2}{\sigma_{n,q}^2(k) + \sum_{r \in \mathcal{N}_q} p_{k,m}^r |H_k^{rq}|^2 + \sigma_{Iq}^2(k,m)} \right) \geq R_q^0 \tag{89}$$

so that we have to verify $\forall k \in \mathcal{V}_N^q$, $m \in \mathcal{V}_M^q$ the following set of inequalities

$$p_{k,m}^q |H_k^{qq}|^2 - \Big(e^{R_q^0} - 1 \Big) \sum_{r \in \mathcal{N}_q} p_{k,m}^r |H_k^{rq}|^2$$
$$\geq \Big(e^{R_q^0} - 1 \Big) \Big(\sigma_{n,q}^2(k) + \sigma_{Iq}^2(k,m) \Big) \quad \forall q \in \Omega. \tag{90}$$

Defining the vector $\boldsymbol{p}_{k,m} = (p_{k,m}^q)_{q=1}^Q$ and the matrices

$$A_k = \begin{bmatrix} |H_k^{11}|^2 & -(e^{R_1^0}-1)|H_k^{12}|^2 & \ldots & -(e^{R_1^0}-1)|H_k^{1Q}|^2 \\ -(e^{R_2^0}-1)|H_k^{21}|^2 & |H_k^{22}|^2 & \ldots & -(e^{R_2^0}-1)|H_k^{2Q}|^2 \\ \vdots & \vdots & \ddots & \vdots \\ -(e^{R_Q^0}-1)|H_k^{Q1}|^2 & -(e^{R_Q^0}-1)|H_k^{Q2}|^2 & \ldots & |H_k^{QQ}|^2 \end{bmatrix}^T \tag{91}$$

we can express the system of inequalities in (90) as

$$A_k \boldsymbol{p}_{k,m} \geq \boldsymbol{v}_{k,m} \quad \forall k \in \mathcal{V}_N^q, \ m \in \mathcal{V}_M^q \tag{92}$$

where the positive entries of the vector $\boldsymbol{v}_{k,m} = (v_{k,m}^q)_{q=1}^Q$ are given by $v_{k,m}^q = (e^{R_q^0}-1)(\sigma_{n,q}^2(k) + \sigma_{Iq}^2(k,m))$. It can be observed that each matrix A_k is a Z-matrix, i.e. a matrix with all off-diagonal elements non positive. Furthermore if we ensure that each A_k is also a P-matrix, i.e. a matrix whose determinant and all principal minors are positive [24], we can deduce that its inverse is well defined. By imposing diagonally dominance on the elements of the matrices A_k, $\forall k = 1,\ldots,N$, we can find the following sufficient conditions for them to be P-matrices, i.e.,

$$\sum_{r \in \mathcal{N}_q} \frac{|H_k^{rq}|^2}{|H_k^{qq}|^2} < \frac{1}{e^{R_q^0}-1} \quad \forall q \in \Omega, \forall k = 1,\ldots,N. \tag{93}$$

Hence, by considering the general case $\mathcal{V}_N^q = \{1,\ldots,N\}$ and $\mathcal{V}_M^q = \{1,\ldots,M\}$, we can deduce from (92) that there exist positive vectors $\boldsymbol{p}_{k,m}, \boldsymbol{p}^{\max}(k) \triangleq (p_q^{\max}(k))_{q=1}^Q \in \mathbb{R}^Q$ such that

$$\boldsymbol{p}^{\max}(k) > \boldsymbol{p}_{k,m} \geq A_k^{-1} \boldsymbol{v}_{k,m} \quad \forall k = 1,\ldots,N, \quad m = 1,\ldots,M \tag{94}$$

or the sets $\mathcal{F}_{k,m} = \{\boldsymbol{p}_{k,m} \in \mathbb{R}^Q : \log(1 + p_{k,m}^q a_I^q(k,m)) \geq R_q^0, 0 \leq p_{k,m}^q \leq p_q^{\max}(k), \forall q\}$ are non-empty. Of course also the product set $\mathcal{F} = \{\prod_{k=1}^N \prod_{m=1}^M \mathcal{F}_{k,m}\} \subseteq \tilde{\mathcal{X}}$ is non-empty, so that the non-emptiness of $\tilde{\mathcal{X}}$ is implied.

Furthermore, $\forall q \in \Omega$ the set $\{\boldsymbol{p} \in \mathbb{R}_+^{MNQ \times 1} : \bar{R}_q(\boldsymbol{p}) \geq R_q^0\}$, is the upper level set of the continuous function $\bar{R}_q(\boldsymbol{p})$, then it is closed for all scalar R_q^0 [25]. Hence, the set $\tilde{\mathcal{X}} = \{\boldsymbol{p} \in \mathbb{R}^{NMQ \times 1} : \bar{R}_q(\boldsymbol{p}) \geq R_q^0, 0 \leq p_{k,m}^q \leq p_q^{\max}(k), \forall k \in \{1,\ldots,N\}, \forall m \in \{1,\ldots,M\}, \forall q \in \Omega\}$, as non-empty intersection of closed sets, is closed [25] and, since it is also bounded, its compactness is proved.

Verification of condition (b) is rather straightforward. In our case, the objective functions do not depend on the other players variables, i.e., $u_q(\boldsymbol{p}_q, \boldsymbol{p}_{-q}) = u_q(\boldsymbol{p}_q)$ so that the interaction of the players takes places only at the level of feasible sets. In this case, it is immediate to see that condition (b) is satisfied with the potential function $\Phi(\boldsymbol{p})$ simply given by the sum of the objective functions of all players, i.e.,

$$\Phi(\boldsymbol{p}) = \sum_{q=1}^Q u_q(\boldsymbol{p}_q) \tag{95}$$

and this concludes the proof. $\qquad\square$

Endnotes

[a] We denote with $\boldsymbol{p}_q$ the *NM*-dimensional power vector with entries $p_{k,m}^q$ and define $\boldsymbol{p}_{-q} \triangleq (\boldsymbol{p}_i)_{i=1,i\neq q}^Q$ where Q is the number of FAPs.

[b] We recall that a set of strategies is *Pareto efficient*, or *Pareto optimal*, if it is not possible to make at least some player better off without making any other player worse off. Given the whole set of feasible strategies, i.e., the strategies satisfying the system constraints, the *Pareto boundary* is defined as the set of choices that are Pareto efficient. If an equilibrium point belongs to the Pareto boundary, the equilibrium is said to be efficient.

[c] We will assume the prices constant with respect to $p_{k,m}^q$. In general, the assumption of $\pi_{k,m}^q$ to be constant with respect to $p_{k,m}^q$ is only an approximation. Nevertheless, the resulting algorithm provides significant performance improvement with respect to the purely competitive game.

[d] In order to prove this result we have exploited the inequality $\frac{a_n(k)a_I(k,m)}{[a_n(k)+a_I(k,m)]} < a_n(k) - [a_n(k) - a_I(k,m)]\gamma_{k,m}$

whose validity can easily be proved.

[e] We assume that $\prod_{q=1}^{Q} \mathbb{R}^{n_q} = \mathbb{R}^n$.

References
1. V Chandrasekhar, JG Andrews, A Gatherer, Femtocell networks: a survey, IEEE Commun. Mag. **46**, 59–67 (2008)
2. O Simeone, E Erkip, SS Shitz, Robust transmission and interference management for femtocells with unreliable network access, IEEE J. Sel. Areas Commun. **28**, 1469–1478 (2010)
3. V Chandrasekhar, JG Andrews, Uplink capacity and interference avoidance for two-tier femtocell networks, IEEE Trans. Wirel. Commun. **8**, 3498–3509 (2009)
4. F Baccelli, B Blaszczyszyn, Stochastic geometry and wireless networks. Vol. II–applications, *Foundations and Trends in Networking*, vol 4, (New York, November 2009), pp. 1–312
5. Y Chen, Q Zhao, A Swami, Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors, IEEE Trans. Inf. Theory. **54**(5), 2053–2071 (2008)
6. J Unnikrishnan, VV Veeravalli, Algorithms for dynamic spectrum access with learning for cognitive radio, IEEE Trans. Signal Process. **58**(2), 750–760 (2010)
7. A Motamedi, A Bahai, Optimal channel selection for spectrum-agile low-power wireless packet switched networks in unlicensed band, EURASIP J. Wirel. Commun. Netw. **2008**, 10 (2008)
8. A Anandkumar, N Michael, AK Tang, A Swami, Distributed algorithms for learning and cognitive medium access with logarithmic regret, IEEE J. Sel. Areas Commun. **29**(4), 731–745 (2011)
9. K Liu, Q Zhao, Distributed learning in multi-armed bandit with multiple players, IEEE Trans. Signal Process. **58**(11), 5667–5681 (2010)
10. P Auer, N Cesa-Bianchi, P Fischer, Finite-time analysis of the multiarmed bandit problem, Mach. Learn. **47**(2), 235–256 (2002)
11. A Galindo-Serrano, L Giupponi, Distributed Q-learning for aggregated interference control in cognitive radio networks, IEEE Trans. Veh. Technol. **59**(4), 1823–1834 (2010)
12. A Galindo-Serrano, L Giupponi, in *71st Veh. Technol. Conf. (VTC 2010-Spring)*. Distributed Q-learning for interference control in OFDMA-based femtocell networks, (Taipei, May 2010, pp. 1–5
13. CJ Watkins, P Dayan, Technical note: Q-learning, Mach. Learn. **8**(3–4), 279–292. (Kluwer Academic Publishers, Boston, 1992)
14. S Geirhofer, L Tong, BM Sadler, in *Proc. of Military Commun. Conf. (MILCOM 2008)*. Interference-aware OFDMA resource allocation: a predictive approach, (San Diego, CA, November 2008, pp. 1-7
15. JS Pang, G Scutari, F Facchinei, C Wang, Distributed power allocation with rate constraints in Gaussian parallel interference channels, IEEE Trans. Inf. Theory. **54**(8), 3471–3489 (2008)
16. S Barbarossa, A Carfagna, S Sardellitti, M Omilipo, L Pescosolido, in *ICASSP 2011*. Optimal radio access in femtocell networks based on Markov modeling of interferers' activity, (Prague, 22–27 May 2011, pp. 3212–3215
17. F Facchinei, J Pang, *Finite-Dimensional Varational Inequalities and Complementarity Problems*. (Springer-Verlag, New York, 2003)
18. J Huang, RA Berry, ML Honig, Distributed interference compensation for wireless networks, IEEE J. Sel. Areas Commun. **24**, 1074–1084 (2006)
19. C Shi, RA Berry, ML Honig, in *Proc. of CISS 2008*. Distributed interference pricing for OFDM wireless networks with non-separable utilities, Princeton, March 2008, pp. 755–760
20. F Facchinei, V Piccialli, M Sciandrone, in *Computational Optimzation and Application*. Decomposition algorithms for generalized potential games (Springer-Verlag, New York, 2010
21. TW Anderson, LA Goodman, Statistical inference about Markov chains, Ann. Math. Stat. **28**, 89–110 (1957)
22. DS Bai, S Kim, Estimation of transition probabilities in a two-state Markov chain, Commun. Stat. Theort Methods. **8**(6), 591–599 (1979)
23. S Boyd, L Vandenberghe, *Convex Optimization*. Cambridge University Press, New York, 2004)
24. RW Cottle, J-S Pang, RE Stone, *The Linear Complementarity Problem*. (Academic Press, Cambridge, UK, 1992)
25. DP Bertsekas, *Convex Analysis and Optimization*. (Athena Scientific, Belmont, MA, 2003)