This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

# Distributed Learning for Energy-Efficient Resource Management in Self-Organizing Heterogeneous Networks

Atefeh Hajijamali Arani, Abolfazl Mehbodniya, *Senior Member, IEEE*, Mohammad Javad Omidi, Fumiyuki Adachi, *Life Fellow Member, IEEE*, Walid Saad, *Senior Member, IEEE*, and Ismail Güvenç, *Senior Member, IEEE*

***Abstract*—In heterogeneous networks, a dense deployment of base stations (BSs) leads to increased total energy consumption, and consequently increased co-channel interference (CCI). In this paper, to deal with this problem, self-organizing mechanisms are proposed for joint channel and power allocation procedures which are performed in a fully distributed manner. A dynamic channel allocation mechanism is proposed, in which the problem is modeled as a noncooperative game, and a no-regret learning algorithm is applied for solving the game. In order to improve the accuracy and reduce the effect of shadowing, we propose another channel allocation algorithm executed at each user equipment (UE). In this algorithm, each UE reports the channel with minimum CCI to its associated BS. Then, the BS selects its channel based on these received reports. To combat the energy consumption problem, BSs choose their transmission power by employing an ON-OFF switching scheme. Simulation results show that the proposed mechanism, which is based on the second proposed channel allocation algorithm and combined with the ON-OFF switching scheme, balances load among BSs. Furthermore, it yields significant performance gains up to about $40.3\%$, $44.8\%$, and $70.6\%$ in terms of average energy consumption, UE's rate, and BS's load, respectively, compared to a benchmark based on an interference-aware dynamic channel allocation algorithm.**

***Index Terms*—Heterogeneous Networks; Energy Efficiency; Co-Channel Interference; Learning Algorithm.**

## I. Introduction

With the anticipated growth in smartphone penetration and demand for ubiquitous information access, next-generation wireless networks face enormous challenges to meet the ever-increasing network capacity [1]. Among them, the growing energy consumption of the networks is a main challenge which will directly result in the increase of carbon footprint and particularly environmental problems [2]. In the networks, base stations (BSs) consume a considerable amount of energy (amounting to about $60\%$–$80\%$ of the whole energy consumption) [3]. Therefore, reducing the energy consumption of BSs can lead to significant energy savings. In this respect, heterogeneous networks (HetNets), in which different macro cell BSs (MBSs) and small cell BSs (SBSs) must coexist, have emerged as a promising approach to increase capacity and improve energy efficiency (EE) of the next-generation wireless networks [2], [4]–[7].

However, the deployment of HetNets introduces significant technical issues which need to be addressed. Some of these issues include radio resource allocation and self-organization [8]. Unlike MBSs, SBSs are likely to be user-deployed in an ad hoc manner in unplanned locations without any operator supervision. Accordingly, self-organizing network (SON) has been recognized as a key approach which improves the network's performance, decreases the networking cost, and enhances the intelligent management [9]. SON allows SBSs to adapt to the changes in the network's conditions, and lead their strategies to provide proper performance in a distributed and flexible manner with minimal human intervention [10].

On the other hand, the traffic load of SBSs and MBSs is dynamic across time and space domain [11]. According to [12], for a large time portion during a day, the traffic load is much below the peak traffic load, and a large number of BSs are often under-utilized. In current cellular networks, BSs' deployment and operation are usually performed based on the peak traffic load. Therefore, in low traffic load situations, the EE of BSs will decrease. This observation requires network operators to design and utilize effective methods for managing the networks in a much more energy efficient way than the existing cellular networks. For instance, techniques such as BS ON-OFF switching (alternatively termed as sleep mode) and cell breathing are suitable solutions to reduce the energy consumption of BSs. In BS ON-OFF switching method, the BS can switch to low power consumption modes in light traffic load conditions. Using cell breathing technique allows the BS to adaptively adjust its cell size according to the traffic load conditions [13], [14]. Therefore, by using these methods, the network can be well adapted to spatial and temporal traffic fluctuations. Additionally, in dense deployment of HetNets, tackling the co-channel interference (CCI) is still a major challenge and a bottleneck towards improving network performance. Thus, the problem of channel sharing among BSs is of utmost importance.



This research is supported by "Towards Energy-Efficient Hyper-Dense Wireless Networks with Trillions of Devices", the Commissioned Research of NICT, JAPAN and KDDI foundation research grant, "Energy-Efficient Radio Resource Management for Next Generation Wireless Network", in part by the U.S. National Science Foundation under CNS-1406968 and CNS-1460333.

A. Hajijamali Arani and MJ.Omidi are with the Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, 84156-83111, Iran (e-mail: atefeh.haji@ec.iut.ac.ir; omidi@cc.iut.ac.ir).

A. Mehbodniya and F. Adachi are with the Research Organization of Electrical Communication Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577 Japan (e-mail: mehbodniya@riec.tohoku.ac.jp; adachi@ecei.tohoku.ac.jp).

W. Saad is with Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, 24061 USA, (e-mail: walids@vt.edu).

I. Güvenç is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC (e-mail: iguvenc@ncsu.edu).
Preliminary results from this work were presented at the IEEE ICC 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

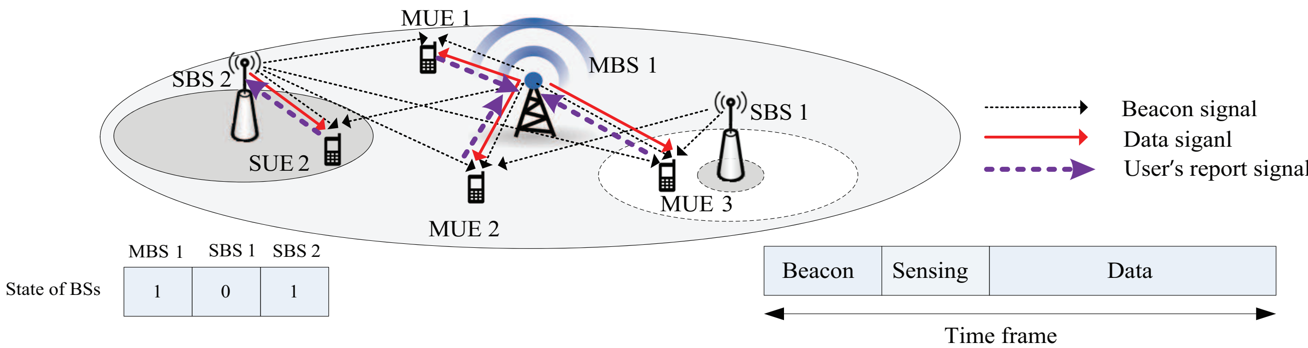


Figure 1. A two-tier HetNet model for the channel assignment scheme based on Gibbs sampling method with BSs which are able to switch between ON and OFF states. In beacon transmission phase, the BSs transmit their estimated loads. In sensing phase, the BSs receive the reports from their associated UEs. In data phase, the BSs send data packets to their associated UEs.

### A. Related Work

Several new techniques have been proposed to enhance the EE of HetNets via various approaches such as sleep mode and cell breathing [12], [15]–[22]. In [12], a user equipment (UE) association mechanism for sleeping cell UEs based on maximum mean channel access probability is proposed. The proposed mechanism adapts to traffic load at BSs, received signal power, and cumulative interference power. In [15], a cooperative optimization problem for maximizing EE, subject to minimum received signal strength with a set of clusters is considered. For solving the problem, each cluster employs simulated annealing search algorithm in a distributed manner. However, this work does not consider each BS's load. The authors in [16] proposed a set of small cell driven, core network driven, and UE driven sleep mode algorithms for SBSs in HetNets.

In [17], a centralized sleep mode technique for HetNets by exploiting the cooperation of cells is proposed. In [18], a multi-objective optimization problem based on sleep mode technique for an orthogonal frequency-division multiple access (OFDMA) based system is developed. In order to find the solution, a genetic algorithm is applied. However, the proposed mechanisms in [17] and [18] rely on centralized methods, and come at the expense of knowing global information. In [19], a random SBS ON-OFF switching scheme is introduced, where the UEs can delay their transmissions for a closer SBS to become available, thereby significantly improving EE.

In [20], network energy consumption is minimized by optimizing the cell sizes. Moreover, it is shown that the optimal cell size from an energy perspective depends on BS technology, data rates, and traffic demands. In [21], the authors concentrated on separating the control plane from the data plane, and propose a probabilistic sleeping approach based on the traffic load for data BSs. In [22], an opportunistic sleep mode scheme based on a noncooperative game for the downlink transmission of a two–tier HetNet is proposed. For solving the proposed game, a learning approach in a distributed manner is applied. Nonetheless, despite the insights glanced from these interesting studies [12], [15]–[18], [20]–[22] in improving the EE of the networks, they focus on only a particular challenge of the HetNets, i.e. energy consumption, and they do not consider channel allocation issues.

Since only a limited number of channels are allocated to the network, efficient assignment of channels among BSs is an important issue. In this regard, several studies have addressed channel allocation problem, and suggest some schemes to share channels among BSs. In [23], a dynamic channel assignment algorithm based on interference information in surrounding cells for a multihop cellular network is proposed. In [24], a centralized dynamic channel assignment based on interference conditions and traffic demands is considered. Several distributed dynamic channel assignment approaches for cellular networks are investigated in [25], [26]. Moreover, a number of schemes based on neural network [27]–[29], tabu search [30]–[33], genetic algorithm [34], and simulated annealing algorithm [35] have been proposed. Although efficient channel assignment is a significant part of resource allocation problem, it is necessary to take into consideration techniques in order to improve the EE of the networks that these works [23]–[35] do not consider EE issue in the networks, especially from the network operator's perspective.

The aforementioned works [12], [15]–[18], [20]–[35] do not jointly address power and channel allocation problems in a distributed manner. Moreover, they do not consider scenarios with UEs' mobility, handoff problems, and the impact of resource allocation mechanisms on UE's quality-of-service (QoS).

### B. Contribution

The main contribution of this paper is to introduce a novel framework for joint channel allocation and BS ON-OFF switching problem in densely deployed HetNets. We propose two novel channel allocation schemes. The first scheme is implemented at the level of BSs, allowing BSs to dynamically choose their channels, and adapt them to the network's conditions. Since the channels themselves do not have their own individual payoff functions or preferences, investigating noncooperative game is more suitable compared to a matching game with two-sided preferences. Moreover, in a matching game, the matching can require additional signaling between two player sets which can lead to overhead in the design, and increase complexity [36]. On the other hand, due to the distributed nature of networks and the competition among BSs, we cast the problem as a noncooperative game. To solve this game, a novel distributed learning-based approach is used, in which the knowledge of received interference on each channel is needed for choosing the appropriate channel. The need for this distributed learning solution is desirable as it has

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

several advantages. First, since BSs can autonomously select their channels based on the environmental information about the channels without information exchange, it can reduce the signaling overhead in the network. Moreover, this information is useful for BSs to select their strategies with better long-term performance. Second, a distributed decision taken by a BS is independent on the number of BSs in the network. Therefore, this approach is a suitable solution for densely deployed HetNets, and when the number of BSs varies over time, and there is no centralized controller. Furthermore, centralized approaches rely on a single controller entity. If the controller entity is compromised, then it can lead to failures at the entire network. Thus, this distributed approach can improve the network's robustness to failures and attacks.

The second proposed scheme is implemented at the level of UEs. In this algorithm, all UEs in the coverage area of a BS choose their priority channels, and report them to the BS. After collecting all the reports, the BS decides to choose its channel based on the reports received from its associated UEs, using Gibbs sampling method [37], [38]. Fig. 1 illustrates our two-tier network model based on BS ON-OFF switching and channel assignment algorithm based on Gibbs sampling method.

The advantage of the first proposed algorithm is its simplicity and low complexity as opposed to the second algorithm which requires feedback from UEs to BSs. However, we can show intuitively that the second algorithm manages the channels among BSs more effectively because of employing additional information from UEs. Furthermore, in order to improve the EE of the network, an ON-OFF switching algorithm similar to [39] is implemented at BSs, in which BSs dynamically choose their transmit power. The transmit power is decided using a machine-learning based game-theoretic algorithm, and by evaluating a payoff function. Moreover, we consider mobility for UEs, and use a Manhattan grid model which is a more realistic mobility model than the random waypoint model. The handoff decision is made according to BS's estimated load which is periodically advertised by BSs through beacon signals. Simulation results show that the proposed approaches improve the network's performance in terms of energy consumption and throughput as compared to the baseline approaches. As a result the proposed methods achieve both EE and spectral efficiency (SE) in a fully distributed manner without any need for information exchanges among BSs.

The rest of this paper is organized as follows. In Section II, we introduce our system model. Section III describes the problem formulation including UE association and BS operation problem. In Section IV, we propose our joint channel and power allocation schemes. Section V presents a no-regret learning approach for solving the proposed games. The simulation results are presented in Section VI, and conclusions are drawn in Section VII.

*Notations*: In this paper, the following notations are used. The regular and boldface symbols refer to scalars and matrices, respectively. For any finite set $\mathcal{A}$, the cardinality of set $\mathcal{A}$ and the set of all probability distributions over it are denoted by $|\mathcal{A}|$ and $\mathfrak{D}(\mathcal{A})$, respectively. $\{f(x)\,|g(x)\}$ represents the subset of all values $f(x)$ for which the assertion $g(x)$ about $x$ is true. $\boldsymbol{X} = [x_{i,j}]_{M\times N}$ and $x_{i,j}$ represent matrix $\boldsymbol{X}$ with dimension $M$-by-$N$ and the element in row $i$ and column $j$ of matrix $\boldsymbol{X}$, respectively, which $i = 1,\ldots,\ M$ and $j = 1,\ldots,\ N$. The function $\mathbb{1}_\phi$ denotes the indicator function which equals 1 if event $\phi$ is true and 0, otherwise. Moreover, $[x]^+ = \max\{0, x\}$ denotes the positive part of $x$. The set of real numbers is denoted by $\mathbb{R}$.

## II. System Model

### A. Deployment Scenario

We consider the downlink of a two-tier HetNet with a set of BSs $\mathcal{B} = \{1,\ldots,|\mathcal{B}|\}$ including a set of MBSs $\mathcal{B}_{\mathrm{M}}$ overlaid with a set of SBSs $\mathcal{B}_{\mathrm{S}}$, i.e. $\mathcal{B} = \mathcal{B}_{\mathrm{M}} \cup \mathcal{B}_{\mathrm{S}}$. For each coverage area, a MBS is located at the center of area, and SBSs are uniformly distributed within the coverage of the MBS. We denote by $\mathcal{K}$ the set of active uniformly distributed mobile UEs with different data rate demands. Since in this paper, we have been focusing more on issues regarding learning schemes for energy-efficient resource management, we assume a single antenna for each UE and BS. However, using multiple input multiple output-orthogonal frequency division multiplexing (MIMO-OFDM) technique can improve the transmission quality (throughput, bit error rate, etc.). Moreover, the power consumption model of BSs depends on the number of transmitting antennas.

We assume that the total bandwidth $W$ is equally divided among several orthogonal channels with bandwidth $W/|\mathcal{Q}|$ where $\mathcal{Q} = \{1,\ldots,|\mathcal{Q}|\}$ is the set of available channels, with $|\mathcal{Q}| < |\mathcal{B}|$. Moreover, MBSs and SBSs can operate over the same channel. OFDM symbols are grouped into a collection of physical resource blocks (RBs). We consider $R_{\mathrm{M}}$ and $R_{\mathrm{S}}$ RBs for MBSs and SBSs, respectively, that are distributed among their associated UEs. We assume an open access scheme for BSs in the system, i.e. the UEs are allowed to associate with any tier's BSs. In order to save power consumption under low traffic load, some BSs can be switched to an OFF state (or sleep mode). The vector of BSs' states at time $t$ in RB $r$ is denoted by $\boldsymbol{I}^{t,r} = [I_b^{t,r}]_{|\mathcal{B}|\times 1}$, with each element $I_b^{t,r}$ being equal to 1 if BS $b$ is in ON state (or active mode) at time $t$, and equal to 0, otherwise. Therefore, the set of active BSs at time $t$ in RB $r$ is given by:

$$\mathcal{B}_{\mathrm{ON}}^r(t) = \{b|I_b^{t,r} = 1, \forall b \in \mathcal{B}\ \}. \tag{1}$$

Let $P_{m,r}^{\mathrm{M}}(t)$ ($P_{s,r}^{\mathrm{S}}(t)$) be the transmit power of MBS $m \in \mathcal{B}_{\mathrm{M}}$ (SBS $s \in \mathcal{B}_{\mathrm{S}}$) over RB $r \in R_{\mathrm{M}}$ ($R_{\mathrm{S}}$) at time $t$. We denote by $q_{b,r}(t)$ the channel over which BS $b$ is transmitting at time $t$. We treat interference as noise. Therefore, the signal-to-interference-and-noise-ratio (SINR) at the receiver of macro cell UE (MUE) $k \in \mathcal{K}$ associated with MBS $m \in \mathcal{B}_{\mathrm{M}}$ transmitting over channel $q_{m,r}(t) \in \mathcal{Q}$, and allocated in RB $r \in R_{\mathrm{M}}$ at time $t$ is defined by (2). In (2), $g_{m,k}^{\mathrm{M}}(t)$ ($g_{s,k}^{\mathrm{S}}(t)$) denotes the total channel gain including path loss and lognormal shadow fading between MBS $m$ (SBS $s$) and UE $k$ at time $t$. Since the time scale for measuring the total channel gain is much larger than the time scale of fast fading, we do not consider fast fading. Let $\sigma^2$ be the additive white Gaussian noise (AWGN) power level per RB at the receiver of UEs, and

$$\gamma_{m,k,r}(t) = \frac{P^{\mathrm{M}}_{m,r}(t)\;g^{\mathrm{M}}_{m,k}(t)}{\underbrace{\sum_{\acute{m}\in\mathcal{B}_{\mathrm{M}},\;\acute{m}\neq m} P^{\mathrm{M}}_{\acute{m},r}(t) I^{t,r}_{\acute{m}}\; g^{\mathrm{M}}_{\acute{m},k}(t)\mathbb{1}_{(q_{m,r}(t)=q_{\acute{m},r}(t))}}_{I^{m,r}_{\mathrm{MBS}\rightarrow\mathrm{MBS}}} + \underbrace{\sum_{\acute{s}\in\mathcal{B}_{\mathrm{S}}} P^{\mathrm{S}}_{\acute{s},r}(t)\, I^{t,r}_{\acute{s}}\; g^{\mathrm{S}}_{\acute{s},k}(t)\,\mathbb{1}_{(q_{m,r}(t)=q_{\acute{s},r}(t))}}_{I^{m,r}_{\mathrm{SBS}\rightarrow\mathrm{MBS}}} + \sigma^2}. \tag{2}$$

$$\gamma_{s,k,r}(t) = \frac{P^{\mathrm{S}}_{s,r}(t)\;g^{\mathrm{S}}_{s,k}(t)}{\underbrace{\sum_{\acute{m}\in\mathcal{B}_{\mathrm{M}}} P^{\mathrm{M}}_{\acute{m},r}(t) I^{t,r}_{\acute{m}}\; g^{\mathrm{M}}_{\acute{m},k}(t)\mathbb{1}_{(q_{s,r}(t)=q_{\acute{m},r}(t))}}_{I^{s,r}_{\mathrm{MBS}\rightarrow\mathrm{SBS}}} + \underbrace{\sum_{\acute{s}\in\mathcal{B}_{\mathrm{S}},\;\acute{s}\neq s} P^{\mathrm{S}}_{\acute{s},r}(t) I^{t,r}_{\acute{s}}\; g^{\mathrm{S}}_{\acute{s},k}(t)\mathbb{1}_{(q_{s,r}(t)=q_{\acute{s},r}(t))}}_{I^{s,r}_{\mathrm{SBS}\rightarrow\mathrm{SBS}}} + \sigma^2}. \tag{3}$$

assumed to be constant for all UEs. $I^{m,r}_{\mathrm{MBS}\rightarrow\mathrm{MBS}}$ ( $I^{s,r}_{\mathrm{MBS}\rightarrow\mathrm{SBS}}$) and $I^{m,r}_{\mathrm{SBS}\rightarrow\mathrm{MBS}}$ ($I^{s,r}_{\mathrm{SBS}\rightarrow\mathrm{SBS}}$) indicate the interference caused by MBSs and SBSs over MBS $m$ (SBS $s$), respectively. The SINR at the receiver of small cell UE (SUE) $k \in \mathcal{K}$ associated with SBS $s \in \mathcal{B}_{\mathrm{S}}$ transmitting over channel $q_{s,r}(t) \in \mathcal{Q}$, and allocated in RB $r \in R_{\mathrm{S}}$ at time $t$ is defined by (3).

From Shannon's capacity formula, the achievable transmission rate of UE $k$ from BS $b$ in RB $r$ at time $t$ in bit/sec/Hz is given by:

$$\mathcal{C}_{k,r}(t) = \frac{W}{|\mathcal{Q}|}\log_2\left(1+\gamma_{b,k,r}(t)\right). \tag{4}$$

We assume that new flows arrive into the system with mean arrival rate $\lambda_k(t)$ and mean packet size $1/\mu_k(t)$ for UE $k$ at time $t$. This assumption can capture different spatial traffic variations by setting different arrival rates or file sizes for different UEs, i.e. heterogeneous UEs. Therefore, the load density of BS $b$ at time $t$ is defined as $\boldsymbol{l}_b(t) = \left\{ l^k_b(t) \,\middle|\, k \in A^{t,r}_b \right\}$, which $l^k_b(t) = \frac{\lambda_k(t)}{\mu_k(t)\mathcal{C}_{k,r}(t)}$ represents the time fraction required to deliver the traffic load density from BS $b$ to UE $k$. Let $A^{t,r}_b$ be the set of UE associated with BS $b$ in RB $r$ at time $t$, which is defined in Section III. Hence, the load of BS $b$ at time $t$ is expressed by:

$$L_b(t) = \sum_{k\in A^{t,r}_b} l^k_b(t). \tag{5}$$

Moreover, for a queue system M/M/1, the average number of flows at BS $b$ is equal to $\frac{L_b}{1-L_b}$ which is proportional to the expected delay at BS $b$ [40]. The load of BS is an important metric for evaluating the UE's QoS. Mathematically speaking, all the UEs which are associated with a BS can be satisfied if the BS's load does not exceed 1. Nevertheless, when the BS's load exceeds 1, some UEs which are associated with it may experience a sudden drop in their received throughput. This phenomenon is referred as *dropped UEs*.

### B. Power Consumption Model

In HetNets, the BSs of different tiers have different power consumption. The total input power of a BS consists of the transmission power and power consumed by the components of the BS, including power amplifier, radio frequency module, cooling system, baseband unit, DC-DC power supply, and main supply. There are several models available for power consumption for the networks. In a simple model, it is assumed that the BSs do not consume any power in a sleep mode. However, such a model is not realistic, and even in the sleep mode, the BSs still consume a certain amount of power for sensing purposes. To address this issue, we consider a linear power model for the BSs in the network [41]. Such a linear model has two parts including static and dynamic part. The static part involves the power consumed in the sleep mode, while the dynamic part depends on the transmit power. Therefore, the total power consumed by the BSs in the network at time $t$ can be expressed as follows [39]:

$$P_{\mathrm{Network}}(t) = \sum_{b\in\mathcal{B}} \sum_{\substack{r\in R_j,\\ j\in\{\mathrm{M,S}\}}} P^{\mathrm{Total}}_{b,r}(t), \tag{6}$$

where

$$P^{\mathrm{Total}}_{b,r}(t) = P^{\mathrm{OFF}}_b + \frac{P^j_{b,r}(t)\mathbb{1}_{(b\in\mathcal{B}^r_{\mathrm{ON}}(t))}}{\eta^{\mathrm{PA}}_b\Lambda(1-\lambda^{\mathrm{Feed}}_b)},\; j\in\{\mathrm{M,S}\}, \tag{7}$$

with

$$P^{\mathrm{OFF}}_b = \frac{P^{\mathrm{RF}}_b + P^{\mathrm{BB}}_b}{\Lambda}, \tag{8}$$

and

$$\Lambda = \left(1-\lambda^{\mathrm{DC}}_b\right)\left(1-\lambda^{\mathrm{MS}}_b\right)\left(1-\lambda^{\mathrm{Cool}}_b\right), \tag{9}$$

where $P^{\mathrm{Total}}_{b,r}(t)$ and $P^{\mathrm{OFF}}_b$ are the total power consumption and the power consumption in sleep mode by BS $b$ in RB $r$ at time $t$, respectively. $P^{\mathrm{RF}}_b$and $P^{\mathrm{BB}}_b$ denote the power of the radio frequency module and the total power of baseband engine consumed by BS $b$, respectively. $\eta^{\mathrm{PA}}_b$ indicates the power amplifier efficiency of BS $b$. $\lambda^{\mathrm{Feed}}_b$,$\lambda^{\mathrm{DC}}_b$, $\lambda^{\mathrm{MS}}_b$, and $\lambda^{\mathrm{Cool}}_b$ represent losses which are incurred by feeder, DC-DC power supply, main supply, and cooling system, respectively [15]. We assume that all parameters except $P^j_{b,r}(t)$ for BSs in any tier are constant over time. However, this model does not capture the embodied energy which comprises the energy consumed in the initial BS manufacturing and the energy associated with maintaining during the effective lifetime of BSs [42].

## III. Problem Formulation

Given this model, our goal is to improve the EE of the network and tackle the CCI problem by focusing on joint power and channel allocation in a distributed manner. Furthermore, the association of the UEs to the BSs is determined by the received signal power at the location of UE and BS's

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

estimated load. At each time $t$, the HetNet can be configured dynamically based on a transmission power vector, $\boldsymbol{P}_r(t)$, a channel vector, $\boldsymbol{Q}_r(t)$, an association matrix between the UEs and BSs, $\boldsymbol{A}^{t,r}$, and a state vector, $\boldsymbol{I}^{t,r}$, as follows:

$$\begin{aligned} &\boldsymbol{P}_r(t) = [P_{b,r}^{j}(t)]_{|\mathcal{B}|\times 1}, \quad r \in R_j,\ b \in \mathcal{B},\ j \in \{\mathrm{M}, \mathrm{S}\}, \\ &\boldsymbol{Q}_r(t) = [q_{b,r}(t)]_{|\mathcal{B}|\times 1}, \qquad\qquad \boldsymbol{A}^{t,r} = [a_{b,k}^{t,r}]_{|\mathcal{B}|\times|\mathcal{K}|}, \\ &\boldsymbol{I}^{t,r} = [I_b^{t,r}]_{|\mathcal{B}|\times 1}. \end{aligned} \tag{10}$$

Each single binary element $a_{b,k}^{t,r}$ in matrix $\boldsymbol{A}^{t,r}$ represents the association relation between UE $k$ and BS $b$ such that $a_{b,k}^{t,r} = 1$ indicates UE $k$ is associated with BS $b$ at time $t$, otherwise $a_{b,k}^{t,r} = 0$. We define the set of UE associated with BS $b$, $A_b^{t,r}$, at time $t$ as follow:

$$A_b^{t,r} = \left\{ k \,\middle|\, a_{b,k}^{t,r} = 1, \forall k \in \mathcal{K} \right\}. \tag{11}$$

We assume that each BS $b \in \mathcal{B}$ broadcasts its estimated load through a beacon signal in the downlink transmission [39]. Then, the UEs periodically assess their actual performance. Thus, the set of dropped UEs, $\mathcal{U}$, and the set of new UEs joined to the network, $\mathcal{N}$, should perform new association processes in order to assign to new BSs. Moreover, we assume that each UE $k \in \mathcal{K}$ is associated with at most one BS at each time $t$, i.e., $\max(\sum_{b\in\mathcal{B}} a_{b,k}^{t,r}) = 1$. Given the set of BSs, we define an optimization problem which involves finding the elements of matrix $\boldsymbol{A}^{t,r}$ (i.e., $a_{b,k}^{t,r}$ $\forall b \in \mathcal{B}$, and $\forall k \in \mathcal{K}$) as follows:

$$\begin{aligned} \max_{\boldsymbol{A}^{t,r}=[a_{b,k}^{t,r}]} \quad & \sum_{\substack{\forall k\in\mathcal{K},\\ k\in(\mathcal{U}\bigcup\mathcal{N})}} \sum_{\substack{\forall b\in\mathcal{B},\\ j\in\{\mathrm{M},\mathrm{S}\}}} \Big\{ a_{b,k}^{t,r}\ P_{b,r}^{j}(t) I_b^{t,r} (\beta_b \mathbb{1}_{(b\in\mathcal{B}_\mathrm{S})} \\ & \qquad + \mathbb{1}_{(b\in\mathcal{B}_\mathrm{M})}) g_{b,k}^{j}(t)(1-\hat{L}_b(t)) \Big\}, \\ \text{subject to} \quad & a_{b,k}^{t,r} \in \{0,1\}, \quad \forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \\ & \sum_{\forall b\in\mathcal{B}} a_{b,k}^{t,r} \le 1, \quad \forall k \in \mathcal{K}, \end{aligned} \tag{12}$$

where $\beta_b$ denotes the cell range expansion (CRE) bias used by SBS $b \in \mathcal{B}_\mathrm{S}$ in order to effectively expand its coverage area. By convention, MBSs have a bias 1 (0 dB) [43]. Moreover, SBSs can adaptively optimize this parameter based on the network's conditions [44]. However, here we assume that the SBSs have a fixed CRE bias. Let $\hat{L}_b(t)$ be the estimated load of BS $b \in \mathcal{B}$ at time $t$, which is obtained according to:

$$\hat{L}_b(t) = \hat{L}_b(t-1) + \left(\frac{1}{t}\right)^{\alpha} \Big( L_b(t-1) - \hat{L}_b(t-1) \Big), \tag{13}$$

where $\alpha > 0$ is learning rate exponent for the load estimation.

For solving the above problem, the UEs choose their associated BSs using exhaustive search within the set of BSs. For each BS $b \in \mathcal{B}$, we define a utility function which is the difference between its benefit and cost, and the BS aims at maximizing it. Since each BS has incentive to increase the number of UEs served by it, we consider the fraction of UEs served by the BS (serving ratio) as the benefit [45]. The cost function for each BS is including its total energy consumption and load. The weighted benefit function $n_b^r(t)$ and cost function $c_b^r(t)$ for BS $b$ in RB $r$ at time $t$ can be expressed as [46]:

$$n_b^r(t) = \omega_b^{\mathrm{n}} \frac{|A_b^{t,r}|}{|\mathcal{K}|}, \tag{14}$$

$$c_b^r(t) = \omega_b^{\mathrm{l}} L_b(t) + \omega_b^{\mathrm{p}} \frac{P_{b,r}^{\mathrm{Total}}(t)}{P_b^{\mathrm{OFF}} + P_b^{\mathrm{TXMax}}}, \tag{15}$$

where

$$P_b^{\mathrm{TXMax}} = \frac{P_b^{\mathrm{Max}}}{\eta^{\mathrm{PA}} \Lambda (1 - \lambda_b^{\mathrm{Feed}})}, \tag{16}$$

where $P_b^{\mathrm{Max}}$ denotes the maximum transmit power of BS $b$, while $\omega_b^{\mathrm{n}}$, $\omega_b^{\mathrm{l}}$, and $\omega_b^{\mathrm{p}}$ denote the weight parameters which indicate the impact of subscription benefit, load, and energy on the utility function for each BS $b \in \mathcal{B}$, respectively. Hence, we define the utility function of BS $b$ in RB $r$ at time $t$ by,

$$\begin{aligned} \psi_b^r(t) =& n_b^r(t) - c_b^r(t) = \\ & \omega_b^{\mathrm{n}} \frac{|A_b^{t,r}|}{|\mathcal{K}|} - \omega_b^{\mathrm{l}} L_b(t) - \omega_b^{\mathrm{p}} \frac{P_{b,r}^{\mathrm{Total}}(t)}{P_b^{\mathrm{OFF}} + P_b^{\mathrm{TXMax}}}. \end{aligned} \tag{17}$$

In order to deal with the CCI problem, each BS aims at selecting the channel with minimum CCI power. Thus, the maximization of network utility and minimization of the total interference to find the optimal resource allocation profile are given by the following problems:

$$\begin{aligned} (\mathbf{P1}) \quad & \max_{\boldsymbol{P}_r(t),\, \boldsymbol{Q}_r(t),\, \boldsymbol{I}^{t,r}} \sum_{\substack{\forall r\in R_j,\\ j\in\{\mathrm{M},\mathrm{S}\}}} \sum_{\forall b\in\mathcal{B}} \psi_b^r(t), \\ \text{subject to} \quad & 0 \le L_b(t) \le 1, \quad \forall b \in \mathcal{B}, \\ & \sum_{r\in R_j} P_{b,r}^{j}(t) \le P_b^{\mathrm{Max}}, \quad \forall b \in \mathcal{B}, j \in \{\mathrm{M},\mathrm{S}\}, \\ & q_{b,r}(t) \in \mathcal{Q}, \quad \forall b \in \mathcal{B}, \\ & I_b^{t,r} \in \{0,1\}, \quad \forall b \in \mathcal{B}, \end{aligned} \tag{18}$$

and

$$\begin{aligned} (\mathbf{P2}) \quad & \min_{\boldsymbol{P}_r(t),\, \boldsymbol{Q}_r(t),\, \boldsymbol{I}^{t,r}} \sum_{\substack{\forall r\in R_j,\\ j\in\{\mathrm{M},\mathrm{S}\}}} \sum_{\forall b\in B} (I_{\mathrm{MBS}\to\mathrm{MBS}}^{b,r} + I_{\mathrm{SBS}\to\mathrm{MBS}}^{b,r}) \\ & \qquad \mathbb{1}_{(b\in\mathcal{B}_\mathrm{M})} + (I_{\mathrm{MBS}\to\mathrm{SBS}}^{b,r} + I_{\mathrm{SBS}\to\mathrm{SBS}}^{b,r}) \mathbb{1}_{(b\in\mathcal{B}_\mathrm{S})}, \\ \text{subject to} \quad & 0 \le L_b(t) \le 1, \quad \forall b \in \mathcal{B} \\ & \sum_{r\in R_j} P_{b,r}^{j}(t) \le P_b^{\mathrm{Max}}, \quad \forall b \in \mathcal{B}, j \in \{\mathrm{M},\mathrm{S}\}, \\ & q_{b,r}(t) \in \mathcal{Q}, \quad \forall b \in \mathcal{B}, \\ & I_b^{t,r} \in \{0,1\}, \quad \forall b \in \mathcal{B}. \end{aligned} \tag{19}$$

In this paper, we deal with above problems with a two-step procedure, including power and channel assignment.

## IV. Opportunistic Resource Allocation for Self-Organized HetNets

In this section, we propose our distributed power and channel assignment schemes. We assume that the BSs independently select their transmission strategies. In this regard, for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

power allocation, each BS individually selects its transmission power according to the following ON-OFF switching approach. For channel allocation, each BS selects its channel by utilizing one of our proposed channel assignment mechanisms. Note that, the traffic pattern dynamically varies in time and space domain, but could be assumed approximately constant for a particular period of time. Therefore, we can make a reasonable assumption of time-scale, in which the variation of network traffic pattern is determined at a slower time-scale compared to the dynamic BS operation. Furthermore, based on the number of iterations required for convergence and the processor's strength, the convergence time for the proposed algorithms would be in the order of minutes, where traffic pattern would not experience a significant change.

### *A. Power Allocation Sub-Problem: Game-Theoretic Framework (ON-OFF Switching)*

The described resource allocation problem can be formulated as noncooperative games. Here, we only consider power allocation sub-problem, and name it ON-OFF switching approach. The normal form of the game is expressed as follows:

$$\mathcal{G}^{(P)} = \left(\mathcal{B}, \{\mathcal{S}_b^{(P)}\}_{b\in\mathcal{B}}, \left\{u_b^{(P)}\right\}_{b\in\mathcal{B}}\right), \tag{20}$$

where $\mathcal{B}$ represents the set of players, $\mathcal{S}_b^{(P)} = \{s_{b,1}^{(P)}, \ldots, s_{b,|\mathcal{S}_b^{(P)}|}^{(P)}\}$ is the strategy set of player $b$ where $s_{b,i}^{(P)}$ denotes the $i$th pure strategy of player $b$. $u_b^{(P)}$ represents the payoff function of player $b$ in game $\mathcal{G}^{(P)}$. The players, strategy set and payoff function are defined as follows:

- **Players:** The set of players $\mathcal{B}$ corresponds to the set of BSs $\mathcal{B}$. For a valid game we require $|\mathcal{B}|\geqslant 1$.
- **Strategy set:** A pure strategy of player $b$ is its transmission power. The available pure strategies for BS $b$ are $\mathcal{S}_b^{(P)} = \{0, \frac{1}{|\mathcal{S}_b^{(P)}|-1}P_b^{\mathrm{Max}}, \ldots, \frac{|\mathcal{S}_b^{(P)}|-1}{|\mathcal{S}_b^{(P)}|-1}P_b^{\mathrm{Max}}\}$ where $|\mathcal{S}_b^{(P)}|\geq 2$. Since the strategy space for player $b$ is including of $s_{b,i}^{(P)} = 0$ and $s_{b,i}^{(P)} \neq 0$, the strategy of BS $b$ not only is composed of transmission power, but also it is composed of the state of BS $b$.
- **Payoff:** The payoff function $u_b^{(P)}$ of player $b$ is defined as (17).

In Section V, we discuss the solution of this game.

### *B. Channel Assignment Sub-Problem*

In this subsection, we focus on channel assignment sub-problem, and propose two novel channel assignment mechanisms. In particular, our proposed mechanisms function in a fully distributed manner. Thus, no central controller is needed. In our first proposed mechanism, we investigate the competition among BSs using a game-theoretic approach. In our proposed game, BSs are the players, in which each player aims at choosing a channel with minimum CCI power. Then, a no-regret learning approach, presented in Section V, is used to solve the game. The algorithm is a probabilistic learning procedure, in which each strategy is played with a probability based on the regret measurement. The proposed algorithm is particularly useful because it does not need to exchange information in the network.

Our second approach leverages reports from UEs to BSs, in which each UE computes the average received CCI power over each channel, and selects the channel having the lowest averaged CCI power. Then, BSs choose their channels based on the reports received from their associated UEs. For this case, BSs apply a probabilistic approach for choosing their channels inspired from softmax selection approach [47], using Gibbs sampling method.

*1) Channel Assignment Mechanism 1: Game-Theoretic Framework:* The channel assignment sub-problem can also be formulated as a noncooperative game. The normal form of game is expressed as follows:

$$\mathcal{G}^{(Q)} = \left(\mathcal{B}, \{\mathcal{S}_b^{(Q)}\}_{b\in\mathcal{B}}, \left\{u_b^{(Q)}\right\}_{b\in\mathcal{B}}\right), \tag{21}$$

where $\mathcal{S}_b^{(Q)}$ is the strategy set of player $b$, and $u_b^{(Q)}$ represents the payoff function of player $b$ in game $\mathcal{G}^{(Q)}$. The differences between $\mathcal{G}^{(Q)}$ and the game described in previous subsection, $\mathcal{G}^{(P)}$, are in the strategy set of players and payoff function. Therefore, the strategy set and payoff function are defined as follows:

- **Strategy set:** A pure strategy of player is its channel. Therefore, the available pure strategies for each player $b$ are $\mathcal{S}_b^{(Q)} = \{s_{b,1}^{(Q)}, \ldots, s_{b,|\mathcal{S}_b^{(Q)}|}^{(Q)}\}$ where $\mathcal{S}_b^{(Q)} = \mathcal{Q}$.
- **Payoff:** The payoff function of player $b$ is defined as follows:

$$\begin{aligned} u_b^{(Q)} &= -(I_{\mathrm{MBS}\to\mathrm{MBS}}^{b,r} + I_{\mathrm{SBS}\to\mathrm{MBS}}^{b,r})\mathbb{1}_{(b\in\mathcal{B}_{\mathrm{M}})} \\ &\quad - (I_{\mathrm{MBS}\to\mathrm{SBS}}^{b,r} + I_{\mathrm{SBS}\to\mathrm{SBS}}^{b,r})\mathbb{1}_{(b\in\mathcal{B}_{\mathrm{S}})}. \end{aligned} \tag{22}$$

To solve the game, we apply the learning algorithm described in Section V, and name it dynamic channel assignment based on learning algorithm (DCA-LA). In this algorithm, each player is interested in minimizing its average regret over time. As a result, it assigns higher probability to the channel with more regret.

*2) Channel Assignment Mechanism 2: Gibbs Sampling-Based UE Interference-Aware:* The performance of networks hinges on CCI experienced at the UE's location which varies by the network's conditions. In light of this, we propose a channel assignment scheme, referred as Gibbs sampling-based UE interference-aware (GUIA) technique. This scheme is based on the interference that the UEs of each BS experience. In the proposed approach, each UE $k$ at the coverage area of BS $b$ computes the average CCI power received over each channel from other BSs in the network. Using the first order filtering with a constant forgetting factor $\lambda$, the average CCI power at the receiver of UE $k$ over channel $q$ at time $t$ in RB $r$ is given by,

$$\bar{I}_{k,b,r}^{q}(t) = (1-\lambda)I_{k,b,r}^{q}(t) + \lambda\ \bar{I}_{k,b,r}^{q}(t-1), \tag{23}$$

where $I_{k,b,r}^{q}(t)$ denotes instantaneous CCI power received by UE $k$ over channel $q$ at time $t$ in RB $r$. Each UE $k \in \mathcal{K}$ has a CCI table which updates for all $q \in \mathcal{Q}$. Then, UE $k$ selects the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

channel with minimum CCI power from this table, and send it (i.e. the channel itself) to its associated BS $b$ according to:

$$q_{k,b,r}(t) = \operatorname*{argmin}_{q \in \mathcal{Q}} \bar{I}^{q}_{k,b,r}(t), \quad \text{subject to} \quad k \in A^{t,r}_{b}. \tag{24}$$

Each BS $b \in \mathcal{B}$ selects its channel using the received reports. In this regard, BS $b$ calculates the repetition of each channel (i.e. channel absolute frequency) $q \in \mathcal{Q}$ in all received reports as follows:

$$f_{b,q,r}(t) = \sum_{\forall k \in A^{t,r}_{b}} \mathbb{1}_{(q_{k,b,r}(t)=q)}. \tag{25}$$

Then, each BS assigns a probability distribution to all available channels according to the reports received from the UEs. The channels which are mostly reported by the UEs have a higher selection probability. A common method for enabling this selection is to use a Boltzmann-Gibbs distribution [37], [38]. A Boltzmann-Gibbs is a probability distribution of particles over states, which can be expressed as: $\Lambda(x) \propto \exp\left(-\frac{E}{\theta}\right)$. Here, $E$ is the energy of state $x$, and $\theta$ is a constant which is proportional to Boltzmann's constant and thermodynamic temperature. This implies the probability which a system will be in state $x$, and it is proportional with the energy of states and system's temperature. In this regard, the Boltzmann-Gibbs distribution for BS $b$, $\mathbf{\Lambda}_b = [\Lambda_{b,1}, \ldots, \Lambda_{b,|\mathcal{Q}|}]$, can be calculated as follows:

$$\Lambda_{b,q}(f_{b,q,r}(t)) = \frac{\exp(\frac{1}{\theta_b} f_{b,q,r}(t))}{\sum_{\forall q' \in \mathcal{Q}} \exp(\frac{1}{\theta_b} f_{b,q',r}(t))}, \tag{26}$$

where $\Lambda_{b,q}(f_{b,q,r}(t))$ is the element of $\mathbf{\Lambda}_b$ which is related to channel $q$ [48]. Let $\frac{1}{\theta_b} > 0$ denotes the temperature parameter for BS $b$ that balances between exploration and exploitation. Note that, by allowing $\theta_b \to 0$, it leads to selecting the channel which is mostly reported by the UEs associated to BS $b$. Therefore, this can lead to the strategies with zero probabilities. On the contrary, by allowing $\theta \to \infty$, it results in a uniform distribution over the strategy set of player $b$. Then, each BS $b \in \mathcal{B}$ updates the probability assigned to each channel $q \in \mathcal{Q}$ at time $t$ according to:

$$\pi_{b,q,r}(t) = \pi_{b,q,r}(t-1) + \left(\frac{1}{t}\right)^{\nu} \left(\Lambda_{b,q}(f_{b,q,r}(t)) - \pi_{b,q,r}(t-1)\right), \tag{27}$$

where $\pi_{b,q,r}(t-1)$ and $\nu$ are the probability assigned to channel $q$ at time $t-1$ and the learning rate exponent, respectively. Then, each BS $b$ chooses its channel using a mapping function which maps the probability distribution $\{\pi_{b,1,r}(t), \ldots, \pi_{b,|\mathcal{Q}|,r}(t)\}$ to an element in the set of channels.

## V. No Regret-Based Learning Algorithm for Meeting Equilibrium

Next, the learning procedure for solving the proposed games is described. Hereinafter, for notational convenience, we drop the indexes $^{(P)}$ and $^{(Q)}$ used in the formulations of game $\mathcal{G}^{(P)}$ and $\mathcal{G}^{(Q)}$ when interchangeably referring to one of the games. Let $\boldsymbol{\pi}_b(t) = \{\pi_{b,1}(t), \ldots, \pi_{b,|\mathcal{S}_b|}(t)\} \in \mathfrak{D}(\mathcal{S}_b)$ be the mixed strategy of player $b$. Here, $\pi_{b,i}(t) = \Pr(s_b(t) = s_{b,i})$ is the probability that player $b$ plays strategy $s_{b,i}$ at time $t$, where $s_b(t)$ is the strategy of player $b$ at time $t$.

We use a no-regret learning approach [49] to solve the BS operation problem based on the game theoretic frameworks in order to allocate power and channel, and consequently obtain $\epsilon$-coarse correlated equilibrium. This algorithm is a probabilistic learning procedure. This type of learning algorithm has a good potential to learn mixed strategy equilibria. In our proposed algorithm, BSs learn their environment, and optimize their performance by modifying their transmission power levels and channels. Moreover, they impact the performance of other neighboring BSs, by minimizing their regrets for not having selected other strategies.

***Definition 1***: ($\epsilon$-coarse correlated equilibrium): A mixed strategy profile $\mathbf{\Gamma} = (\Gamma_{\boldsymbol{s}}, \forall \boldsymbol{s} \in \mathcal{S} = \mathcal{S}_1 \times \ldots \times \mathcal{S}_{|\mathcal{B}|})$ is an $\epsilon$-coarse correlated equilibrium if no player can gain more $\epsilon$ by deviating to a pure strategy. More formally, a mixed strategy profile $\mathbf{\Gamma}$ on the set of strategy profiles $\mathcal{S}$ is an $\epsilon$-coarse correlated equilibrium for game $\mathcal{G}$, if and only if, for all $b \in \mathcal{B}$, $s'_b \in \mathcal{S}_b$ we have,

$$\sum_{\boldsymbol{s} \in \mathcal{S}} \Gamma_{\boldsymbol{s}} u_b(s'_b, \boldsymbol{s}_{-b}) \leqslant \sum_{\boldsymbol{s} \in \mathcal{S}} \Gamma_{\boldsymbol{s}} u_b(\boldsymbol{s}) + \epsilon, \tag{28}$$

where $\boldsymbol{s}_{-b}$ is the strategy of players other than player $b$. In the following, we use a no-regret procedure in a distributed manner, in order to obtain $\epsilon$-coarse correlated equilibrium. We assume that each BS $b \in \mathcal{B}$ has only local information, and aims at maximizing its payoff function $u_b$, and minimizing its regret for not having played other strategies. The regret of player $b$ is defined as the difference between average payoff obtained up to time $t$ when playing strategy $s_{b,i}$, and average payoff obtained up to time $t$ when the player changes its strategies [50]. Specially, for any strategy $s_{b,i} \in \mathcal{S}_b$, the regret of $s_{b,i}$ at time $t$ is:

$$R_{s_{b,i}}(t) := [D_{s_{b,i}}(t)]^{+}, \tag{29}$$

where

$$D_{s_{b,i}}(t) = \frac{1}{t} \sum_{\tau \leqslant t} u_b(s_{b,i}, \boldsymbol{s}_{-b}(\tau)) - \tilde{u}_b(\tau), \tag{30}$$

where $\tilde{u}_b(\tau)$ denotes the time average of player $b$'s payoff. We consider the behavioral rule proposed in [49], in which BSs can balance the tradeoff between minimizing their regrets and estimating average payoffs. Moreover, they play each strategy with non-zero probability and based on their regrets. The solution which captures this behavioral rule is $\Lambda_{b,s_{b,i}}(R_{s_{b,i}}(t))$. Based on the sole knowledge of the value of the obtained payoff, each BS $b$ learns and estimates its expected payoff for all its strategies. Therefore, for each BS $b \in \mathcal{B}$ and each $s_{b,i} \in \mathcal{S}_b$, payoff estimation is updated as follows:

$$\hat{u}_{b,s_{b,i}}(t+1) = \hat{u}_{b,s_{b,i}}(t) + \left(\frac{1}{t+1}\right)^{\gamma} \mathbb{1}_{(s_b(t+1)=s_{b,i})} (u_b(t+1) - \hat{u}_{b,s_{b,i}}(t)), \tag{31}$$

where $\hat{u}_{b,s_{b,i}}(t)$ and $\gamma$ denote the estimated payoff at time $t$ and the learning rate exponent for updating the payoff estimation, respectively. The point here is that the strategy played at the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

last iteration sees the corresponding estimated payoff updated, while the estimated payoff for the other strategies were not changed. For calculating the regret, each player needs the knowledge of other players' strategies. Therefore, we use a learning tool for updating the estimated regret. In order to obtain distributed approach, each BS $b \in \mathcal{B}$ estimates its regret for each $s_{b,i} \in \mathcal{S}_b$ as follows:

$$\hat{R}_{s_{b,i}}(t+1) = \hat{R}_{s_{b,i}}(t) + \left(\frac{1}{t+1}\right)^{\zeta} \left(\hat{u}_{b,s_{b,i}}(t+1) - u_b(t+1) - \hat{R}_{s_{b,i}}(t)\right), \tag{32}$$

and then, the probability assigned to each strategy $s_{b,i} \in \mathcal{S}_b$ is updates as follows:

$$\pi_{b,s_{b,i}}(t+1) = \pi_{b,s_{b,i}}(t) + \left(\frac{1}{t+1}\right)^{\nu} \left(\Lambda_{b,s_{b,i}}(\hat{R}_{s_{b,i}}(t+1)) - \pi_{b,s_{b,i}}(t)\right), \tag{33}$$

where $\zeta$ and $\nu$ are the learning rate exponent for updating the estimation regret and probability distribution, respectively. We assume that for all BSs $\mathcal{B}$, there is a mapping function which maps the mixed strategy to a pure strategy.

We combine DCA-LA and GUIA with ON-OFF switching approach, and refer them to hereinafter as DCA-LA/ON-OFF switching and GUIA/ON-OFF switching, respectively. The pseudo code for the proposed SON mechanisms, i.e. DCA-LA/ON-OFF switching and GUIA/ON-OFF switching are summarized in Algorithm 1 and Algorithm 2, respectively. The procedures continue until reaching a maximum number of iterations, $T^{\mathrm{Max}}$, which is required for the mechanisms to converge. In Appendix C, a discussion about the computational requirements of the proposed mechanisms is presented.

**Algorithm 1** : Proposed DCA-LA/ON-OFF switching algorithm.

**Input**: $\mathcal{U}$, $\mathcal{N}$, $\hat{u}_{b,s_{b,i}}(t)$, $\hat{R}_{s_{b,i}}(t)$, $\pi_{b,s_{b,i}}(t)$ $\forall b \in \mathcal{B}$, and $s_{b,i} \in \mathcal{S}_b$<br>
**Output**: $\mathbf{A}^t$, $u_b^{(Q)}(t)$, $u_b^{(P)}(t)$, $\hat{u}_{b,s_{b,i}}(t+1)$, $\hat{R}_{s_{b,i}}(t+1)$, $\pi_{b,s_{b,i}}(t+1)$ $\forall b \in \mathcal{B}$, and $s_{b,i} \in \mathcal{S}_b$<br>
**Initialization**: $t = 1$, $\mathcal{B} = \{1, ..., |\mathcal{B}|\}$, $\mathcal{K} = \{1, ..., |\mathcal{K}|\}$, $\mathcal{S}_b^{(P)}$, $\mathcal{S}_b^{(Q)}$ $\forall b \in \mathcal{B}$<br>
**while** $t \leq T^{\mathrm{Max}}$ **do**<br>
  **for** $\forall r \in R_j$, $j \in \{\mathrm{M}, \mathrm{S}\}$ **do**<br>
    **for** $\forall b \in \mathcal{B}$, and $j \in \{\mathrm{M}, \mathrm{S}\}$ **do**<br>
      Advertise estimated load $\hat{L}_b(t)$<br>
      Select $P_{b,r}^{j}(t)$ using $\pi_{b,s_{b,i}}(t)$ for game $\mathcal{G}^{(P)}$<br>
      Select $q_{b,r}(t)$ using $\pi_{b,s_{b,i}}(t)$ for game $\mathcal{G}^{(Q)}$<br>
    **end for**<br>
    **for** $\forall k \in \mathcal{K}$ **do**<br>
      **if** $(k \in \mathcal{U}) \vee (k \in \mathcal{N})$ **then**<br>
        Find $a_{b,k}^{t,r}$ (12)<br>
      **end if**<br>
    **end for**<br>
    Updating: $\boldsymbol{A}^{t,r}$<br>
    **for** $\forall b \in \mathcal{B}$ **do**<br>
      Calculations: $L_b(t)$, $u_b^{(Q)}(t)$, $u_b^{(P)}(t)$<br>
    **end for**<br>
    **for** $\forall b \in \mathcal{B}$ **do**<br>
      Updating: $\hat{u}_{b,s_{b,i}}(t+1)$, $\hat{R}_{s_{b,i}}(t+1)$, $\pi_{b,s_{b,i}}(t+1)$ for both game $\mathcal{G}^{(P)}$, and $\mathcal{G}^{(Q)}$ (31)- (33)<br>
    **end for**<br>
  **end for**<br>
  $t \leftarrow t+1$,<br>
**end while**

### A. Convergence Analysis

In this subsection, we investigate the convergence of the proposed SON mechanisms. Since both mechanisms use the Boltzmann-Gibbs distribution in order to allocate the resources, we use $\boldsymbol{x}_b(t) = (x_b^{(P)}(t), x_b^{(Q)}(t))$ as the transmission strategy of BS $b$ composed of its transmission power and channel at time $t$. Let $x_b^{(P)}(t)$ and $x_b^{(Q)}(t)$ denote the transmission power level and channel selected by BS $b$ at time $t$, respectively. Therefore, the vector $\boldsymbol{x}(t) = [\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_{|\mathcal{B}|}(t)] \in \mathcal{X}$ denotes the transmission strategies selected by the BSs, where $\mathcal{X} = \prod_{b=1}^{|\mathcal{B}|} \mathcal{S}_b^{(P)} \times \mathcal{S}_b^{(Q)}$. For DCA-LA/ON-OFF switching mechanism, let $\Psi_b$ be a vector comprising average regret estimation for power and channel. For GUIA/ON-OFF switching mechanism, $\Psi_b$ comprises average regret estimation for power, and channel absolute frequency. Let $1/t^{\imath}$ for all $\imath = \{\gamma, \zeta, \nu\}$ be the learning rates. For converging the mechanisms, using stochastic approximation theory and the ordinary differential equation, all learning rate exponents $\imath = \{\gamma, \zeta, \nu\}$ are chosen according to the following conditions [47], [51]:

$$\lim_{t\to\infty} \sum_{n=1}^{t} \frac{1}{n^{\imath}} = +\infty, \tag{34}$$

and

$$\lim_{t\to\infty} \sum_{n=1}^{t} \left(\frac{1}{n^{\imath}}\right)^2 < +\infty. \tag{35}$$

The condition expressed in (34) ensures that the learning rates are large enough to overcome any random fluctuations or initial conditions. The condition expressed in (35) guarantees that the learning rates diminish with iterations, and eventually become small enough to assure convergence. For satisfying the above conditions, we choose all $\imath = \{\gamma, \zeta, \nu\} \in (0.5, 1)$. Furthermore, we consider the utility estimation procedure as a fast process relative to the regret estimation, and the regret estimation procedure as a fast process relative to the strategy distribution process. Therefore, the learning rate exponents meet the following criteria:

$$\lim_{t\to\infty} \frac{(1/t^{\zeta})}{(1/t^{\gamma})} = 0, \tag{36}$$

and

$$\lim_{t\to\infty} \frac{(1/t^{\nu})}{(1/t^{\zeta})} = 0. \tag{37}$$

Therefore, the learning rate exponents follow $\zeta > \gamma$, $\nu > \zeta$. Moreover, we assume that $\theta > 0$, unless when we mention $\theta \to 0$. In the following theorem, we show the proposed SON mechanisms converge to stationary distributions using similar argument as [38].

**Theorem 1.** *Starting from an arbitrary initial configuration, the SON mechanisms correspond to Markov chains,*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

**Algorithm 2** : Proposed GUIA/ON-OFF switching algorithm.

**Input**: $\mathcal{U}, \mathcal{N}, \hat{u}_{b,s_{b,i}}(t), \hat{R}_{s_{b,i}}(t), \pi_{b,s_{b,i}}(t)\ \forall b \in \mathcal{B}$, and $s_{b,i} \in \mathcal{S}_b$
**Output**: $\mathbf{A}^t$, $u_b^{(Q)}(t)$, $u_b^{(P)}(t)$, $\hat{u}_{b,s_{b,i}}(t+1)$, $\hat{R}_{s_{b,i}}(t+1)$, $\pi_{b,s_{b,i}}(t+1)\ \forall b \in \mathcal{B}$, and $s_{b,i} \in \mathcal{S}_b$
**Initialization**: $t = 1$, $\mathcal{B} = \{1, ..., |\mathcal{B}|\}$, $\mathcal{K} = \{1, ..., |\mathcal{K}|\}$, $\mathcal{Q} = \{1, \ldots, |\mathcal{Q}|\}$, $\mathcal{S}_b^{(P)}$, $\forall b \in \mathcal{B}$
**while** $t \leq T^{\text{Max}}$ **do**
  **for** $\forall r \in R_j, j \in \{\text{M}, \text{S}\}$ **do**
    **for** $\forall b \in \mathcal{B}$, and $j \in \{\text{M}, \text{S}\}$ **do**
      Advertise estimated load $\hat{L}_b(t)$
      Select $P_{b,r}^j(t)$ using $\pi_{b,s_{b,i}}(t)$ for game $\mathcal{G}^{(P)}$
      Select $q_{b,r}(t)$ according to $f_{b,q,r}(t)$ (28)
    **end for**
    **for** $\forall k \in \mathcal{K}$ **do**
      **if** $(k \in \mathcal{U}) \vee (k \in \mathcal{N})$ **then**
        Find $a_{b,k}^{t,r}$ (12)
      **end if**
    **end for**
    Updating: $\boldsymbol{A}^{t,r}$
    **for** $\forall b \in \mathcal{B}$ **do**
      **for** $\forall k \in A_b^{t,r}$ **do**
        **for** $\forall q \in \mathcal{Q}$ **do**
          Calculations: $I_{k,b,r}^q(t)$
          Calculations: $\bar{I}_{k,b,r}^q(t)$
        **end for**
        Select $q_{k,b,r}$ (24)
      **end for**
    **end for**
    **for** $\forall b \in \mathcal{B}$ **do**
      Calculations: $L_b(t)$, $u_b^{(P)}(t)$, $f_{b,q,r}(t)$
    **end for**
    **for** $\forall b \in \mathcal{B}$ **do**
      Updating: $\pi_{b,q,r}(t)$ (27),$\hat{u}_{b,s_{b,i}}(t+1)$, $\hat{R}_{s_{b,i}}(t+1)$, $\pi_{b,s_{b,i}}(t+1)$ for game $\mathcal{G}^{(P)}$ (31)- (33)
    **end for**
  **end for**
  $t \leftarrow t+1$,
**end while**

*in which the system converges to a stationary distribution* $\mathbf{\Gamma} = (\Gamma_{\boldsymbol{x}}, \forall \boldsymbol{x} \in \mathcal{X})$*, i.e., [38]*

$$\Gamma_{\boldsymbol{x}} = \frac{\exp(\frac{1}{\theta}\bar{\Psi}_{\boldsymbol{x}})}{\sum_{\forall \boldsymbol{x}' \in \mathcal{X}} \exp(\frac{1}{\theta}\bar{\Psi}_{\boldsymbol{x}'})}, \tag{38}$$

*and*

$$\lim_{\theta \to 0} \sup(\Gamma_{\boldsymbol{x}} - \Gamma_0) = 0, \tag{39}$$

*where* $\Gamma_0$ *is the uniform distribution over the set of global optimal solutions.*

The proof of Theorem 1 is presented in Appendix A.

The following theorem characterizes the convergence of the proposed games under the no-regret learning algorithm.

**Theorem 2.** *In the behavioral rule defined in* (31)-(33)*, the stationary distribution* $\mathbf{\Gamma}$ *comprises an* $\epsilon$*-coarse correlated equilibrium for* $\mathcal{G}^{(P)}$ *and* $\mathcal{G}^{(Q)}$.

Table I
SYSTEM-LEVEL SIMULATION PARAMETERS

| System Parameters | | |
|---|---|---|
| **Parameter** | | **Value** |
| Physical link type | | Downlink |
| Carrier frequency/ Channel bandwidth | | 2 GHz/ 10 MHz |
| Noise power spectral density | | -174 dBm/Hz |
| Mean packet arrival rate | | 1800 Kbps |
| $\theta_b$ | | 0.1 |
| Weights $\omega_b^{\text{n}}$ , $\omega_b^{\text{l}}$, $\omega_b^{\text{P}}$ | | 1, 0.5, 0.5 |
| Learning rate exponent $\gamma$, $\zeta$, $\nu$, $\alpha$ | | 0.6, 0.7, 0.8, 0.9 |
| $P^{\text{Threshold}}$ | | -60 dBm |
| **BSs Parameters** | | |
| **Parameter** | **MBS** | **SBS** |
| Maximum power | 46 dBm | 30 dBm |
| Feeder loss | 3 dB | 0 dB |
| DC-DC loss | 7.5% | 9% |
| Mains supply loss | 9% | 11% |
| Cooling loss | 10% | 0% |
| $\eta_b^{\text{PA}}$ | 31.1% | 6.7% |
| $P_b^{\text{RF}}$ | 12.9 W | 0.8 W |
| $P_b^{\text{BB}}$ | 29.6 W | 3 W |
| $\lvert\mathcal{S}_b^{(P)}\rvert$ | 2 | 4 |
| Shadowing standard deviation | 8 dB | 10 dB |
| Radius cell | 250 m | 40 m |
| Distance-dependent path loss model ($d$ in Km) | 128.1+37.6 $\log_{10}(d)$ | 140.7+36.7 $\log_{10}(d)$ |
| Minimum distance | MBS-SBS: 75 m<br>MBS-UE: 35 m | SBS-SBS: 40 m<br>SBS-UE: 10 m |
| Load threshold ($L_b^{\text{Threshold}}$) | 0.9 | 0.7 |

The proof of Theorem 2 is presented in Appendix B. Moreover, a proof of the convergence of the set of coupled learning algorithms relies on the stochastic approximation algorithms by using ordinary differential equations [49].

## VI. SIMULATION RESULTS

For our simulations, the proposed mechanisms are validated in a single cell served by one MBS and a set of SBSs with the maximum transmit power of 46 dBm and 30 dBm as presented in [9], respectively. We assume the number of channels to be 4, i.e. $|\mathcal{Q}| = 4$. Our proposed self-organizing mechanisms, i.e. proposed DCA-LA/ON-OFF switching and proposed GUIA/ON-OFF switching, are compared with two following benchmark references:

- *Interference-Aware Dynamic Channel Selection (IADCS)*: Each BS transmits with its maximum power, and evaluates averages CCI power over each channel, and finally selects the channel with minimum average CCI power [52].
- *IADCS/ON-OFF Switching*: Each BS transmits according to the ON-OFF switching mechanism, and evaluates average CCI power over each channel, and finally selects the channel with minimum average CCI power [53].

Here, the solid curves belong to the benchmark solutions, and the dashed curves refer to the proposed mechanisms. We present simulation results for two scenarios, including stationary environment, i.e. without considering mobility for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

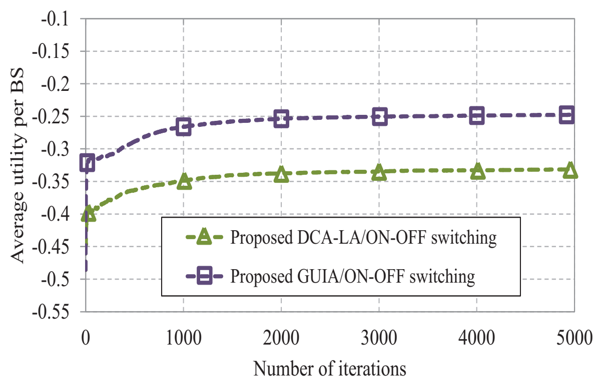


Figure 2. Convergence of average utility per BS of the proposed algorithms for a network with 20 SBSs and 40 UEs.

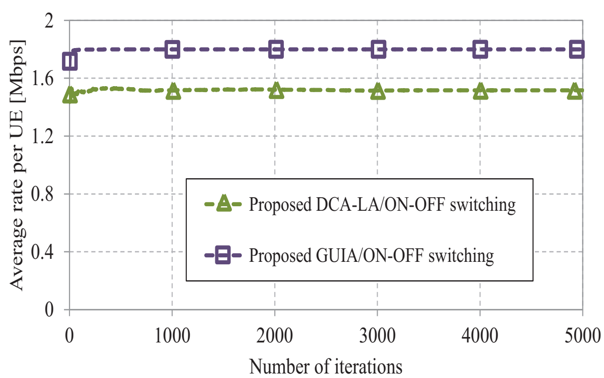


Figure 3. Convergence of average rate per UE of the proposed algorithms for a network with 20 SBSs and 40 UEs.

UEs, and a non-stationary environment, i.e. with considering UEs' mobility. The communications are carried out in full buffer [54], [55] in accordance to the system parameters shown in Table I.

### A. Convergence of the Proposed Algorithms

Here, we investigate the convergence behavior of our proposed SON mechanisms. First, we consider a network with 40 UEs and 20 SBSs. Fig. 2 and Fig. 3 show the convergence behavior of BS's utility, defined in (17), and UE's rate, respectively. We can observe that DCA-LA/ON-OFF switching algorithm needs more iterations to converge compared to GUIA/ON-OFF switching, i.e. the tradeoff between convergence speed and computational complexity. Furthermore, the proposed GUIA/ON-OFF switching yields better performance compared to DCA-LA/ON-OFF switching algorithm.

Fig. 4 and Fig. 5 show convergence behavior of our proposed mechanisms, respectively, in terms of BS's utility and UE's rate, for a network with 5 SBSs, and the number of UEs $=20$ and 50. From Fig. 4 and Fig. 5, it can be observed that two proposed algorithms converge within a small number of iterations. Furthermore, by increasing the number of UEs, the iterations for convergence slightly increases.

### B. Stationary Environment

Fig. 6 shows average energy consumption per BS as the number of SBSs varies for 60 active UEs. As the number of SBSs in the network increases, the path loss between the BS and the UE degrades. Thus, it induces a decreasing in the required transmit power to provide the certain

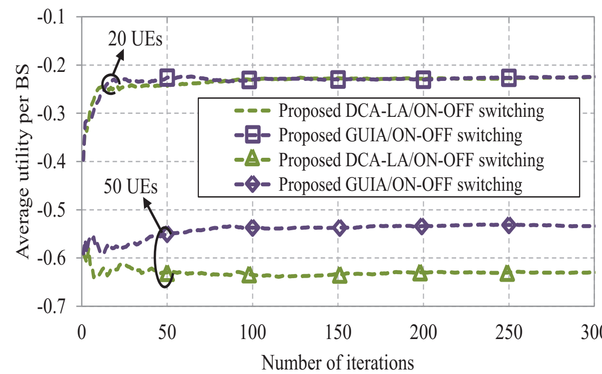


Figure 4. Convergence of average utility per BS of the proposed algorithms for a network with 5 SBSs, and the number of UEs $=20$ and 50.

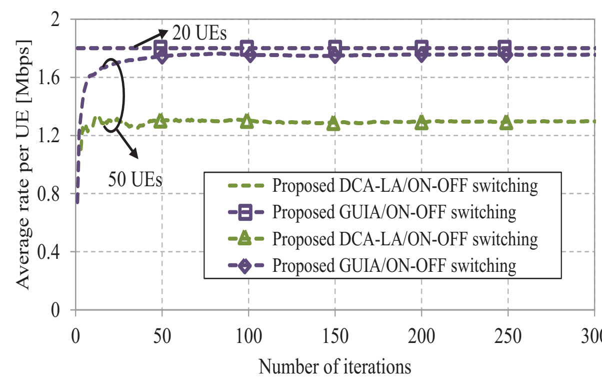


Figure 5. Convergence of average rate per UE of the proposed algorithms for a network with 5 SBSs, and the number of UEs $=20$ and 50.

received power at the UE's receiver. From Fig. 6, we can observe that with increasing the number of SBSs, the average energy consumption per BS decreases. At the number of SBSs $=50$, the reduction of average energy consumption per BS in IADCS/ON-OFF switching, DCA-LA/ON-OFF switching, and proposed GUIA/ON-OFF switching compared to IADCS approach are $33.8\%$, $35.6\%$, and $40.3\%$, respectively. Moreover, for a given number of SBSs, proposed GUIA/ON-OFF switching mechanism consumes less energy compared to the other approaches. However, it has almost the same performance compared to the approaches based on ON-OFF switching. For instance, it reduces the average energy consumption per BS by $7.2\%$, and $9.8\%$ compared to DCA-LA/ON-OFF switching, and IADCS/ON-OFF switching, respectively. The main reason is that the DCA-LA/ON-OFF switching and IADCS/ON-OFF switching utilize the ON-OFF switching mechanism for unnecessary BSs, whereas in IADCS algorithm, each BS transmits with its maximum power. On the other hand, in GUIA/ON-OFF switching, the BS selects the better downlink channel, in terms of CCI, compared to the other mechanisms based on the ON-OFF switching. This comes because of receiving the reports from its associated UEs. Therefore, the interference over the selected channel using GUIA/ON-OFF switching mechanism is less than the other mechanisms. This leads to a decrease in the required transmit power to guarantee the same received power at the UE's receiver. Since the proposed mechanisms are simulated for the number of SBSs $=50$, it can imply a dense SBSs deployment. Furthermore, in the proposed mechanisms, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

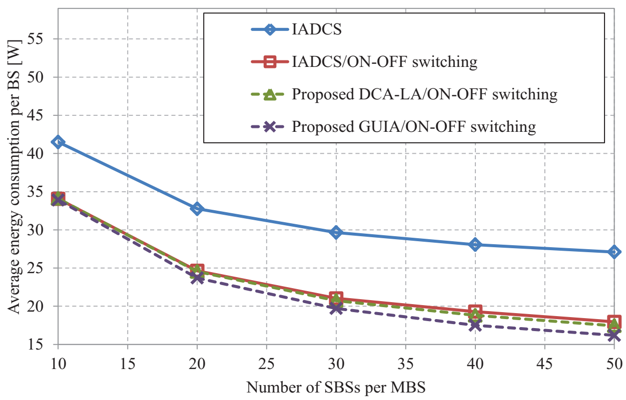


Figure 6. Average energy consumption per BS versus the number of SBSs, given 60 UEs.

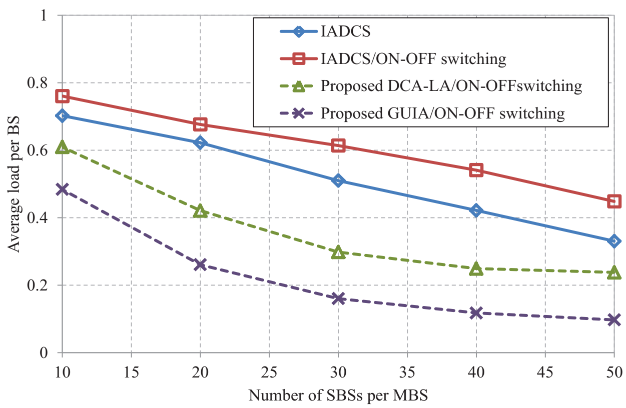


Figure 7. Average load per BS versus the number of SBSs, given 60 UEs.

average CCI power over each channel is evaluated without the distinguishing the source of interference, which may consist of interference from MBSs and SBSs. As a result, without loss of generality, increasing the number of SBSs can increase interference in the network, and thus, it can imply a dense deployment. Fig. 7 shows the average load per BS for 60 active UEs. As the number of SBSs increases, average load per BS decreases through offloading UEs associated with highly loaded BSs to lightly loaded BSs. Since GUIA/ON-OFF switching mechanism assigns the resources in the efficient manner, it can reduce interference on downlink channel at the UE's receiver. As a result, it improves the achievable transmission rate of the UE, and thus, it reduces the BS's load. From Fig. 7, it can be seen that the proposed GUIA/ON-OFF switching method balances the load among BSs, and yields $78.3\%$, $70.6\%$, and $59.2\%$ average load reduction per BS compared to IADCS/ON-OFF switching, IADCS, and DCA-LA/ON-OFF switching approaches, respectively.

Fig. 8 shows the average utility per BS, defined in (17), versus the number of UEs for 5 SBSs. As the number of UEs increases, more BSs are working in active mode, and less BSs choose OFF mode. Therefore, the BS's load and energy consumption increase, and thus, the utility of the BS decreases. Since GUIA/ON-OFF switching balances load, and saves more energy compared to the other mechanisms, it improves the BS's utility. However, for high number of UEs, the behavior of mechanisms based on the ON-OFF switching method becomes closer to each other. This is mainly due to the fact that, the load

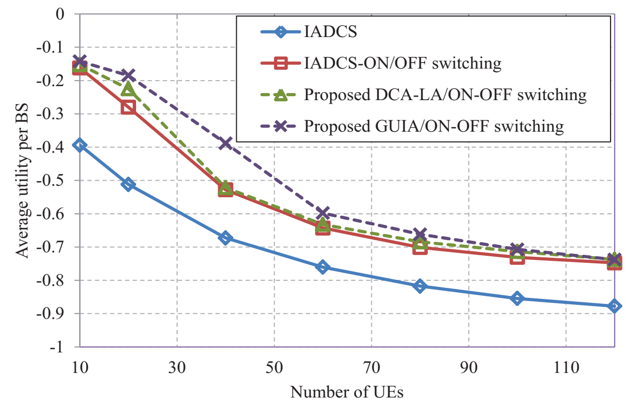


Figure 8. Average utility per BS versus the number of UEs, given 5 SBSs.

Table II
PROBABILITY DISTRIBUTION FOR UE'S MOVEMENT

| Movement direction | probability |
|---|---|
| Current direction | $p_k^{\mathrm{c}} = 0.5$ |
| Turning right | $p_k^{\mathrm{r}} = 0.25$ |
| Turning left | $p_k^{\mathrm{l}} = 0.25$ |

of BSs increases, and approaches to the maximum BS's load. On the other hand, the average energy consumed by BSs in the mechanisms based on the ON-OFF switching method is almost same. Fig. 8 shows that the average utility of the proposed GUIA/ON-OFF switching yields, respectively, $17.8\%$, $34\%$, and $63.9\%$ improvement over DCA-LA/ON-OFF switching, IADCS/ON-OFF switching, and IADCS approaches for 20 UEs.

### C. Non-Stationary Environment

We now turn to the scenario, in which the UEs can freely move, and perform handoff according to their received signal strength (RSS), distance, and BS's estimated load. Since in the random waypoint model, UEs' movements occur in an open field, and they can be located at every point of the region, it may lead to inaccurate results. The models based on the streets of urban areas are more realistic mobility models, and provide more accurate simulation results. In the streets of urban areas, the UEs' movements are limited to streets often separated by buildings, trees, and different obstructions.

Hereby, we use a Manhattan grid model, in which the UEs move in a straight line, and they can change their directions at each intersection with a given probability. One way is by considering a reduced size for the grid squares, which allows us to make the model closer to the irregular situation. At each intersection, each UE selects its movement direction according to the probability distributions in Table II. Here $p_k^{\mathrm{c}}$, $p_k^{\mathrm{r}}$, and $p_k^{\mathrm{l}}$ denote the probabilities for moving on the current direction, turning right, and turning left for UE $k$, respectively. We assume that the UEs move with constant speed, and when a UE goes out of a boundary, another UE enters on the other side. We consider a two-fold handoff algorithm: handoff between MBS and SBS, and handoff between SBS and SBS. Let $d_{k,b}(t)$ and $R_b$ be estimated distance between UE $k$ and serving BS $b$, and the cell radius of BS $b$, respectively. The handoff algorithm employed by the UEs is given in Algorithm 3. Here, $P_{b,k,r}\,(t)$ denotes the received power at UE $k$ associated with serving

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

**Algorithm 3** : Handoff algorithm based on BS's estimated load.

**if** $(d_{k,b}(t+1) > d_{k,b}(t)) \wedge (d_{k,b}(t+1) > 0.8 \times R_b)$
  **if** $10 \log P_{b,k,r}(t)(1-\hat{L}_b) < P^{\text{Threshold}} + 10\log(1 - L_b^{\text{Threshold}})$ **then**
    Search for another BS better than serving BS according to (12)
  **else**
    Continue with serving BS
  **end if**
**end if**

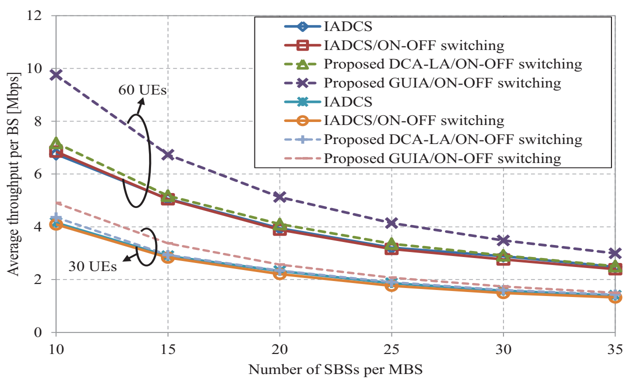


Figure 9. Average throughput per BS versus the number of SBSs, given 30 and 60 UEs.

BS $b$ in RB $r$ at time $t$. For this subsection, we consider the UEs move with velocity 5 m/sec on the layout.

In Fig. 9, we depict average throughput per BS when increasing the number of SBSs for 30 and 60 UEs. We observe that the proposed GUIA/ON-OFF switching mechanism achieves the highest throughput compared to the other approaches. For instance, for a network with 60 UEs, it yields up to $35.9\%$, $42.4\%$, and $44.4\%$ improvement in terms of throughput, relative to the proposed DCA-LA/ON-OFF switching, IADCS/ON-OFF switching, and IADCS, respectively. We can observe that as the number of SBSs increases, average throughput per BS decreases. The main reason is that with increasing the number of SBSs, the BS's load decreases. On the other hand, for a given number of UEs in the network, the number of associated UEs to the BS decreases. Moreover, in our scenario, the decreasing rate of associated UEs to the decreasing rate of load is less than one. Furthermore, at a given number of SBSs, with an increase in the number of UEs, the average throughput per BS increases.

In Fig. 10, we show the average rate per UE versus the number of SBSs for 30 and 60 UEs. As the number of SBSs increases, the BS's load decreases, and thus, the UE's rate increases. Moreover, with increasing the number of UEs in the area, UE's rate decreases. This is due to the fact that the BS's load increases, and it may lead to overload some BSs, and thus, decreasing UE's rate. Since GUIA/ON-OFF switching mechanism exhibits a load balancing behavior, it improves the UE's rate. In this respect, in Fig. 10, we can see that the proposed GUIA/ON-OFF switching improves the average rate, respectively, $35.7\%$, $43.2\%$, and $44.8\%$ when compared to proposed DCA-LA/ON-OFF switching, IADCS/ON-OFF switching, and IADCS for the network with 60 UEs and 10 SBSs.

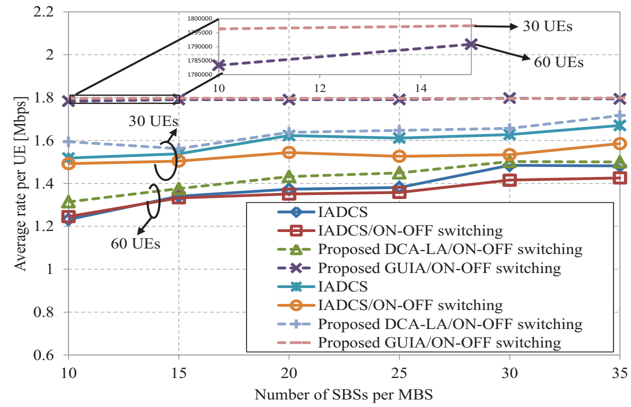


Figure 10. Average rate per UE versus the number of SBSs, given 30 and 60 UEs.

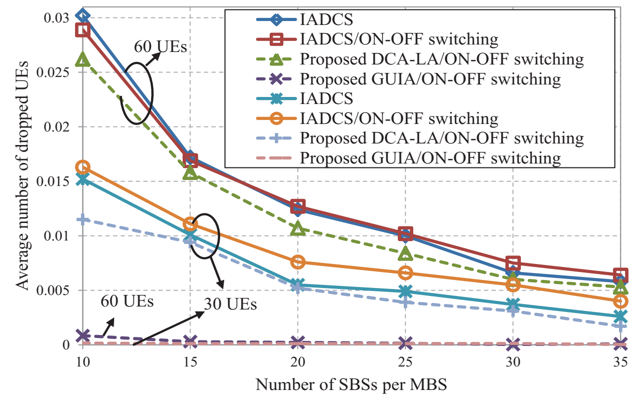


Figure 11. Average number of dropped UEs versus the number of SBSs, given 30 and 60 UEs.

Fig. 11 shows the average number of dropped UEs, which indicates the UE's QoS, versus the number of SBSs for 30 and 60 UEs. As the number of SBSs increases, the number of dropped UEs decreases because of load balancing behavior. Furthermore, it can be observed that increasing the number of UEs results in increasing the average number of dropped UEs. Since GUIA/ON-OFF switching mechanism has a good performance in terms of average BS's load, it yields insignificant dropped UEs. For a network with 30 UEs and 10 SBSs, the reductions of average number of dropped UEs in the proposed GUIA/ON-OFF switching compared to the proposed DCA-LA/ON-OFF switching, IADCS/ON-OFF switching, and IADCS are $96.8\%$, $97.1\%$, and $97.2\%$, respectively.

## VII. Conclusion

In this paper, we have proposed two dynamic channel assignment mechanisms, i.e. DCA-LA and GUIA. Later, we have combined them with a BS ON-OFF switching algorithm in order to reduce the energy consumption of the network. The proposed DCA-LA/ON-OFF switching algorithm uses a game-theoretic approach, in which each BS selects its channel and power based on a no-regret learning algorithm. In GUIA/ON-OFF switching algorithm, BSs utilize some information from their associated UEs to improve the performance of the network. Then, they select their channels based on a Gibbs-sampler. GUIA/ON-OFF switching algorithm balances the load among BSs. Therefore, it improves

Table III
COMPUTATIONAL REQUIREMENTS FOR THE PROPOSED CHANNEL ASSIGNMENT ALGORITHMS

| **Operations** | **Required Instructions for** | | | |
|---|---|---|---|---|
| | **Proposed DCA-LA** | | **Proposed GUIA** | |
| | at BSs | at UEs | at BSs | at UEs |
| Sum | $(6\times\lvert\mathcal{Q}\rvert+1)\times\lvert\mathcal{B}\rvert$ | 0 | $(3\times\lvert\mathcal{Q}\rvert-1)\times\lvert\mathcal{B}\rvert+\lvert\mathcal{K}\rvert$ | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{K}\rvert$ |
| Multiplication | $(3\times\lvert\mathcal{Q}\rvert+1)\times\lvert\mathcal{B}\rvert$ | 0 | $2\times\lvert\mathcal{Q}\rvert\times\lvert\mathcal{B}\rvert$ | $2\times\lvert\mathcal{Q}\rvert\times\lvert\mathcal{K}\rvert$ |
| Division | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{B}\rvert$ | 0 | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{B}\rvert$ | 0 |
| Exponential | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{B}\rvert$ | 0 | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{B}\rvert$ | 0 |
| Comparison | $(\lvert\mathcal{Q}\rvert-1)\times\lvert\mathcal{B}\rvert$ | 0 | $\lvert\mathcal{Q}\rvert\times\lvert\mathcal{K}\rvert$ | $(\lvert\mathcal{Q}\rvert-1)\times\lvert\mathcal{K}\rvert$ |
| Total number of instructions | $(63\times\lvert\mathcal{Q}\rvert+1)\times\lvert\mathcal{B}\rvert$ | | $5\times\lvert\mathcal{Q}\rvert\times\lvert\mathcal{K}\rvert+(58\times\lvert\mathcal{Q}\rvert-1)\times\lvert\mathcal{B}\rvert$ | |

system throughput, and consequently yields a better SE. As a result, our proposed algorithm achieves both high EE and SE. The proposed algorithms are all executed in a distributed manner. Simulation results have shown that GUIA/ON-OFF switching algorithm provides a better performance over the other algorithms, and significantly outperforms them in terms of average energy consumption, average load, average BS's utility, average throughput, average number of dropped UEs, and average UE's rate.

## APPENDIX A
## PROOF OF THEOREM 1

Transition from $\boldsymbol{x}(t)$ to $\boldsymbol{x}(t+1)$ takes place when some BSs update their resource values according to their observations on $\boldsymbol{x}(t)$. Therefore, the transition depends only on current configuration vector. As a result, each proposed SON mechanism can be modeled as a Markov chain with the set of cliques $\mathcal{B}$, global energy $\bar{\Psi}:\mathcal{X}\rightarrow\mathbb{R}^2$ derived from the potential function $\Psi_b$ where $\bar{\Psi}_{\boldsymbol{x}}=\sum_{\forall b\in\mathcal{B}}\Psi_b(\boldsymbol{x}_b)$ and finite configuration space $\mathcal{X}$. Since for each $\boldsymbol{x}\in\mathcal{X}$, $\Gamma_{\boldsymbol{x}}$ is non-zero and positive, the Markov chain is ergodic (aperiodic and positive recurrent). Therefore, a stationary distribution exists. Now, we prove the second part of the theorem. Let $\mathcal{X}^*$ be the set of global optimal solutions. For each $\boldsymbol{x}^*\in\mathcal{X}^*$,

$$(\Gamma_{\boldsymbol{x}}-\Gamma_0)=\frac{\exp(\frac{1}{\theta}\bar{\Psi}_{\boldsymbol{x}})}{\sum_{\forall\boldsymbol{x}'\in\mathcal{X}}\exp(\frac{1}{\theta}\bar{\Psi}_{\boldsymbol{x}'})}-\Gamma_0=\frac{\exp(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}}-\bar{\Psi}_{\boldsymbol{x}^*}))}{\sum_{\forall\boldsymbol{x}'\in\mathcal{X}}\exp(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}'}-\bar{\Psi}_{\boldsymbol{x}^*}))}-\Gamma_0\leq\frac{1}{|\mathcal{X}^*|+\sum_{\forall\boldsymbol{x}'\notin\mathcal{X}^*}\exp(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}'}-\bar{\Psi}_{\boldsymbol{x}^*}))}-\Gamma_0,\quad(40)$$

When $\theta\rightarrow 0$, we have,

$$\lim_{\theta\rightarrow 0}\Big(\frac{1}{|\mathcal{X}^*|+\sum_{\forall\boldsymbol{x}'\notin\mathcal{X}^*}\exp(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}'}-\bar{\Psi}_{\boldsymbol{x}^*}))}-\Gamma_0\Big)=\frac{1}{|\mathcal{X}^*|}-\Gamma_0=0.\quad(41)$$

This concludes the proof.

## APPENDIX B
## PROOF OF THEOREM 2

First, we use the following lemma to show the existence of the equilibrium.

**Lemma 1**: In every finite game, the set of correlated equilibria is nonempty, closed and convex. Since the proposed games are finite, they have at least one mixed strategy equilibrium.

For $\boldsymbol{x}'_b\in\mathcal{S}_b^{(P)}\times\mathcal{S}_b^{(Q)}$, we define,

$$\vartheta=\max_{x\in\mathcal{X}\backslash\mathcal{X}^*}(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x})),\quad(42)$$

and

$$\delta=\min_{\boldsymbol{x}\in\mathcal{X}\backslash\mathcal{X}^*}((\bar{\Psi}_{\boldsymbol{x}^*}-\bar{\Psi}_{\boldsymbol{x}})).\quad(43)$$

Let $0\leq\theta\leq\frac{\delta}{\ln\left(\frac{\vartheta}{\epsilon}(\frac{|\mathcal{X}|}{|\mathcal{X}^*|}-1)\right)}$. Now, we prove (28), thus, we have,

$$\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))=\sum_{\boldsymbol{x}\in\mathcal{X}}\frac{\exp(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}}-\bar{\Psi}_{\boldsymbol{x}^*}))(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))}{|\mathcal{X}^*|+\sum_{\boldsymbol{x}''\in\mathcal{X}\backslash\mathcal{X}^*}\exp\left(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}''}-\bar{\Psi}_{\boldsymbol{x}^*})\right)}\leq\frac{|\mathcal{X}^*|(u_b(\boldsymbol{x}'_b,\boldsymbol{x}^*_{-b})-u_b(\boldsymbol{x}^*))}{|\mathcal{X}^*|}+\frac{\sum_{\boldsymbol{x}\in\mathcal{X}\backslash\mathcal{X}^*}\exp\left(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}}-\bar{\Psi}_{\boldsymbol{x}^*})\right)(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))}{|\mathcal{X}^*|},\quad(44)$$

Since $(u_b(\boldsymbol{x}',\boldsymbol{x}^*_{-b})-u_b(\boldsymbol{x}^*))\leq 0$, we have,

$$\frac{|\mathcal{X}^*|(u_b(\boldsymbol{x}'_b,\boldsymbol{x}^*_{-b})-u_b(\boldsymbol{x}^*))}{|\mathcal{X}^*|}+\frac{\sum_{\boldsymbol{x}\in\mathcal{X}\backslash\mathcal{X}^*}\exp\left(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}}-\bar{\Psi}_{\boldsymbol{x}^*})\right)(u_b(\boldsymbol{x}',\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))}{|\mathcal{X}^*|}\leq\frac{\vartheta\sum_{\boldsymbol{x}\in\mathcal{X}\backslash\mathcal{X}^*}\exp\left(\frac{1}{\theta}(\bar{\Psi}_{\boldsymbol{x}}-\bar{\Psi}_{\boldsymbol{x}^*})\right)}{|\mathcal{X}^*|}\leq\vartheta(\frac{|\mathcal{X}|}{|\mathcal{X}^*|}-1)\exp(\frac{-1}{\theta}\delta)\leq\vartheta(\frac{|\mathcal{X}|}{|\mathcal{X}^*|}-1)\frac{\epsilon}{\vartheta(\frac{|\mathcal{X}|}{|\mathcal{X}^*|}-1)}\leq\epsilon,\quad(45)$$

Therefore, this concludes the proof of (28). Moreover, using similar argument as [56], we can assume that (28) does not hold for $\epsilon=0$ when $\theta\rightarrow 0$. Therefore,

$$\limsup_{\theta\rightarrow 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))=\frac{1}{|\mathcal{X}^*|}\sum_{\boldsymbol{x}\in\mathcal{X}}(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})-u_b(\boldsymbol{x}))>0.\quad(46)$$

Thus, for some $\boldsymbol{x}\in\mathcal{X}^*$, $(u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b})>u_b(\boldsymbol{x}))$, and it implies the regret of player $b$ for $\boldsymbol{x}'_b$. Therefore, player $b$ interests to deviate from it, and this yields $\boldsymbol{x}\notin\mathcal{X}^*$. Thus, the assumption (46) is not true, and $\Gamma_{\boldsymbol{x}}$ converges to an $\epsilon$-coarse

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

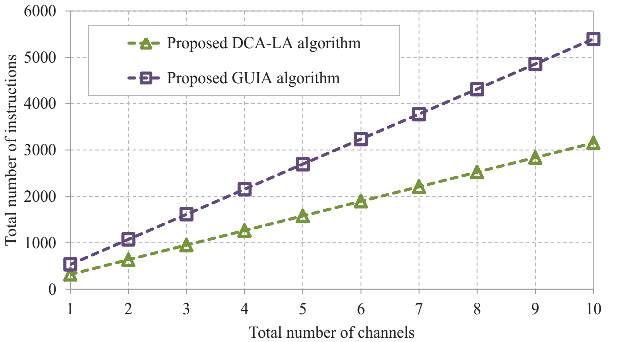


Figure 12. Total required instructions of proposed channel assignment algorithms versus the number of channels for a network with $|\mathcal{B}|= 5$ and $|\mathcal{K}|= 50$.

correlated equilibrium. Since $\Gamma_{\boldsymbol{x}}$ and the expected value of $\bar{\Psi}_{\boldsymbol{x}}$ monotonically decrease with $\theta$ [38], we assume that:

$$(\lim_{\theta\to 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x})) - \sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}) \leq \epsilon. \tag{47}$$

Thus, for $\theta > 0$,

$$\lim_{\theta\to 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b}) - \lim_{\theta\to 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}) \geq \sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b}) - \sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}) - \epsilon. \tag{48}$$

Since

$$\lim_{\theta\to 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}'_b,\boldsymbol{x}_{-b}) - \lim_{\theta\to 0}\sum_{\boldsymbol{x}\in\mathcal{X}}\Gamma_{\boldsymbol{x}}u_b(\boldsymbol{x}) \leq 0. \tag{49}$$

we conclude that $\boldsymbol{\Gamma}$ is an $\epsilon$-coarse correlated equilibrium.

## Appendix C
## Computational Requirements

In the proposed DCA-LA/ON-OFF switching and GUIA/ON-OFF switching mechanisms, the BSs select their transmit power by using the same ON-OFF switching approach. Therefore, in order to compare the computational requirements of the proposed mechanisms, we only consider the computational requirements of the proposed channel assignment approaches. Here, with considering digital signal processors (DSPs), we present the computational requirements of the proposed DCA-LA and GUIA approaches. We assume that addition, multiplication, and comparison operations require one DSP cycle [57]. For division operation and exponential computation, 42 and 11 operations are considered, respectively. Since the computational analysis does not take into account the compiler optimizations and the ability of DSPs to execute various instructions, the analysis provides an upper bound on computations. Table III summarizes the computational instructions for the proposed DCA-LA and GUIA approaches. From UE's perspective, the computational instructions of channel assignment algorithms are given by the number of available channels. Accordingly, the proposed GUIA algorithm requires more computational instructions compared to DCA-LA algorithm.

Fig. 12 shows the total required instructions over different number of channels for the proposed channel assignment algorithms, assuming $|\mathcal{B}|= 5$ and $|\mathcal{K}|= 50$. We can observe that the proposed GUIA algorithm requires more instructions

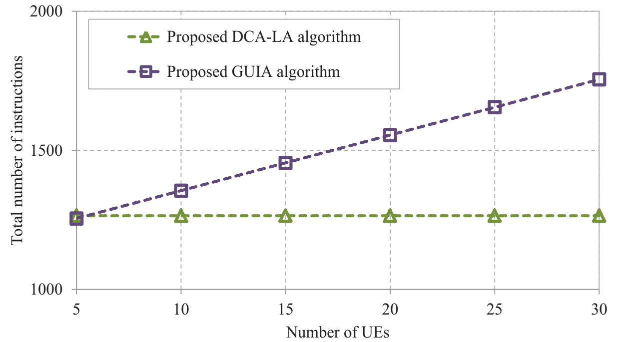


Figure 13. Total required instructions of proposed channel assignment algorithms versus the number of UEs for a network with $|\mathcal{B}|= 5$ and $|\mathcal{Q}|= 4$.

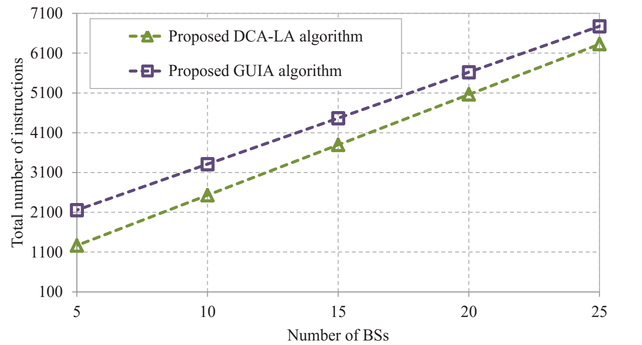


Figure 14. Total required instructions of proposed channel assignment algorithms versus the number of BSs for a network with $|\mathcal{K}|= 50$ and $|\mathcal{Q}|= 4$.

compared to DCA-LA algorithm. Fig. 13 presents the total required instructions over different number of UEs for a network with $|\mathcal{B}|= 5$ and $|\mathcal{Q}|= 4$. It can be seen that with increasing the number of UEs, the total required instructions of the proposed GUIA algorithm increases, while the required instructions of DCA-LA algorithm does not change. Fig. 14 shows the total required instructions over different number of BSs for a network with $|\mathcal{K}|= 50$ and $|\mathcal{Q}|= 4$. We see that with increasing the number of BSs, the required instructions increases for both proposed algorithms.

## References


[1] S. Navaratnarajah, A. Saeed, M. Dianati, and M. A. Imran, "Energy efficiency in heterogeneous wireless access networks," *IEEE Wireless Commun.*, vol. 20, no. 5, pp. 37–43, 2013.

[2] R. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 94–101, 2014.

[3] B. Wang, Q. Kong, W. Liu, and L. T. Yang, "On efficient utilization of green energy in heterogeneous cellular networks," *IEEE Syst. J.*, no. 99, pp. 1–12, 2015.

[4] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, 2014.

[5] H. Zhang, S. Chen, X. Li, H. Ji, and X. Du, "Interference management for heterogeneous networks with spectral efficiency improvement," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 101–107, 2015.

[6] J. Tang, D. So, E. Alsusa, K. Hamdi, and A. Shojaeifard, "Resource allocation for energy efficiency optimization in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2104–2117, 2015.

[7] C. Yang, Z. Chen, B. Xia, and J. Wang, "When ICN meets C-RAN for HetNets: an SDN approach," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 118–125, 2015.

[8] M. Peng, C. Wang, J. Li, H. Xiang, and V. Lau, "Recent advances in underlay heterogeneous networks: Interference control, resource allocation, and self-organization," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 700–729, 2015.

[9] D. Lopez-Perez, I. Güvenç, G. De la Roche, M. Kountouris, T. Q. Quek, and J. Zhang, “Enhanced intercell interference coordination challenges in heterogeneous networks,” *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, 2011.

[10] R. Hernandez-Aquino, S. A. R. Zaidi, D. McLernon, M. Ghogho, and A. Imran, “Tilt angle optimization in two-tier cellular networks-a stochastic geometry approach,” *IEEE Trans.Commun.*, vol. 63, no. 12, pp. 5162–5177, 2015.

[11] J. G. Andrews, “Seven ways that HetNets are a cellular paradigm shift,” *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, 2013.

[12] H. Tabassum, U. Siddique, E. Hossain, and M. J. Hossain, “Downlink performance of cellular systems with base station sleeping, user association, and scheduling,” *IEEE Trans.Wireless Commun.*, vol. 13, no. 10, pp. 5752–5767, 2014.

[13] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, “A survey of energy-efficient wireless communications,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, 2013.

[14] A. Merwaday and I. Güvenç, “Optimisation of FeICIC for energy efficiency and spectrum efficiency in LTE-advanced HetNets,” *IET Electronics Letters*, vol. 52, no. 11, pp. 982–984, 2016.

[15] G. P. Koudouridis and H. Li, “Distributed power on-off optimisation for heterogeneous networks-a comparison of autonomous and cooperative optimisation,” in *Proc. IEEE 17th Int. Workshop CAMAD*, 2012, pp. 312–317.

[16] I. Ashraf, F. Boccardi, and L. Ho, “Sleep mode techniques for small cell deployments,” *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 72–79, 2011.

[17] G. Wu, G. Feng, and S. Qin, “Cooperative sleep-mode and performance modeling for heterogeneous mobile network,” in *Proc. IEEE WCNCW*, 2013, pp. 6–11.

[18] P. Chandhar and S. S. Das, “Energy saving in OFDMA cellular networks with multi-objective optimization,” in *Proc. IEEE ICC*, 2014, pp. 3951–3956.

[19] H. Celebi, N. Maxemchuk, Y. Li, and I. Güvenç, “Energy reduction in small cell networks by a random on/off strategy,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2013, pp. 176–181.

[20] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, “Breathe to stay cool: adjusting cell sizes to reduce energy consumption,” in *Proceedings of the first ACM SIGCOMM workshop on Green networking*, 2010, pp. 41–46.

[21] S. Zhang, J. Wu, J. Gong, S. Zhou, and Z. Niu, “Energy-optimal probabilistic base station sleeping under a separation network architecture,” in *Proc. IEEE GLOBECOM*, 2014, pp. 4239–4244.

[22] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-aho, “Opportunistic sleep mode strategies in wireless small cell networks,” in *Proc. IEEE ICC*, 2014, pp. 2707–2712.

[23] X. J. Li and P. H. J. Chong, “A dynamic channel assignment scheme for TDMA-based multihop cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 1999–2003, 2008.

[24] G. F. Marias, D. Skyrianoglou, and L. Merakos, “A centralized approach to dynamic channel assignment in wireless ATM LANs,” in *Proc. IEEE INFOCOM’99.*, vol. 2, 1999, pp. 601–608.

[25] G. Cao and M. Singhal, “Distributed fault-tolerant channel allocation for cellular networks,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1326–1337, 2000.

[26] Y. Furuya and Y. Akaiwa, “Channel segregation, a distributed adaptive channel allocation scheme for mobile communication systems,” *IEICE Transactions on Communications*, vol. 74, no. 6, pp. 1531–1537, 1991.

[27] Z. He, Y. Zhang, C. Wei, and J. Wang, “A multistage self-organizing algorithm combined transiently chaotic neural network for cellular channel assignment,” *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1386–1396, 2002.

[28] C. Zhao and L. Gan, “Dynamic channel assignment for large-scale cellular networks using noisy chaotic neural network,” *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 222–232, 2011.

[29] J. Wang, Z. Tang, X. Xu, and Y. Li, “A discrete competitive Hopfield neural network for cellular channel assignment problems,” *Neurocomputing*, vol. 67, pp. 436–442, 2005.

[30] R. Montemanni, J. N. Moon, and D. H. Smith, “An improved tabu search algorithm for the fixed-spectrum frequency-assignment problem,” *IEEE Trans. Veh. Technol.*, vol. 52, no. 4, pp. 891–901, 2003.

[31] D. Gözüpek, G. Genç, and C. Ersoy, “Channel assignment problem in cellular networks: A reactive tabu search approach,” in *Computer and Information Sciences, 2009. ISCIS 2009. 24th International Symposium on*, 2009, pp. 298–303.

[32] Y. Peng, L. Wang, and B. H. Soong, “Optimal channel assignment in cellular systems using tabu search,” in *Proc. IEEE PIMRC.*, vol. 1, 2003, pp. 31–35.

[33] C. Y. Lee and H. G. Kang, “Cell planning with capacity expansion in mobile communications: A tabu search approach,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1678–1691, 2000.

[34] S. C. Ghosh, B. P. Sinha, and N. Das, “Channel assignment using genetic algorithm based on geometric symmetry,” *IEEE Trans. Veh. Technol.*, vol. 52, no. 4, pp. 860–875, 2003.

[35] J. Chen, S. Olafsson, and X. Gu, “Observations on using simulated annealing for dynamic channel allocation in 802.11 WLANs,” in *Proc. IEEE VTC Spring*, 2008, pp. 1801–1805.

[36] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, “Matching theory for future wireless networks: fundamentals and applications,” *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 52–59, 2015.

[37] P. Bremaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, ser. Texts in Applied Mathematics (Book 31). Springer, 1999.

[38] L. P. Qian, Y. J. Zhang, and M. Chiang, “Distributed nonconvex power control using Gibbs sampling,” *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3886–3898, 2012.

[39] A. H. Arani, A. Mehbodniya, M. J. Omidi, and F. Adachi, “Learning-based joint power and channel assignment for hyper dense 5G networks,” in *Proc. IEEE ICC*, 2016.

[40] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, “Distributed $\alpha$ -optimal user association and cell load balancing in wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 177–190, 2012.

[41] M. Ismail, W. Zhuang, E. Serpedin, and K. Qaraqe, “A survey on green mobile networking: From the perspectives of network operators and mobile users,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1535–1556, 2015.

[42] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, “Rethinking energy efficiency models of cellular networks with embodied energy,” *IEEE Netw.*, vol. 25, no. 2, pp. 40–49, 2011.

[43] J. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, “An overview of load balancing in HetNets: Old myths and open problems,” *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 18–25, 2014.

[44] M. Simsek, M. Bennis, and I. Güvenç, “Learning based frequency- and time-domain inter-cell interference coordination in HetNets,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4589–4602, 2015.

[45] M. H. Manshaei, P. Marbach, and J. P. Hubaux, “Evolution and market share of wireless community networks,” in *Game Theory for Networks, 2009. GameNets ’09. International Conference on*, 2009, pp. 508–514.

[46] A. H. Arani, M. J. Omidi, A. Mehbodniya, and F. Adachi, “A handoff algorithm based on estimated load for dense green 5G networks,” in *Proc. IEEE GLOBECOM*, 2015, pp. 1–7.

[47] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, ser. Adaptive Computation and Machine Learning series. MIT Press, 1998.

[48] 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. 5813–5825, 2013.

[49] M. Bennis, S. M. Perlaza, and M. Debbah, “Learning coarse correlated equilibria in two-tier wireless networks,” in *Proc. IEEE ICC*, 2012, pp. 1592–1596.

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

[51] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[52] Y. Matsumura, S. Kumagai, T. Obara, T. Yamamoto, and F. Adachi, “Channel segregation based dynamic channel assignment for WLAN,” in *Proc. IEEE ICCS*, 2012, pp. 463–467.

[53] A. Mehbodniya, K. Temma, R. Sugai, W. Saad, I. Güvenç, and F. Adachi, “Energy-efficient dynamic spectrum access in wireless heterogeneous networks,” in *Proc. IEEE ICCW*, 2015, pp. 2775–2780.

[54] W. Guo and T. O’Farrell, “Dynamic cell expansion with self-organizing cooperation,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 851–860, 2013.

[55] I. Aydin, H. Yanikomeroglu, and . Aygl, “User-aware cell switch-off algorithms,” in *Proc. 2015 IWCMC*, 2015, pp. 1236–1241.

[56] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-aho, “Dynamic clustering and on/off strategies for wireless small cell networks,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2164–2178, 2016.

[57] A. Galindo-Serrano, “Self-organized femtocells: A time difference learning approach,” Ph.D. dissertation, Universitat Politecnica de Catalunya (UPC), Barcelona, Spain, 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2017.2696974, IEEE Transactions on Vehicular Technology

**Atefeh Hajijamali Arani** received her M.Sc. degree in Telecommunication engineering from Tarbiat Modares University, Tehran, Iran, in 2012. Currently, she is a Ph.D. candidate at Isfahan University of Technology, Isfahan, Iran. She was a visiting researcher at Tohoku University, Japan, in 2015. Her research interests include heterogeneous networks, resource management, game theory, and self-organizing networks.

**Abolfazl Mehbodniya** (SM'16) he has been an Assistant Professor at department of communications engineering, Tohoku University since Jan. 2013. He has 10+ years of experience in electrical engineering, wireless communications, and project management. He has over 40 published conference and journal papers in the areas of radio resource management, sparse channel estimation, interference mitigation, short-range communications, 4G/5G design, OFDM, heterogeneous networks, artificial neural networks (ANNs) and fuzzy logic techniques with applications to algorithm and protocol design in wireless communications. He is the recipient of JSPS fellowship for foreign researchers, JSPS young faculty startup grant and KDDI foundation grant.

**M. Javad Omidi** received his Ph.D. from University of Toronto in 1998. He worked in industry by joining a research and development group designing broadband communication systems for 5 years in US and Canada. In 2003 he joined the Department of Electrical and Computer Engineering, at Isfahan University of Technology (IUT), Iran; and then served as the chair of Information Technology Center and chair of the ECE department at this university. Currently he is the Director of IRIS a UNESCO organization in Iran, and the VP for Research and Development at Isfahan Science and Technology Town .He has numerous publications and more than 15 US and international patents in the area of telecommunications. He is supervising a software radio lab in IUT and his scientific research interests are in the areas of mobile computing, wireless communications, digital communication systems, software radio, cognitive radio, and VLSI architectures for communication algorithms.

**Fumiyuki Adachi** (Fellow'02) is currently a specially-appointed Professor at Research Organization of Electrical Communication (ROEC), Tohoku University. His research interest includes wireless signal processing including wireless access, equalization, transmit/receive antenna diversity, adaptive transmission, and channel coding. He is an IEEE Fellow and an IEICE Fellow. He was a recipient of the IEEE Vehicular Technology Society Avant Garde Award 2000, IEICE Achievement Award 2002, Thomson Scientific Research Front Award 2004, Ericsson Telecommunications Award 2008, Telecom System Technology Award 2010, the Prime Minister Invention Award 2010, the KDDI Foundation Excellent Research Award 2012, and C&C Prize 2014.

**Walid Saad** (S'07, M'10, SM'15) received his Ph.D degree from the University of Oslo in 2010. Currently, he is an Assistant Professor and the Steven O. Lane Junior Faculty Fellow at the Department of Electrical and Computer Engineering at Virginia Tech, where he leads the Network Science, Wireless, and Security (NetSciWiS) laboratory, within the Wireless@VT research group. His research interests include wireless networks, game theory, cybersecurity, unmanned aerial vehicles, and cyber-physical systems. Dr. Saad is the recipient of the NSF CAREER award in 2013, the AFOSR summer faculty fellowship in 2014, and the Young Investigator Award from the Office of Naval Research (ONR) in 2015. He was the author/co-author of five conference best paper awards at WiOpt in 2009, ICIMP in 2010, IEEE WCNC in 2012, IEEE PIMRC in 2015, and IEEE SmartGridComm in 2015. He is the recipient of the 2015 Fred W. Ellersick Prize from the IEEE Communications Society. Dr. Saad serves as an editor for the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, and IEEE Transactions on Information Forensics and Security.

**Ismail Güvenç** (SM'10) has been an Associate Professor at North Carolina State University since August 2016. His recent research interests include 5G and mmWave wireless networks, heterogeneous networks, and UAV communications. He has published more than 170 conference/journal papers and book chapters, several standardization contributions, three books, and over 30 U.S. patents. He is a recipient of the 2016 FIU College of Engineering Faculty Research Award, 2015 NSF CAREER Award, 2014 Ralph E. Powe Junior Faculty Enhancement Award, and 2006 USF Outstanding Dissertation Award.