Received October 19, 2017, accepted November 19, 2017, date of publication November 27, 2017,
date of current version December 22, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2777914

Minimizing Base Stations' ON/OFF Switchings in Self-Organizing Heterogeneous Networks: A Distributed Satisfactory Framework

ATEFEH HAJIJAMALI ARANI[1], MOHAMMAD JAVAD OMIDI[1], ABOLFAZL MEHBODNIYA[2], (Senior Member, IEEE), AND FUMIYUKI ADACHI[2], (Life Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran
[2]Research Organization of Electrical Communication, Tohoku University, Sendai 980-8577, Japan

Corresponding author: Atefeh Hajijamali Arani (atefeh.haji@ec.iut.ac.ir)

This work was supported in part by the Commissioned Research of NICT, JAPAN, Towards Energy-Efficient Hyper-Dense Wireless Networks with Trillions of Devices, and in part by the KDDI Foundation Research Grant, Energy-Efficient Radio Resource Management for Next Generation Wireless Network.

ABSTRACT The deployment of heterogeneous networks (HetNets) can significantly boost the network capacity. However, the large number of small cell base stations (SBSs) deployed in HetNets can result in an increased total energy consumption. One of the promising techniques to reduce the energy consumption of networks is base station (BS) ON/OFF switching (sleeping) approaches. Due to device lifetime and energy waste by unnecessary switchings, the number of switchings is considered as an important problem. In this paper, we formulate the ON/OFF switching problem as a satisfaction game, where BSs seek to meet certain performance constraints in order to avoid the frequent BS switchings. Furthermore, BSs can choose their transmission power levels according to the network conditions in a distributed manner. The proposed satisfaction game involves a multi-step process. In the first step, we aim at satisfying the players with the high satisfaction threshold in a predefined time interval. To measure a BS's satisfaction, a utility function is used that includes BS's load and power consumption, in which the load of each BS is coupled with the load of other BSs. Since all players cannot be simultaneously satisfied, unsatisfied players decide to reduce their thresholds, and form a game with the redefined thresholds. To solve the game, a regret-based satisfaction algorithm and a satisfaction equilibrium search algorithm are applied. Simulation results show that the proposed schemes can achieve significant reductions in the number of switchings compared with the benchmark methods.

INDEX TERMS Heterogeneous networks, game theory, learning algorithm, self-organizing networks, sleep mode.

I. INTRODUCTION

Heterogeneous networks (HetNets), which consist of traditional macro cell overlaid with small cell base stations (SBSs), have been considered as a promising approach to meet the explosive mobile data traffic demand [1]. HetNets are envisioned to enable next-generation wireless networks to deliver higher rate data services, and offload traffic load from macro base station (MBS) to SBSs [2], [3]. Compared to MBS, SBSs cover much smaller areas, and hence they operate with lower transmit power [4], [5]. However, the growing number of SBSs gives rise to several challenges. Especially, densely deployed SBSs can significantly increase energy consumption of the networks, which will directly result in the increase of carbon footprint and particularly environmental problems [6], [7]. It has been estimated that the information and communication technology (ICT) sector is responsible for about 2% of global CO_2 emissions, in which the mobile communication industry contributes 15-20% [8]. Even with the technological advancements in the ICT infrastructure, 6% growth rate is expected in CO_2 emissions every year till 2020 [9]. From the economical perspective, the energy cost accounts to a large portion of operational expenditure (OPEX) of the network operators [10]. Therefore, it is essential to consider a paradigm shift to reduce the energy consumption of the networks.

According to some surveys on the energy consumption (e.g., see [8], [11], [12]), base stations (BSs) consume a significant proportion of energy in the network.

Furthermore, based on the results from China Mobile Communications Corporation, a BS consumes 100% and about 50%-60% energy at the peak traffic load and zero traffic load, respectively [8]. Therefore, switching OFF BSs (alternatively termed as sleep mode) and keeping them in energy-saving mode are effective solutions to achieve energy saving in low traffic load situations [13], [14]. The main advantages of these approaches are easier to test and implement as they do not require upgrading the equipments and changing the network architecture [11].

On the other hand, a major priority in HetNets is self-organization issue. In particular, 3GPP has defined self-organizing network (SON) as a key standardization feature [15]. SON allows BSs adjust their configurations with minimal human intervention in a distributed and flexible manner [16]. Therefore, SON can significantly reduce capital expenditure (CAPEX) and OPEX for network operators [17]. In HetNets, SONs can gain more importance when the HetNets become denser and more heterogeneous. Accordingly, ON/OFF switching schemes based on SON for BSs have been extensively studied in the literatures [18]–[21]. In LTE and LTE-advanced standards, a discontinuous transmission (DTX) and discontinuous reception (DTR) approaches are introduced where transceivers can be switched to a sleep mode whenever there is no data to transmit or receive [11]. Several predefined BS sleeping schemes according to traffic variation patterns are proposed in [22]–[24]. In [23], optimal power saving schemes by reducing the number of active cells in low traffic load situations are analyzed. It is proved that a 25%-30% energy saving is possible. A random ON/OFF switching approach for SBSs is proposed in [25]. There is a tradeoff between delay and transmit power where user equipments (UEs) can save energy by delaying their transmissions in order to connect closer SBSs. In [26], a UE offloading algorithm to balance load for the sleeping BSs is proposed. To minimize power consumption, an integer optimization problem is formulated. A sleep mechanism for IEEE 802.11 wireless local area network access points is developed in [27]. To determine the length of sleep interval, a dynamic sleep boundary decision algorithm based on the rate and delay is proposed. A sleep mode approach for BSs based on a binary social spider algorithm is proposed in [28]. In [29], two low complexity approximation algorithms are proposed to select the active BSs set that can preserve the UEs' minimal rate requirements. For the BS ON/OFF switching problem, two simple greedy-on and greedy-off algorithms are proposed in [30]. Furthermore, Son *et al.* [30] develop the greedy heuristic algorithms based on the distances between BSs without any additional signaling overhead. In [31], a BS switch-OFF algorithm based on an artificial neural network is developed. In order to offload UEs from MBSs to SBSs or neighboring MBSs, a decentralized sleep mechanism for MBSs is introduced in [32]. In this mechanism, it is assumed that the SBSs are always ON. By exploiting the cooperation of cells in HetNets, a sleep mode technique for SBSs is proposed in [33]. In [34], a centralized greedy-add algorithm is proposed for the BS switch-OFF approach. To investigate the effect of cell sorting on the energy saving, three different BS sorting criteria are compared. For BS ON/OFF switching problem in the HetNets, a centralized algorithm based on the simulated-annealing search is proposed in [35]. A game theoretic approach for switching OFF BSs in a network with two operators is proposed in [36], where a cooperation between the operators is required. To find the optimal BS sleeping strategies, a distributed cooperative approach is proposed in [37]. The problem is formulated as a constrained graphical game, in which the set of players corresponds to the set of BSs, and the solution converges to a generalized Nash equilibrium. In this approach, the neighboring BSs cooperate with each other based on the load to reduce the energy consumption, while guaranteeing UE's quality of service (QoS). Furthermore, it only requires local information exchange among the neighboring BSs in the network. The proposed mechanisms in [33]–[37], rely on centralized methods or require cooperation among BSs in the network. Thus, they come at the expense of knowing local or global information. This means that, some information is needed to exchange among BSs. Therefore, they may increase the signaling overhead in the networks. However, such approaches have their own advantages, and the outcomes obtained from the fully distributed approaches might not always be as good as the outcomes obtained from the approaches that utilize the local or global information.

Noncooperative game theory plays a fundamental role for enabling SON feature. In [38], an energy-efficient resource management for SONs using game theoretic framework is proposed. To save energy consumption, BSs employ an ON/OFF switching algorithm, where there is a tradeoff between energy consumption reduction and load. To solve the game, a regret based learning algorithm is applied, in which BSs select their transmission power in a distributed manner. The proposed algorithm converges to an ϵ-coarse correlated equilibrium. From this point of view, players in the game aim at optimizing their utility functions by selfishly choosing theirs strategies. For most widely used applications in wireless networks, players seek strategies that meet certain performance constraints, instead of optimizing the individual performance [39]. Games in satisfaction-form enable players to adapt their strategies to guarantee these constraints. The idea of satisfaction game and the corresponding notion of equilibrium in decentralized self-configuring networks are introduced in [40]–[42]. In [43], the problem of QoS provisioning in decentralized networks is modeled as a satisfaction game approach. Moreover, a comparison between the concept of generalized Nash equilibrium and the satisfaction equilibrium is provided.

However, when a BS switches between ON and OFF modes, there is a switching cost which cannot be ignored. This issue has not been addressed properly in the aforementioned research works. In this paper, we use the framework of satisfaction equilibrium, and propose satisfaction based approaches for BS ON/OFF switching problem. In these approaches, the ON/OFF switching problem is modeled as a noncooperative game in satisfaction-form, in which players only need to achieve target utility constraints. Our proposed framework considers a multi-level threshold mechanism. In order to meet an equilibrium, we introduce adaptive satisfaction thresholds for players that are unable to achieve satisfaction. There are two issues involved in this problem: 1) for given satisfaction thresholds, forming a satisfaction game and applying a learning algorithm, such that the set of satisfied players is maximized, and 2) decreasing satisfaction thresholds for the set of unsatisfied players after a predefined time interval. If the player is satisfied, it has no inclination to change its strategy. To solve the game in a distritbuted manner, we apply a regret based satisfaction algorithm and a satisfaction equilibrium search algorithm (SESA). In the regret based satisfaction algorithm, players play each strategy with a non-zero probability and based on their regrets. In the SESA, each player only requires the value of its individual utility to assign a probability distribution to its strategies.

The proposed satisfaction based approaches can help to reduce the number of switches in BSs, and thus switching costs. Furthermore, they do not require a central controller and information exchange. Therefore, they are executed in a fully distributed manner, and suitable for dense BSs deployments scenarios. Since players can only target satisfactory performance levels, they can reduce the computational complexity.

The rest of the paper is organized as follows. Section II introduces the system model. Section III presents the problem statement, and formulates the ON/OFF switching problem as a satisfaction-form game. Furthermore, a no-regret learning algorithm and a learning algorithm based on the observed utility are described to learn a satisfaction equilibrium. In section IV, we evaluate the performance of the proposed approach, and Section V concludes the paper.

Notations: 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 $\Delta(\mathcal{A})$, respectively. The function $\mathbb{1}_\phi$ denotes the indicator function which equals 1 if event ϕ is true and 0, otherwise. The set of real numbers is denoted by $\mathbb{R}$. Furthermore, the single- and n-dimension space of all real non-negative numbers are denoted by $\mathbb{R}_+$ and $\mathbb{R}_+^n$, respectively.

II. SYSTEM MODEL

We consider the downlink of a two-tier HetNet consisting of the set of MBSs $\mathcal{B}_{\mathrm{M}}$ overlaid with a set of SBSs $\mathcal{B}_{\mathrm{S}}$. For each MBS's coverage area, a MBS is located at the center of area, and the SBSs are uniformly distributed within the coverage of the MBS as shown in Fig. 1. The set of UEs is denoted by $\mathcal{K}$, which are uniformly distributed over the area. Assume that each UE $k \in \mathcal{K}$ can be associated with at most one BS at each time.

Let $p_b(t)$ be the transmit power of BS b at time instant t. In order to save energy, BSs can switch to an OFF mode.

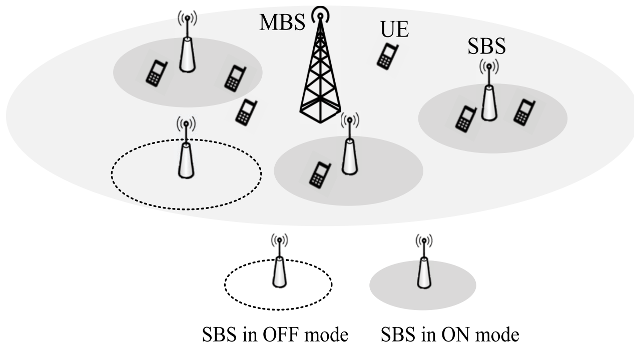

FIGURE 1. An illustration of a two-tier HetNet.

In this respect, each BS $b \in \mathcal{B}$ is either in an ON mode (i.e. $p_b(t) > 0$) or an OFF mode (i.e. $p_b(t) = 0$) at time instant t. To characterise the power consumption at a BS, we need to consider the efficiency of the power amplifier, the consumed power in the components such as baseband unit, radio frequency module, and cooling system. The consumed power of various components in BSs depends on the BS type. We adopt a linear power consumption model proposed in [44] which can be expressed as:

$$P_b^{\mathrm{total}}(t) = \frac{P_b^{\mathrm{PA}}(t) + P_b^{\mathrm{RF}} + P_b^{\mathrm{BB}}}{\Lambda_b} \text{ [W]}, \tag{1}$$

with

$$P_b^{\mathrm{PA}}(t) = \frac{p_b(t)}{\eta_b^{\mathrm{PA}} \cdot (1 - \lambda_b^{\mathrm{feed}})} \text{ [W]}, \tag{2}$$

and

$$\Lambda_b = (1 - \lambda_b^{\mathrm{DC}})(1 - \lambda_b^{\mathrm{MS}})(1 - \lambda_b^{\mathrm{cool}}), \tag{3}$$

where $P_b^{\mathrm{PA}}(t)$, P_b^{RF}, and P_b^{BB} indicate the power consumed by power amplifier, radio frequency module, and baseband engine, respectively. Parameter η_b^{PA} denotes the power amplifier efficiency of BS b. Parameters $\lambda_b^{\mathrm{feed}}$, λ_b^{DC}, λ_b^{MS}, and $\lambda_b^{\mathrm{cool}}$ represent the losses which are incurred by feeder, DC-DC power supply, main supply, and cooling system, respectively. Typically, the feeder losses can be ignored for SBSs. Note that there are no cooling equipments in SBSs. In order to select different amounts for the power levels, there are several practical limitations that are imposed by standards and hardware, such as minimum and maximum transmission power. Practically, there are only limited number of discrete power levels available at the transmitter. Therefore, we consider finite and predetermined number of discrete power levels for BSs [45].

Suppose UE k is served by BS b, and σ^2 denotes the additive white Gaussian noise (AWGN) power. We assume that all BSs transmit over the same channel, i.e. co-channel deployment. To model signal to interference plus noise ratio (SINR), a load coupling model is considered, in which the interference from the other BSs are scaled by their loads [46]. The load of each BS is defined as the fraction of the available resource that are being used [47]. Then, SINR experienced by UE k can be formulated as [48], [49]:

$$\gamma_{b,k}(t) = \frac{p_b(t)g_{b,k}(t)}{\sum_{b' \in \mathcal{B}\setminus b} p_{b'}(t)g_{b',k}(t)\rho_{b'}(t) + \sigma^2}, \tag{4}$$

where $g_{b,k}(t)$ and $\rho_b(t)$ are the channel gain between BS b and UE k and the load of BS b at time t, respectively. The term $\sum_{b' \in \mathcal{B}\setminus b} p_{b'}(t)g_{b',k}(t)\rho_{b'}(t)$ denotes the time averaged interference power [50]. This model provides a good approximation for multi-cell systems [49], [51]. From Shannon's capacity formula, the achievable transmission rate of UE k is:

$$R_{b,k}(t) = \omega \log_2(1 + \gamma_{b,k}(t)) \text{ [bits/sec]}, \tag{5}$$

where ω is the bandwidth. Let ϑ_k be the traffic influx rate of UE k. This assumption can capture different UEs' QoS demands (i.e. heterogeneous UEs) depending on hardware capabilities and UE's activity [52]. The fraction of time required to serve the traffic load from BS b to the location of UE k is defined as $\frac{\vartheta_k}{R_{b,k}(t)}$ [53]. Therefore, the load of BS b at time t is expressed by:

$$\begin{aligned}\rho_b(t) &= \sum_{k \in \mathcal{K}_b} \frac{\vartheta_k}{R_{b,k}(t)} \\ &= \sum_{k \in \mathcal{K}_b} \frac{\vartheta_k}{\omega \log_2\left(1 + \frac{p_b(t)g_{b,k}(t)}{\sum_{b' \in \mathcal{B}\setminus b} p_{b'}(t)g_{b',k}(t)\rho_{b'}(t)+\sigma^2}\right)} \\ &\triangleq h_b(\boldsymbol{\rho}),\end{aligned} \tag{6}$$

where $\mathcal{K}_b$ and $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_{|\mathcal{B}|}]$ denote the set of UEs associated with BS b and the network load vector, respectively. The function $h_b(\boldsymbol{\rho}), b = 1, \ldots, |\mathcal{B}|$, represents the $\mathbb{R}_+^{|\mathcal{B}|-1} \to \mathbb{R}_+$ function as defined by (6). The load vector is obtained (if a solution exists) by solving the following nonlinear equations:

$$\begin{aligned}\rho_1 &= h_1(\boldsymbol{\rho}) \\ &\vdots \\ \rho_{|\mathcal{B}|} &= h_{|\mathcal{B}|}(\boldsymbol{\rho}).\end{aligned} \tag{7}$$

Let $\boldsymbol{h}(\boldsymbol{\rho}) = [h_1(\boldsymbol{\rho}), \ldots, h_{|\mathcal{B}|}(\boldsymbol{\rho})]$. In vector form, we have:

$$\boldsymbol{\rho} = \boldsymbol{h}(\boldsymbol{\rho}). \tag{8}$$

Since the load $\boldsymbol{\rho}$ appears in both sides of (8), it cannot be solved in closed-form. Note that a solution $\boldsymbol{\rho}^*$ is feasible if $\boldsymbol{\rho}^*$ satisfies (8) and $\boldsymbol{\rho}^* \geq 0$. Due to the limited resources available in the network, the load cannot exceed one. Thus, if the load of a BS exceeds one, the BS will drop some UEs and/or decrease the UEs' throughput [54]. Since the function $h_b(\boldsymbol{\rho})$ is a standard interference function (SIF) as a function of $\boldsymbol{\rho}$, the solution can be computed by using fixed point iterations.

Definition 1: A function $I(\boldsymbol{x})$ is a SIF if for all $x \geq 0$ the following properties hold [55]:

1) *Positivity: $I(\boldsymbol{x}) > 0$,*
2) *Monotonicity: $x \geq x' \Rightarrow I(x) \geq I(x')$,*
3) *Scalability: $\alpha I(x) > I(\alpha x)$ for $\alpha > 1$,*

Starting from an initial load $\boldsymbol{\rho}^0 > 0$, the unique fixed point solution (if exists) of (8) can be found iteratively by the following algorithm:

$$\boldsymbol{\rho}^m = \min\left(\boldsymbol{h}(\boldsymbol{\rho}^{m-1}), 1\right), \tag{9}$$

for $m = 1, \ldots, M$, where M and $\boldsymbol{\rho}^m$ are the total number of iterations and the output for iteration m, respectively.

Lemma 1: If a feasible load $\boldsymbol{\rho}^*$ exists for (8), then it is unique, and $\boldsymbol{\rho}^M$ converges to $\boldsymbol{\rho}^*$ as $M \to \infty$ [51].

Proof: Proving that $h_b(.)$ is a SIF is presented in [49]. Using Theorem 7 in [55], if the function $h_b(\boldsymbol{\rho})$ is a SIF, then the function $\min(h_b(\boldsymbol{\rho}), 1)$ is a SIF. Then, by applying Theorem 2 in [55], the convergence is proved. ■

From (6), with increasing the transmission power of BS b, the SINR of its associated UEs increases. Therefore, given the traffic influx rate of its UEs, the load of BS b decreases. Thus, there is the tradeoff between BS's load and energy consumption reduction. Moreover, the load grows with the traffic demand and the amount of interference [48]. In the following section, the problem of selecting BS's power level is presented.

III. PROBLEM FORMULATION

In most existing works that consider the problem of switching BS's power level, the number of ON/OFF switchings is not considered. In practical systems, switchings in a BS impose an additional power consumption. In a specific case, switching between ON and OFF modes incurs an energy cost that is needed to switch ON and OFF some components of a BS. Moreover, a portion of energy is consumed for signaling for re-association of the dropped UEs by BSs that switch to a sleep mode [56]. For future generation cellular networks, cloud radio access network (C-RAN) is considered as a novel mobile network architecture [57]. In C-RAN, BSs' switchings will cause the high variations of resource demand in the servers of C-RAN. As a result, it will yield the higher switching cost, including: the energy used, the delay in migration connections/data, increased wear-and-tear on the servers, and the risk associated with server toggling [58].

Our objective is to reduce the energy consumption of BSs while satisfying QoS requirement and avoiding frequent switchings. Due to the limited device lifetime and switching energy consumption, it is necessary to minimize the number of switchings [59]. Further, ON/OFF switching in BSs must be coupled with UE association problem: when a BS switches to an OFF mode, its associated UEs need to migrate to another BSs. Therefore, UE association problem should be considered jointly with the BS operation problem [30].

A. UE ASSOCIATION

At time t, UE k is associated with BS $b(k,t)$ based on the received signal power and BS's estimated load as follows [60]:

$$b(k,t) = \underset{b \in \mathcal{B}}{\operatorname{argmax}} \left\{p_b(t)g_{b,k}(t)\big(1 - \hat{\rho}_b(t)\big)\right\}, \tag{10}$$

where $\hat{\rho}_b(t)$ denotes the estimated load of BS b at time t, which is advertised by BS b through a broadcast control message. The estimated load $\hat{\rho}_b(t)$ is calculated based on history as follows [61]:

$$\hat{\rho}_b\left(t\right) = \hat{\rho}_b\left(t-1\right) + \tau(t)\Big(\rho_b(t-1) - \hat{\rho}_b(t-1)\Big), \quad (11)$$

where $\tau(t)$ is the learning rate for the load estimation.

B. ON/OFF SWITCHING GAME IN SATISFACTION-FORM

In this section, we solve the ON/OFF switching in two steps. In the first step, for a given satisfaction threshold, a learning algorithm is used to solve the game such that the set of satisfied players is maximized. In the second step, for the set of unsatisfied players, the satisfaction thresholds decrease. Now, we define a utility function including the energy consumption and load, as follows:

$$u_b(t) = -\left(\omega_b \cdot \rho_b(t) + \phi_b \cdot \frac{P_b^{\text{total}}(t)}{P_b^{\text{TM}}}\right), \quad (12)$$

where ω_b and ϕ_b are the weight parameters for BS b that indicate the impact of energy and load on the utility function, respectively. Parameter P_b^{TM} denotes the maximum power consumed by BS b when it transmits with its maximum power. Then, the energy saving problem can be formulated as:

$$\max_{p_1(t),\ldots,p_{|\mathcal{B}|}(t)} \sum_{b\in\mathcal{B}} u_b(t) \quad (13a)$$
$$\text{subject to } p_b(t) \le P_b^{\max}, \quad \forall b \in \mathcal{B} \quad (13b)$$
$$\rho_b(t) = h_b(\boldsymbol{\rho}), \quad \forall b \in \mathcal{B} \quad (13c)$$
$$0 \le \rho_b(t) \le 1, \quad \forall b \in \mathcal{B}, \quad (13d)$$

where $P_b^{\max}$ denotes the maximum transmit power of BS b. The constraint in (13b) corresponds to the limits on the BSs' transmission power. The constraints in (13c)-(13d) are related to the definition of load. In general, to solve the problem (13), a game $\mathcal{G}_{\text{NF}}$ in normal-form can be used which is described by the following triplet:

$$\mathcal{G}_{\text{NF}} = \langle \mathcal{B}, \quad \{\mathcal{A}_b\}_{b\in\mathcal{B}}, \ \{u_b\}_{b\in\mathcal{B}} \rangle \,, \quad (14)$$

where $\mathcal{B}$ and $\mathcal{A}_b$ are the set of BSs as the players and the strategy set of player b, respectively. The utility function of player b is given by $u_b : \mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_{|\mathcal{B}|} \to \mathbb{R}$. To solve the problem (13) with avoiding frequent mode switchings, we model the $\mathcal{G}_{\text{NF}}$ in satisfaction-form which can be defined as follows:

$$\mathcal{G}_{\text{SF}} = \langle \mathcal{B}, \quad \{\mathcal{A}_b\}_{b\in\mathcal{B}}, \ \{f_b\}_{b\in\mathcal{B}} \rangle \,, \quad (15)$$

where $f_b : \mathcal{A}_{-b} \to 2^{\mathcal{A}_b}$ represents a correspondence function for satisfaction of the constraint, where $\mathcal{A}_{-b} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_{b-1} \times \mathcal{A}_{b+1} \times \cdots \times \mathcal{A}_{|\mathcal{B}|}$. The correspondence $f_b(\boldsymbol{a}_{-b}) \subseteq \mathcal{A}_b$ determines the set of strategies of player b which allows its satisfaction given the strategies of all other players $\boldsymbol{a}_{-b}$, and $2^{\mathcal{A}_b}$ is the set of all subsets of the set $\mathcal{A}_b$. Let $a_b \in \mathcal{A}_b$ and $\boldsymbol{a}_{-b} = (a_1, \ldots, a_{b-1}, a_{b+1}, \ldots, a_{|\mathcal{B}|})$ denote the strategy of player b and the strategies of all players except player b, respectively. More specifically, given a strategy profile $\boldsymbol{a} = (a_b, \boldsymbol{a}_{-b})$, player b is satisfied if $a_b \in f_b(\boldsymbol{a}_{-b})$. The strategy set and correspondence are defined as follows:

- **Strategy set:** A pure strategy of player b is its transmission power. The available pure strategies for BS b are $\mathcal{A}_b = \{0, \frac{1}{|\mathcal{A}_b|-1}P_b^{\text{Max}}, \ldots, \frac{|\mathcal{A}_b|-1}{|\mathcal{A}_b|-1}P_b^{\text{Max}}\}$ where $|\mathcal{A}_b| \ge 2$.
- **Correspondence:** The correspondence f_b is defined as follows:

$$f_b(\boldsymbol{a}_{-b}) = \{a_b \in \mathcal{A}_b : u_b(t) \ge \Gamma_b\}, \quad (16)$$

where Γ_b denotes a threshold value for player b, which represents the satisfactory threshold for the player.

Definition 2 (Satisfaction Equilibrium): A strategy profile $\boldsymbol{a}' = (a'_1, \ldots, a'_{|\mathcal{B}|})$ is an equilibrium for the game $\mathcal{G}_{\text{SF}}$ if

$$a'_b \in f_b(\boldsymbol{a}_{-b}), \quad \forall b \in \mathcal{B}. \quad (17)$$

In other words, a strategy profile $\boldsymbol{a}'$ is an equilibrium if the strategy at all players corresponds to a strategy that yields satisfaction given all other players' strategies. In wireless networks, simultaneously satisfying all players might not always be feasible, and thus a satisfaction equilibrium does not exist. In such a case, there are two options [62]:

- find the largest subset of satisfied players
- redefine the satisfaction thresholds for unsatisfied players.

To implement the first option, if a player is unsatisfied for a certain time period, it will select a *do-nothing* strategy for a period of time. In the second option, the satisfaction threshold is decreased for each unsatisfied player. Here, we consider the second option. In this regard, we divide time period (i.e. total number of iterations) T into N time intervals with duration T/N. In each time interval $[T_i, T_{i+1})$, the satisfaction threshold is constant, and it is lower than previous threshold, i.e. $\Gamma_b(T_i) < \Gamma_b(T_{i-1})$. Therefore, we consider the adaptive threshold which is defined as follows:

$$f_b(t, \boldsymbol{a}_{-b}) = \{a_b \in \mathcal{A}_b : u_b(t) \ge \Gamma_b(T_i)\}, \quad T_i \le t < T_{i+1} \quad (18)$$

where T_i is i^{th} time interval of time period T. Moreover, in each interval T_i, the player keeps a constant $\Gamma_b(T_i)$. Note that T_i should be large enough such that the player is able to try all strategies in order to satisfy. When all players are simultaneously satisfied, an equilibrium is observed.

For each player b, we choose a proper choice $\Gamma_b(T_i)$ for $i = 0$. If none of its strategies lead to meeting its satisfaction threshold $\Gamma_b(T_0)$, the player need to decrease the threshold. We redefine the threshold $\Gamma_b(T_i)$ as follows:

$$\Gamma_b(T_i) = \Gamma_b(T_{i-1}) - \delta \cdot |\Gamma_b(T_0)|, \quad (19)$$

where $0 < \delta < 1$ denotes a decrement coefficient. If the player chooses a strategy that satisfies it, then it has no incentive to deviate from it.

In order to find a strategy that satisfies the constraint, the player can adopt the following algorithm:

i) For a time interval T_0, the player b starts by choosing a strategy from its strategy set $\mathcal{A}_b$.
ii) If the player is satisfied, it has no incentive to deviate from it.
iii) If the player is not satisfied, it selects another strategy according a learning tool.

To learn at least one satisfaction equilibrium, players utilize one of the behavioral rules described in the next subsection.

C. LEARNING THE SATISFACTION EQUILIBRIUM

In this subsection, two learning algorithms are presented to distributively achieve a satisfaction equilibrium.

1) REGRET BASED SATISFACTION ALGORITHM

Each player $b \in \mathcal{B}$ chooses its strategy as follows. In the first iteration, the player selects its strategy following an arbitrary probability distribution $\boldsymbol{\pi}_b(0) \in \triangle(\mathcal{A}_b)$. Let $\tilde{v}_b(t)$ be the satisfaction indicator of BS b such that $\tilde{v}_b(t) = 1$ indicates that BS b is satisfied at time t and $\tilde{v}_b(t) = 0$, otherwise. At time t, player b keeps playing the same action played at time $t-1$, if it is satisfied (i.e. $\tilde{v}_b(t-1) = 1$). If the player is not satisfied (i.e. $\tilde{v}_b(t-1) = 0$), it may change its strategy. In this case, it chooses its strategy based on a probability distribution $\boldsymbol{\pi}_b(t)$. Therefore, the player selects the strategy according to [43]:

$$a_b(t) = \begin{cases} a_b(t-1), & \text{if } \tilde{v}_b(t-1) = 1 \\ \sim \boldsymbol{\pi}_b(t), & \text{if } \tilde{v}_b(t-1) = 0. \end{cases} \tag{20}$$

where $\sim \boldsymbol{\pi}_b(t)$ means according to the probability distribution $\boldsymbol{\pi}_b(t)$. In order to evaluate the probability distribution $\boldsymbol{\pi}_b(t)$, each player evaluates its regret for not having played its strategies, and aims at minimizing its regret [63]. In order to calculate the regret, the player needs the knowledge of other players' strategies. Thus, a learning tool is applied. For each player $b \in \mathcal{B}$ and $a_b \in \mathcal{A}_b$, the estimated utility $\widehat{u}_{b,a_b}(t)$, the estimated regret $\widehat{R}_{b,a_b}(t)$, and the probability $\pi_{b,a_b}(t)$ are updated as follows [63]:

$$\begin{aligned}\widehat{u}_{b,a_b}(t) &= \widehat{u}_{b,a_b}(t-1) \\ &\quad + (\frac{1}{t^{\kappa}})\mathbb{1}_{\{a_b(t)=a_b\}} \left(u_b(t) - \widehat{u}_{b,a_b}(t)\right), \quad (21)\\ \widehat{R}_{b,a_b}(t) &= \widehat{R}_{b,a_b}(t-1) \\ &\quad + (\frac{1}{t^{\zeta}}) \left(\widehat{u}_{b,a_b}(t) - u_b(t) - \widehat{R}_{b,a_b}(t-1)\right), \quad (22)\\ \pi_{b,a_b}(t) &= \pi_{b,a_b}(t-1) \\ &\quad + (\frac{1}{t^{\nu}}) \left(G_{b,a_b}(\widehat{\boldsymbol{R}}_b(t)) - \pi_{b,a_b}(t-1)\right) \quad (23)\end{aligned}$$

where $\kappa, \zeta > 0$ and $\nu > 0$ denote the learning rate exponents. Let $\widehat{\boldsymbol{R}}_b(t) = [\widehat{R}_{b,1}(t), \ldots, \widehat{R}_{b,|\mathcal{A}_b|}(t)]$ represents the vector of regret estimation values. Here, $\mathbf{G}_b(\widehat{\boldsymbol{R}}_b(t)) = [G_{b,1}(\widehat{\boldsymbol{R}}_b(t)), \ldots, G_{b,|\mathcal{A}_b|}(\widehat{\boldsymbol{R}}_b(t))]$ is the Boltzmann-Gibbs distribution vector defined as follows:

$$G_{b,a_b}\left(\widehat{\boldsymbol{R}}_b(t)\right) = \frac{\exp\left(\frac{1}{\theta_b}\widehat{R}_{b,a_b}(t)\right)}{\sum_{\forall \acute{s}_b \in \mathcal{A}_b} \exp\left(\frac{1}{\theta_b}\widehat{R}_{b,\acute{a}_b}(t)\right)}, \tag{24}$$

where $\frac{1}{\theta_b} > 0$ denotes the temperature parameter for player b. In the learning algorithm, 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 [63]. Thus, the learning rate exponents are chosen as $\nu > \zeta > \kappa$. Since players may not update (21)-(23) at each iteration, thus the proposed approach can reduce the computational complexity with respect to a normal-form game $\mathcal{G}_{\text{NF}}$ which uses the learning algorithm (21)-(23) to solve the game. The behavioral rule in (20)-(23) are summarized in Algorithm 1.

Theorem 1: Algorithm 1 is guaranteed to converge to a satisfaction equilibrium of game $\mathcal{G}_{\text{SF}}$ in finite time.

Proof: The proof of Theorem 1 is presented in Appendix A. ■

2) SATISFACTION EQUILIBRIUM SEARCH ALGORITHM (SESA)

This approach is the modified version of the algorithm proposed in [64], which is based on the observed utilities. It uses the knowledge of the individual utility to assign the probabilities to the strategies. In the SESA approach, each player $b \in \mathcal{B}$ selects its strategy according to (20). If the player is not satisfied, it updates its strategy according to a probability update function $d_b\Big(\pi_{b,a_b}(t)\Big)$ which is defined as follows [41]:

$$\begin{aligned}d_b\Big(\pi_{b,a_b}(t)\Big) &= \pi_{b,a_b}(t) \\ &\quad + \alpha_b(t) r_b(t)\Big(\mathbb{1}_{\{a_b(t)=a_b\}} - \pi_{b,a_b}(t)\Big),\end{aligned} \tag{25}$$

where $\alpha_b(t) = \frac{1}{t+1}$ is the learning rate of BS b. The parameter $r_b(t)$ is computed as follows:

$$r_b(t) = \frac{1 + u_b(t) - \Gamma_b(T_i)}{2}, \quad T_i \leq t < T_{i+1}. \tag{26}$$

Thus, the probability distribution for each player $b \in \mathcal{B}$ is described as follows:

$$\pi_{b,a_b}(t+1) = \begin{cases} \pi_{b,a_b}(t), & \text{if } \tilde{v}_b(t) = 1 \\ d_b\Big(\pi_{b,a_b}(t)\Big), & \text{if } \tilde{v}_b(t) = 0. \end{cases} \tag{27}$$

The pseudocode for SESA is presented in Algorithm 2. The point here is that, each player only requires the value of its individual utility at time t. Moreover, the increment in the probability of each strategy depends on the observed utility and the learning rate [65].

Theorem 2: The proposed ON/OFF switching based on the SESA converges to an equilibrium of the game $\mathcal{G}_{\text{SF}}$ in finite time, if it holds that $\pi_{b,a_b}(t) > 0$ for all $b \in \mathcal{B}$ and all $a_b \in \mathcal{A}_b$.

Proof: This theorem is proved similar to Theorem 1. However, please note that we need the probability distribution in (27) to fulfill the constraint $\pi_{b,a_b}(t+1) > 0$ for

Algorithm 1 (Regret Based satisfaction): Learning the Satisfaction Equilibrium of the Game $\mathcal{G}_{\mathrm{SF}} = \langle \mathcal{B}, \{\mathcal{A}_b\}_{b\in\mathcal{B}}, \{f_b\}_{b\in\mathcal{B}} \rangle$

Input: $\widehat{u}_{b,a_b}(t)$, $\widehat{R}_{b,a_b}(t)$, $\pi_{b,a_b}(t)$, $\tilde{v}_b(t-1)$, $\Gamma_b(T_0)$, $\forall b \in \mathcal{B}$ and $\forall a_b \in \mathcal{A}_b$
Output: $\widehat{u}_{b,a_b}(t+1)$, $\widehat{R}_{b,a_b}(t+1)$, $\pi_{b,a_b}(t+1)$, $\tilde{v}_b(t)$, $\forall b \in \mathcal{B}$ and $\forall a_b \in \mathcal{A}_b$
Initialization: $\boldsymbol{\pi}_b(0)$, $\forall b \in \mathcal{B}$
for $\forall b \in \mathcal{B}$ **do**
Select a strategy $a_b(0) \sim \boldsymbol{\pi}_b(0)$,
Calculate $u_b(0)$, {using (12).}
if $u_b(t) \geq \Gamma_b(T_0)$ **then**
$\tilde{v}_b(0) = 1$ {BS b is satisfied.}
else
$\tilde{v}_b(0) = 0$ {BS b is not satisfied.}
for $\forall a_b \in \mathcal{A}_b$ **do**
Update $\widehat{u}_{b,a_b}(1)$, $\widehat{R}_{b,a_b}(1)$, and $\pi_{b,a_b}(1)$, {using (21)-(23).}
end for
end if
end for
while $0 \leq t < T$ **do**
$t \leftarrow t+1$
for $\forall b \in \mathcal{B}$ **do**
if $\tilde{v}_b(t-1) = 1$ **then**
$a_b(t) = a_b(t-1)$ {BS b is satisfied and it does not change its strategy.}
else
Select a strategy $a_b(t) \sim \boldsymbol{\pi}_b(t)$ {BS b is not satisfied, and it selects its strategy according to the probability distribution $\boldsymbol{\pi}_b(t)$.}
end if
Calculate $u_b(t)$, {using (12).}
if $u_b(t) \geq \Gamma_b(T_i)$ **then**
$\tilde{v}_b(t) = 1$ {BS b is satisfied.}
else
$\tilde{v}_b(t) = 0$ {BS b is not satisfied.}
for $\forall a_b \in \mathcal{A}_b$ **do**
Update $\widehat{u}_{b,a_b}(t+1)$, $\widehat{R}_{b,a_b}(t+1)$, and $\pi_{b,a_b}(t+1)$, {using (21)-(23).}
end for
end if
end for
end while

$\forall b \in \mathcal{B}$, $\forall a_b \in \mathcal{A}_b$ and $t \geq 0$. Thus, at least one strategy profile exists that satisfies all players, and as a result, it converges to a satisfaction equilibrium. ■

Note that a lower satisfaction threshold leads to finding a solution faster, while a higher satisfaction threshold can lead to a better performance.

IV. SIMULATION RESULTS

To evaluate the performance of our proposed schemes, we consider a HetNet with a hexagonal layout with one MBS

Algorithm 2 (SESA) : Learning the satisfaction equilibrium of the game $\mathcal{G}_{\mathrm{SF}} = \langle \mathcal{B}, \{\mathcal{A}_b\}_{b\in\mathcal{B}}, \{f_b\}_{b\in\mathcal{B}} \rangle$

Input: $\pi_{b,a_b}(t)$, $\tilde{v}_b(t-1)$, $\Gamma_b(T_0)$, $\forall b \in \mathcal{B}$ and $\forall a_b \in \mathcal{A}_b$
Output: $\pi_{b,a_b}(t+1)$, $\tilde{v}_b(t)$, $\forall b \in \mathcal{B}$ and $\forall a_b \in \mathcal{A}_b$
Initialization: $\boldsymbol{\pi}_b(0)$, $\forall b \in \mathcal{B}$
for $\forall b \in \mathcal{B}$ **do**
Select a strategy $a_b(0) \sim \boldsymbol{\pi}_b(0)$
Calculate $u_b(0)$, {using (12).}
if $u_b(0) \geq \Gamma_b(T_0)$ **then**
$\tilde{v}_b(0) = 1$ {BS b is satisfied.}
else
$\tilde{v}_b(0) = 0$ {BS b is not satisfied.}
for $\forall a_b \in \mathcal{A}_b$ **do**
Update $\pi_{b,a_b}(1)$, {using (25).}
end for
end if
end for
while $0 \leq t < T$ **do**
$t \leftarrow t+1$
for $\forall b \in \mathcal{B}$ **do**
if $\tilde{v}_b(t-1) = 1$ **then**
$a_b(t) = a_b(t-1)$ {BS b is satisfied and it does not change its strategy.}
else
Select a strategy $a_b(t) \sim \boldsymbol{\pi}_b(t)$ {BS b is not satisfied, and it selects its strategy according to the probability distribution $\boldsymbol{\pi}_b(t)$.}
end if
Calculate $u_b(t)$, {using (12).}
if $u_b(t) \geq \Gamma_b(T_i)$ **then**
$\tilde{v}_b(t) = 1$ {BS b is satisfied.}
else
$\tilde{v}_b(t) = 0$ {BS b is not satisfied.}
for $\forall a_b \in \mathcal{A}_b$ **do**
Update $\pi_{b,a_b}(t+1)$, {using (25).}
end for
end if
end for
end while

located in the center of network. The set of SBSs and UEs are uniformly distributed in the coverage of MBS. The maximum transmit power of the MBS and the SBSs are 46 dBm and 30 dBm, respectively [66]. Some other factors which contribute to calculate the energy consumption of the BSs are chosen based on [67]. The channel is represented as a path loss fading according to [68]. The parameters used for the simulations are summarized in Table 1. Moreover, the presented results are averaged over a large number of independent runs (Monte Carlo simulations).

We consider an initial satisfaction threshold value ($\Gamma_b(T_0)$), for all BSs in the network. It is assumed that the BSs have a lower initial satisfaction threshold ($\Gamma_0' = -0.5$) or a higher initial satisfaction threshold ($\Gamma_0'' = -0.3$).

The proposed satisfaction based ON/OFF switching mechanisms are compared with the no-regret learning algorithm

TABLE 1. 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
ϑ_k		1800 Kbps
Number of UEs ($\lvert\mathcal{K}\rvert$)		50
θ_b		0.1
δ		1/8
Total number of iterations (T)		7000
Time interval between iterations		1 ms
T/N		100
Weights ω_b, ϕ_b		0.5, 0.5
Learning rate exponent for τ		0.9
Learning rate exponent κ, ζ, ν		0.6, 0.7, 0.8
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%
η_b^{PA}	31.1%	6.7%
P_b^{RF}	12.9 W	0.8 W
P_b^{BB}	29.6 W	3 W
$\lvert\mathcal{A}_b\rvert$	2	4
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 MBS-UE: 35 m	SBS-SBS: 40 m SBS-UE: 10 m

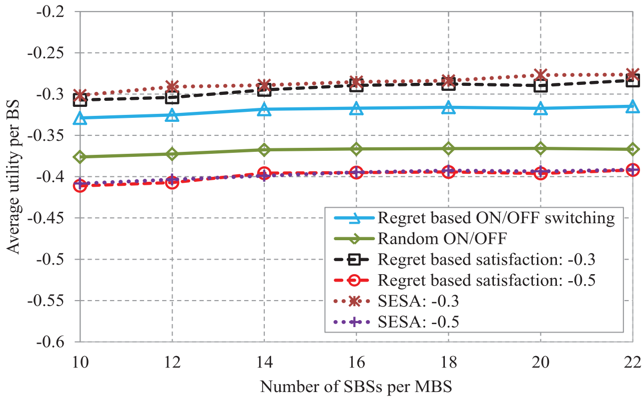

FIGURE 2. Average utility per BS versus the number of SBSs.

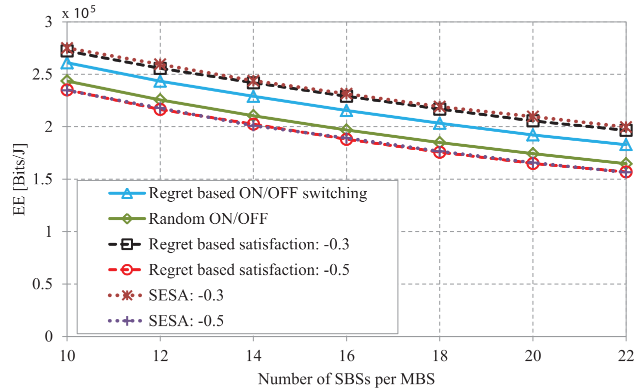

FIGURE 3. EE versus the number of SBSs.

without considering satisfaction for players, which is presented in [53], referred to hereinafter as "*regret based ON/OFF switching*". In this approach, the players utilize the learning algorithm described in (21)–(23), at each time instant (i.e. solving the game $\mathcal{G}_{\mathrm{NF}}$). For further comparisons, we consider a random BS ON/OFF switching approach, in which each player selects its strategy with equal probability (i.e. $\pi_{b,a_b}(t) = \frac{1}{|\mathcal{A}_b|}$ for $\forall b \in \mathcal{B}$ and $\forall a_b \in \mathcal{A}_b$). This approach is referred to as "*random ON/OFF*".

Fig. 2 shows the changes in average utility per BS as the number of SBSs varies. As the number of SBSs increases, average energy consumption per BS decreases, and the average utility per BS increases. Since the proposed approaches with initial threshold vector Γ_0'' save more energy compared to the other approaches, they improve the average utility per BS. For Γ_0', two proposed satisfaction approaches have almost the same performance, while SESA slightly improves average utility compared with the regret based satisfaction approach for Γ_0''. In this respect, the regret based satisfaction approach with initial threshold vector Γ_0'' yields, respectively, up to 13%, 16.5%, and 38% of utility improvement, relative to the regret based ON/OFF switching, random ON/OFF, and the satisfaction based approaches with initial threshold Γ_0', for a network with 16 SBSs.

We consider an energy efficiency (EE) metric defined as the ratio between average throughput per BS and average energy consumption per BS. Fig. 3 illustrates the EE versus different number of SBSs. We can see that the proposed approaches with initial threshold vector Γ_0'' improve the EE. For instance, when the number of SBS $= 22$, the regret based satisfaction approach improves the EE, respectively, up to 10.7%, 19.4%, and 44.9% as compared to the regret based ON/OFF switching, random ON/OFF, and the satisfaction approaches with initial threshold Γ_0'. This is due to the fact that the proposed approaches save more energy compared to the other approaches.

In General, it is more desirable to have a stable condition for the network, and avoid too many changes in the network settings. By reducing the number of switchings and power level changes, each receiver will experience a more stable channel condition and less change in the interference level. Hence, less effort is required for the receivers for adaptation to new channel conditions. The normalized amount of change in the power level can be shown by $\frac{|a_b(t)-a_b(t-1)|}{P_b^{\mathrm{Max}}}$. The total amount for this normalized change in power level for BS b is $\mathcal{N}_b \triangleq \sum_{t\in T} \frac{|a_b(t)-a_b(t-1)|}{P_b^{\mathrm{Max}}}$, and then it is averaged over the set of BSs ($\mathcal{N} \triangleq \frac{1}{|\mathcal{B}|}\sum_{b\in\mathcal{B}} \mathcal{N}_b$). Comparing Fig. 4 and Fig. 5, we observe that the proposed approaches

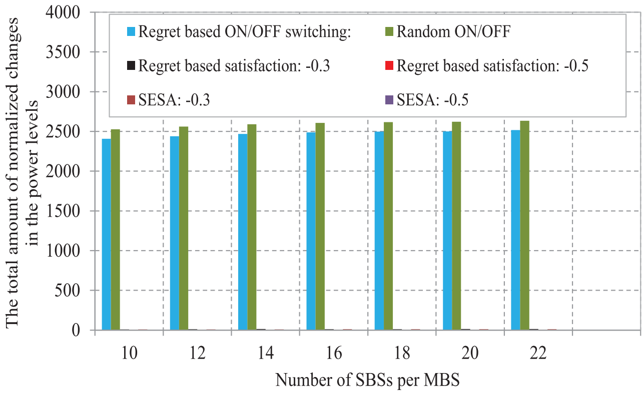

FIGURE 4. The total amount of normalized changes in the power levels.

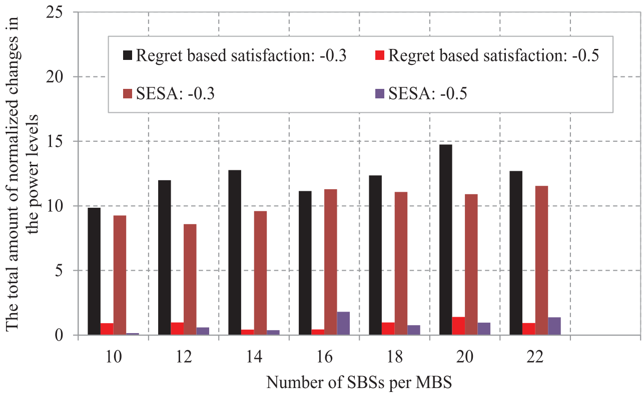

FIGURE 5. The total amount of normalized changes in the power levels.

significantly reduce $\mathcal{N}$. In Fig. 5, the regret based satisfaction approach with initial threshold vector Γ_0'' yields higher $\mathcal{N}$ comparing the other satisfaction approaches.

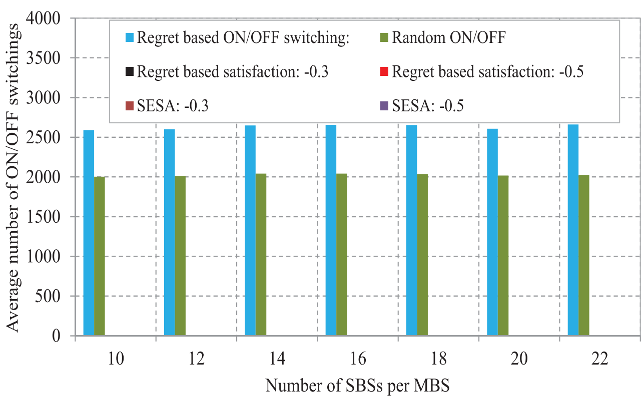

FIGURE 6. Average number of ON/OFF switchings during 7000 iterations.

Fig. 6 and Fig. 7 show the average number of switchings between an ON and OFF mode during 7000 iterations. In Fig. 6, we can observe that the regret based ON/OFF switching and random ON/OFF approaches switch frequently. Practically, such a frequent ON/OFF switching is not acceptable for BSs. In contrast, in the proposed satisfaction approaches, BSs switch less than the other approaches. The reason is that, in the proposed approaches, BSs tend to keep modes, unless they are not satisfied. Fig. 7 illustrates the average number of ON/OFF switchings for the proposed approaches. It can be seen that for Γ_0'', the regret based satisfaction approach has lower number of switchings compared to the SESA. Moreover, the satisfaction approach with higher threshold switches between ON and OFF modes more than the satisfaction approach with lower threshold. Therefore, the tradeoff between the energy saved from the ON/OFF switching and the number of switchings can be balanced by adjusting the threshold values.

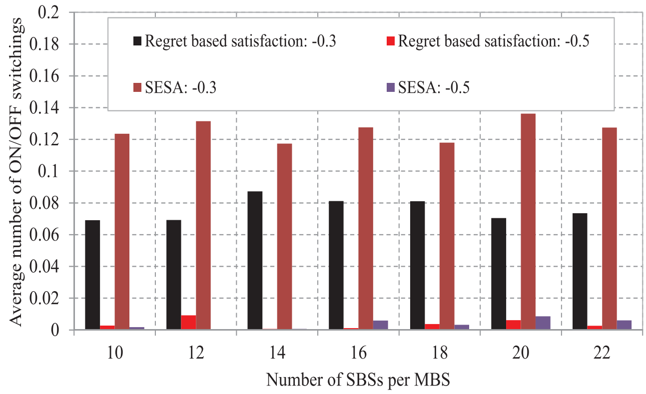

FIGURE 7. Average number of ON/OFF switchings for the proposed approach during 7000 iterations.

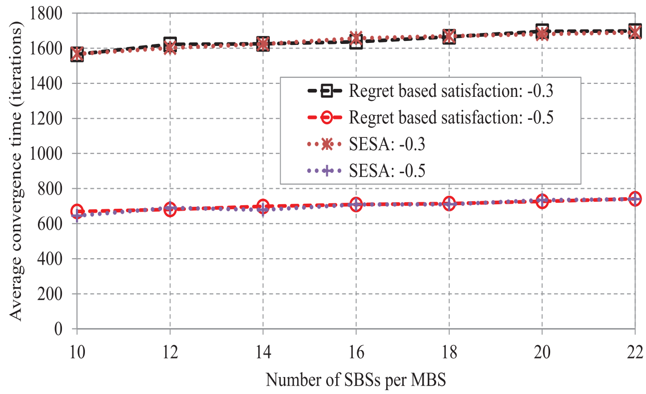

FIGURE 8. The convergence time of the proposed approach.

In Fig. 8, we show the convergence time of the proposed approaches to a satisfaction equilibrium versus the number of SBSs for $\Gamma_0' = -0.5$ and $\Gamma_0'' = -0.3$. We can see that with increasing the number of SBSs, the average number of iterations for convergence increases. For a given satisfaction threshold, two proposed approaches have almost the same convergence time. Moreover, Fig. 8 shows that reducing the satisfaction threshold leads to a faster convergence time. For instance, it shows that for regret based satisfaction approach, the average number of iterations for convergence for a network with 14 SBSs reaches up to about 1707 and 733 iterations with satisfaction threshold Γ_0'' and Γ_0', respectively.

V. CONCLUSION

In this paper, we have proposed two low-complexity BS ON/OFF switching mechanisms, i.e. regret based satisfaction algorithm and SESA, for HetNets, where BSs can choose their transmission power to save energy. In the SESA and the regret based satisfaction algorithm, BSs select each strategy based on the individual utilities and their regrets, respectively. The proposed approaches use the satisfaction based learning algorithms to avoid frequent BS switchings. In the proposed approaches, BSs are interested in achieving certain levels of satisfaction. By adjusting the threshold values, the tradeoff between the energy saved from the ON/OFF switching and the number of switchings is balanced. To operate in a practical environment, where simultaneously satisfying all players might not always be feasible, an option is provided to ensure convergence. This option allows the unsatisfied BSs redefine their satisfaction thresholds. The proposed approaches do not impose any signaling overhead, and thus can implement in a distributed way. Simulation results have shown that the proposed approaches provide better performance over other benchmark algorithms, and significantly outperform them in terms of the EE and average BS's utility. Furthermore, they significantly reduce the number of switchings at BSs.

APPENDIX
PROOF OF THEOREM 1

Proof: As [43], we consider the following hypothesis of game $\mathcal{G}_{\mathrm{SF}} = \langle \mathcal{B}, \{\mathcal{A}_b\}_{b\in\mathcal{B}}, \{f_b\}_{b\in\mathcal{B}}\rangle$:

- The game $\mathcal{G}_{\mathrm{SF}}$ has at least one equilibrium.
- For all $b \in \mathcal{B}$ the set f_b is not empty.
- The sets $\mathcal{B}$ and $\{\mathcal{A}_b\}_{b\in\mathcal{B}}$ are finite.

The first hypothesis ensures that the players are assigned a feasible task. The second hypothesis refers to the fact that each player is able to find a strategy to satisfy it, given the strategies of all other players. The third hypothesis is considered to ensure the convergence of the algorithm in finite time. Since $\pi_{b,a_b} > 0, \forall b \in \mathcal{B}, a_b \in \mathcal{A}_b, t > 1$, thus every strategy profile can be played at least once during sufficiently large iterations. According to (23), we have:

$$\pi_{b,a_b}(t) = \pi_{b,a_b}(t-1)\cdot(1-\frac{1}{t^{\nu}}) + (\frac{1}{t^{\nu}})\cdot G_{b,a_b}(\widehat{\boldsymbol{R}}_b(t)), \tag{28}$$

Since $0 < \frac{1}{t^{\nu}} \leq 1$ and $G_{b,a_b}(\widehat{\boldsymbol{R}}_b(t)) > 0$ for $\theta_b > 0$. Even if $\pi_{b,a_b}(0)$ be zero, then $\pi_{b,a_b}(t)$ at time $t > 1$ will be non-zero. Furthermore, after a large number of iterations, if a satisfaction equilibrium is not observed, then the target utility (satisfaction threshold) decreases. Thus, at least one strategy profile exists that satisfies the satisfaction thresholds, and no player changes its strategy. ■

ACKNOWLEDGEMENT

The authors would like to thank to Dr. Hamidou Tembine for fruitful discussions and comments.

REFERENCES

[1] K. Vasudeva, S. Dikmese, I. Güvenç, A. Mehbodniya, W. Saad, and F. Adachi, "Fuzzy-based game theoretic mobility management for energy efficient operation in hetnets," *IEEE Access*, vol. 5, pp. 7542–7552, 2017.
[2] A. R. Ekti, M. Z. Shakir, E. Serpedin, K. A. Qaraqe, and M. A. Imran, "On the traffic offloading in Wi-Fi supported heterogeneous wireless networks," *J. Signal Process. Syst.*, vol. 83, no. 2, pp. 225–240, 2016.
[3] M. Z. Shakir, H. Tabassum, and M.-S. Alouini, "Analytical bounds on the area spectral efficiency of uplink heterogeneous networks over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2306–2318, Jun. 2014.
[4] D. Liu, Y. Chen, K. K. Chai, T. Zhang, and M. Elkashlan, "Two-dimensional optimization on user association and green energy allocation for HetNets with hybrid energy sources," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4111–4124, Nov. 2015.
[5] D. Muirhead, M. A. Imran, and K. Arshad, "A survey of the challenges, opportunities and use of multiple antennas in current and future 5G small cell base stations," *IEEE Access*, vol. 4, pp. 2952–2964, 2016.
[6] R. Q. 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, May 2014.
[7] A. Abrol and R. K. Jha, "Power optimization in 5G networks: A step towards green communication," *IEEE Access*, vol. 4, pp. 1355–1374, 2016.
[8] F. Han, S. Zhao, L. Zhang, and J. Wu, "Survey of strategies for switching off base stations in heterogeneous networks for greener 5G systems," *IEEE Access*, vol. 4, pp. 4959–4973, 2016.
[9] M. Z. Shakir, K. A. Qaraqe, H. Tabassum, M.-S. Alouini, E. Serpedin, and M. A. Imran, "Green heterogeneous small-cell networks: Toward reducing the CO_2 emissions of mobile communications industry using uplink power adaptation," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 52–61, Jun. 2013.
[10] J. Xu, L. Duan, and R. Zhang, "Cost-aware green cellular networks with energy and communication cooperation," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 257–263, May 2015.
[11] J. Wu, Y. Zhang, M. Zukerman, and E. K. N. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 803–826, 2nd Quart., 2015.
[12] P. Chandhar and S. S. Das, "Multi-objective framework for dynamic optimization of OFDMA cellular systems," *IEEE Access*, vol. 4, pp. 1889–1914, 2016.
[13] C. Liu, B. Natarajan, and H. Xia, "Small cell base station sleep strategies for energy efficiency," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1652–1661, Mar. 2016.
[14] L. Tang, W. Wang, Y. Wang, and Q. Chen, "An energy-saving algorithm with joint user association, clustering, and on/off strategies in dense heterogeneous networks," *IEEE Access*, vol. 5, pp. 12988–13000, 2017.
[15] 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, Oct. 2015.
[16] 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, Dec. 2015.
[17] L. Jorguseski, A. Pais, F. Gunnarsson, A. Centonza, and C. Willcock, "Self-organizing networks in 3GPP: Standardization and future trends," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 28–34, Dec. 2014.
[18] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.
[19] S. Zhang, N. Zhang, S. Zhou, J. Gong, Z. Niu, and X. Shen, "Energy-aware traffic offloading for green heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1116–1129, May 2016.
[20] J. B. Rao and A. O. Fapojuwo, "A survey of energy efficient resource management techniques for multicell cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 154–180, 1st Quart., 2014.
[21] L. Budzisz *et al.*, "Dynamic resource provisioning for energy efficiency in wireless access networks: A survey and an outlook," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2259–2285, 4th Quart., 2014.
[22] C. Peng, S.-B. Lee, S. Lu, and H. Luo, "GreenBSN: Enabling energy-proportional cellular base station networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 11, pp. 2537–2551, Nov. 2014.
[23] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *Proc. IEEE ICCW*, Jun. 2009, pp. 1–5.
[24] J. Gong, S. Zhou, Z. Niu, and P. Yang, "Traffic-aware base station sleeping in dense cellular networks," in *Proc. IEEE 18th Int. Workshop Quality Service (IWQoS)*, Jun. 2010, pp. 1–2.

[25] H. Celebi, N. Maxemchuk, Y. Li, and I. Guvenc, "Energy reduction in small cell networks by a random on/off strategy," in *Proc. IEEE GC Workshops*, Dec. 2013, pp. 176–181.

[26] Y. Kang, S. Dong, and J. Zhang, "An energy-efficient user offloading algorithm for cellular network in base station sleep mode," in *Proc. 8th Int. Conf. WCSP*, Oct. 2016, pp. 1–4.

[27] H. Oh, S. H. S. Newaz, S. Park, and J. K. Choi, "Maximizing energy efficiency in off-peak hours: A novel sleep scheme for WLAN access points," in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2016, pp. 342–348.

[28] J. J. Q. Yu and V. O. K. Li, "Base station switching problem for green cellular networks with social spider algorithm," in *Proc. IEEE CEC*, Jul. 2014, pp. 2338–2344.

[29] Y. Yang, L. Chen, W. Dong, and W. Wang, "Active base station set optimization for minimal energy consumption in green cellular networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5340–5349, Nov. 2015.

[30] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1525–1536, Sep. 2011.

[31] R. Zeng, S. Zhu, H. Yang, and J. Zhu, "An artificial neural network based cell switch-off algorithm in cellular system," in *Proc. 2nd IEEE ICCC*, Oct. 2016, pp. 1434–1439.

[32] P. Ren and M. Tao, "A decentralized sleep mechanism in heterogeneous cellular networks with QoS constraints," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 509–512, Oct. 2014.

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

[34] T. Beitelmal and H. Yanikomeroglu, "A set cover based algorithm for cell switch-off with different cell sorting criteria," in *Proc. IEEE ICC*, Jun. 2014, pp. 641–646.

[35] G. P. Koudouridis, H. Gao, and P. Legg, "A centralised approach to power on-off optimisation for heterogeneous networks," in *Proc. IEEE VTC Fall*, Sep. 2012, pp. 1–5.

[36] A. Bousia, E. Kartsakli, A. Antonopoulos, L. Alonso, and C. Verikoukis, "Game theoretic approach for switching off base stations in multi-operator environments," in *Proc. IEEE ICC*, Jun. 2013, pp. 4420–4424.

[37] J. Zheng, Y. Cai, X. Chen, R. Li, and H. Zhang, "Optimal base station sleeping in green cellular networks: A distributed cooperative framework based on game theory," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4391–4406, Aug. 2015.

[38] A. H. Arani, A. Mehbodniya, M. J. Omidi, F. Adachi, W. Saad, and I. Güvenç, "Distributed learning for energy-efficient resource management in self-organizing heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9287–9303, Oct. 2017.

[39] M. Goonewardena, S. M. Perlaza, A. Yadav, and W. Ajib, "Generalized satisfaction equilibrium for service-level provisioning in wireless networks," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2427–2437, Jun. 2017.

[40] S. Ross and B. Chaib-Draa, "Satisfaction equilibrium: Achieving cooperation in incomplete information games," in *Proc. Conf. Can. Soc. Comput. Stud. Intell.*, 2006, pp. 61–72.

[41] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Satisfaction equilibrium: A general framework for QoS provisioning in self-configuring networks," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[42] S. Ross and B. Chaib-Draa, "Learning to play a satisfaction equilibrium," in *Proc. Workshop Evol. Models Collaboration*, 2007.

[43] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Quality-of-service provisioning in decentralized networks: A satisfaction equilibrium approach," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 104–116, Apr. 2012.

[44] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[45] W. Ikram, S. Petersen, P. Orten, and N. F. Thornhill, "Adaptive multi-channel transmission power control for industrial wireless instrumentation," *IEEE Trans Ind. Informat.*, vol. 10, no. 2, pp. 978–990, May 2014.

[46] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning wireless cellular networks of future: Outlook, challenges and opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.

[47] D. G. G, J. Hämäläinen, H. Yanikomeroglu, M. García-Lozano, and G. Senarath, "A novel multiobjective cell switch-off framework for cellular networks," *IEEE Access*, vol. 4, pp. 7883–7898, 2016.

[48] I. Siomina and D. Yuan, "Analysis of cell load coupling for LTE network planning and optimization," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2287–2297, Jun. 2012.

[49] A. J. Fehske and G. P. Fettweis, "Aggregation of variables in load models for interference-coupled cellular data networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 5102–5107.

[50] H. Klessig, A. Fehske, G. Fettweis, and J. Voigt, "Cell load-aware energy saving management in self-organizing networks," in *Proc. IEEE 78th VTC Fall*, Sep. 2013, pp. 1–6.

[51] C. K. Ho, D. Yuan, and S. Sun, "Data offloading in load coupled networks: A utility maximization framework," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1921–1931, Apr. 2014.

[52] R. Southwell, X. Chen, and J. Huang, "Quality of service games for spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 589–600, Mar. 2014.

[53] 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, Mar. 2016.

[54] R. L. G. Cavalcante, S. Stańczak, J. Zhang, and H. Zhuang, "Low complexity iterative algorithms for power estimation in ultra-dense load coupled networks," *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 6058–6070, Nov. 2016.

[55] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[56] N. Yu, Y. Miao, L. Mu, H. Du, H. Huang, and X. Jia, "Minimizing energy cost by dynamic switching on/off base stations in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7457–7469, Nov. 2016.

[57] A. Checko *et al.*, "Cloud RAN for mobile networks—A technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 1st Quart., 2015.

[58] M. Lin, A. Wierman, L. L. H. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1378–1391, Oct. 2013.

[59] G. Jie, Z. Sheng, and N. Zhisheng, "A dynamic programming approach for base station sleeping in cellular networks," *IEICE Trans. Commun.*, vol. 95, no. 2, pp. 551–562, Feb. 2012.

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

[61] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-Aho, "Opportunistic sleep mode strategies in wireless small cell networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 2707–2712.

[62] B. Ellingsæter, "Frequency allocation game in satisfaction form," *Trans. Emerg. Telecommun. Technol.*, vol. 25, no. 12, pp. 1238–1251, 2014.

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

[64] Y. Xing and R. Chandramouli, "Stochastic learning solution for distributed discrete power control game in wireless data networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 932–944, Aug. 2008.

[65] H. Tembine, *Distributed Strategic Learning for Wireless Engineers*. Boca Raton, FL, USA: CRC Press, 2012.

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

[67] 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*, Sep. 2012, pp. 312–317.

[68] *3GPP; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, document 3GPP TR 36.814 V9.0.0, 2010.

ATEFEH HAJIJAMALI ARANI received the M.Sc. degree in telecommunication engineering from Tarbiat Modares University, Tehran, Iran, in 2012. She is currently pursuing the Ph.D. degree with the Isfahan University of Technology, Isfahan, Iran. She was a Visiting Researcher with Tohoku University, Japan, in 2015. Her research interests include heterogeneous networks, resource management, game theory, and self-organizing networks.

MOHAMMAD JAVAD OMIDI received the Ph.D. degree from the University of Toronto in 1998. He was with the industry by joining the Research and Development Group designing broadband communication systems for five years in U.S. and Canada. In 2003, he joined the Department of Electrical and Computer Engineering, Isfahan University of Technology (IUT), Iran, and then served as the Chair of Information Technology Center and the Chair of the ECE Department. He is currently the Director of IRIS a UNESCO organization in Iran, and the Vice President for Research and Development with Isfahan Science and Technology Town. He is currently the supervising the Software Radio Laboratory, 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. He has authored over 15 U.S. and International patents in the area of telecommunications.

ABOLFAZL MEHBODNIYA (SM'16) has been an Assistant Professor with the Department of Communications Engineering, Tohoku University, since 2013. He has over 10 years of experience in electrical engineering, wireless communications, and project management. He has authored 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, and fuzzy logic techniques with applications to algorithm and protocol design in wireless communications. He was a recipient of the Japan Society for the Promotion of Science (JSPS) fellowship for foreign researchers, JSPS Young Faculty Startup Grant and KDDI Foundation Grant.

FUMIYUKI ADACHI (F'02–LF'16) is currently a specially-appointed Professor with the Research Organization of Electrical Communication, Tohoku University. His research interest lies in the area of wireless signal processing including wireless access, equalization, antenna diversity, adaptive transmission, and channel coding. He is an IEICE Fellow. He was a recipient of the IEEE VTS Avant Garde Award 2000, the 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, C&C Prize 2014, and the IEEE VTS Stuart Meyer Memorial Award 2017.