Game Theory and Femtocell Communications: Making Network Deployment Feasible

Mehdi Bennis, Samir Medina Perlaza, Mérouane Debbah

Game Theory and Femtocell Communications: Making Network Deployment Feasible

M. Bennis, S. M. Perlaza, and M. Debbah

October 8, 2010

Contents

Chapter 1

Game Theory and Femtocell Communications: Making Network Deployment Feasible

Mehdi Bennis, Centre for Wireless Communications, University of Oulu, Finland
Samir Medina Perlaza, France Telecom R&D, Orange Labs, France
Merouane Debbah, Alcatel-Lucent Chair in Flexible Radio, SUPELEC, France

1.1 Introduction

Femtocell base stations (FBSs), also known as Home eNodeBs (HeNBs), are short-range low-power and *low-cost* base stations with third party backhaul (e.g., DSL or cable modem). They are usually deployed and controlled by end-users who desire better indoor signal transmission and reception. With the help of such FBSs, the network operator is able to extend the high quality coverage inside peoples' homes without the need of additional expensive cellular towers. At the same time, FBSs offload traffic from the overlay macrocell network and subsequently improve the overall network capacity. Nevertheless, despite FBSs promise, many concerns still remain, especially *cross-tier* interference. Two particular aspects of FBSs give rise to serious interference issues: *(a)* the co-channel spectrum sharing between femtocells and macrocells as well as among femtocells, *(b)* the "random" location of user-installed FBSs (they can be deployed anywhere inside the macrocell area with no-prior warning). First, unlike Wi-Fi access points, FBSs serve users in licensed spectrum, to guarantee Quality-of-Service (QoS). Moreover, instead of allocating dedicated licensed channels to FBSs, sharing the spectrum is preferable from an operator's perspective (higher spectral efficiency). Secondly, FBSs are installed by end-users in a "plug-and-play" manner, which translates into "randomness" in their locations. For these two reasons, interference in two-tier networks is quite different than in conventional cellular networks, which endangers their successful co-existence. A typical scenario is the "Dead Zone" or "Loud Neighbor" problem, where mobile users transmit and receive signals at positions near FBSs but far from the macrocell BSs, causing significant macro-to-femto interference in the uplink. In the downlink, these users likewise suffer from low signal to interference ratios (SIRs) because of the strong interference from the FBSs. These affects are akin to the well known *near-far* problem, but exacerbated by the de-centralization and lack of coordinated power control inherent in two-tier networks. In addition, because of the non-existent coordination between FBSs and macrocell BSs, centralized cooperation to mitigate cross-tier interference is deemed infeasible, hence two-tier networks need to adopt decentralized strategies for interference management.

Game Theory (GT) is a mathematical tool that analyzes interactions among decision makers. GT is seen as a natural paradigm to study and analyze wireless networks where players compete for the same resources. The importance of studying the coexistence between macro- and femtocells from a game theoretical perspective is multi-fold: First, by modeling the dynamic spectrum sharing among network players (MBSs, FBSs, MUEs, FUEs) as games, the behaviours and actions of players can be analyzed in a formalized structure, by which the theoretical achievements in GT can be fully utilized. Second, GT equips us with various optimality criteria for the spectrum sharing problem, which are of key importance when it comes to analyzing the equilibria of the game. Third, the application of GT enables us to derive efficient distributed algorithms for self-organized networks relying only on partial information. In order to achieve this, the theory of learning in games is of high importance allowing players to choose the right strategies and gradually learn from their environment through trials and errors until convergence. The expected contribution is this chapter is as follows:

In the first section of the chapter, we show that the deployment of femtocell networks is cast into a dynamic spectrum sharing problem with different interference scenarios such as femtocell-to-femtocell and femtocell-to-macrocell interference. Femto base stations and their associated femtocell user equipments (FUEs) need to to autonomously determine their transmission strategies in order to maximize their own performance metrics. Often, this configuration is described by the power allocation scheme, channel/carrier selection, modulation scheme, etc. However, the transmission configuration of one femto-base station affects the performance of other FBS and all their associated FUEs. Hence, this scenario is clearly a competition of selfish autonomous devices for radio resources. Under these conditions, GT appears as one of the paradigms to study these types of networks. Finally, in this section, we provide the fundamental concepts of game theory required to understand the results presented in the rest of this chapter.

The second section tackles one of the most important issues in self organized networks, namely learning. Due to their non-coordinated nature, femtocells opportunistically share the spectrum with the macrocell network while mitigating their interference towards among others. In order to do so, femtocells need to self-organize and optimize their strategies/actions (power level, carrier allocation etc) taking into account their side information. Learning is an inherent part of self-organization paradigms and is crucial for developing and adapting multi-agent strategies based on their perceived reward. The problem boils down as to how femtocells gradually learn from their environment (through trials-and-errors) while at the same time not interfering with the overlay macrocell network. Furthermore, learning is adamantly driven by the type of information available at every agent where the possibility of exchanging their local information drives the learning efficiency (and convergence) and thereby their respective payoffs. For this purpose, we will investigate both private and public information exchange among femtocells.

The third section looks into the QoS provisioning of femtocell networks, in which femtocells mitigate their interference towards the overlay macrocell network. Tools from game theory and Stochastic geometry are used to come up with decentralized algorithms to achieve the equilibrium of the game.

The fourth section investigates the spectral efficiency of femtocell networks using the framework of potential games. Here, typical problems such as power allocation and channel selection are tackled in detail. We put in evidence the existence of the following paradoxes (observable only in decentralized networks): In a multi-carrier system, increasing the number of channels each device can use highly reduces the global performance of the system. Hence, it is more beneficial for a given network to constraint the devices to use a single channel rather than allowing them to simultaneously use several channels. Finally, we discuss several decentralized algorithms which allow self-configuring femtocell networks to achieve the equilibrium.

1.2 System Model

Consider a set of $\mathcal{M} = \{1, \ldots, M\}$ macrocell base stations (MBS) each one operating over an exclusive fixed frequency band and serving their respective macrocell user equipments (MUEs) following a time division multiple access (TDMA) policy. At each time interval, each MBS serves one of its corresponding MUE aiming to guaranteeing a minimum time-average signal to interference plus noise ratio (SINR) over their communication duration.

We assume that there exists a set $\mathcal{N} = \{1, \dots, N\}$ of N frequency bands over which MBS can operate. Let $\Gamma_0^{(m)}$, $m \in \mathcal{M}$, denote the minimum time-average SINR offered by MBS m over its corresponding fixed frequency band. Consider now a set $\mathcal{K} = \{1, \dots, K\}$ of K femtocells underlaying the M-cell N-frequency band macrocell system. Each femtocell can use any of the available frequency bands to serve its corresponding femto end-users (FUE) as long as it does not induce a lower time-average SINR than the minimum required by the MUE, i.e., $\Gamma_0^{(1)}, \dots, \Gamma_0^{(M)}$. At each time interval each FBS serves one FUE over one of the available channels following a TDMA policy.

Let $t \in \{1, \dots, \infty\}$ be a discrete time index. For all $(j,k,m) \in \mathcal{M}^2 \times \mathcal{N}$, $h_{1,j,k}^{(n)}$ represents the channel realization between MBS k and MUE j over channel m at time t. For all $(j,k,m) \in \mathcal{K} \times \mathcal{M} \times \mathcal{N}$, $h_{2,j,k}^{(n)}$ represents the channel realization between MBS k and FUE j over channel n at time t. For all $(j,k,m) \in \mathcal{M} \times \mathcal{K} \times \mathcal{N}$, $h_{3,j,k}^{(n)}$ represents the channel realization between FBS k and MUE j over channel n at time t. Finally, for all $(j,k,m) \in \mathcal{K}^2 \times \mathcal{N}$, $h_{4,j,k}^{(n)}$ represents the channel realization between FBS k and FUE j over channel n at time t. Denote by $\boldsymbol{h}(t)$ the vector of all channel realizations at time t. All channel realizations, i.e., each component of $\boldsymbol{h}(t)$, are independent and identically distributed following a probability distribution which is a parameter of the network. Let the finite set denoted by $\mathcal{H}$ be the set of all possible vectors $\boldsymbol{h}(t)$, for all $t > 0$. Finally, channel realizations at time t are independent of those at time $t-1$, for all $t > 0$.

Let $p_{k,\max}$ and $p_{0,m}$, with $k \in \mathcal{K}$ and $m \in \mathcal{M}$, be the maximum transmit power of FBS k and MBS m, respectively. For all $k \in \mathcal{K}$, let the N-dimensional vector $\boldsymbol{p}_k(t) = \left(p_k^{(1)}(t), ..., p_k^{(N)}(t)\right)$ denote the power allocation vector of FBS $k \in \mathcal{K}$ at time t. Here $p_k^{(n)}(t)$ is the transmit power of femtocell k over frequency band n at time t. In the channel selection problem, all FBSs are assumed to transmit only over one frequency band at each time t at a given power level not exceeding $p_{k,max}$. Let L_k be the number of power levels of FBS k, i.e., $\frac{p_{k,\max}}{L_k}, \dots, p_{k,\max}$. For all $(k, \ell, s) \in \mathcal{K} \times \{1, \dots, L_K\} \times \mathcal{N}$, denote by the N-dimensional vector

$$\boldsymbol{q}_k^{(\ell,n)} = \frac{\ell}{L} p_{k,\max} \boldsymbol{e}_n^{(N)}, \tag{1.1}$$

the power allocation (PA) vector when FBS transmits over channel n at power level ℓ. Denote also by $\boldsymbol{q}_k^{(0,0)}$, with $k \in \mathcal{K}$, the N-dimensional null vector, i.e., $\boldsymbol{q}^{(0,0)} = (0, \ldots, 0)$. Thus, FBS k has $N_k = L_k \cdot S + 1$ possible PA vectors, $\boldsymbol{q}^{(0,0)}, \boldsymbol{q}_k^{(1,1)}, \ldots, \boldsymbol{q}_k^{(L_k,N)}$.

For all $(k, n) \in \mathcal{K} \times \mathcal{N}$, let $\gamma_k^{(n)}$ be the SINR of FUE k at time t and for all $m \in \mathcal{M}$, let $\gamma_{0,m}^{(n_m)}$ be the SINR of the MUE in the macrocell m at time t. Let also the set $\mathcal{M}_n \subset \mathcal{K}$, with $n \in \mathcal{N}$, be the set of MBS using channel n. Then, we can write that

$$\gamma_k^{(n)}(t) = \frac{p_k^{(n)}(t) \left| h_{4,k,k}^{(n)}(t) \right|^2}{\sigma_k^{(n)^2} + \sum_{m \in \mathcal{M}_n} p_{0,m} \left| h_{2,k,m}^{(n)}(t) \right|^2 + \sum_{j \in \mathcal{K} \setminus \{k\}} p_j^{(n)}(t) \left| h_{4,k,j}^{(n)}(t) \right|^2} \tag{1.2}$$

and for all $m \in \mathcal{M}$,

$$\gamma_{0,m}^{(n_m)}(t) = \frac{p_{0,m} \left| h_{1,m,m}^{(n_m)}(t) \right|^2}{\sigma_{0,m}^2 + \sum_{j \in \mathcal{M}_{n_m} \setminus \{m\}} p_{0,j} \left| h_{1,m,j}^{(n_m)}(t) \right|^2 + \sum_{i \in \mathcal{K} \setminus \{k\}} p_{i,max} \left| h_{3,m,i}^{(n_m)}(t) \right|^2}, \tag{1.3}$$

where for all $m \in \mathcal{M}$, n_m is the channel used by MBS m and $\sigma_{0,m}^{(n_m)^2}$ and $\sigma_k^{(n_m)^2}$ is the noise power over MUE m and the noise power over FUE k on the frequency band n.

All FBSs are interested in optimizing a given interference mitigation metric denoted by $u : \mathbb{R}^{N \cdot K + M} \to \mathbb{R}$, which determines at each instant t the impact of the interference on the macro system based on the observation of all the SINR levels $\gamma_k^{(n)}$ and $\gamma_{0,m}^{(n_m)}$, with $(k, n) \in \mathcal{K} \times \mathcal{N}$ and $m \in \mathcal{M}$.

In the case of multiple access channels (MACs), femtocell base stations communicate with a single receiver using a common channel. This case can happen when femto user equipments compete for the same resources in the uplink, or when FBSs compete for backhaul resources. We assume that there exists a set $\mathcal{N} = \{1, ..., N\}$ of N frequency bands over which transmitters operate. If $N > 1$ channels are available, then there exists N independent or parallel MACs, where transmitters in different MACs do not interfere each other. Moreover, the channel gain from transmitter $k \in \mathcal{K}$ to the receiver over channel $n \in \mathcal{N}$ is denoted by $h_k^{(n)}$. We assume a block flat-fading channel model such that channel realizations remain constant during the transmission of M consecutive symbols. All the channel realizations,

$\forall k \in \mathcal{K}$ and $\forall n \in \mathcal{N}$ are drawn from a Gaussian distribution with zero mean and unit variance. The power allocated by transmitter k to channel n is denoted by $p_k^{(n)}$. Each transmitter is power-limited wherein for the i^{th} transmitter, its transmit power cannot exceed $p_{k,max}$, i.e., $\forall k \in \mathcal{K}, \sum_{n=1}^{N} p_k^{(n)} \leq p_{k,max}$. The received SINR on channel n for transmitter k, denoted by $\gamma_k^{(n)}$ for all $\forall (k,n) \in \mathcal{K} \times \mathcal{N}$ is:

$$\gamma_k^{(n)} = \frac{p_k^{(n)} |h_k^{(n)}|^2}{\sum_{j \neq k}^{K} p_j^{(n)} |h_j^{(n)}|^2 + \sigma^2} \tag{1.4}$$

1.3 Game Theory and Self-Organized Femtocell Networks

In the following, we introduce basic concepts and definitions of game theory which will be used throughout the chapter.

1.3.1 Strategic-Form Games

The spectrum sharing between the macro and femtocells can be respectively modeled by the following two non-cooperative static games in strategic form (with $i \in \{a, b\}$):

$$\mathcal{G}_{(i)} = \left(\mathcal{K}, \left(\mathcal{P}_k^{(i)} \right)_{k \in \mathcal{K}}, (u_k)_{k \in \mathcal{K}} \right). \tag{1.5}$$

Let $\mathcal{K}$ be the set of players (i.e., FBSs). An action of a given FBS $k \in \mathcal{K}$ is a particular power allocation (PA) scheme, i.e., an N-dimensional PA vector $\boldsymbol{p}_k = \left(p_k^{(1)}, \ldots, p_k^{(N)} \right) \in \mathcal{P}_k^{(i)}$, where $\mathcal{P}_k^{(i)}$ is the set of all possible PA vectors which FBS k can use either in the game $\mathcal{G}_{(a)}$ $(i = a)$ or in the game $\mathcal{G}_{(b)}$ $(i = b)$. An action profile of the game $i \in \{a, b\}$ is a super vector

$$\boldsymbol{p} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K) \in \mathcal{P}^{(i)},$$

where $\mathcal{P}^{(i)}$ is a set obtained from the Cartesian product of the action sets $\mathcal{P}_k^{(i)}$, for all $k \in \mathcal{K}$, i.e., $\mathcal{P}^{(i)} = \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_K^{(i)}$, where,

$$\begin{aligned} \mathcal{P}_k^{(a)} &= \Big\{ \left(p_k^{(1)}, \ldots p_k^{(N)} \right) \in \mathbb{R}^N : \forall n \in \mathcal{N}, p_k^{(n)} \geqslant 0, \\ &\quad \text{and } \sum_{n \in \mathcal{N}} p_k^{(n)} \leqslant p_{k,\max} \}, \text{ and} \end{aligned} \tag{1.6}$$

$$\begin{aligned} \mathcal{P}_k^{(b)} &= \big\{ p_{k,\max}\, \boldsymbol{e}_n : \forall n \in \mathcal{N},\ \boldsymbol{e}_n = \left(e_n^{(1)}, \ldots, e_n^{(N)} \right) \\ &\quad \text{and } \forall r \in \mathcal{N} \setminus n, e_n(r) = 0, \text{ and } e_n^{(n)} = 1 \big\}. \end{aligned} \tag{1.7}$$

In this chapter, we respectively refer to the games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ as the PA game and CS game. Let us denote by $\boldsymbol{p}_{-k}$ any vector in the set

$$\mathcal{P}_{-k}^{(i)} \triangleq \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_{k-1}^{(i)} \times \mathcal{P}_{k+1}^{(i)} \times \ldots \times \mathcal{P}_K^{(i)} \tag{1.8}$$

with $(i,k) \in \{a,b\} \times \mathcal{K}$. For a given $k \in \mathcal{K}$, the vector denoted by $\boldsymbol{p}_{-k}$ represents the strategies adopted by all the players other than player k. With a slight abuse of notation, we write any vector $\boldsymbol{p} \in \mathcal{P}^{(i)}$, with $i \in \{a,b\}$, as $\left(\boldsymbol{p}_k, \boldsymbol{p}_{-k}\right)$, in order to emphasize the k-th vector component of the super vector $\boldsymbol{p}$. The utility for player k in the game $\mathcal{G}_{(i)}$ is its spectral efficiency $u_k :\ \mathcal{P}^{(i)} \rightarrow \mathbb{R}_+$, with $i \in \{a,b\}$ and

$$u_k(\boldsymbol{p}_k, \boldsymbol{p}_{-k}) = \sum_{n \in \mathcal{N}} \frac{B_n}{B} \log_2 \left(1 + \gamma_k^{(n)} \right) \ \text{[bps/Hz]} \tag{1.9}$$

where $\gamma_k^{(n)}$ is the SINR of Femto base station k over channel n.

1.3.2 Potential Games

Potential games (PG) [38] are a class of games for which existence of pure NE is guaranteed.

Definition 1 (Exact Potential Game) Any game in strategic form defined by the 3-tuple $\left(\mathcal{K}, (\mathcal{P}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right)$ is an exact potential game if there exists a function $\phi(\boldsymbol{p})$ for all $\boldsymbol{p} \in \mathcal{P}$

such that for all players $k \in \mathcal{K}$ and for all $\boldsymbol{p}'_k \in \mathcal{P}_k$, it holds that

$$u_k(\boldsymbol{p}_k, \boldsymbol{p}_{-k}) - u_k(\boldsymbol{p}'_k, \boldsymbol{p}_{-k}) = \phi(\boldsymbol{p}_k, \boldsymbol{p}_{-k}) - \phi(\boldsymbol{p}'_k, \boldsymbol{p}_{-k}).$$

Definition 2 (Exact Constrained Potential Game) Any game in normal form defined by the 4-tuple $\left(\mathcal{K}, \{\mathcal{P}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ is an exact constrained potential game (PG) if there exists a function $\phi(\boldsymbol{s})$ for all $\boldsymbol{p} \in \mathcal{P}_{\mathrm{SE}}$ such that for all players $k \in \mathcal{K}$ and for any pair of actions $(\boldsymbol{p}_k, \boldsymbol{p}'_k) \in \left\{f_k\left(\boldsymbol{p}_{-k}\right)\right\}^2$, it holds that

$$c_k(p_k, \boldsymbol{p}_{-k}) - c_k(p'_k, \boldsymbol{p}_{-k}) = \phi(p_k, \boldsymbol{p}_{-k}) - \phi(p'_k, \boldsymbol{p}_{-k}).$$

Using the spectral efficiency utility function (1.9), the following proposition holds:

Proposition 1 The strategic games $\mathcal{G}_{(i)}$, with $i \in \{a, b\}$, are exact potential games with potential function

$$\phi(\boldsymbol{p}) = \sum_{n\in\mathcal{N}} \frac{B_n}{B} \log_2\left(\sigma_n^2 + \sum_{k=1}^{K} p_k^{(n)} g_k^{(n)}\right). \tag{1.10}$$

In fact, the games $\mathcal{G}_{(i)}$, $i \in \{a, b\}$ are not only potential games but also best-response potential games [40]. A BRPG is a PG which verifies:

$$\arg\max_{\boldsymbol{q}_k\in\mathcal{P}_k^{(i)}} u_k\left(\boldsymbol{q}_k, \boldsymbol{p}_{-k}\right) = \arg\max_{\boldsymbol{q}_k\in\mathcal{P}_k^{(i)}} \phi\left(\boldsymbol{q}_k, \boldsymbol{p}_{-k}\right). \tag{1.11}$$

Indeed, the utility function (1.9) can be written in terms of the potential function (1.10) as follows

$$\forall k \in \mathcal{K}, \quad u_k\left(\boldsymbol{p}_k, \boldsymbol{p}_{-k}\right) = \phi\left(\boldsymbol{p}_k, \boldsymbol{p}_{-k}\right) - v_k\left(\boldsymbol{p}_{-k}\right),$$

where

$$v_k\left(\boldsymbol{p}_{-k}\right) = \sum_{n\in\mathcal{N}} \frac{B_n}{B} \log_2\left(\sigma_n^2 + \sum_{j\in\mathcal{K}\setminus\{k\}} p_j^{(n)} g_j^{(n)}\right),$$

and by inspection, it becomes clear that both $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ satisfy the condition (1.11). In other words, condition (1.11) shows that the individual utility maximization problem is *equivalent* to maximizing a common function.

1.3.3 The Concept of Equilibrium

Our interest is to find a strategy profile $\boldsymbol{p}$ such that no player is interested in changing its own strategy. Once the network configuration $\boldsymbol{p}^*$ is reached, any unilateral deviation of a given player decreases its own utility. A network configuration $\boldsymbol{p}^*$ is known as a Nash equilibrium [19].

The Nash equilibrium (NE) is an important concept in the field of game theory wherein an N.E corresponds to a profile of strategies $p^* = (p_1^*, ..., p_K^*)$ for which each player's strategy $p_1^* \in \mathcal{P}$ is an N.E if it satisfies:

$$\forall k \in \mathcal{K} \quad \text{and} \quad \forall p_k \in \mathcal{P}_K, \quad u_k(p_k^*, p_{-k}^*) \geq u_k(p_k, p_{-k}^*) \tag{1.12}$$

That is, at the NE, any unilateral deviation from the strategy profile p_k of player k, $\forall k \in \mathcal{K}$ will not increase its utility function. Hence, at the NE there does not exist any motivation for a player to deviate from the NE strategy profile. As players are selfish and decide by themselves their strategy, one question arises: does an NE lead to an efficient game outcome? The NE is generally inefficient and non-optimal, however it is a lower bound in the case of non-cooperation between players.

Definition 3 (Satisfaction Equilibrium) An action profile $\boldsymbol{p}^+$ is a satisfaction equilibrium for the game $\mathcal{G}' = \left(\mathcal{K}, \{\mathcal{P}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ if

$$\forall k \in \mathcal{K}, \quad p_k^+ \in f_{-k}\left(\boldsymbol{p}_k^+\right). \tag{1.13}$$

Definition 4 (Efficient Satisfaction Equilibrium) Define a function $c_k : \mathcal{P}_k \rightarrow [0,1]$ for all $k \in \mathcal{K}$ and consider the game $\mathcal{G}'$. For all $(k, p_k^*, p_k') \in \mathcal{K} \times \mathcal{P}_k^2$, the action p_k' is said to be more costly that action p_k^* if $c_k\left(p_k'\right) > c_k\left(p_k\right)$. An action profile $\boldsymbol{p}^* \in \mathcal{P}$ is an ESE if and only if:

$$\forall k \in \mathcal{K}, \quad \boldsymbol{p}_k^* \in \arg \min_{p_k \in f_k\left(\boldsymbol{p}_{-k}^*\right)} c_k\left(p_k\right). \tag{1.14}$$

Then, $\boldsymbol{p}^*$ is one of the efficient SE (ESE) of the game $\mathcal{G}'$.

Definition 5 ***(Logit Equilibrium)****:* Consider the Markov game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t \geqslant 0}$. The mixed strategy profile $\boldsymbol{\pi}^* = (\boldsymbol{\pi}_1^*, ..., \boldsymbol{\pi}_L^*) \in \triangle(\mathcal{A}_1), \times \ldots, \times \triangle(\mathcal{A}_K)$ is a logit equilibrium, if $\forall k \in \mathcal{K}$,

$$\boldsymbol{\pi}_k^* = \boldsymbol{\beta}_k\left(\hat{\boldsymbol{u}}_k\left(\boldsymbol{\pi}_{-k}^*\right)\right), \tag{1.15}$$

where the N_k dimensional vector $\hat{\boldsymbol{u}}_k\left(\boldsymbol{\pi}_{-k}\right) = \left(\hat{u}_{k,\boldsymbol{q}_k^{(0,0)}}\left(\boldsymbol{\pi}_{-k}\right), \hat{u}_{k,\boldsymbol{q}_k^{(1,1)}}\left(\boldsymbol{\pi}_{-k}\right), \ldots, \hat{u}_{k,\boldsymbol{q}_k^{(L_k,N)}}\left(\boldsymbol{\pi}_{-k}\right)\right)$ is the expected interference minimization metric, i.e., for all $k \in \mathcal{K}$ and for all $(\ell_k, n_k) \in \{1, \ldots, L_k\} \times \mathcal{N} \cup \{(0,0)\})$,

$$u_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(\boldsymbol{\pi}_{-k}) = \mathbb{E}\sum_{\boldsymbol{p}_{-k} \in \mathcal{P}_{-k}} \left(\prod_{j \in \mathcal{K} \setminus \{k\}} \pi_{j,\boldsymbol{p}_j}^*\right) u(\mathbf{h}, \boldsymbol{q}_k^{(\ell_k,n_k)}, \mathbf{p}_{-k}) \boldsymbol{h}.$$

and:

$$\beta_k(\hat{u}_k) = \frac{\exp\left(\frac{1}{\kappa_k}\hat{u}_k\right)}{\sum_{(i,j) \in \mathcal{N} \times \mathcal{L}} \exp\left(\frac{1}{\kappa_k}\hat{u}_k\right)} \tag{1.16}$$

Where $\forall k \in \mathcal{K}$ *and* $\forall (s,l) \in \mathcal{N} \times \mathcal{L}$, $\beta_k(\hat{\mathbf{u}}_{\mathbf{k}}) > 0$ *regardless of the estimation vector* $\hat{\mathbf{u}}_{\mathbf{k}}$.

Definition 6 (Best-Response Correspondence) In a non-cooperative game described by the 3-tuple $\left(\mathcal{K}, (\mathcal{P}_k)_{\forall k \in \mathcal{K}}, (u_k)_{\forall k \in \mathcal{K}}\right)$, the relation $\mathrm{BR}_k : \mathcal{P}_{-k} \rightarrow \mathcal{P}_k$ such that

$$\mathrm{BR}_k\left(\boldsymbol{p}_{-k}\right) = \arg\max_{\boldsymbol{q}_k \in \mathcal{P}_k} u_k\left(\boldsymbol{q}_k, \boldsymbol{p}_{-k}\right), \tag{1.17}$$

is defined as the best-response correspondence of player $k \in \mathcal{K}$, given the actions $\boldsymbol{p}_{-k}$ adopted by all the other players.

Definition 7 (Best Response Dynamics) Let the action profile $\boldsymbol{p}(t) = (\boldsymbol{p}_1(t), \ldots, \boldsymbol{p}_K(t))$ be the result of a best-response dynamics at time t. Then, for all $k \in \mathcal{K}$, and for all $t \in \mathbb{N}$ and $t > 0$, the vector $\boldsymbol{p}_k(t)$ can be obtained as follows: (1) In the sequential best-response dynamics (round-Robin order):

$$\boldsymbol{p}_k(t) \in \mathrm{BR}_k\left(\boldsymbol{p}_1(t),\ldots,\boldsymbol{p}_{k-1}(t),\boldsymbol{p}_{k+1}(t-1),\ldots,\boldsymbol{p}_K(t-1)\right), \tag{1.18}$$

(2) in the simultaneous best-response dynamics:

$$\boldsymbol{p}_k(t) \in \mathrm{BR}_k\left(\boldsymbol{p}_{-k}(t-1)\right), \tag{1.19}$$

where $\boldsymbol{p}(0)$ can be any vector $\boldsymbol{p} \in \mathcal{P}$.

1.4 Learning mechanisms for femtocell networks

Two mechanisms for interference mitigation, inspired by evolutionary game theory and reinforcement-based learning [21] to support the coexistence of a macrocell network and underlaid self-organized femtocells are herein compared in terms of achievable performances as well as needed signalling. In both approaches, femto base stations rely on the measurements fed back by their user equipments in order to update their strategy until an eventual convergence.

1.4.1 Evolutionary-based approach

The first interference mitigation mechanism is based on the concept of evolutionary game theory, where each FBS chooses its strategy against other FBSs located within the same network. FBSs observe the behavior of other competitors, learn from the observations, and make the best decision based on their instantaneous payoff, as well as the average payoff of all other femtocells. The game theoretic model $\mathcal{G}^{(ev)} = \left(\mathcal{K}, \left(\mathcal{A}_k\right)_{k\in\mathcal{K}}, \left(u_k\right)_{k\in\mathcal{K}}\right)$ is formulated as follows:

- $\mathcal{K} = \{1, ..., K\}$ is the set of players (i.e., femtocell base stations),
- $\mathcal{A}_k^t = \{a_t^{k,s}\}_{k\in\mathcal{K}}$ where $a_t^{k,n}$ is the action of FBS k at time t is to transmit over a carrier

n.

- $R_k^t = \sum_n \log_2 \Big(1 + SINR\Big)$ is the reward of FBS k transmitting in sub-carrier n at time t. Moreover, the average payoff of the entire population is defined as: $\bar{R}_k^t = \frac{\sum_k R_k^t}{K}$.

In the aforementioned mechanism, an entity which is referred to as HNB-Gateway [25] collects the payoffs for all femtocells and calculates the average rate of the entire femtocell network. The payoff R_k^t of FBS k is then compared with the average payoffs $\bar{R}_k^t$ and in the case when it is less that the average rate of the femtocell network, a random strategy is chosen and the whole process is repeated again.

1.4.2 Q-learning based approach

The Q-learning model consists of a set of states $n \in \mathcal{N}$ and actions $a \in \mathcal{A}$ aiming at finding a policy that maximizes the observed rewards over the interaction time of the agents/players (i.e., femtocells). Every FBS $k \in \mathcal{K}$ explores its environment, observes its current state s, and takes a subsequent action a, according to a decision policy $\pi : s \rightarrow a$. With their ability to learn, the knowledge about other players' strategies is not explicitly needed by FBS k. Instead, a Q-function maintains the knowledge about other players in the network locally, based on which decisions are made. Several authors have shown that Q-learning converges to optimal values in Markov decision process environment. Thus, the goal of the agent is to find an optimal policy $\pi^*(s)$ for each state s, which maximizes a cumulative measure of the rewards over time.

The *expected* discounted reward over an infinite horizon is given by:

$$V^\pi(s) = \mathbb{E}\left\{\gamma^t \times r(s_t, \pi^*(s_t))|s_0 = s\right\}, \tag{1.20}$$

where $0 \leq \gamma \leq 1$ is a discount factor and r is the agent's reward at time t. Furthermore, Equation (1.20) can be rewritten as:

$$V^{\pi}(s) = R(s, \pi^*(s)) + \gamma \sum_{v \in S} P_{s,v}(\pi(s)) V^{\pi}(v), \tag{1.21}$$

where $R(s, \pi^*(s)) = \mathbb{E}\{r(s, \pi(s))\}$ is the mean value of reward $r(s, \pi(s))$, and $P_{s,v}$ is the transition probability from state s to v. Moreover, the optimal policy π^* satisfies the optimality criterion:

$$V^*(s) = V^{\pi^*}(s) = \max_{a \in \mathcal{A}} \left(R(s, a) + \gamma \sum_{v \in \mathcal{N}} P_{s,v}(a) V^*(v) \right), \tag{1.22}$$

It is generally difficult to explicitly calculate the reward $R(s, a)$ and transition probability $P_{s,v}(a)$. However, through Q-learning, the knowledge of these values can be gradually learnt and reinforced with time. For a given policy π, define a Q-value as:

$$Q^*(s, a) = R(s, a) + \gamma \sum_{v \in S} P_{s,v}(a) V^{\pi}(v), \tag{1.23}$$

which is the expected discounted reward when executing action a at state s and then following policy π thereafter.

We make use of the Q-learning algorithm to iteratively approximate the state-action value function $Q(s, a)$. Here, the FBS keeps trying all actions in all states with non-zero probability and must sometimes explore by choosing at each step a random action with probability $\beta \in (0, 1)$. Alternatively, FBSs can make use of the Boltzmann exploration strategy with the temperature parameter κ, where the action a in state s is taken with a probability $P(a|s)$, and the femtocell receives a reinforcement r. The actions are chosen according to their Q-values as:

$$p(a|s) = \frac{e^{Q(s^i, a)/\kappa}}{\sum_{a' \neq a} e^{Q(s^i, a')/\kappa}}. \tag{1.24}$$

The game is defined as follows: $\mathcal{G}^{(rl)} = \left(\mathcal{K}, (\mathcal{P}_k)_{k \in \mathcal{K}}, (u_k)_{k \in \mathcal{K}}\right)$ where

- $\mathcal{K} = \{1, ..., K\}$ is the set of players (i.e., femtocell base stations),

- $\mathcal{S}_k^t = \{s_t^{k,n}\}_{k \in \mathcal{K}}$ where $s_t^{k,n} \in \{0,1\}$. The state of every FBS k at time t in sub-carrier n indicates whether FBS k generates interference towards MUE and the QoS of MUE is violated, (i.e., $C_0 < \Gamma_0$).

- $\mathcal{A}_k^t = \{a_t^{k,n}\}_{k \in \mathcal{K}}$ where $a_t^{k,n}$ is the action of FBS k at time t is to transmit over a set of sub-carrier $\mathcal{C} \in \{1, ..., N_{sub}\}$.

- $R_k^t = 1_{\{\Gamma_0 - C_0\}} \sum_{n \in \mathcal{C}} \log(1 + SINR)$ is the reward of FBS k transmitting in the set $\mathcal{C}$ at time t.

In the Q-learning process, in which FBS k performs the exploration step with probability β. A new Q-value, i.e., $Q(S_k^{t+1}, A_k^{t+1})$, which is the expected payoff for the future iterations, is obtained based on previous value, i.e., $Q(S_k^t, A_k^t)$, along with the new observed payoff $R_{t+1}^k(S_{t+1}^k)$. Moreover, the Q-learning is updated as follows:

$$Q(S_k^t, A_k^t) = (1-\alpha) Q(S_k^t, A_k^t) + \alpha \left[R_{t+1}^k(S_{t+1}^k) + \gamma \max_{B_k^{t+1}} Q(S_k^{t+1}, B_k^{t+1}) \right] \tag{1.25}$$

where α is the players' willingness to learn from its environment and γ is the discount factor.

1.5 QoS provisioning for femtocell network

The focus is on a fully decentralized approach for solving interference mitigation from the FBS to the macrocell user equipments (MUEs). The underlying assumption is that MUEs feed back messages to their corresponding macrocell BS containing their instantaneous SINR (decoded by all FBSs). The repetitive observation of the SINR is used by the FBS to dynamically configure their frequency band and power levels such that a minimum time-average SINR can be guaranteed at the macrocells. We make use of recent results in game theory and stochastic approximation applied to the problem of femto-to-macrocell cross-tier interference.

The interference minimization problem is modeled by a stochastic game made of a sequence of strategic games played at different states, e.g., channel realizations. Let us denote by $\mathcal{G}(\boldsymbol{h}(t)) = (\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{\phi\}_{k\in\mathcal{K}})$ the static strategic game and let us denote by $G = \{\mathcal{G}(\boldsymbol{h}(t))\}_{t>0}$ the stochastic game where at each time t, the game $\mathcal{G}(\boldsymbol{h}(t))$ is played, with $t \in \{1, \ldots, \infty\}$. In order to define the stochastic game, a short and long-term formulation is given:

Let us describe the network during the interval from $t-1$ to t by the game $\mathcal{G}(\boldsymbol{h}(t)) = (\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u\}_{k\in\mathcal{K}})$. Here, $\mathcal{K}$ represents the set of FBS in the network. For all $k \in \mathcal{K}$, the set of actions of FBS k is the set of PA vectors, i.e., $\mathcal{A}_k = \{\boldsymbol{q}_k^{(\ell,n)} : \ell \in \{0, \ldots, L_k\},$ and $n \in \mathcal{N}\}$. Finally, $u : \mathcal{H} \times \mathcal{A}_1 \times \mathcal{A}_K \rightarrow \mathbb{R}_+$ is the payoff or interference minimization metric of all femtocells.

At each time $t > 0$ and for all $k \in \mathcal{K}$, FBS k chooses its action from the finite set $\mathcal{A}_k$ following a probability distribution $\boldsymbol{\pi}_k(t) = \left(\pi_{k,\boldsymbol{q}_k^{(1,1)}}(t), ..., \pi_{k,\boldsymbol{q}_k^{(L_k,N_k)}}(t)\right)$ where $\pi_{k,\boldsymbol{q}_k^{(l_k,n_k)}}$ is the probability that femtocell k plays action $\boldsymbol{q}_k^{(l_k,n_k)}$ at time t, i.e.,

$$\pi_{k,\boldsymbol{q}_k^{(l_k,n_k)}} = \Pr\left(\boldsymbol{p}_k(t) = \boldsymbol{q}_k^{(l_k,n_k)}\right). \tag{1.26}$$

where $(l_k, n_k) \in \{1, ..., L_K\} \times \mathcal{S} \cup \{(0,0)\}$ In the following, we describe the long-term game $\mathcal{G}$, and we introduce the method which each FBS uses to choose the probability distribution $\boldsymbol{\pi}_k(t)$, at each time t.

Long-Term Formulation

The long-term behaviour of the network is modeled by the succession of static strategic games $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t\geqslant 0}$. This succession produces a Markov game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t\geqslant 0}$, where at each stage t the game $\mathcal{G}(\boldsymbol{h}(t))$ is played assuming that the network is described by the vector $\boldsymbol{h}(t)$. According to the system model, the actual state of the network $\boldsymbol{h}(t)$ follows a Markov chain with transitions following the rule, $\forall\ (\boldsymbol{h}', \boldsymbol{h}'') \in \mathcal{H}^2$, $\Pr(\boldsymbol{h}(t) = \boldsymbol{h}' \,|\boldsymbol{h}(t-1) = \boldsymbol{h}'') = \Pr(\boldsymbol{h}(t) = \boldsymbol{h}') = \pi_{\boldsymbol{h}'}$. Here, $\pi_{\boldsymbol{h}'}$, for all $\boldsymbol{h}' \in \mathcal{H}$, are parameters obtained from previous channel modeling studies. Note that transitions between states are independent of the actions of the transmitters. The game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t\geqslant 0}$ proceeds in infinitely many stages. At each stage $t \in \{0, \ldots, \infty\}$, FBSs choose their corresponding actions $\boldsymbol{p}_1(t), \ldots, \boldsymbol{p}_K(t)$. When doing so, each FBS k observes a noisy sample $\tilde{u}_k(t)$ of the corresponding instantaneous interference minimization metric $u(\boldsymbol{h}(t), \boldsymbol{p}_k(t), \boldsymbol{p}_{-k}(t))$, i.e.,

$$\tilde{u}_k(t) = u(\boldsymbol{h}(t), \boldsymbol{p}_k(t), \boldsymbol{p}_{-k}(t)) + \varepsilon_{k,\boldsymbol{p}_k(t)}(t), \tag{1.27}$$

where, $\forall(\ell_k, n_k) \in \{1, \ldots, L_k\} \times \mathcal{S} \cup \{(0,0)\}$, and $\forall k \in \mathcal{K}$, $\varepsilon_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t)$ is the realization at time t of a random variable $\varepsilon_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}$ which represents the additive noise on the observation of the instantaneous performance $u(t)$ when FBS k plays action $\boldsymbol{q}_k^{(\ell_k,n_k)}$. Here, we assume that $\mathbb{E}\varepsilon_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}} = 0$.

Our behavioral assumption is that all FBS are interested on choosing the probability distribution $\boldsymbol{\pi}_k(t) \in \triangle(\mathcal{A}_k)$ to optimize the time-average interference minimization metric at each time $t > 0$, i.e., $\bar{u}_k(t)$, which is calculated empirically based on the observations $\tilde{u}_k(t)$ as follows,

$$\bar{u}_k(t) = \frac{1}{t}\sum_{n=1}^{t} \widetilde{u}_k(n). \tag{1.28}$$

To choose the optimal probability distribution $\boldsymbol{\pi}_k(t)$, the FBS relies on estimations of the time-average interference minimization metric obtained with each of its actions. For all $(\ell_k, n_k) \in \{1, \ldots, L_k\} \times \mathcal{S} \cup \{(0,0)\}$, let $\hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t)$, be the estimation of time-average interference minimization metric obtained by playing action $\boldsymbol{q}_k^{(\ell_k,n_k)}$. This estimation is calculated as follows,

$$\hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(n) = \frac{1}{T_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t)} \sum_{n=1}^{t} \tilde{u}_k(n) 1_{\left\{\boldsymbol{p}_k(n)=\boldsymbol{q}_k^{(\ell_k,n_k)}\right\}}, \tag{1.29}$$

where, $T_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t) = \sum_{n=1}^{t} 1_{\left\{\boldsymbol{p}_k(n)=\boldsymbol{q}_k^{(\ell_k,n_k)}\right\}}$. Once the N_k-dimensional vector of estimations of FBS k is obtained, i.e., $\hat{\boldsymbol{u}}_k = \left(\hat{u}_{k,\boldsymbol{q}_k^{(0,0)}}, \hat{u}_{k,\boldsymbol{q}_k^{(1,1)}}, \ldots, \hat{u}_{k,\boldsymbol{q}_k^{(L_k,S)}}\right)$ for all $k \in \mathcal{K}$, it is used to determine the optimal probability distribution $\boldsymbol{\pi}_k(t) = \left(\pi_{k,\boldsymbol{q}_k^{(0,0)}}, \pi_{k,\boldsymbol{q}_k^{(1,1)}}, \ldots, \pi_{k,\boldsymbol{q}_k^{(L_k,S)}}\right)$ at each time t. For doing so, we define the function $\boldsymbol{\beta}_k : \mathbb{R}^{N_k} \rightarrow \triangle(\mathcal{A})$. Note that the probability distribution $\boldsymbol{\beta}_k(\hat{\boldsymbol{u}}_k(t))$ must take into consideration that, FBSs must experiment between different actions such that the estimation vector $\hat{\boldsymbol{u}}_k(t)$ is improved at each time t, but also FBSs must optimize their respective interference minimization metric.

1.5.1 Exploration vs. Performance Optimization

Femtocells face a trade-off between optimizing their time-average utility by taking the action that does it at each time t, and trying out different actions so as to improve the estimation of the time-average interference mitigation metric obtained with each action. This implies that a reasonable behavioral rule would be to choose the actions which yield high payoffs more likely than actions yielding low payoffs, but in any case, always letting a non-null probability of playing any of the actions. Following the results in [20], the behavioral rule described above can be modeled by the probability distribution $\beta_k(\hat{\boldsymbol{u}}_k(t))$ satisfying:

$$\beta_k(\hat{\boldsymbol{u}}_k(t)) \in \arg\max_{\boldsymbol{\pi}_k \in \triangle(\mathcal{A}_k)} \left[\sum_{\boldsymbol{p}_k \in \mathcal{A}_k} \pi_{k,\boldsymbol{p}_k} \hat{u}_{k,\boldsymbol{p}_k}(t) + \kappa_k H(\boldsymbol{\pi}_k) \right] \tag{1.30}$$

where H represents the Shannon entropy function. For all $k \in \mathcal{K}$, the parameter $\kappa_k > 0$ represents the interest of FBS k to choose other actions rather than the optimal one in order to improve the time-average interference minimization metric. The unique solution to the right hand side of the optimization problem in (1.30) is written as (1.16):

$$\beta_k(\hat{\boldsymbol{u}}_{\boldsymbol{k}}(t)) = \left(\beta_{k,\boldsymbol{q}_k^{(0,0)}}(\hat{\boldsymbol{u}}_{\boldsymbol{k}}(t)), \beta_{k,\boldsymbol{q}_k^{(1,1)}}(\hat{\boldsymbol{u}}_{\boldsymbol{k}}(t)), \ldots, \beta_{k,\boldsymbol{q}_k^{(L_k,S)}}(\hat{\boldsymbol{u}}_{\boldsymbol{k}}(t)) \right), \tag{1.31}$$

In order to achieve the LE of game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t \geqslant 0}$. Assume that the estimation of the time-average interference minimization metric and the mixed strategy of FBS k are calculated as follows, $\forall k \in \mathcal{K}$ and $\forall (\ell_k, n_k) \in \{1, \ldots, L_k\} \times \mathcal{N} \cup \{(0,0)\}$,

$$\begin{cases} \hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t) &= \hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t-1) + \\ & \alpha(t) \frac{\mathbb{1}_{\left\{\boldsymbol{p}_k(t) = \boldsymbol{q}_k^{(\ell_k,n_k)}\right\}}}{\pi_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t)} \left(\tilde{u}(t) - \hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t-1) \right), \\ \pi_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t) &= \pi_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t-1) + \\ & \lambda(t) \left(\boldsymbol{\beta}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(\hat{\boldsymbol{u}}_k(t)) - \pi_{k,\boldsymbol{q}_k^{(\ell_k,n_k)}}(t-1) \right), \end{cases} \tag{1.32}$$

where, $\hat{\boldsymbol{u}}_k(0) \in \mathbb{R}^{N_k}$ and $\boldsymbol{\pi}_k(0) \in \triangle(\mathcal{A}_k)$ are arbitrary initializations and λ and α are learning rates chosen such that

$$\lim_{T\to\infty}\sum_{t=0}^{T}\alpha(t)+\lambda(t) = +\infty \tag{1.33}$$

$$\lim_{T\to\infty}\sum_{t=0}^{T}\alpha(t)^2+\lambda(t)^2 < +\infty, \text{ and,} \tag{1.34}$$

$$\lim_{t\to\infty}\frac{\lambda(t)}{\alpha(t)} = 0. \tag{1.35}$$

Then, both learning processes (1.32) converge for all $k \in \mathcal{K}$. Additionally, and it holds that,

$$\lim_{t\to\infty}\boldsymbol{\pi}_k(t) = \boldsymbol{\pi}_k^*, \tag{1.36}$$

$$\lim_{t\to\infty}\hat{u}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t) = \boldsymbol{u}_k(\boldsymbol{\pi}_{-k}^*), \tag{1.37}$$

where $\boldsymbol{\pi}^* = (\boldsymbol{\pi}_1^*,\ldots,\boldsymbol{\pi}_K^*)$ is a LE of the game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t\geqslant 0}$.

1.6 Spectral Efficiency in Femtocell Networks

Spectral efficiency in femtocell networks is instrumental for the successful deployment of femtocells. Here, Femtocells access the spectrum in a parallel multiple access mode. Two particular deployment scenarios are hereafter studied: (a) FBSs strategically split their transmit power between the available channels, and (b) FBSs use only one channel at a time. In both cases, FBSs maximize their individual spectral efficiency. Problems (a) and (b), which respectively correspond to a decentralized power allocation and channel selection problem, are modeled by *strategic* form games.

1.6.1 Power Allocation Problems

The PA game $\mathcal{G}_{(a)}$ models the scenario where FBSs allocate any power level to any of their own channels subject to the power constraints given in (1.6)

It is proved in [24] that the game $\mathcal{G}_{(a)}$ is an exact potential game, and hence at least one NE in pure strategies exists. Similarly, at least one NE in mixed strategies exist because the action spaces, $\mathcal{P}_k^{(a)}$ for all $k \in \mathcal{K}$ are convex and compact in their respective finite dimensional spaces, and the utility is jointly continuous [43]. Furthermore, the game $\mathcal{G}_{(a)}$ is shown to have a *unique* NE strategy profile in pure strategies, $\boldsymbol{p}^{\dagger} \in \mathcal{P}^{(a)}$ with probability one, solution of the following optimization problem:

$$\boldsymbol{p}^{\dagger} \in \arg\max_{p \in \mathcal{P}^{(a)}} \phi(\boldsymbol{p}). \tag{1.38}$$

The components of the vector $\boldsymbol{p}^{\dagger}$ in (1.38) are for all $(k, s) \in \mathcal{K} \times \mathcal{N}$,

$$(p_k^{(n)})^{\dagger} = \left[\frac{B_n}{B} \frac{1}{\beta_k} - \frac{\sigma_n^2 + \sum_{j \in \mathcal{K} \setminus \{k\}} (p_j^{(n)})^{\dagger} g_j^{(n)}}{g_k^{(n)}} \right]^{+}, \tag{1.39}$$

where, β_k is a Lagrangian multiplier chosen to saturate the power constraints (1.16).

Assuming that each transmitter knows its actual channel gains $\boldsymbol{g}_k = \left(g_k^{(1)}, \ldots, g_k^{(N)}\right)$, the bandwidth of all channels $\boldsymbol{b} = (B_1, \ldots, B_N)$, and its own actual PA vector $\boldsymbol{p}_k$, each transmitter can determine its best response (1.19) based on a common message from the receiver, and hence obtain the unique outcome of the game using the best response dynamics (BRD) in which the sequential best response dynamics of any potential game converges to an NE [36].

It is worth noting that the sequential BRD in the game $\mathcal{G}_{(a)}$ lead to the same result as the iterative water-filling algorithm (IWFA) presented in [33] but under a different setting. One of the main drawbacks of the iterative BRD (and IWFA) is its large time for convergence as well as its required signaling (message $\boldsymbol{\kappa}(t)$). To overcome this problem, other algorithms such as the simultaneous water-filling algorithm (SWFA), which follows the simultaneous best-response dynamics (Def. 7) was proposed [34].

Different information drive different convergence behaviors. In terms of convergence, the

IWFA (iterative BRD) always converges. The SWFA (simultaneous BRD), at least on the interference channel, converges to an NE depending on the channel realizations and the number of active transmitters [34]. In the interference channel PA game in [34], when the SWFA converges, hence it converges faster than the IWFA. Finally, the IWFA requires a (broadcast) signaling message (the multiple access interference over each channel) every time that a given transmitter updates its PA vector. Conversely, SWFA requires one message after all transmitters have updated their PA vector.

1.6.2 Channel Selection Problems

Here, a constraint is imposed onto the femtocells, in which they can only transmit over one channel at a time. It is ready to check that every femto BS has to saturate its transmit power to maximize its utility. Indeed, $\forall k \in \mathcal{K}$ and $\forall \boldsymbol{p}_{-k} \in \mathcal{P}_{-k}^{(b)}$, $\phi\left(p_k\, \boldsymbol{e}_n, \boldsymbol{p}_{-k}\right) < \phi\left(p_{k,\max}\, \boldsymbol{e}_s, \boldsymbol{p}_{-k}\right)$ where $\boldsymbol{e}_n = \left(e_n^{(1)}, \ldots, e_n^{(N)}\right) \in \mathbb{R}^N$, $\forall r \in \mathcal{N} \setminus n$, $e_n^{(r)} = 0$, and $e_n^{(n)} = 1$. The problem under investigation is therefore a channel selection problem. Technically, the main difference between $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ is that the latter is a finite game ($|\mathcal{K} \times \mathcal{N}| < +\infty$). As a consequence, the number of pure NE is generally more than 1.

The game $\mathcal{G}_{(b)}$ is an exact potential game (Prop. 1) and thus, following *Lemma 2.3* in [38], the game $\mathcal{G}_{(b)}$ has always at least one NE in pure as well as mixed strategies. To identify an upper-bound of the number of NE, basic tools from graph theory in which $L \in \mathbb{N}$ multiple pure NE strategy profiles exist, and $1 \leqslant L \leqslant S^{K-1}$.

Assuming that the game is played repeatedly, one can exploit Prop. **??** to state that BRD converges to an NE. The best response of transmitter $k \in \mathcal{K}$ in the game $\mathcal{G}_{(b)}$ is

$$\mathrm{BR}_k\left(\boldsymbol{p}_{-k}\right) = \{\boldsymbol{p}_k \in \mathcal{P}_k : \boldsymbol{p}_k = p_{k,\max}\boldsymbol{e}_{n_k^*} \text{ and } n_k^* = \arg\max_{n \in \mathcal{N}} \tfrac{B_N}{B} \log_2\!\left(1+\tfrac{p_{k,\max} g_{k,n}}{\sigma_n^2+\sum_{j\in\mathcal{K}\setminus\{k\}} p_{j,\max} g_{j,n}}\right)\}, \quad (1.40)$$

and can be determined locally by each transmitter using the feedback message, the knowledge of its own channel realizations and the bandwidth of each channel.

Besides the BRD (iterative and sequential, Def. 7), when the set of actions is discrete, other dynamics such as fictitious play [42], reinforcement learning (RL) [36], among some others, converge to equilibrium in the game $\mathcal{G}_{(b)}$, and more importantly, the information assumptions can be significantly relaxed [41].

Let us denote by $\boldsymbol{p}_k(t) \in \mathcal{P}_k^{(b)}$ and $\boldsymbol{h}(t) \in \mathcal{H}$ the action taken by player k and the vector of channel realizations at time t. Hence, for all $k \in \mathcal{K}$, $u_k : \mathcal{H} \times \mathcal{P}^{(b)} \rightarrow \mathbb{R}$, and $u_k\left(\boldsymbol{h}(t), \boldsymbol{p}_k(t), \boldsymbol{p}_{-k}(t)\right)$ represents the instantaneous utility obtained by player k at time t. For the ease of notation, we use the equality $u_k(t) = u_k\left(\boldsymbol{h}(t), \boldsymbol{p}_k(t), \boldsymbol{p}_{-k}(t)\right)$. At each time t, each player $k \in \mathcal{K}$ chooses its action $\boldsymbol{p}_k(t)$ following a probability distribution $\boldsymbol{\pi}_k(t) = \left(\pi_{k,p_k^{(1)}}(t), \ldots, \pi_{k,p_k^{(N)}}(t)\right)$, where $\forall k \in \mathcal{K}$ and $\forall \boldsymbol{p}_k \in \mathcal{P}_k^{(b)}$, $\pi_{k,p_k}(t)$ represents the probability that player k plays action $\boldsymbol{p}_k \in \mathcal{P}_k^{(b)}$ at time t, i.e., $\pi_{k,p_k}(t) = \Pr\left(\boldsymbol{p}_k(t) = \boldsymbol{p}_k\right)$. In the fictitious play (FP) [42], player $k \in \mathcal{K}$ must observe the actions of all the other players at each time t and be able to calculate the empirical probability distributions $\pi_j(t)$, for all $j \in \mathcal{K} \setminus \{k\}$. At each time t, the action of player k is:

$$p_k(t) \in \arg \max_{\boldsymbol{p}_k \in \mathcal{P}_k^{(b)}} \mathbb{E}_{\boldsymbol{\pi}_{-k}} u_k(\boldsymbol{p}_k, \boldsymbol{p}_{-k}). \tag{1.41}$$

In the reinforcement learning (RL), e.g. replicator dynamics [37], transmitters $k \in \mathcal{K}$ observes a numerical sample $\tilde{u}_k(t)$ of its obtained utility at each time t and its set of actions $\mathcal{P}_k^{(b)}$ are supposed to be known. At each time t, transmitter k chooses its PA vector following the probability distribution $\boldsymbol{\pi}_k(t)$, which is updated based on $\tilde{u}_k(t)$, as follows,

$$\begin{aligned}\pi_{k,\boldsymbol{p}_k^{(n)}}(t) \quad = \quad & \pi_{k,\boldsymbol{p}_k^{(n)}}(t-1) \\ & + \lambda_k(t)\tilde{u}_k(t)\left(\mathbb{1}_{\{\boldsymbol{p}_k(t)=\boldsymbol{p}_k^{(n)}\}} - \pi_{k,\boldsymbol{p}_k^{(n)}}(t-1)\right),\end{aligned} \tag{1.42}$$

where $\lambda_k(t)$ is such that $0 \leqslant \lambda_k(t)\tilde{u}_k(t) \leqslant 1$, for all $k \in \mathcal{K}$ and $t > 0$.

The FP converges (at least in probability) to NE in potential games [39]. On the contrary, the convergence of the RL is conditioned on the initial strategies of the dynamics $\boldsymbol{\pi}_k(0)$, for all $k \in \mathcal{K}$ [37]. Moreover, in the FP, players must be able to solve at each time t, the optimization problem (1.41). Hence, the computational capabilities depend on the complexity of the utility function. Conversely, in RL only arithmetic operations are required, since the utility is assumed to be observed at each time t. Besides, In the FP players are supposed to know the structure of the game and observe the actions of all the other players. In RL, players know only their own actions and are able to observe a numerical value of the obtained utility. Finally, in the FP players are assumed to be myopically rational, i.e., players take the action which maximizes their instant utility. In contrast, the rationality assumption is not required in the RL case. Here, players take randomly (subject to a probability distribution) an action which might or not maximize its own utility. Note that unlike RL, FP is too demanding in terms of information assumptions.

1.7 Conclusions and open issues

In this chapter, we have looked at the *strategic* interaction between the macrocell and underlaid femtocells from a game theoretic perspective. Different game theoretic formulations were investigated along with different information assumption. It turns out that game theory along with learning theory are two key enabling frameworks for the analysis and design of femtocell networks. Yet, more work is needed for addressing the strategic interaction of femtocells such as in the case of hierarchy and cooperation among femtocells.

Contents

References

[1] H. Claussen et al., "An Overview of the Femtocell Concept", Bell Labs Technical Journal, vol.13, issue 1, March 2008.

[2] S. Carlaw, C. Wheelock, "Femtocell market challenges and opportunities", ABI research, research report, 2007.

[3] Femto forum - driving convergence worldwide.

[4] V. Chandrasekhar and J. G. Andrews, "Femtocell networks: A survey," IEEE Communication Magazine, 46(9): 59-67, September 2008.

[5] H. Claussen, L. T. W. Ho, and L. G. Samuel, "Self-optimization coverage for femtocell deployments," In Wireless Telecommunications Symposium, 24-26, pp. 278-285, April 2008.

[6] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA Femtocells: A Roadmap on Interference Avoidance," IEEE Communications Magazine, vol. 47, no. 9, pp. 41-48, Jun. 2009.

[7] S. M. Perlaza, S. Lasaulce and M. Debbah and J-M Chaufray, "Game Theory for Dynamic Spectrum Access," Cognitive Radio Networks: Architectures, Protocols and Standards, Auerbach Publications, 2010.

[8] M. Bennis, S. Lasaulce and M. Debbah, "Hierarchy in power allocation games," Game Theory for Wireless Communications and Networking, Auerbach Publications, Taylor and Francis Group, CRC Press.

[9] M. Bennis, S. Lasaulce, and M. Debbah, "Inter-Operator Spectrum Sharing from a Game

Theoretical Perspective," EURASIP Journal of Advances in Signal Processing (JASP), special issue on dynamic spectrum access for wireless networking, August 2009.

[10] M. Nazir, M. Bennis and A. Mackenzie, "Learning mechanisms for interference mitigation in self-organized femtocells," Accepted to appear at the IEEE ASILOMAR 2010, November 2010.

[11] S. Lasaulce, M. Debbah and E. Altman, "Methodologies for Analyzing Equilibria in Wireless Games," IEEE Signal Processing Magazine, special issue on Game Theory, 26(5):51-52, 2009, September 2009.

[12] S. M. Perlaza, H. Tembine and S. Lasaulce, "How can Ignorant but Patient Cognitive Terminals Learn Their Strategy and Utility?," in proc. of IEEE Intl. Workshop on Signal Processing Advances for Wireless Communications (SPAWC), Marrakesh, Morroco, June, 2010.

[13] Y. Hayel, S. Lasaulce, R. El-Azouzi, and M. Debbah, "Introducing Hierarchy in Energy Games," IEEE Transactions on Communications, Volume: 8, Issue: 7, page(s): 3833-3843, July 2009.

[14] S. M. Perlaza, H. Tembine, S. Lasaulce and M. Debbah, "Satisfaction Equilibrium: A General Framework for QoS Provisioning in Self-Configuring Networks," IEEE GLOBECOM 2009, Miami, Florida.

[15] M. LeTreust, H. Tembine, S. Lasaulce and M. Debbah, "Coverage games in small cells networks," 2010, invited paper, IEEE Proc. of the Future Network and Mobile Summit (FUNEMS'10).

[16] M. Bennis, M. LeTreust and S. Lasaulce and M. Debbah, "Spectrum sharing games in the interference channel," invited paper, IEEE Intl. Conference on Game Theory for Networks (Gamenets), Istanbul, Turkey, May 2009.

[17] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks," to be presented in IEEE International Workshop on Femtocell Networks (FEMnet) in conjunction with IEEE GLOBECOM, Miami, FL, USA, 6 December 2010.

[18] M. LeTreust, S. Lasaulce and M. Debbah, "Implicit cooperation in distributed energy-efficient networks," 2010, 1-4, Proc. of the 4th IEEE International Symposium on Communications, Control and Signal Processing (ISCCSP).

[19] J. F. Nash, "Equilibrium points in n-person games," National Academy of Sciences of the United States of America, 1950.

[20] S. M. Perlaza, H. Tembine, S. Lasaulce and M. Debbah, "Radio resource sharing in decentralized wireless networks: a dynamical system approach," submitted to IEEE Transaction in Signal Processing, 2010.

[21] J. Hofbauer and K. Sigmund, "Evolutionary Games and Population Dynamics," Cambridge University Press, UK, 1998,

[22] D. McFadden, "Quantal choice analysis: a survey," Annals of Economic and Social Measurement, vol. 5, no. 4, pp. 363-390, June 1976.

[23] R. D. Mckelvey and T. Palfrey, "Quantal response equilibria for normal form games," Games and Economic Behavior, vol. 10, no. 1, pp. 6-38, 1995.

[24] S. M. Perlaza, H. Tembine, S. Lasaulce and M. Debbah, "Spectral Efficiency of Decentralized Parallel Multiple Access Channels," Submitted to IEEE Trans. in Signal Processing, 2010.

[25] 3GPP, "TR:3GPP-TR25.820, 3G Home Node B Study Item," TSG-RAN, Sept. 2008.

[26] H. Li, "Multi-agent Q-Learning of Channel Selection in Multi-user Cognitive Radio Systems: A Two by Two Case," IEEE Conference on System, Man and Cybernetics (SMC), 2009.

[27] J. Hofbauer and K. Sigmund, Evolutionary Games and Population Dynamics," Cambridge University Press, UK, 1998.

[28] B. Wang, K. J. R. Liu and T. C. Clancy, "Evolutionary game framework for behavior dynamics in cooperative spectrum sensing," in Proc. of IEEE Globecom, New Orleans, USA, 2008.

[29] D. Braess, "Uber ein paradoxon aus der verkehrsplanung," Unternehmensforschung, 24(5):258 268, May 1969.

[30] T.Ui "Correlated Equilibrium and Concave Games," International Journal of Game Theory, 1(37), pp. 1–13, 2008.

[31] R. Aumann, "Subjectivity and correlation in randomized strategies," Journal of Mathematical Economics, pp. 67–96, 1974.

[32] A. Neyman, "Correlated Equilibrium and Potential Games," International Journal of Game Theory, 26(2), pp. 223–227, 1997.

[33] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," IEEE Transactions on Information Theory, 50(1), pp. 145–152, Jan. 2004.

[34] G. Scutari, D. R. Palomar, and S. Barbarossa, "Simultaneous Iterative Water-Filling for Gaussian Frequency-Selective Interference Channels," IEEE International Symposium on Information Theory, pp. 600-604, July 2006.

[35] "Learning dynamics in game with stochastic perturbations," Games and Economic Behavior, vol. 11, no. 2, pp. 330-363, Nov. 1995.

[36] H. P. Young, "Strategic Learning and Its Limits (Arne Ryde Memorial Lectures Sereis)," Oxford University Press, January 2004.

[37] J. Hofbauer and K. Sigmund, "Evolutionary Games and Population Dynamics," Cambridge University Press, UK ,1998.

[38] D. Monderer and L. S. Shapley, "Potential games," Games and Economic Behavior, 14, pp. 124–143, 1996.

[39] D. Monderer and L. S. Shapely, "Fictitious Play Property for Games with Identical Interests," Int. J. Economic Theory, 68, 258–265, 1996.

[40] M. Voorneveld, "Best-response potential games," Economics Letters, 66(3), pp. 289-295, March 2000.

[41] S. M. Perlaza, H. Tembine, S. Lasaulce and M. Debbah, "How can Ignorant but Patient Cognitive Terminals Learn Their Strategy and Utility?"," IEEE SPAWC, Marrakech, Morocco, June 2010.

[42] George W. Brown "Iterative Solution of Games by Fictitious Play," Activity Analysis of Production and Allocation, 1(13), pp. 374–376, 1951.

[43] K. Fan, "Fixed-point and Minimax Theorems in Locally Convex Topological Linear Spaces," Proceedings of the National Academy of Sciences of the United States of America, 38(2), pp. 121-126, 1952.