A Stochastic Game-Theoretic Approach for Interference Mitigation in Small Cell Networks

Jianchao Zheng, Yueming Cai, and Alagan Anpalagan

***Abstract*—In this letter, we investigate the distributed channel selection for interference mitigation in small cell networks considering the random deployment and dynamic on–off activity. Game theory is adopted to analyze the distributed and interactive decision making, and the problem is formulated as an exact potential game, in which the Nash equilibrium (NE) minimizes the expected network interference, either globally or locally. Furthermore, we design a fully distributed, no-regret channel selection algorithm to find the NE solution in dynamic environment.**

***Index Terms*—Small cell networks, interference mitigation, distributed channel allocation, potential game, no-regret learning.**

I. Introduction

SMALL cell networks, a mixture of low-power nodes such as micro-, pico- and femtocells, has emerged as a promising evolution of future wireless cellular system to boost network capacity and lower capital and operational expenditures [1]. However, their deployment faces a number of key technical challenges, among which self-organization and interference management are very crucial [1], [2]. Small cells (e.g., femtocells) are connected to capacity-limited Internet protocol (IP) backhauls, thus the increased delay for the control data exchange will impede any centralized control [3]. Moreover, due to the exponentially increasing complexity with the number of small cells, a centralized control may be infeasible in the hyper-densely deployed small cell networks [3], [4]. Besides, small cells are often user-deployed, and can be turned on/off autonomously by the end users according to their traffic variation [2], [4]. In this setting, fully distributed, scalable and self-organizing strategies for interference management in small cell networks are urgently required.

Game theory, a mathematical tool that analyzes strategic interactions among distributed decision makers, is seen as a natural paradigm to study this problem due to the pronounced mutual interference and coupling among small cells' strategies [6]. However, the random deployment and dynamic on–off activity impose great challenges on the existing game-theoretic framework. In this letter, we incorporate the no-regret learning automata into the stochastic game to solve the interference mitigation problem in self-organizing small cell networks. The main contributions of this letter are threefold: *i*) we formulate a stochastic game to investigate the distributed and interactive channel selection for the dynamic small cells networks; *ii*) the formulated game is proved to be an exact potential game, in which the Nash equilibrium (NE) solution minimizes the expected network interference, either globally or locally; and *iii*) we design a fully distributed, no-regret learning algorithm to find the NE in dynamic environment.

It is worth noting that most existing works perform centralized interference management which creates enormous signaling overhead [2], [5]. Besides, the existing distributed solutions using game theory mainly focus on the static network [1], [7], [8], and pay little attention to fully distributed scenarios without information exchange [7], [8]. Our work pioneers the investigation of interference mitigation in the dynamic and distributed small cell networks.

Manuscript received June 28, 2014; revised November 17, 2014; accepted November 19, 2014. Date of publication December 4, 2014; date of current version February 6, 2015. This work was supported in part by the National Natural Science Foundation of China under Grants 61301163 and 61301162 and in part by the Jiangsu Provincial Natural Science Foundation of China under Grant BK20130067. The associate editor coordinating the review of this paper and approving it for publication was Q. Du.

J. Zheng and Y. Cai are with the College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China (e-mail: longxingren.zjc.s@163.com; caiym@vip.sina.com).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: alagan@ee.ryerson.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LCOMM.2014.2377156

II. System Model and Preliminaries

We consider the downlink of an OFDMA-based heterogeneous cellular network, where multiple femtocells underlay each macrocell. Each macro- or femtocell is equipped with a macrocell base station (MBS) or femtocell base station (FBS). The set of MBSs and FBSs is denoted by $\mathcal{M} = \{m_1, m_2, \ldots, m_L\}$ and $\mathcal{F} = \{1, 2, \ldots, F\}$, respectively. We study the full frequency reuse deployment in which every macrocell shares the whole spectrum. The available spectrum is divided into K channels (i.e., subcarriers), denoted by $\mathcal{K} = \{1, 2, \ldots, K\}$. At each time slot, each MBS serves one macrocell user equipment (MUE) over each channel, and the downlink transmit power of MBS $m \in \mathcal{M}$ on channel $k \in \mathcal{K}$ is denoted by P_m^k. Besides, each FBS dynamically chooses one of the K available channels to serve its corresponding femtocell user equipment (FUE), and the transmit power of FBS $i \in \mathcal{F}$ on the selected channel is denoted by P_i.

The frequency reuse among macrocells and femtocells leads to severe co-channel interference. The instantaneous power gain from transmitter i to receiver[1] j on channel k is denoted by $H_{ij}^k = (d_{ij})^{-\alpha}\beta_{ij}^k$, where d_{ij} is the distance between transmitter i and receiver j, α is the path loss exponent, and β_{ij}^k is the random fading coefficient. In the two-tier coexisting network, the performance of the macrocell tier should be first guaranteed. Thus, it is intuitive to impose an interference temperature limit on the femtocell operations, as specified by

$$\sum_{i \in \mathcal{F}^k} P_i H_{im}^k \leq \upsilon_{m,k}, \qquad \forall k \in \mathcal{K},\ m \in \mathcal{M}, \tag{1}$$

where $\mathcal{F}^k$ represents the set of femtocells that accesses channel k, and $\upsilon_{m,k}$ denotes the interference temperature limit over channel k by the macrocell m.

[1] Receiver j refers to the user served by the corresponding BS j on channel k. With a little abuse of notation, they are denoted by the same index j.

Besides, the channels are modeled to be time-varying, since the instantaneous fading coefficient β_{mn}^k varies from time to time. Moreover, FBSs are considered to be dynamically on–off according to their traffic variations. For the convenience of analysis, we define a probability space as $(\Omega, \mathbb{P})$, where Ω is a sample space, $\mathbb{P}$ is a probability measure on Ω, and an event in the space Ω is denoted by ω. Then, a random vector $\mathbf{X}(\omega) = [\mathbf{H}(\omega), \mathbf{C}(\omega)]$ is used to characterize the dynamic environment, where $\mathbf{H} = [H_{ij}^k]_{\forall i,j,k}$ is the channel gain vector, and $\mathbf{C} = [c_i]_{\forall i \in \mathcal{F}}$ denotes the state of FBSs ($c_i \in \{0,1\}$, 0 for off, and 1 for on). In addition, define $C(t)$ as the active user set at time t, $C(t) = \{i \in \mathcal{F} : c_i^t = 1\}$, where c_i^t is the state of node i at time t.

III. Interference Mitigation Game

In this section, we formulate a stochastic game to investigate the optimal channel selection in dynamic environment from the perspective of minimizing interference.

A. Game Model

Since self-organizing small cells can sense, learn from the environment and autonomously tune their transmission strategies towards an optimal performance, it is intuitive to model each FBS as a game player, who independently selects its best channel for transmission. We denote the channel selection strategy of FBS i by a_i, and denote the channel selection profile of all the active (i.e., on-working) FBSs excluding FBS i by a_{-i}. Due to the dynamic variation of active FBSs and the dynamic channel environment, the generated interference is a random variable at each time and can vary from time to time. For a realization $\omega[t] \in \Omega$ at time t, the received interference by the user of FBS i at time t is given by:

$$\begin{aligned} I_i(a_i, a_{-i}, \omega[t]) &= I_i^F + I_i^M \\ &= -\sum_{j \in C(t)\setminus\{i\}} P_j H_{ji}^{a_i} \delta_{ji} - \sum_{m \in \mathcal{M}} P_m^{a_i} H_{mi}^{a_i}, \end{aligned} \tag{2}$$

where I_i^F denotes the interference from other FBSs (i.e., intra-layer interference), I_i^M denotes the interference from MBSs (i.e., cross-layer interference), and δ_{ji} is an indicator function given by $\delta_{ji} = \begin{cases} 1, & a_j = a_i, \\ 0, & \text{otherwise}. \end{cases}$

To minimize the received interference, the state-based utility function of FBS i is defined as:

$$\hat{u}_i(a_i, a_{-i}, \omega[t]) = -P_i I_i(a_i, a_{-i}, \omega[t]), \tag{3}$$

where the transmission power P_i is a weight factor to help the following analysis.[2] Considering that $\omega[t]$ is random during each time slot, we define the expected utility function as $u_i(a_i, a_{-i}) = \mathbf{E}_{\mathbf{X}}[\hat{u}_i(a_i, a_{-i}, \mathbf{X})] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \{\hat{u}_i(a_i, a_{-i}, \omega[t])\} = -p_n \mathbf{E}[I_i]$. Then, we formulate the following stochastic game as $\mathcal{G} = [\mathcal{F}, \{\mathcal{A}_i\}_{i\in\mathcal{F}}, \{u_i\}_{i\in\mathcal{F}}]$, where $\mathcal{F}$ is the set of femtocells who act as game players, $\mathcal{A}_i$ is the set of available channels for player i that is determined by (1). Each player autonomously tunes its channel strategy to maximize its expected utility, i.e.,

$$(\mathcal{G}): \quad \max_{a_i \in \mathcal{A}_i} u_i(a_i, a_{-i}), \qquad \forall i \in \mathcal{F}. \tag{4}$$

[2]Since power control is not discussed in this letter, P_i is a constant which does not affect the formulated game.

B. Analysis of Nash Equilibrium (NE)

Theorem 1: $\mathcal{G}$ is an exact potential game which has at least one pure strategy NE point that maximizes the potential function either globally or locally.

Proof: First, we construct a state-based potential function for an arbitrary realization $\omega[t] \in \Omega$ as

$$\begin{aligned} \hat{\Phi}(a_i, a_{-i}, \omega[t]) &= \hat{\Phi}_1(a_i, a_{-i}, \omega[t]) + \hat{\Phi}_2(a_i, a_{-i}, \omega[t]) \\ &= -\frac{1}{2}\sum_{i\in C(t)} P_i I_i^F - \sum_{i \in C(t)} P_i I_i^M, \end{aligned} \tag{5}$$

and the corresponding expected potential function is

$$\begin{aligned} \Phi(a_i, a_{-i}) &= \mathbf{E}_{\mathbf{X}}\left[\hat{\Phi}(a_i, a_{-i}, \mathbf{X})\right] \\ &= -\frac{1}{2}\sum_{i\in\mathcal{F}} P_i \mathbf{E}\left[I_i^F\right] - \sum_{i\in\mathcal{F}} P_i \mathbf{E}\left[I_i^M\right]. \end{aligned} \tag{6}$$

In (5),

$$\begin{aligned} &\hat{\Phi}_1(a_i, a_{-i}, \omega[t]) \\ &= -\frac{1}{2}\sum_{i\in C(t)} P_i I_i^F = -\frac{1}{2}\sum_{i\in C(t)} \sum_{j\in C(t)\setminus\{i\}} P_i P_j H_{ji}^{a_i}\delta_{ji} \\ &= -\frac{1}{2}\Bigg(\sum_{j\in C(t)\setminus\{i\}} P_i P_j H_{ji}^{a_i}\delta_{ji} \\ &\qquad + \sum_{n\in C(t)\setminus\{i\}} \sum_{j\in C(t)\setminus\{n\}} P_n P_j H_{jn}^{a_n}\delta_{jn}\Bigg) \\ &= -\frac{1}{2}\Bigg(\sum_{j\in C(t)\setminus\{i\}} P_i P_j H_{ji}^{a_i}\delta_{ji} \\ &\qquad \times \sum_{n\in C(t)\setminus\{i\}} \sum_{j\in C(t)\setminus\{n,i\}} P_n P_j H_{jn}^{a_n}\delta_{jn} \\ &\qquad + \sum_{n\in C(t)\setminus\{i\}} P_n P_i H_{in}^{a_n}\delta_{in}\Bigg). \end{aligned} \tag{7}$$

Under the assumption of interference symmetry [9], [10], i.e., $H_{ji}^k = H_{ij}^k$, we have

$$\sum_{j\in C(t)\setminus\{i\}} P_i P_j H_{ji}^{a_i}\delta_{ji} = \sum_{n\in C(t)\setminus\{i\}} P_n P_i H_{in}^{a_n}\delta_{in}, \tag{8}$$

and thus

$$\hat{\Phi}_1(a_i, a_{-i}, \omega[t]) = -\sum_{j\in C(t)\setminus\{i\}} P_i P_j H_{ji}^{a_i}\delta_{ji} - \Psi_1(a_{-i}), \tag{9}$$

where $\Psi_1(a_{-i}) = \frac{1}{2}\sum_{n\in C(t)\setminus\{i\}}\sum_{j\in C(t)\setminus\{n,i\}} P_n P_j H_{jn}^{a_n}\delta_{jn}$ is independent of player i's strategy a_i.

In addition,

$$\begin{aligned} \hat{\Phi}_2(a_i, a_{-i}, \omega[t]) &= -\sum_{i\in C(t)} P_i I_i^M = -\sum_{i\in C(t)}\sum_{m\in\mathcal{M}} P_i P_m^{a_i} H_{mi}^{a_i} \\ &= -\sum_{m\in\mathcal{M}} P_i P_m^{a_i} H_{mi}^{a_i} - \sum_{n\in C(t)\setminus\{i\}}\sum_{m\in\mathcal{M}} P_n P_m^{a_n} H_{mn}^{a_n} \\ &= -\sum_{m\in\mathcal{M}} P_i P_m^{a_i} H_{mi}^{a_i} - \Psi_2(a_{-i}), \end{aligned} \tag{10}$$

where $\Psi_2(a_{-i}) = \sum_{n\in C(t)\setminus\{i\}}\sum_{m\in\mathcal{M}} P_n P_m^{a_n} H_{mn}^{a_n}$ is also independent of player i's strategy.

TABLE I
NO-REGRET DYNAMIC CHANNEL SELECTION ALGORITHM

Initialization: Set the iteration $t = 0$ and each player arbitrarily selects one channel .

Loop for $t = 0, 1, 2, \ldots$

1) **Utility Update:** At the time t, each active player $i \in \mathcal{C}(t)$, calculates the utility of the current strategy $a_i \in \mathcal{A}_i$ and the utility of choosing a different strategy $a_i' \in \mathcal{A}_i$. Then the $|\mathcal{A}_i| \times |\mathcal{A}_i|$ instantaneous regret matrix Q_i^t is calculated by $Q_i^t(a_i, a_i') = \delta\left(a_i^{(t)} = a_i\right) \cdot \left[\hat{u}_i\left(a_i', a_{-i}^{(t)}, \omega[t]\right) - \hat{u}_i\left(a_i^{(t)}, a_{-i}^{(t)}, \omega[t]\right)\right]$.
2) **Average Regret Update:**

$$D_i^t(a_i, a_i') = D_i^{t-1}(a_i, a_i') + \varepsilon^t \left(Q_i^t(a_i, a_i') - D_i^{t-1}(a_i, a_i')\right), \quad (13)$$

$$R_i^t(a_i, a_i') = \left[D_i^t(a_i, a_i')\right]^+ = \max\left\{D_i^t(a_i, a_i'), 0\right\}, \quad (14)$$

where R_i^t represents the average regret matrix at time t and $\varepsilon^t = 1/(t+1)$ is the step size of update.
3) **Strategy Decision:** Assume a_i is the channel selected by player i at time t, i.e., $a_i^{(t)} = a_i$. Then at time $t+1$, player i updates its decision strategy according to the probability distribution:

$$\begin{cases} \Pr_i^{t+1}(a_i') = \frac{1}{\mu} R_i^t(a_i, a_i'), \forall a_i' \neq a_i \\ \Pr_i^{t+1}(a_i) = 1 - \sum_{a_i' \neq a_i} \Pr_i^{t+1}(a_i'), \end{cases} \quad (15)$$

where μ is a normalization factor.

End loop until the maximum number of iterations is reached.

Therefore,

$$\begin{aligned} &\hat{\Phi}\left(a_i', a_{-i}, \omega[t]\right) - \hat{\Phi}\left(a_i, a_{-i}, \omega[t]\right) \\ &= \hat{\Phi}_1\left(a_i', a_{-i}, \omega[t]\right) + \hat{\Phi}_2\left(a_i', a_{-i}, \omega[t]\right) \\ &\quad - \hat{\Phi}_1\left(a_i, a_{-i}, \omega[t]\right) - \hat{\Phi}_2\left(a_i, a_{-i}, \omega[t]\right) \\ &= \sum_{j \in C(t) \setminus \{i\}} P_i P_j H_{ji}^{a_i} \delta_{ji} + \sum_{m \in \mathcal{M}} P_i P_m^{a_i} H_{mi}^{a_i} \\ &\quad - \sum_{j \in C(t) \setminus \{i\}} P_i P_j H_{ji}^{a_i'} \delta_{ji} - \sum_{m \in \mathcal{M}} P_i P_m^{a_i'} H_{mi}^{a_i'} \\ &= \hat{u}_i\left(a_i', a_{-i}, \omega[t]\right) - \hat{u}_i\left(a_i, a_{-i}, \omega[t]\right). \end{aligned} \quad (11)$$

By taking expectation for the above equation, we can derive

$$u_i\left(a_i', a_{-i}\right) - u_i(a_i, a_{-i}) = \Phi\left(a_i', a_{-i}\right) - \Phi(a_i, a_{-i}). \quad (12)$$

It is seen from (12) that the change in individual utility function caused by any player's unilateral deviation equals to the change in the potential function. According to the definition given in [11], $\mathcal{G}$ is an exact potential game with Φ serving as a potential function. Moreover, any pure strategy NE point of the game $\mathcal{G}$ maximizes the potential function either globally or locally [11]. Therefore, Theorem 1 is proved. ■

Remark 1: The potential function defined in (6) is comprised of two parts: one is the negative of intra-layer interference, and the other is the negative of cross-layer interference. Thus, the NE points of the game minimize the intra-layer interference and cross-layer interference at the same time.

C. Achieving Nash Equilibrium in Dynamic Environment

In this subsection, we present a fully distributed, no-regret channel selection algorithm to achieve NE for the dynamic small cell networks, as shown in Table I.

In the implementation of the algorithm, the utility values, $\hat{u}_i(a_i', a_{-i}^{(t)}, \omega[t])$ and $\hat{u}_i(a_i^{(t)}, a_{-i}^{(t)}, \omega[t])$, are required to update the regret matrix as well as make strategy decision. However, because the utility function is designed as the weighted interference in (3), the utility values can be achieved by each player through scanning all the frequency bands and measuring the interference it experiences in each channel. Therefore, the implementation of this algorithm is fully distributed among each player who can independently and autonomously determine its own channel strategy. There is no information exchange needed.

Besides, since the no-regret procedure is originally designed by Hart and Mas-Colell [12] for the static environment, the convergence of this algorithm and achievable solution in our dynamic case need to be re-investigated.

Convergence Analysis: Denote the joint channel allocation strategy space by $\mathcal{A} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_F$, and let $e_j = [0, \ldots, 0, 1, 0, \ldots, 0]$ denote the $|\mathcal{A}|$ dimensional unit vector with a "1" in the j^{th} position. Thus, the empirical distribution of the F-tuple strategy up to time t can be defined by $\bar{z}^t = \frac{1}{t} \sum_{\tau \leq t} e_{A^\tau}$, where $A^\tau \in \mathcal{A}$ is the joint strategy at time τ. According to (13), we have

$$D_i^t(a_i, a_i') = \frac{1}{t} \sum_{\tau \leq t: a_n^{(\tau)} = a_i} \left[\hat{u}_i\left(a_i', a_{-i}^{(\tau)}, \omega[\tau]\right) - \hat{u}_i\left(a_i^{(\tau)}, a_{-i}^{(\tau)}, \omega[\tau]\right)\right]. \quad (16)$$

When $t \to \infty$, asymptotically,

$$D_i^t(a_i, a_i') \to \sum_{A \in \mathcal{A}: A_i = a_i} \bar{z}^t(A)\left[u_i\left(a_i', a_{-n}\right) - u_i(a_i, a_{-i})\right]. \quad (17)$$

Moreover, following Blackwell's Approachability Theorem, authors in [9] have proved that $\forall i \in \mathcal{F}$, $\forall a_i, a_i' \in \mathcal{A}_i$, each player's regret $R_i^t(a_i, a_i')$ converges to zero almost surely, i.e., $\lim_{t \to \infty} R_i^t(a_i, a_i') = [D_i^t(a_i, a_i')]^+ = 0$. Therefore, when $t \to \infty$, $\forall \alpha > 0$, $D_i^t(a_i, a_i') \leq \alpha$. Based on the definition of correlated equilibrium given in [12], we can derive that the empirical distributions of play $\bar{z}^t$ converge as $t \to \infty$ to the set of correlated equilibria of our game almost surely. Besides, since each player's regret $R_i^t(a_i, a_i')$ converges to zero when t is sufficiently large, each player will keep its current pure strategy. Because any pure strategy point in the set of correlated equilibria constitutes the pure strategy NE, we can conclude that the proposed no-regret algorithm converges to the NE of the game.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, we present simulation results to evaluate the proposed no-regret channel adaptation algorithm in a two-tier heterogeneous cellular network. The number of macrocells is 4, and there are 4 MUEs and 3 femtocells randomly located in each macrocell. The number of available channels is set to be 4, and the bandwidth of each channel is 200 KHz. The radius of the macrocell and femtocell is 650 m and 25 m, respectively [5]. The transmitting power of the macrocell and femtocell is 10 W and 1 W, respectively. Besides, to address the dynamic on–off, the active probability of each FBS is randomly set in [0, 1] at each time. Each active FBS selects the best channel to serve its corresponding FUE that is randomly located in its coverage. Rayleigh fading model is considered in the simulation, where the channel gains are exponentially distributed with unit-mean. The path loss exponent is set to be $\alpha = 4$, and the noise power experienced at each receiver is -130 dBm. Additionally, the interference temperature threshold $\upsilon_{m,k}$ is set to be 5×10^{-8} W, and the normalization factor of the proposed algorithm is set to be $\mu = 2 \times 10^{-8}$.

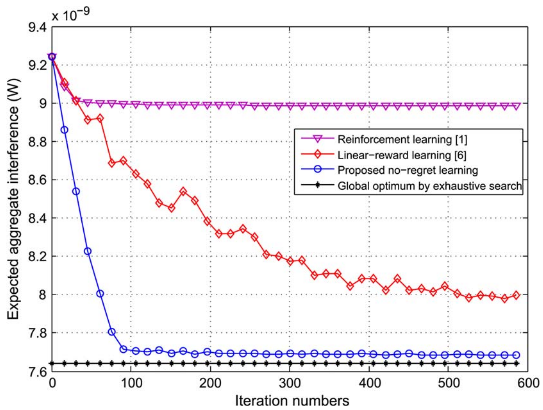

Fig. 1. Convergence behavior.

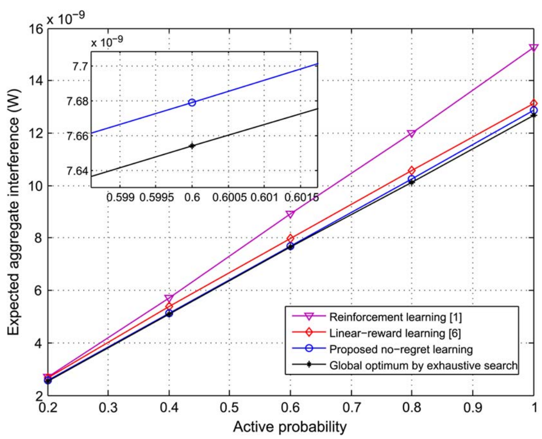

Fig. 2. Performance comparison of the expected aggregate interference.

The convergence behavior of proposed algorithm is shown in Fig. 1, in which the reinforcement learning approach [1], the linear-reward learning algorithm [6] and the global optimum by exhaustive search are plotted for comparison. The presented results are obtained by simulating 1000 independent trials and then taking the expected value. Specifically, the proposed algorithm can converge to the near-optimal solution within 100 iterations. The performance of the linear-reward learning algorithm follows, but it requires many more iterations for convergence. Besides, the convergence speed of the reinforcement learning algorithm is near to that of the proposed algorithm, but it converges to a worse solution generating more severe interference.

Fig. 2 plots performance comparison results for the different solutions when the active probability of the FBSs varies from 0.2 to 1. As shown in the figure, the expected aggregate interferences achieved by all the algorithms increase almost linearly with the active probability of the FBSs, which can be intuitively understood since a high active probability leads to more frequent spectrum sharing. As for the performance comparison, the proposed algorithm achieves a near-optimal solution, the linear-reward learning algorithm follows, and the reinforcement learning algorithm is worst, which parallel the results in Fig. 1. Moreover, the gap between the reinforcement learning algorithm and the global optimum increases with the active probability of the FBSs, while the proposed algorithm can always lead to near-optimal performance irrespective of the active probability. Besides, Fig. 3 presents performance comparison of the expected network sum-rate, which shows that the proposed algorithm minimizing the expected aggregate interference can also achieve near-optimal network sum-rate.

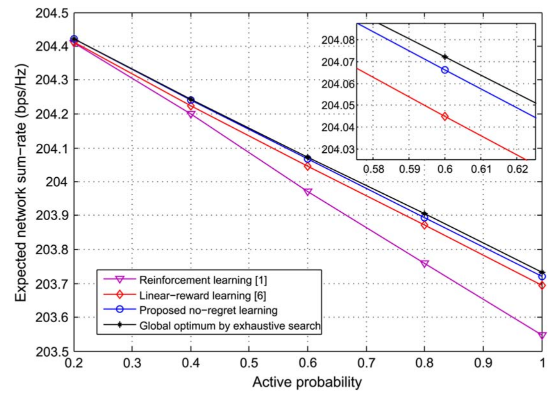

Fig. 3. Performance comparison of the expected network sum-rate.

V. Conclusion

In this letter, we used game theory to analyze the distributed channel selection for interference mitigation in randomly deployed and dynamically on–off small cell networks. The formulated game was proved to be an exact potential game, in which the Nash equilibrium (NE) solution minimizes the expected network interference, either globally or locally. Then, based on the no-regret procedure, we designed a fully distributed and dynamic channel selection algorithm that converges to the NE solution. Finally, simulations were conducted to validate the effectiveness of our proposed algorithm.

References

[1] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, "Self-organization in small cell networks: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3202–3212, Jul. 2013.

[2] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[3] H. ElSawy, E. Hossain, and D. I. Kim, "HetNets with cognitive small cells: User offloading and distributed channel access techniques," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 28–36, Jun. 2013.

[4] J. Xu *et al.*, "Cooperative distributed optimization for the hyper-dense small cell deployment," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 61–67, May 2014.

[5] F. Pantisano, M. Bennis, W. Saad, M. Debbah, and M. Latva-aho, "Improving macrocell–small cell coexistence through adaptive interference draining," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 942–955, Feb. 2014.

[6] J. Zheng, Y. Cai, N. Lu, Y. Xu, and X. Shen, "Stochastic game-theoretic spectrum access in distributed and dynamic environment," *IEEE Trans. Veh. Technol*, DOI: 10.1109/TVT.2014.2366559, to be published.

[7] J. Zheng, Y. Cai, Y. Liu, Y. Xu, B. Duan, and X. Shen, "Optimal power allocation and user scheduling in multicell networks: Base station cooperation using a game-theoretic approach," *IEEE Trans. Wireless Commun*, vol. 13, no. 12, pp. 6928–6942, Dec. 2014.

[8] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-aho, "Backhaul-aware interference management in the uplink of wireless small cell networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 942–955, Nov. 2013.

[9] J. Zheng, Y. Cai, Y. Xu, and A. Anpalagan, "Distributed channel selection for interference mitigation in dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4757–4762, Nov. 2014.

[10] J. Zheng, Y. Cai, W. Yang, Y. Wei, and W. Yang, "A fully distributed algorithm for dynamic channel adaptation in canonical communication networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 491–494, Oct. 2013.

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

[12] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.