Learning-Based Small Cell Traffic Balancing Over Licensed and Unlicensed Bands

M. G. S. Sriyananda, *Member, IEEE*, and Mehdi Bennis, *Senior Member, IEEE*

Abstract—Unlicensed spectrum can be utilized by long term evolution (LTE) cellular systems to satisfy high throughput requirements. In this letter, a regret-based learning aided downlink traffic balancing scheme for licensed and unlicensed bands is proposed while ensuring fair coexistence of LTE-unlicensed (LTE-U) and Wi-Fi devices in the same band. It is further improved with the optimization of energy efficiency (EE) for small cell (SC) and macrocell scenarios followed by an inter-SC interference management mechanism with better performance over the existing literature. Compared to the cases with fixed airtime, up to 8%–10% superior results are shown for the scenarios of EE and rate maximization, respectively.

Index Terms—Energy efficiency, LTE-U, WiFi.

I. Introduction

A TIGHTER interaction between WiFi and Long Term Evolution (LTE) networks, has recently attracted much attention. In this regard, traffic balancing over the licensed band (LB) and unlicensed band (UB) with integrated WiFi and LTE interfaces [1], [2], handover of WiFi stations (STAs) served by LTE [3], carrier scheduling [4] and numerous lis-ten before talk (LBT) [5] schemes are being studied. The performance of traffic offloading and resource sharing are compared in [6]. Further, a time-domain multiplexing based UB inter-cell interference co-ordination mechanism for multi-operator small cells (SCs) is given in [7]. In contrast to those approaches, a regret based learning (RL) approach to man-age radio resources (power, spectrum and time) among STAs and LTE-Unlicensed (LTE-U) user equipments (UEs) in a SC scenario while guaranteeing a proper coexistence is studied in this letter. The approach is further extended to optimize energy efficiency (EE) at the macrocell (MC) while managing inter-SC interference. All WiFi users are served by the LTE system and as a performance measure the trade-off between EE and throughput is analyzed while clearly distinguishing the contribution of this letter from the existing literature.

The rest of this letter is organized as follows. System model with problem formulation, traffic balancing algorithms and simulation results are presented in Sections II–IV, respectively while concluding this letter in Section V.

II. System Model and Problem Formulation

A cellular layout consisting a MC and SCs are considered where the serving base stations (BSs) are represented by mBS and sBS, respectively. UEs served by mBS and sBS are named as mUE and sUE, accordingly. LTE, LTE-U and WiFi devices are served by the same sBS in both LB and UB [1], [2] where LTE traffic is handled in either bands by both sBS and sUE. Then SCs and sBSs are introduced as Integrated Femto-WiFi (IFW) cells and IFW BSs, respectively. To a sBS in a SC q, $q \in Q$, one sUE and a set of STAs $\mathcal{W}^q$ are connected where Q is the set of IFW [1], [2] cells in the MC. STAs are only served through a WiFi interface in UB and there is no separate wireless local area network for them. Only downlinks (DLs) for all the devices are considered [1] except for STAs. For any STA, the DL contends with uplink (UL) in a random fashion.

III. EE Through Traffic Balancing

The EE of SCs and MC is optimized using a traffic balancing scheme for UB of IFW cells [1]. Usage of the UB in IFW cell q is governed by the fractions of channel times allocated for the sUE and the set of STAs $\mathcal{W}^q$ denoted by t_s^q and t_W^q, respectively, where t_W^q is used for bidirectional traffic. Under full-buffer conditions, t_W^q is related to t_s^q, as $t_{\max} - t_s^q = t_W^q$. $t_{\max}$ is the maximum fraction of time that the UB can be used. $t_W^q \leq \bar{t}_W$ and $\bar{t}_W$ is used to set the maximum value for t_W^q. $\bar{t}_W$ is the aggregate load and the buffer status (e.g., full-buffer or not) of the STAs [1]. With optimal traffic balancing in a IFW cell, devices are in full-buffer status irrespective of their traffic loads. The transmit power of an sBS in SC q on the subcarrier k_L^q, $p_{L,\lambda_q}^{q,k}$, $k_L^q \in K_L^q$, $K_L^q = \{1, 2, \ldots, |\mathcal{K}_L^q|\}$, $\mathcal{K}_L^q \in \mathcal{K}_L$ is adjusted while that is a constant for the mBS with $k_L \in K_L$, $K_L = \{1, 2, \ldots, |\mathcal{K}_L^q|, \ldots, |\mathcal{K}_L|\}$. λ_q, $\lambda_q \in \Lambda$ is the step index for power levels for LB of sBS q with the values set Λ, $\Lambda = \{1, 2, \ldots, \lambda_q, \ldots, \lambda_{\max}\}$ and λ_q is an independent value for each subcarrier. For notational simplicity $\lambda_q^{k_L^q}$ and k_L^q are indexed as λ_q and k, accordingly. $\lambda_{\max}$ is the index for the maximum value. $\mathcal{K}_L$ and $\mathcal{K}_L^q$ are the sets of all and currently used LB subcarriers for a sUE, accordingly. $\gamma_L^{q,k}$ is the path loss of the desired signal divided by the interference plus noise power on subcarrier k in LB from sBS to sUE. Then the received signal-to-interference-plus-noise ratio (SINR) of sUE connected to the sBS q is given by $p_{L,\lambda_q}^{q,k}\gamma_L^{q,k}$. The interference on subcarrier k from an sBS in SC q to a mUE in the LB is kept below a maximum level $\bar{I}_{k,L}^q$. For the operation of sBS in the UB, a similar set of parameters are defined as $k_U^q \in K_U$, $K_U = \{1, 2, \ldots, |\mathcal{K}_U|\}$, $p_{U,a_q}^{q,k}$ and $\gamma_U^{q,k}$. a_q is the step index for power levels. For STAs, parameters are defined as $k_W^q \in K_W$, $K_W = \{1, 2, \ldots, |\mathcal{K}_W|\}$,

$p_{W,a_q}^{q,k}$ and $\gamma_{W,n}^{q,k}$, $n \in N_W^q$, $N_W^q = \{1, 2, \ldots, |\mathcal{W}^q|\}$. k is used for both k_U^q and k_W^q. $\gamma_U^{q,k}$ and $\gamma_{W,n}^{q,k}$ are due to inter SC interference and noise. P_U^q is the set of DL transmit power value for UB where $\sum_{k=1}^{|\mathcal{K}_U|} p_{U,a_q}^{q,k} = \sum_{k=1}^{|\mathcal{K}_W|} p_{W,a_q}^{q,k} = p_{a_q}^U$, $p_{U,a_q}^{q,k} = \frac{p_{a_q}^U}{|\mathcal{K}_U|}$, $\forall k \in K_U$, $p_{W,a_q}^{q,k} = \frac{p_{a_q}^U}{|\mathcal{K}_W^q|}$, $\forall k \in K_W$, and $p_{a_q}^U \in P_U^q$. $p_{a_q}^U$ is the total maximum allowed transmission power. For the simplicity, mobility is not considered.

For a IFW q, STAs in full-buffer and optimal traffic balancing without other UB users, the transmission time of STA n, $t_{W,n}^q$, as a fraction of t_W^q is given by $t_{W,n}^q = \alpha_n^q t_W^q$ [1], $n \in N_W^q$. α_n^q is the parameter deciding the fraction for the STA n. Then the weighted EE of STA n, $\mathcal{E}_{t,W,n}^q$ [1], is given as $\mathcal{E}_{t,W,n}^q = \mathcal{E}_{W,n}^q t_{W,n}^q = \mathcal{E}_{W,n}^q \alpha_n^q t_W^q$. The EE of STA n, $\mathcal{E}_{W,n}^q$, is defined as $\mathcal{E}_{W,n}^q = \sum_{k=1}^{|\mathcal{K}_W|} \mathcal{E}_U^q(p_{W,a_q}^{q,k}, \gamma_{W,n}^{q,k})$. t_s^q is used to control the EE of the sUE $\mathcal{E}_s^q$, $\mathcal{E}_s^q = \sum_{k=1}^{|\mathcal{K}_L^q|} \mathcal{E}_L^q(p_{L,\lambda_q}^{q,k}, \gamma_L^{q,k}) + t_s^q \sum_{k=1}^{|\mathcal{K}_U|} \mathcal{E}_U^q(p_{U,a_q}^{q,k}, \gamma_U^{q,k})$, where $\mathcal{E}_L^q(\cdot)$ and $\mathcal{E}_U^q(\cdot)$ are the LB and UB EE functions. Multiple modulation and coding schemes (MCSs) are used for both LTE and WiFi in which their MCSs are adapted to the instantaneous SINRs. The EE function $\mathcal{E}_X^q(p, y)$ for band X, $X \in \{L, U\}$, with transmit power p and SINR expression py is given by $\mathcal{E}_X^q(p, y) = \frac{B_X \log_2(1+py)}{p}$. y and B_X are the pathloss-to-interference-plus-noise ratio and the relevant bandwidth, respectively. LB and UB are represented by L and U. Similarly, the rate function $\mathcal{R}_X^q(p, y)$ is given by $\mathcal{R}_X^q(p, y) = B_X \log_2(1+py)$. As of $\mathcal{E}_{t,W,n}^q$ and $\mathcal{E}_s^q$ the rate functions for WiFi and LTE-U are given as $\mathcal{R}_{t,W,n}^q$ and $\mathcal{R}_s^q$.

A. Problem Formulation

In this letter, two optimization problems are studied for throughput capacity and EE. In the case of LTE and LTE-U, DL is considered where both DL and UL are used in WIFi.

1) Optimization of SC EE: The objective is to maximize the sum EE of a SC $q \in Q$ subject to a set of constrains as:

$$\text{maximize} \quad U_{\mathcal{E},\text{sum}}^q = U\left(\sum_{n=1}^{|\mathcal{W}^q|} \mathcal{E}_{t,W,n}^q\right) + U(\mathcal{E}_s^q) \quad (1)$$

$$\text{subject to} \quad t_s^q + t_W^q \leq t_{\max}, \quad (2)$$
$$t_W^q \leq \bar{t}_W, \quad (3)$$
$$p_{L,\lambda_q}^{q,k}\left|h_{L,m}^{q,k}\right|^2 \leq \bar{I}_{k,L}^q, \quad \forall k \in K_L^q, \; \lambda_q \in \Lambda \quad (4)$$
$$\sum_{k=1}^{|\mathcal{K}_L^q|} p_{L,\lambda_q}^{q,k} \leq P_{\max}^L, \quad \lambda_q \in \Lambda \quad (5)$$
$$\left\{t_s^q, t_W^q, P_L^{q,k}\right\} \geq 0, \quad \forall k \in K_L^q, \quad (6)$$

where $P_{\max}^L$ is the maximum total LB transmit power for a SC. It is assumed that the inter-cell interference channel gain from the sBS to the mUE m on the subcarrier k of LB $h_{L,m}^{q,k}$ is available for the sBS. The function $U(\cdot)$ is given as $U(\cdot) = \ln(\cdot)$ where $\ln(\cdot)$ is the natural logarithm function. Similar to (1), the overall user utility given by $U_{\mathcal{R},\text{sum}}^q$, $U_{\mathcal{R},\text{sum}}^q = U(\sum_{n=1}^{|\mathcal{W}^q|} \mathcal{R}_{t,W,n}^q) + U(\mathcal{R}_s^q)$, is used for the trade-off analysis without UB PC.

2) Optimization of Overall EE: The average and individual interferences to a sUE or a STA in a SC q', $\forall q' \in Q \setminus q$, from the sBS in q are measured as

$$\frac{P_{\max}^{q,U}(l')}{|\mathcal{W}^{q'}|+1}\left\{\frac{\left|h_s^{q,q'}\right|^2}{\mathcal{K}_L^q} + \frac{\sum_{n=1}^{|\mathcal{W}^{q'}|}\left|h_n^{q,q'}\right|^2}{\mathcal{K}_W}\right\} \leq \bar{I}_{\max}^{q,q'}, \quad (7)$$

$$\left\{\frac{P_{\max}^{q,U}(l')\left|h_s^{q,q'}\right|^2}{\mathcal{K}_L^q}, \frac{P_{\max}^{q,U}(l')\left|h_n^{q,q'}\right|^2}{\mathcal{K}_W}\right\} \leq I_{\max}^{q,q'}, \; n \in \mathcal{W}^{q'}, \quad (8)$$

accordingly. The maximum power value $P_{\max}^{q,U}(l')$ for the time instance l' of MC is decided by the mBS as $P_{\max}^{q,U}(l') = p_{v_q}^U$, $p_{v_q}^U \in P_U^q$, $P_U^q \subset P_U$, $v_q \in \mathcal{V}^q$. $P_U^q = \{p_1^U, p_2^U, \ldots, p_{v_q}^U\}$, $P_U = \{p_1^U, p_2^U, \ldots, p_{v_q}^U, \ldots, p_V^U\}$, $\mathcal{V}^q = \{1, 2, \ldots, v_q\}$, $\mathcal{V}^q \subset \mathcal{V}$, $\mathcal{V} = \{1, 2, \ldots, v_q, \ldots, V\}$ and $V = |\mathcal{V}|$. $P_{\max}^U = p_V^U$ is the maximum power allowed for the ULB. The set P_U^q is used for the PC in ULB in SC q. $h_s^{q,q'}$ and $h_n^{q,q'}$ are channel gains from sBS in q to sUE and STA n in q' on a mid subcarrier, respectively. $\bar{I}_{\max}^{q,q'}$ and $I_{\max}^{q,q'}$ are the maximum allowed total and individual interference power for the selected subcarrier from sBS in q to all the devices and any individual device in q' during an LTE-U transmission, accordingly. Average EE in the MC is maximized under a set of constrains as:

$$\text{maximize} \quad U_{\mathcal{E},\text{sum}} = \frac{\sum_{q \in Q} U_{\mathcal{E},\text{sum}}^q}{|Q|} \quad (9)$$

$$\text{subject to} \quad (7), (8), \; \forall q \in Q, \; \forall q' \in Q \setminus q$$
$$P_{\max}^{q,U}(l') \leq P_{\max}^U, \; \forall q \in Q, \quad (10)$$
$$\left\{P_{\max}^{q,U}(l'), P_{\max}^U\right\} \geq 0, \; \forall q \in Q. \quad (11)$$

Same way as in (9) and using rate functions, average utility $U_{\mathcal{R},\text{sum}}$, $U_{\mathcal{R},\text{sum}} = \sum_{q \in Q} U_{\mathcal{R},\text{sum}}^q$ is used for performance evaluation under full-buffer condition. In the absence of LBT style of channel access schemes, a power control mechanism like this is with a high importance to manage interference.

B. RL Based Approach for Traffic Balancing

The RL game $G^q = \{\mathcal{P}^q, \{\mathcal{A}_i^q\}_{\{i \in \mathcal{P}^q\}}, \{U_{\mathcal{E},\text{sum}}^{q,i}\}_{\{i \in \mathcal{P}^q\}}\}$ is developed as a solution for (1) and $\mathcal{P}^q = \{S^q, \mathcal{W}^q\}$, $|S^q| = 1$. $U_{\mathcal{E},\text{sum}}^{q,i}$ is the same as in (1). i, $i \in \mathcal{P}^q$ is the dominant player and each element of $\mathcal{P}^q$ is named for i in every other play. There are D, $d \in \mathcal{D}$, $\mathcal{D} = \{1, 2, \ldots, D\}$, compatible and suitable set of values $\mathcal{T}_C$, $\mathcal{T}_C = \{t_1, t_2, \ldots, t_D\}$, for the fractional time value for game round l, $t_d^{i,q}(l)$, $t_d^{i,q}(l) \in \mathcal{T}_C$ for a player i of LTE-U or WiFi. $\pi_q^i(l)$ is the probability value set for the elements in $\mathcal{T}_C$. The probability of t_d is selected for $t_d^{i,q}(l)$ is $\pi_d^{i,q}(l)$, $\pi_d^{i,q}(l) \in \pi_q^i(l)$, $\pi_q^i(l) = \{\pi_1^{i,q}(l), \pi_2^{i,q}(l), \ldots, \pi_D^{i,q}(l)\}$. $P_L^{q,k}$ is the transmit power levels set for the LB subcarriers of an sBS $P_L^{q,k} = \{p_{L,1}^{q,k}, p_{L,2}^{q,k}, \ldots, p_{L,\lambda_q}^{q,k}, \ldots, p_{L,\Lambda_q}^{q,k}\}$ with $p_{L,\Lambda_q}^{q,k} = P_{L,\max}^{q,k}$, $P_{L,\max}^{q,k} = \frac{P_{\max}^L}{|\mathcal{K}_L^q|}$. For (9), the RL game is given as $G = \{Q, \{\mathcal{A}^q\}_{\{q \in Q\}}, \{U_{\mathcal{E},\text{sum}}^{q,U}\}_{\{q \in Q\}}\}$. $P_{\max}^{q,U}(l')$, $P_{\max}^{q,U}(l') \in P_U$ are picked by q with L_{G^q} rounds of G^q for a G indexed

by l'. $\mu^q(l')$ is the probability value set for the elements in P_U. Probability for $p_{v_q}^U$ is given by $\mu_{v_q}^q(l')$, $\mu_{v_q}^q(l) \in \mu^q(l)$, $\mu^q(l') = \{\mu_1^q(l'), \mu_2^q(l'), \ldots, \mu_{v_q}^q(l'), \ldots, \mu_V^q(l')\}$.

- *Players:* In a SC q two user groups of the UB $\mathcal{P}^q$ are the players. A game round or an iteration l is played by player i, $i \in \mathcal{P}^q$. In the MC, IFWs Q are the players.
- *Action:* The action set $\mathcal{A}_i^q$ is defined for player i of q, as $\mathcal{A}_i^q = \{t_d^{i,q}(l), p_{a_q}^U, p_{L,\lambda_q}^{q,k}\}_{\{d \in \mathcal{D}, a_q \in \mathcal{V}^q, \forall k \in K_L^q\}}$, $p_{a_q}^U \in P_U^q$ and $\lambda_q \in \Lambda$. For the round l, player i is allowed to select his own fraction of transmission time $t_d^{i,q}(l)$. For the non-dominant player, the time portion $t_d^{j,q}(l)$, $t_d^{j,q}(l) \leq t_{\max} - t_d^{i,q}(l)$ is as in (2) where $j \in \mathcal{P}^q \setminus i$. $t_{\max}$ is a constant determined by the system. When the target EE is $U_{\mathcal{E},\text{tar}}^q$ and $U_{\mathcal{E},\text{sum}}^{q,i} > U_{\mathcal{E},\text{tar}}^q$, a_q is decreased by one ($a_q \leftarrow a_q - 1$) reducing $p_{a_q}^U$ by one level and vice versa. If (4) is not satisfied, λ_q is decreased by one ($\lambda_q \leftarrow \lambda_q - 1$) reducing $p_{L,\lambda_q}^{q,k}$ by a step and vice versa. The action set of player q, $\mathcal{A}^q$ is defined as $\mathcal{A}^q = \{p_{v_q}^U\}_{\{v_q \in \mathcal{V}\}}$ leading to $P_{\max}^{q,U}(l') = p_{v_q}^U$.

When $d = d'$ with $\pi_d^{q,i}(l)$, $\pi_{d'}^{q,i}(l+1)$ is given [8] as

$$\pi_{d'}^{q,i}(l+1) = \pi_{d'}^{q,i}(l) + \tau_i^q(l)b_i^q(l)\Big(\mathbb{1}_{\{d'=d\}} - \pi_{d'}^{q,i}(l)\Big)\mathbb{1}_{\{U_{\mathcal{E},\text{sum}}^{q,i} > U_{\mathcal{E},\text{tar}}^q\}}, \quad \forall d' \in \mathcal{D}, \quad (12)$$

where $\tau_i^q(l) = \frac{1}{l+1}$ and $b_i^q(l) = \frac{U_{\mathcal{E},\max}^q + U_{\mathcal{E},\text{sum}}^{q,i} - U_{\mathcal{E},\text{tar}}^q}{2U_{\mathcal{E},\text{tar}}^q}$. $\mathbb{1}_c$ is the indicator function with condition c. $t_d^{i,q}(l+1) = t_d$ is chosen with $d = \arg\max_{d' \in \mathcal{D}}(\pi_{d'}^{q,i}(l))$. In $\mathcal{G}$, if $U_{\mathcal{E},\text{sum}} > U_{\mathcal{E},\text{tar}}$, (7) and (8) are met, when $v_q = v_q'$ with $\mu_{v_q'}^q(l')$, then $\mu_{v_q'}^q(l'+1)$ is

$$\mu_{v_q'}^q(l'+1) = \mu_{v_q'}^q(l') + \tau^q(l)b^q(l)\Big(\mathbb{1}_{\{v_q'=v_q\}} - \mu_{v_q'}^q(l')\Big)\mathbb{1}_{\{U_{\mathcal{E},\text{sum}} > U_{\mathcal{E},\text{tar}}\}}, \quad \forall v_q' \in \mathcal{V}, \quad (13)$$

where $\tau^q(l') = \frac{1}{l'+1}$ and $b^q(l') = \frac{U_{\mathcal{E},\max} + U_{\mathcal{E},\text{sum}}^q - U_{\mathcal{E},\text{tar}}^q}{2U_{\mathcal{E},\text{tar}}}$. $U_{\mathcal{E},\max}^q$ and $U_{\mathcal{E},\max}$ are the maximum achievable EEs for (12) and (13), respectively. $U_{\mathcal{E},\text{tar}}$ is the target EE. Then, according to $v_q' = \arg\max_{v_q \in \mathcal{V}^q}(\mu_{v_q}^q(l'))$, $P_{\max}^{q,U}(l'+1) = p_{v_q}^U$ is selected. $L_{\mathcal{G}^q}$ is larger than the iterations needed for convergence of $\mathcal{G}^q$. EE and utility values after convergence are used for $\mathcal{G}$ as explained in Algorithm 1. RL belongs to a category in which agents jointly learn their utility while trying to minimize their regret for selecting a given action. Subsequently, a probability distribution over time is built. The interaction among small cells is cast as a strategic game whereby players (agents) minimize their expected utility over time by selecting judicious actions. The used RL algorithm is known to outperform best response algorithms and fictitious play which are based on beliefs, and converge to Nash (under some conditions). In contrast RL algorithms leverage correlation among players actions and converge to the socalled correlated equilibrium.

IV. SIMULATION RESULTS

A MC of radius 700 m with two SCs of radius 20 m situated at two random locations at 600 m away

---

Algorithm 1 Algorithm for EE Optimization

1: **Initialization;** $\mathcal{G}$:
2: Set $\mu_{v_q}^q(1) = \frac{1}{V}$, $\forall q \in Q$, $\forall v_q \in \mathcal{V}$. For P_U^q, set $v_q = V$, $p_{v_q}^U = p_V^U$, $P_{\max}^{q,U}(1) = p_V^U$, $\forall q \in Q$. Set $U_{\mathcal{E},\max}$ and $U_{\mathcal{E},\text{tar}}$.
3: **for** $l' = 1, 2, 3, \ldots, L_{\mathcal{G}}$ **do**
4:     **Initialization;** $\mathcal{G}^q$:
5:     Set $\pi_d^{q,i}(1) = \frac{1}{D}$, $U_{\mathcal{E},\max}^q$, $U_{\mathcal{E},\text{tar}}^q$, $\forall i \in \mathcal{P}^q$, $\forall d \in \mathcal{D}$, $\forall q \in Q$.
6:     **for** $l = 1, 2, 3, \ldots, L_{\mathcal{G}^q}$, $\forall q \in Q$ **do**
7:         Execute the action $\mathcal{A}_i^q$, update (12).
8:         **if** $U_{\mathcal{E},\text{sum}}^{q,i} > U_{\mathcal{E},\text{tar}}^q$ **then**
9:             Update $(a_q \leftarrow a_q - 1)$, $p_{a_q}^U$.
10:         **else**
11:             Update $(a_q \leftarrow a_q + 1)$, $p_{a_q}^U$.
12:         **end if**
13:         **if** (4) **then**
14:             Update $(\lambda_q \leftarrow \lambda_q + 1)$, $p_{L,\lambda_q}^{q,k}$, $\forall k \in K_L^q$.
15:         **else**
16:             Update $(\lambda_q \leftarrow \lambda_q - 1)$, $p_{L,\lambda_q}^{q,k}$, $\forall k \in K_L^q$.
17:         **end if**
18:         Select $t_d^{i,q}(l+1) = t_d$, $d = \arg\max_{d' \in \mathcal{D}}\Big(\pi_{d'}^{q,i}(l)\Big)$.
19:     **end for**
20:     Execute the action $\mathcal{A}^q$, update (13).
21:     **if** (7) and (8) **then**
22:         Select $v_q' = \arg\max_{v_q \in \mathcal{V}^q}\Big(\mu_{v_q}^q(l')\Big)$.
23:     **else**
24:         Update $\Big(v_q' \leftarrow v_q - 1\Big)$.
25:     **end if**
26:     Update $v_q \leftarrow v_q'$, $P_{\max}^{q,U}(l'+1) = p_{v_q}^U$, P_U^q.
27: **end for**

---

from the center or mBS is considered. Single sUE with STAs are in the IFW cell randomly and uniformly where there are no mUEs [1]. mUEs are also randomly and uniformly placed in the MC. Pathloss (PL) of [mBS↔UE] is given by $PL = 15.3 + 37.6\log_{10}(L)$. Intra SC PLs, $PL = 38.46 + 20\log_{10}(L) + 0.7(L)$ is for the links [sBS↔{sUE, STA}]. With $L_{\text{ow}} = 10$ dB, PLs for the links [mBS↔{sBS, sUE}, UE↔{sBS, sUE}] are given as $PL = 15.3 + 37.6\log_{10}(L) + L_{\text{ow}}$. Same model is used for inter SC losses of links [sBS↔{sUE, STA}] with outer wall penetration loss $L_{\text{ow}} = 20$ dB. PL is in dB and distance L is given in meters. The maximum total transmit powers of mBS and sBSs are set to 40 dBm and 15 dBm, accordingly. There are ten power levels for sBS PC and the minimum total power is 1.5 dBm. Noise power over 20 MHz bandwidth is -95 dBm. $\bar{I}_{\max}^{q,q'}$, $I_{\max}^{q,q'}$ and $\bar{I}_{k,L}^q$ are set as -100 dBm, -110 dBm and -100 dBm, respectively. The same LB to UB bandwidth ratio $\frac{R_L}{R_U}$ is set to all the SCs with the same UE and STA densities. For LTE and LTE-U, TDD-LTE [9] is considered and it is assumed that all the STAs, UEs are synchronized in both time and frequency domains with the serving BSs. All the SCs are synchronized with the MC. The maximum WiFi rate is 72 Mbps. t_W^q is in multiples of 10 ms and it is equally divided among STAs. The remainder is assigned randomly.

In Fig. 1, results are organized under two categories viz. Fig. 1(a) to Fig. 1(d) is for a selected SC and Fig. 1(e) to Fig. 1(h) is for the MC. Performance of Algorithm 1 executed in the SC and MC for two $\bar{t}_W$ values are shown in
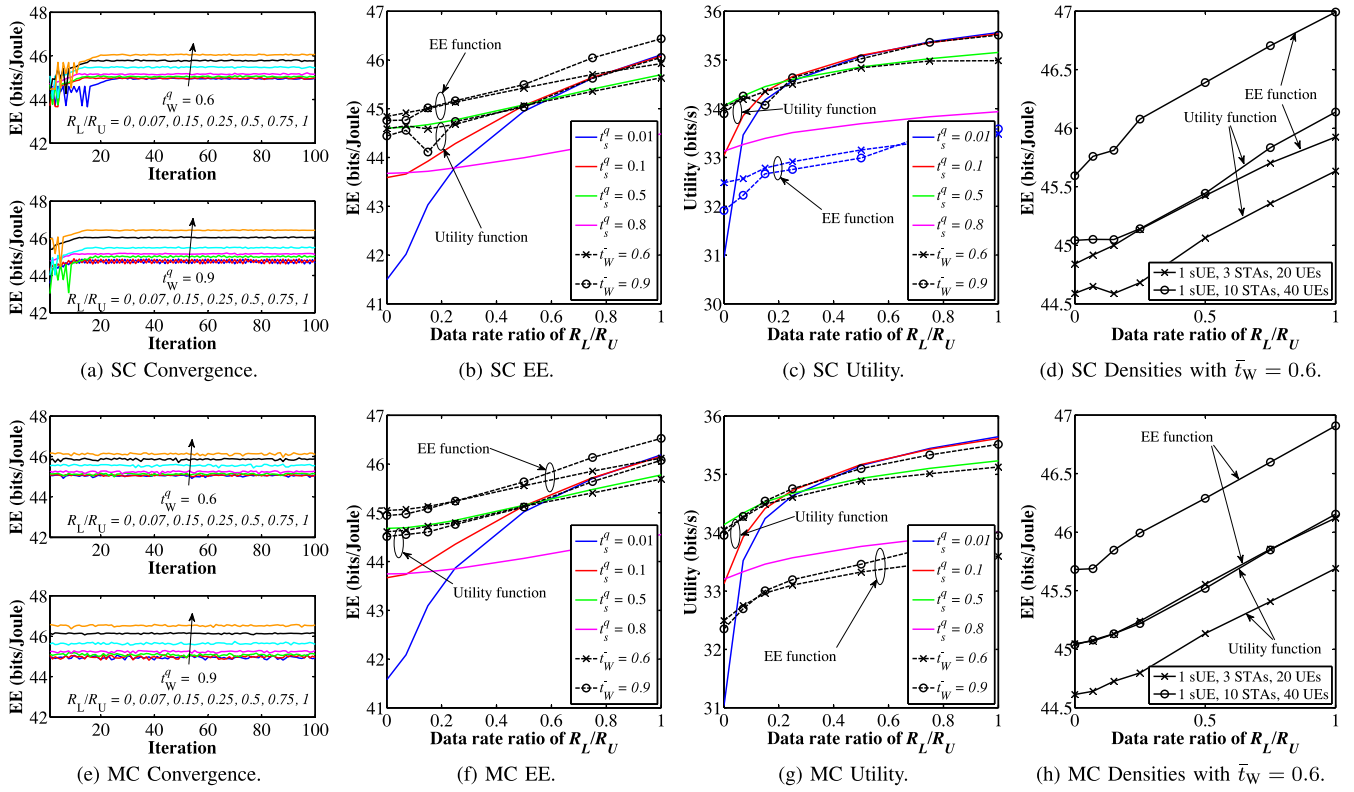
Fig. 1. Performance results for a SC ($q = 1$) and the MC.

Fig. 1(a) and Fig. 1(e), respectively. In both cases, the stability and convergence of the algorithms are clearly observed. If the players play according to the adaptive procedure of regret matching game, as time goes infinity, most probably the empirical distribution of the play converges to the set of correlated equilibrium distributions of the game [10]. Further, in the converged regions, high EE values are shown for the high $\frac{R_L}{R_U}$ values. EE performance for the SC and MC is compared in Fig. 1(b) and Fig. 1(f) using functions $U_{\mathcal{E},\text{sum}}^q$ and $U_{\mathcal{E},\text{sum}}$ in (1) and (9) as the objective functions, accordingly. Up to 1% - 10% performance increase is shown compared to the scenarios with different fixed t_s^q values. Corresponding EE results, when $U_{\mathcal{R},\text{sum}}^q$ and $U_{\mathcal{R},\text{sum}}$ are used as the objective functions, are also included and compared. In Fig. 1(c) and Fig. 1(g), $U_{\mathcal{R},\text{sum}}^q$ and $U_{\mathcal{R},\text{sum}}$ are used as the objective functions and up to 9% better results could be reached against certain cases of fixed t_s^q values. Plots for the utility values when $U_{\mathcal{E},\text{sum}}^q$ and $U_{\mathcal{E},\text{sum}}$ are used as the objective function are also included in the same figures. In Fig. 1(d) and Fig. 1(h), EE outcomes with different STA and UE densities are shown. EE values are also given when the objective functions are $U_{\mathcal{R},\text{sum}}^q$ and $U_{\mathcal{R},\text{sum}}$. Due to effective use of interference management mechanism, better performance is shown for both EE and utility maximization functions for high densities.

V. CONCLUSION

In this letter, principles of RL were used to optimize EE of each IFW cell and the MC while guaranteeing coexistence of STAs and sUEs. The concept of traffic offloading and balancing based on transmission times between sUEs and STAs was used for that against ratios of LB to UB. An inter SC

interference management criteria was also integrated for the optimization formulas and RL algorithm. EE versus utility trade-off was studied with different parameter settings and it was shown our proposed solution outperforms not only for EE but also for individual SC and overall utility.

REFERENCES

[1] F. Liu, E. Bala, E. Erkip, M. C. Beluri, and R. Yang, "Small-cell traffic balancing over licensed and unlicensed bands," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5850–5865, Dec. 2015.

[2] Q. Chen, G. Yu, and Z. Ding, "Optimizing unlicensed spectrum sharing for LTE-U and WiFi network coexistence," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2562–2574, Oct. 2016.

[3] Q. Chen, G. Yu, A. Maaref, G. Y. Li, and A. Huang, "Rethinking mobile data offloading for LTE in unlicensed spectrum," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4987–5000, Jul. 2016.

[4] Z. Luo, M. Ding, and H. Luo, "CC on/off scheduling using learning-based prediction for LTE in the unlicensed spectrum," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2158–2161, Dec. 2015.

[5] R. Yin, G. Yu, A. Maaref, and G. Ye Li, "LBT-based adaptive channel access for LTE-U systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6585–6597, Oct. 2016.

[6] Q. Chen *et al.*, "Cellular meets WiFi: Traffic offloading or resource sharing?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3354–3367, May 2016.

[7] M. R. Khawer, J. Tang, and F. Han, "usICIC—A proactive small cell interference mitigation strategy for improving spectral efficiency of LTE networks in the unlicensed spectrum," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2303–2311, Mar. 2015.

[8] M. Simsek, M. Bennis, and İ. 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.

[9] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. West Sussex, U.K.: Wiley, 2011.

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