Decentralized Cross-Tier Interference Mitigation in Cognitive Femtocell Networks

Mehdi Bennis, Samir M. Perlaza

Decentralized Cross-Tier Interference Mitigation in Cognitive Femtocell Networks

M. Bennis* and S. M. Perlaza†
*Centre for Wireless Communications, University of Oulu, Finland
†Orange Labs, France Telecom R&D, Issy Les Moulineaux, France
E-mail: bennis@ee.oulu.fi, samir.medina@lss.supelec.fr

***Abstract*—In this paper, recent results in game theory and stochastic approximation are brought together to mitigate the problem of femto-to-macrocell cross-tier interference. The main result of this paper is an algorithm which reduces the impact of interference of femtocells over the existing macrocells. Such algorithm relies on the observations of the signal to interference plus noise ratio (SINR) of all active communications in both macro and femtocells when they are fed back to the corresponding base stations. Based on such observations, femto base stations learn the probability distributions over the feasible transmit configurations (frequency band and power levels) such that a minimum time-average SINR can be guaranteed in the macrocells, at the equilibrium. In this paper, we introduce the concept of logit equilibrium (LE) and present its interpretation in terms of the trade-off faced by femtocells when experimenting several actions to discover the network, and taking the action to maximize their instantaneous performance. Finally, numerical results are given to validate our theoretical findings.**

I. Introduction

Recently, a new type of indoor Base Station (BS), called femtocell, has gained the attention of the industry [1] due to the enormous benefits it brings to both end-users and network operators. For instance, end-users can enjoy better signal qualities due to the reduced distance between the transmitter and the receiver, resulting in higher throughputs, power and battery savings. From the operator's point of view, femtocells will extend the indoor coverage, enhance system capacity, and share the spectrum in a more efficient manner [2]. However, these benefits are not easy to accomplish, and there are challenges that mobile operators must address before successfully deploying femtocell networks. Among these challenges, there is the *cross-tier* interference between macrocells and femtocells which highly impact the quality-of-service (QoS) of the already existing networks. Therefore, distributed and efficient self-organization strategies need to be designed in order to make the deployment of femtocell networks feasible. Many results exist along this direction, e.g., see [3], [5] among others. In [3] and [5], a Q-learning based algorithm was investigated in the context of network selection for heterogeneous wireless networks, and channel selection in multiuser cognitive radios, respectively. In [6], a reinforcement-learning framework based on Q-learning was studied for interference mitigation among femtocells. Nevertheless, the above-mentioned works often require information exchange among transmitters, which represents a non-affordable increment of signaling messages.

In this paper, we propose a fully decentralized method for interference minimization/mitigation from the femtocell BS (FBS) to the macrocell user equipments (MUEs), i.e., our interest focuses in the downlink interaction between femto- and macrocell systems, as shown in Fig. 1. The underlying assumption over which our work relies on is that, the feedback messages from MUEs to their corresponding macrocell BS containing their instantaneous signal to interference plus noise ratio (SINR) can be decoded by all FBSs. The repetitive observation of the SINR is used by the FBS to dynamically configure how often different frequency bands are used such that, a minimum time-average SINR level can be guaranteed to the MUEs. Our proposal combines recent results in game theory, learning theory, and stochastic approximation to address such an issue.

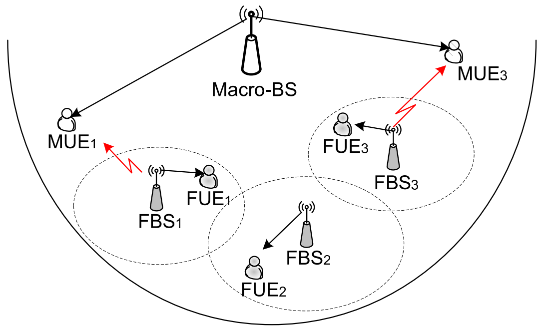

Fig. 1. Network topology with one macrocell underlaid with three femtocell networks. MUE and FUE stand for macro/femtocell user equipment, respectively. MBS and FBS stand for macro and femtocell base station, respectively.

This paper is organized as follows. In Section II, the system and game model are presented. Section III describes how FBSs can build a reliable image of the average state of the network based on noisy observations of the SINR of the active user terminals in the downlink. Section IV describes a learning algorithm for interference mitigation, which is the main result of this paper. Numerical results are presented in Section V. Finally, Section VI concludes this paper.

II. Models

A. Notations

Boldface lower case and lower case symbols represent vectors and scalars. Given a random variable z, the expectation with respect to z is denoted by $\mathbb{E}_z[.]$. The indicator function is denoted by $\mathbf{1}_{\{condition\}}$ and it equals 1 (resp. 0) when *condition* is true (resp. false). Given a finite set $\mathcal{A}$, $\triangle(\mathcal{A})$ represents the set of all probability distributions over the elements of the finite set $\mathcal{A}$. Let the vector $\boldsymbol{e}_s^{(S)} = \left(e_{s,1}^{(S)}, ..., e_{s,S}^{(S)}\right) \in \mathbb{R}^S$ denote the s-th vector of the canonical base spanning the space of real vectors of dimension S. Here, $\forall n \in \{1, ..., S\} \setminus \{s\}$, $e_{s,n}^{(S)} = 0$ and $e_{s,s}^{(S)} = 1$.

B. 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) using 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 SINR over their communication duration. We assume that there exists a set $\mathcal{S} = \{1, \ldots, S\}$ of S 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 $\mathcal{K} = \{1, \ldots, K\}$ of K femtocells underlaying the M-cell S-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., $\Gamma_0^{(1)}, \ldots, \Gamma_0^{(M)}$. At each time interval each FBS serves one FUE over one of the available channels following a TDMA policy.

Let $t \in \{1, \ldots, \infty\}$ be a discrete time index. For all $(j, k, m) \in \mathcal{M}^2 \times \mathcal{S}$, $h_{1,j,k}^{(s)}$ represents the channel realization between MBS k and MUE j over channel m at time t. For all $(j, k, m) \in \mathcal{K} \times \mathcal{M} \times \mathcal{S}$, $h_{2,j,k}^{(s)}$ represents the channel realization between MBS k and FUE j over channel s at time t. For all $(j, k, m) \in \mathcal{M} \times \mathcal{K} \times \mathcal{S}$, $h_{3,j,k}^{(s)}$ represents the channel realization between FBS k and MUE j over channel s at time t. Finally, for all $(j, k, m) \in \mathcal{K}^2 \times \mathcal{S}$, $h_{4,j,k}^{(s)}$ represents the channel realization between FBS k and FUE j over channel s 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 S-dimensional vector $\boldsymbol{p}_k(t) = \left(p_k^{(1)}(t), ..., p_k^{(S)}(t)\right)$ denote the power allocation vector of FBS $k \in \mathcal{K}$ at time t. Here $p_k^{(s)}(t)$ is the transmit power of femtocell k over frequency band s at time t. All FBS are assumed to transmit only over one frequency band at each time t at a given power level not exceeding $p_{k,\max}$. Let $L_k \in \mathbb{N}$ be the number of discrete power levels of FBS k, i.e., $\frac{p_{k,\max}}{L_k}, \ldots, p_{k,\max}$. For all $(k, \ell, s) \in \mathcal{K} \times \{1, \ldots, L_K\} \times \mathcal{S}$, denote by the S-dimensional vector

$$\boldsymbol{q}_k^{(\ell,s)} = \frac{\ell}{L} p_{k,\max} \boldsymbol{e}_s^{(S)}, \tag{1}$$

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

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

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

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

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

where for all $m \in \mathcal{M}$, s_m is the channel used by MBS m and ${\sigma_{0,m}^{(s_m)}}^2$ and ${\sigma_k^{(s_m)}}^2$ is the noise power over MUE m and the noise power over FUE k on the frequency band s.
All FBSs are interested in optimizing a given interference minimization/mitigation metric denoted by $\phi : \mathbb{R}^{S \cdot K + M} \rightarrow \mathbb{R}$, which determines at each instant t the impact of the interference on the macro system based on the observation of all the SINR levels $\gamma_k^{(s)}$ and $\gamma_{0,m}^{(s_m)}$, with $(k, s) \in \mathcal{K} \times \mathcal{S}$ and $m \in \mathcal{M}$. Later, we provide explicit expressions for ϕ depending on the interest of all FBSs.

C. Game Theoretic Model

The interference minimization problem described in the previous section can be 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 $t \in \{1, \ldots, \infty\}$. We describe in detail both formulations.

1) Short Term Formulation: Let us describe the network during the interval from $t-1$ to t by the game $\mathcal{G}(\boldsymbol{h}(t)) = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{\phi\})$. 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 power allocation vectors, i.e., $\mathcal{A}_k = \{\boldsymbol{q}_k^{(\ell,s)} : \ell \in \{0, \ldots, L_k\}$, and $s \in \mathcal{S}\}$. Finally, $\phi : \mathcal{H} \times \mathcal{A}_1 \times \mathcal{A}_K \rightarrow \mathbb{R}_+$ 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 $\boldsymbol{\pi}_k(t) = \left(\pi_{k,\boldsymbol{q}_k^{(0,0)}}(t), \pi_{k,\boldsymbol{q}_k^{(1,1)}}(t), ..., \pi_{k,\boldsymbol{q}_k^{(L_k,S_k)}}(t)\right)$ where $\pi_{k,\boldsymbol{q}_k^{(l_k,s_k)}}$ is the probability that femtocell k plays action $\boldsymbol{q}_k^{(l_k,s_k)}$ at time t, i.e.,

$$\pi_{k,\boldsymbol{q}_k^{(l_k,s_k)}} = \Pr\left(\boldsymbol{p}_k(t) = \boldsymbol{q}_k^{(l_k,s_k)}\right). \tag{4}$$

where $(l_k, s_k) \in \{1, ..., L_K\} \times \mathcal{S} \cup \{(0,0)\}$. 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.

2) Long-Term Formulation: The long-term behavior 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, $\forall \ \left(\boldsymbol{h}', \boldsymbol{h}''\right) \in \mathcal{H}^2$,

$\Pr\left(\boldsymbol{h}(t)=\boldsymbol{h}' \,\middle|\, \boldsymbol{h}(t-1)=\boldsymbol{h}''\right) = \Pr\left(\boldsymbol{h}(t)=\boldsymbol{h}'\right) = \pi_{\boldsymbol{h}'}$. 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. This assumption might appear restrictive, however, it perfectly models the time-varying nature of wireless channels, which are independent of the transmit configurations of radio devices.

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 $\tilde{\phi}_k(t)$ of the corresponding instantaneous interference minimization metric $\phi(\boldsymbol{h}(t),\boldsymbol{p}_k(t),\boldsymbol{p}_{-k}(t))$, i.e.,

$$\tilde{\phi}_k(t) = \phi(\boldsymbol{h}(t),\boldsymbol{p}_k(t),\boldsymbol{p}_{-k}(t)) + \varepsilon_{k,\boldsymbol{p}_k(t)}(t), \qquad (5)$$

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

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., $\bar{\phi}_k(t)$, which is calculated empirically based on the observations $\tilde{\phi}_k(t)$ as follows,

$$\bar{\phi}_k(t) = \frac{1}{t}\sum_{n=1}^{t}\tilde{\phi}_k(n). \qquad (6)$$

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 $(\ell_k,s_k)\in\{1,\ldots,L_k\}\times\mathcal{S}\cup\{(0,0)\}$, let $\hat{\phi}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t)$, be the estimation of time-average interference minimization metric obtained by playing action $\boldsymbol{q}_k^{(\ell_k,s_k)}$. This estimation is calculated as follows,

$$\hat{\phi}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t) = \frac{1}{T_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t)}\sum_{n=1}^{t}\tilde{\phi}_k(n)\mathbb{1}_{\left\{\boldsymbol{p}_k(n)=\boldsymbol{q}_k^{(\ell_k,s_k)}\right\}}, \qquad (7)$$

where, $T_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t) = \sum_{n=1}^{t}\mathbb{1}_{\left\{\boldsymbol{p}_k(n)=\boldsymbol{q}_k^{(\ell_k,s_k)}\right\}}$. Once the N_k-dimensional vector of estimations of FBS k is obtained, i.e., $\hat{\boldsymbol{\phi}}_k = \left(\hat{\phi}_{k,\boldsymbol{q}_k^{(0,0)}},\hat{\phi}_{k,\boldsymbol{q}_k^{(1,1)}},\ldots,\hat{\phi}_{k,\boldsymbol{q}_k^{(L_k,S)}}\right)$ 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 $\boldsymbol{\beta}_k\left(\hat{\boldsymbol{\phi}}_k(t)\right)$ must take into consideration that, FBSs must experiment between different actions such that the estimation vector $\hat{\boldsymbol{\phi}}_k(t)$ is improved at each time t, but also FBSs must optimize their respective interference minimization metric. In the following, we present the existing trade-off between both goals which might appear at a first glance as two antagonic processes.

III. Exploration vs. Performance Optimization

As shown in Sec II-C, all FBSs 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 [8], [9], the behavioral rule described above can be modeled by the probability distribution $\beta_k(\hat{\boldsymbol{\phi}}_k(t))$ satisfying,

$$\beta_k(\hat{\boldsymbol{\phi}}_k(t)) \in \arg\max_{\boldsymbol{\pi}_k\in\triangle(\mathcal{A}_k)}\left[\sum_{\boldsymbol{p}_k\in\mathcal{A}_k}\pi_{k,\boldsymbol{p}_k}\hat{\phi}_{k,\boldsymbol{p}_k}(t)+\kappa_k H(\boldsymbol{\pi}_k)\right] \qquad (8)$$

where H represents the Shannon entropy function. In general, given a probability measure $\pi_1,...,\pi_N$ over an set of N elements, it follows that

$$H(\pi_1,...,\pi_N) = -\sum_{n=1}^{N}\pi_n\ln(\pi_n). \qquad (9)$$

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 (8) is written as:

$$\beta_k(\hat{\boldsymbol{\phi}}_{\boldsymbol{k}}(t))=\left(\beta_{k,\boldsymbol{q}_k^{(0,0)}}(\hat{\boldsymbol{\phi}}_{\boldsymbol{k}}(t)),\beta_{k,\boldsymbol{q}_k^{(1,1)}}(\hat{\boldsymbol{\phi}}_{\boldsymbol{k}}(t)),\ldots,\beta_{k,\boldsymbol{q}_k^{(L_k,S)}}(\hat{\boldsymbol{\phi}}_{\boldsymbol{k}}(t))\right), \qquad (10)$$

where for all $k\in\mathcal{K}$ and for all $(\ell_k,s_k)\in\mathcal{L}_k\times\mathcal{S}\cup\{(0,0)\}$ and $k\in\mathcal{K}$:

$$\beta_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(\hat{\boldsymbol{\phi}}_k(t)) = \frac{\exp\left(\frac{1}{\kappa_k}\hat{\phi}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(t)\right)}{\sum_{\boldsymbol{p}_k\in\mathcal{A}_k}\exp\left(\frac{1}{\kappa_k}\hat{\phi}_{k,\boldsymbol{p}_k}(t)\right)}, \qquad (11)$$

where $\beta_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(\hat{\phi}_k(t))>0$, with strict inequality regardless of the estimation vector $\hat{\boldsymbol{\phi}}_k(t)$, with $t>0$. Equation (11) is known in the game theoretic jargon as *smooth* best response [7] and implies a different concept of equilibrium with respect to the classical Nash equilibrium. This equilibrium, known as logit equilibrium or Boltzmann-Gibbs equilibrium [12] is defined as follows,

Definition 1: **(Logit Equilibrium)**: *Consider the Markov game $\mathcal{G}=\{G(\boldsymbol{h}(t))\}_{t\geqslant 0}$. The mixed strategy profile $\boldsymbol{\pi}^*=(\boldsymbol{\pi}_1^*,...,\boldsymbol{\pi}_K^*)\in\triangle(\mathcal{A}_1),\times\ldots,\times\triangle(\mathcal{A}_K)$ is a logit equilibrium, if $\forall k\in\mathcal{K}$,*

$$\boldsymbol{\pi}_k^* = \boldsymbol{\beta}_k\left(\widehat{\boldsymbol{\phi}}_k\left(\boldsymbol{\pi}_{-k}^*\right)\right), \qquad (12)$$

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

$$\widehat{\phi}_{k,\boldsymbol{q}_k^{(\ell_k,s_k)}}(\boldsymbol{\pi}_{-k})=\mathbb{E}_{\boldsymbol{h}}\left[\sum_{\boldsymbol{p}_{-k}\in\mathcal{A}_{-k}}\left(\prod_{j\in\mathcal{K}\setminus\{k\}}\pi^*_{j,\boldsymbol{p}_j}\right)\phi(\mathbf{h},\boldsymbol{q}_k^{(\ell_k,s_k)},\mathbf{p}_{-k})\right].$$

Note that Def. 1 implies a fixed point equation as in the case of the classical Nash equilibrium [10], e.g., let $\boldsymbol{\beta} : \mathbb{R}^{\sum_{j=1}^{K} N_j} \rightarrow \triangle(\mathcal{A}_1) \times, \ldots \times \triangle(\mathcal{A}_K)$ be defined as follows,

$$\boldsymbol{\beta}(\boldsymbol{\pi}) = \left(\boldsymbol{\beta}_1\left(\widehat{\phi}_1(\boldsymbol{\pi}_{-1})\right), \ldots, \boldsymbol{\beta}_K\left(\widehat{\phi}_K(\boldsymbol{\pi}_{-K})\right)\right).$$

Thus, if $\boldsymbol{\pi}^*$ is a logit equilibrium then, $\boldsymbol{\pi}^* = \boldsymbol{\beta}(\boldsymbol{\pi}^*)$.

It is also important to remark that when $\kappa_k \rightarrow 0$, the resulting probability distribution approaches the best response (BR) correspondence. First, let us define the best response correspondence as follows:

Definition 2: **(Best Response Correspondence)**: *Consider the Markov game* $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t \geqslant 0}$. *For all* $k \in \mathcal{K}$, *the best response correspondance in pure strategies* $\mathrm{BR}_k : \mathcal{H} \times \prod_{i \in \mathcal{K} \setminus \{k\}} \triangle(\mathcal{A}_i) \rightarrow \mathcal{A}_k$ *is defined as follows:*

$$BR_k(\boldsymbol{h}(t), \boldsymbol{\pi}_{-k}(t)) = \arg\max_{\boldsymbol{q}_k \in \mathcal{A}_k} \sum_{\boldsymbol{p}_{-k} \in \mathcal{A}_{-k}} \left(\phi(\mathbf{h}(t), \boldsymbol{q}_k, \boldsymbol{p}_{-k}) \prod_{j \in \mathcal{K} \setminus \{k\}} \pi_{j, \boldsymbol{p}_j}(t) \right). \quad (13)$$

Now, we show how both smooth best response (11) and the best-response (13) are related when the perturbation (entropy term in (8)) vanishes.

Theorem 1: **(Theorem 2 in [9])** Consider the Markov game $\mathcal{G} = \{G(\boldsymbol{h}(t))\}_{t \geqslant 0}$. Then, for all $k \in \mathcal{K}$ and for all $(\ell_k, s_k) \in \mathcal{L}_k \times \mathcal{S} \cup \{(0,0)\}$, it holds that:

$$\lim_{\kappa_k \rightarrow 0} \beta_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}\left(\widehat{\boldsymbol{\phi}}_k(\boldsymbol{\pi}_{-k})\right) = \frac{\mathbb{1}_{\left\{\boldsymbol{q}_k^{(\ell_k, s_k)} \in \mathrm{BR}_k\left(\boldsymbol{\pi}_{-\boldsymbol{k}}\right)\right\}}}{|BR_k(\boldsymbol{\pi}_{-k})|}. \quad (14)$$

The proof of Theorem 1 follows the same arguments that the proof of *Theorem* 2 *in [9]*. Theorem 1 implies that as long as the perturbation vanishes, the mixed strategies obtained by using the smooth best-response approach a uniform probability distribution over all the actions of the best response at a given time t.

Conversely to the vanishing perturbation case, when $\kappa_k \rightarrow \infty$, the resulting probability distribution approaches the uniform probability distribution over all the set of feasible frequency-bands and power-levels sets. In the following section, we introduce a novel technique for achieving LE for a given set of constants $\kappa_1, \ldots, \kappa_K$.

IV. Achieving The Logit Equilibrium (LE)

In this section, we focus on the procedure used by FBSs to achieve a logit equilibrium for a given constant set of parameters κ_k, with $k \in \mathcal{K}$. We present the main result of this paper in the following proposition.

Proposition 1 (***Achieving the LE***)*: Consider the 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* $\forall(\ell_k, s_k) \in \{1, \ldots, L_k\} \times \mathcal{S} \cup \{(0,0)\}$,

$$\begin{cases} \hat{\phi}_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t) = \hat{\phi}_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t-1) + \alpha(t) \frac{\mathbb{1}_{\left\{\boldsymbol{p}_k(t) = \boldsymbol{q}_k^{(\ell_k, s_k)}\right\}}}{\pi_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t)} \left(\tilde{\phi}(t) - \hat{\phi}_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t-1)\right), \\ \pi_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t) = \pi_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t-1) + \lambda(t)\left(\boldsymbol{\beta}_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}\left(\hat{\boldsymbol{\phi}}_k(t)\right) - \pi_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t-1)\right), \end{cases} \quad (15)$$

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

$$\lim_{T \rightarrow \infty} \sum_{t=0}^{T} \alpha(t) + \lambda(t) = +\infty \quad (16)$$

$$\lim_{T \rightarrow \infty} \sum_{t=0}^{T} \alpha(t)^2 + \lambda(t)^2 < +\infty, \text{ and,} \quad (17)$$

$$\lim_{t \rightarrow \infty} \frac{\lambda(t)}{\alpha(t)} = 0. \quad (18)$$

Then, both learning processes in (15) converge for all $k \in \mathcal{K}$, *and it holds that,*

$$\lim_{t \rightarrow \infty} \boldsymbol{\pi}_k(t) = \boldsymbol{\pi}_k^*, \quad (19)$$

$$\lim_{t \rightarrow \infty} \hat{\phi}_{k, \boldsymbol{q}_k^{(\ell_k, s_k)}}(t) = \widehat{\boldsymbol{\phi}}_k(\boldsymbol{\pi}_{-k}^*), \quad (20)$$

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

The proof of Prop. 1 follows the same steps of the proof of Prop. 3 in [9] and can be described in three steps. First, the processes $\hat{\boldsymbol{\phi}}_1(t), \ldots, \hat{\boldsymbol{\phi}}_K(t)$ and the processes $\boldsymbol{\pi}_1(t), \ldots, \boldsymbol{\pi}_K(t)$ can be written as two stochastic approximation (SA) algorithms $\hat{\boldsymbol{\phi}}(t)$ and $\boldsymbol{\pi}(t)$ by stacking them as a single vectors. Second, both SA algorithms satisfy the standard conditions to approximate them by two ordinary differential equations (ODE) [11]. Finally, using (18), it can be assumed that the process $\hat{\boldsymbol{\phi}}(t)$ sees the process $\boldsymbol{\pi}(t)$ as almost time-invariant, and the process $\boldsymbol{\pi}(t)$ sees the process $\hat{\boldsymbol{\phi}}(t)$ as always calibrated to the current value of the former. Applying this reasoning to the asymptotic analysis of the ODEs leads to the proof of Prop. 1.

V. Simulation Results

Consider only one macrocell per frequency band and assume $S = 4$ available frequency bands, i.e., $M = S = 4$. The macrocell system is underlaid with $K = 2$ femtocells. We assume that femtocells have $L = 3$ transmit power levels and the average signal-to-noise ratio for the macro- and femtocells are 20 dB and 10 dB per frequency band, respectively. More precisely, $\forall(k,s) \in \mathcal{K} \times \mathcal{S}$, $\mathrm{SNR}_k^{(s)} = \frac{p_{k,\max}}{\sigma_{k,s}^2} = 10$dBs and $\forall m \in \mathcal{M}$, $\mathrm{SNR}_m^{(s_m)} = \frac{p_{0,m}}{\sigma_{0,m}^2} = 20$dBs. The minimum SINR of the macrocells are given by $\boldsymbol{\Gamma}_0 = (8, 9, 12, 13)$ dB. The interference minimization/minimization metric adopted in this numerical example is the following:

$$\phi(\mathbf{h}(\mathbf{t}), \boldsymbol{p}_k(t), \boldsymbol{p}_{-k}(t)) = \sum_{k=1}^{K} \sum_{s=1}^{S} \gamma_k^{(s)}(t) . \prod_{m=1}^{M} \mathbb{1}_{\left\{\{\gamma_{0,m}^{(s_m)}(t) > \Gamma_{0,m}^{(s_m)}\}\right\}}. \quad (21)$$

This metric at a given instant t is different from zero only if all the MBSs satisfy at time t the minimum SINR level required for their own communications. The non-zero value represents the sum of the achieved SINR of all the FBS in the system. Hence, as long as all MBSs see their QoS requirement satisfied, FBSs obtain a positive reward. This models a certain altruism from the behavior of the FBSs which sacrifice their performance to guarantee the QoS of the macrocell system. Many other interference minimization functions can be devised to model the problem depending on particular features of the network. However this is out of the scope of this paper.

In Figure 2, the evolution of the probability distribution over the set of actions taken by both FBSs is shown. As time goes by, FBS $k = 1$ increases the probability to transmit with the maximum power level on frequency band $s = 1$, while the probability of transmitting on other bands decreases. On the other hand, the probability that FBS $k = 2$ transmits with maximum transmit power level on carrier $s = 2$ increases with time, whereas the probabilities of transmitting over the other frequency bands decrease. It can also be seen that although femtocells do not communicate with each other, they coordinate their access to the spectrum by using different frequency bands with very high probability. Note also that $s = 1$ and $s = 2$ are the frequency bands where the corresponding MBS demands the lowest SINR level. Figure 3 shows the evolution of the corresponding ergodic transmission rate for both FBS. Finally, the evolution of the ergodic transmission rate of the macrocells is shown in Figure 4. Clearly, all requirements are satisfied for the macrocell system.

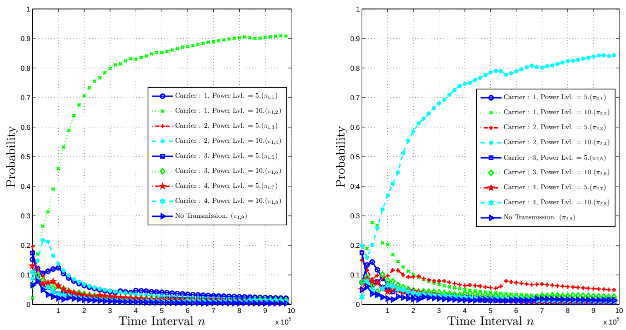

Fig. 2. Evolution of the probability distribution over the set of actions of femtocell 1 (left) and femtocell 2 (right).

VI. Conclusions and Future Work

In this paper, the cross-tier interference mitigation was studied in the framework of cognitive femtocells. Here, the behavior of femtocells aiming to avoid interference over the existing macrocells has been tackled using the fact that SINR messages fed-back from the end user terminals to their corresponding base stations can be decoded by all femtocells. Based on those observations, the behavior of femtocells is modeled by a smooth best response with respect to their expected interference minimization metric. This behavioral rule has been shown to lead the network to converge towards a logit equilibrium, where FBSs perfectly balance their willingness of experimenting several actions to discover the network, and adopting the actions for performance optimization. In our future work, we will investigate the possibility of improving the convergence speed as well as studying other learning mechanisms in the context of femtocell networks.

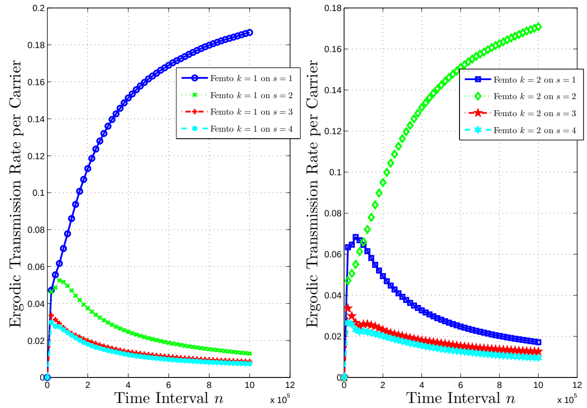

Fig. 3. Evolution of the ergodic transmission rate of femtocell 1 (left) and femtocel 2 (right).

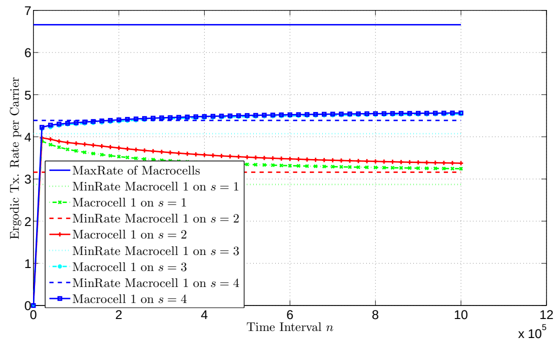

Fig. 4. Evolution of the ergodic transmission rate of the macrocell.

References

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

[2] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: A dynamic game approach," IEEE Transactions on Wireless Communications, vol. 7, no. 7, 2651-2660, July 2008.

[3] D. Niyato and E. Hossain, "Dynamic of network selection in heterogeneous wireless networks: an evolutionary game approach," IEEE Transactions on Vehicular Communications, vol. 58, no. 4, 2651-2660, May 2009.

[4] D. Fudenberg and D. K. Levine, "The Theory of Learning in Games," The MIT Press, Cambridge, MA, 1998.

[5] 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.

[6] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks," to appear in IEEE FEMnet (co-located with IEEE GLOBECOM), Miami, USA, Dec. 2010.

[7] Y. M. Kaniovski and H. P. Young, "Learning dynamics in games with stochastic perturbations," Games and Economic Behavior, vol. 11, no. 2, pp. 330-363, Nov. 1995.

[8] S. M. Perlaza, H. Tembine and S. Lasaulce, "How can ignorant but patient cognitive terminals learn their strategy and utility?," in proc. of the IEEE Intl. Workshop on Signal Processing Advances for Wireless Communications (SPAWC), Marrakesh, Morroco, June 2010.

[9] 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.

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

[11] V. S. Borkar, "Stochastic Approximation: A Dynamical System Viewpoint," Cambridge University Press, 2008.

[12] 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.