Efficient LTE/WiFi Coexistence in Unlicensed Spectrum using Virtual Network Entity: Optimization and Performance Analysis

Atoosa Dalili Shoaei*, Mahsa Derakhshani†, Tho Le-Ngoc*

*Department of Electrical & Computer Engineering, McGill University, Montreal, QC, Canada
† Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, UK
Email: atoosa.dalilishoaei@mail.mcgill.ca; m.derakhshani@lboro.ac.uk; tho.le-ngoc@mcgill.ca

Abstract—Long term evolution (LTE) operation in the unlicensed spectrum is a promising solution to address the scarcity of licensed spectrum for cellular networks. Although this approach brings higher capacity for LTE networks, the WiFi performance operating in this band can be significantly degraded. To address this issue, we consider a coordinated structure, in which both networks are controlled by a higher-level network entity. In such a model, LTE users can transmit in the assigned time-slots, while WiFi users can compete with each other by using p-persistent CSMA in their exclusive time-share. In an unsaturated network, at each duty cycle, the TDMA scheduling for LTE users and p values for WiFi users should be efficiently updated by the central controller. The corresponding optimization problem is formulated and an iterative algorithm is developed to find the optimal solution using complementary geometric programming (CGP) and monomial approximations. Aiming to address Quality-of-Service (QoS) assurance for LTE users, an upper bound for average delay of these users are obtained. This analysis could be a basis for admission control of LTE users in unlicensed bands. The simulation results reveal the performance gains of the proposed algorithm in preserving the WiFi throughput requirement.

I. INTRODUCTION

A. Background & Motivation

LTE operation on unlicensed bands, so-called unlicensed LTE (U-LTE), is considered by third-generation partnership project (3GPP) as a promising solution to meet the growing wireless data demand and to improve the spectrum efficiency. Although transmission across both unlicensed and licensed bands can boost LTE, such an approach may jeopardize the performance in WiFi systems solely operating on unlicensed bands for data transmission. The reason is that LTE networks exploit a schedule-based channel access, while in WiFi a contention-based scheme is applied, in which the user would randomly access the channel once it is detected idle. Therefore, in a coexistence scenario that both systems share the same channel, starvation may happen for WiFi as the whole airtime may be occupied by the LTE network [1]–[6].

In order to address this issue, two approaches have been so far proposed, LTE-Unlicensed (LTE-U) and licensed-assisted access (LAA). In LTE-U, developed in 3GPP Releases 10/11/12, a duty-cycle-based approach is used in which at each duty cycle, LTE transmits over only a portion of a duty cycle, securing the rest of cycle for WiFi [7]. The main problem of this approach stems from no carrier sensing before LTE

transmissions, as WiFi transmissions occurring in the LTE cycle might be interrupted by LTE transmissions. On the other hand, in LAA which has been featured in 3GPP Release 13, the LTE base station (BS) is equipped with the listen-before-talk (LBT) mechanism, i.e., carrier sensing is performed before any transmission [8]–[10]. Although this approach may lead to enhanced performance for WiFi compared to the LTE-U, the utilization still cannot reach to the optimal point due to the lack of coordination between two networks. Thus, an efficient structure along with a proper MAC needs to be designed.

B. Related Works

In the literature, the LTE and WiFi coexistence in unlicensed spectrum has been mostly studied where LTE employs this band for *downlink scenarios* [10]–[17].

Assuming *saturated* LTE and WiFi networks, in [11], an analytical work is presented to provide proportional fairness between two networks. To reach this goal, it is assumed that LTE transmits with probability of q for a burst transmission and this probability is optimized such that the proportional fairness can be achieved. The work in [12] also uses the LBT-based method in which the aim is to guarantee the required rate of LTE users while collision probability of Wi-Fi users is minimized. In [10], assuming that the number of WiFi users is unknown, a further analysis is presented to estimate the population of the WiFi system. Moreover, in [15], [18], the fair coexistence problem is addressed for multi-cell scenario comprised of multiple WiFi access points (APs) and LTE BSs.

The authors of [13] investigate the fair airtime allocation for both scenarios of LTE-U/WiFi coexistence and Integrated Femto-WiFi (IFW) small cell in which LTE users can access the unlicensed band via WiFi interface. While the work in [19] considers different scenario called hybrid scheme in which traffic offloading and resource sharing are used simultaneously by the LTE.

In [14], a hyper access point is introduced which is able to operate as both LTE base station and WiFi access point. However, this work also considers saturated scenarios.

Under assumption of *unsaturated* LTE and WiFi networks, in [16], an optimization problem is studied with the aim of maximizing overall throughput, while maintaining WiFi throughput. In [17], maximum allowable packet arrival rates

of both LTE and WiFi are derived under which WiFi delay requirement is guaranteed. However, in these proposed schemes, it is required that the LTE BS performs the LBT operation before transmissions. Furthermore, they only consider the scenarios that LTE uses the unlicensed band for the downlink traffic.

For *uplink scenarios*, [20] proposes an approach in which LTE users need to perform clear channel assessment (CCA) before transmission on the unlicensed band. In this approach, each user requested a resource batch is assigned multiple resource batches. The user performs CCA over all assigned resource batches and transmits its data over the one which is detected as idle. The number of resource batches assigned to each user is optimized in this work. In [21], it is assumed that LTE users can access the channel through both scheduled and random schemes. This problem is formulated by a game that divides users to the two groups: one with users employing the scheduled scheme and the other with users accessing the channel in a random manner.

There are few works that address uplink resource allocation for U-LTE [20], [21]. However, in these works, to access the channel, LTE users should perform channel assessment, which brings modifications on the current deployed LTE scheme. Moreover, none of these works analyzes delay performance for LTE users. Such analysis is indispensable to offer satisfactory experience for LTE users.

C. Contributions

To address the aforementioned challenges, in this paper [1] we propose a coordinated structure for coexistence of these two networks over the unlicensed spectrum [23]. In such a coordinated model, the control of spectrum access between the two systems are governed by virtualized network entities and higher-level management. The benefits gained from this architecture are three folds. First, the unlicensed spectrum utilization can be improved since the interference between the U-LTE and WiFi systems are avoided. Second, no LBT operation is required for U-LTE system. Third, as the central entity has the authorities to control over both systems, WiFi throughput requirement can be kept satisfied.

More specifically, in the proposed model, the central network entity facilitates separation of LTE and WiFi transmissions in two different phases. With no major MAC modification required, LTE users can access the channel in the TDMA manner, while in the second phase, WiFi users can opportunistically transmit their packets using p-persistent CSMA. Moreover, such network entity can enable dynamic scheduling by assigning time-slots to LTE users and adjusting p for WiFi users. This can improve the network throughput and preserve the WiFi throughput requirement. Using such approach, the unlicensed band can be used for both uplink and downlink transmissions.

In this work, assuming an unsaturated network for both LTE and WiFi systems, the goal is to maximize the overall throughput over each duty cycle, while the WiFi throughput does not fall below a target threshold. In other words, this scheme acts like a duty-cycle-based approach, where a period of a time-frame with a variable length is assigned exclusively to the LTE system, which cannot be used by WiFi users. In order to solve this optimization problem, we formulate it by complementary geometric programming, which can be solved by applying an iterative algorithm.

Furthermore, delay analysis is performed for LTE users to provide an analytical framework that can be used for designing the LTE admission control policy, e.g., a specific number of users that can be supported, while meeting their QoS constraints. We consider a system of homogeneous LTE users, in which all users have the same packet arrival probabilities. For this system, we derive an analytical upper-bound of average delay by modeling the time-slot assignment to the user with a queuing system which consists of two servers. Consequently, the probability mass function of packet delay is derived and average packet delay is calculated.

D. Organization

The rest of this paper is organized as follows. We first introduce the system model in Section II. Section III presents the problem formulation and transformation of the optimization problem into CGP. The performance analysis for LTE users in terms of average packet delay is presented in Section IV. Furthermore, Section V presents the simulation results. Finally, we provide some concluding remarks in Section VI.

II. SYSTEM MODEL

A. Coordinated Structure for LTE/WiFi Coexistence

We consider an IEEE 802.11-based WLAN with N_W users sharing the channel with an U-LTE network serving N_L users. In order to enable efficient coexistence between these two networks, we assume that there is a virtual network entity for software-defined wireless networking, which can control both the WiFi AP and the U-LTE BS in a central manner, where users from different wireless technologies form different slices: LTE slice and WiFi slice.

In the virtualized scheme, users of different systems access through their distributed access points which are controlled by a central entity. More specifically, this architecture supports separating the data plane from the control plane. The control plane responsible for managing resources is centralized while separated data planes are considered, i.e., data are being forwarded independently from LTE BS and WiFi AP [23].

We use the duty-cycle-based approach in order to access the channel. In this approach, at each duty cycle t, the time-frame is divided into two phases. In the first phase, LTE users can access the channel in the TDMA manner, while in the second phase, WiFi users can opportunistically transmit their packets using p-persistent CSMA. To realize this model, we deploy point coordination function (PCF) mechanism for the WiFi. In this mechanism, each superframe consists of a contention-free period (CFP) followed by a contention period (CP). Thus, by

activating the PCF, the CFP can be used by LTE users, while the CP can be assigned to the WiFi users [24].

Each duty cycle has a fixed duration of T. Although the length of each duty cycle is fixed, the length of each phase assigned to WiFi or U-LTE network varies over different cycles. The duration of U-LTE transmission is denoted by $C(t)$ and $T - C(t)$ that indicates the length of second phase for WiFi transmissions. In this setting, the smallest unit of time is called a backoff unit and each time-slot with duration of T_s is divided into backoff units.

This architecture allows each network to use its current deployed MAC protocol, i.e., LTE users can transmit in TDMA phase, while WiFi users compete with each other in the CSMA phase. Thus, the model leads to minimal modification requirement for both systems. It also eliminates the LBT requirement for LTE network. Specifically, it facilitates uplink transmissions for LTE network, since in the uplink scenario, the user has to perform the LBT operation while in downlink only BS listens to the channel before transmission. However, as in this approach, users of different networks access in separated time, unlicensed band can be also used for LTE uplink transmissions with no need for LBT operation. To emphasize the advantage of employing this structure, consider a scenario using a non-LBT approach without any coordination among LTE and WiFi systems, where at the beginning of each duty cycle, the LTE base station schedules the users for the uplink transmission. If one or some of these scheduled LTE users have no packets for transmission considering an unsaturated network, the assigned time-slot will be left empty and during that time the WiFi user will find the channel as idle and it may transmit a packet. If the packet transmission continues until the next time-slot and the next scheduled LTE user has a packet for transmission, collision will happen leading to performance degradation. Furthermore, having the centralized control, it brings the benefit of spectrum efficiency while meeting the WiFi requirements.

B. Traffic and Channel Model

We assume that each user n regardless of its network, has a queue with maximum length of Q_{max}. At each duty cycle, with probability of a_n, a new packet is added to the queue of user n if its length is smaller than Q_{max}. Otherwise, the packet is discarded. Furthermore, we assume that the central controller is aware of packet arrival probabilities of users and it keeps a vector denoted by $\boldsymbol{V}(t)$, where its component $v_n(t)$ denotes the time that it has received the last packet from user n. Moreover, each time the user sends a packet, it piggybacks an extra bit (denoted by $q_n(t)$) telling if its queue is empty ($q_n(t) = 0$) or non-empty ($q_n(t) = 1$, i.e., it has packets backlogged in the queue to transmit). Therefore, at the duty cycle t, the controller updates $\theta_n(t)$, which indicates the probability that user n has a non-empty queue at t

$$\theta_n(t) = \begin{cases} 1 - (1 - a_n)^{t - v_n(t)}, & \text{if } q_n(v_n(t)) = 0 \\ 1 & \text{if } q_n(v_n(t)) = 1. \end{cases} \quad (1)$$
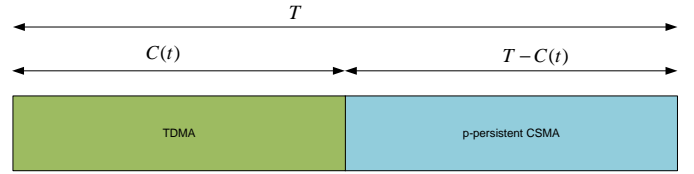

Fig. 1: Duty-cycle-based structure for WiFi and U-LTE coexistence

Regarding the wireless channel model, we consider path loss and small scale fading. The instantaneous received SNR of user n at BS/AP is equal to $GP_t h_n d_n^{-\xi}/\sigma^2$, where P_t is the transmission power, σ^2 is the noise power, h_n is the small-scale Rayleigh fading component of the link from the user n to the BS/AP, d_n is the link distance between user n and the BS/AP, ξ is the path-loss exponent, and G is a constant dependent on the frequency and transmitter/receiver antenna gain. For simplicity, without loss of generality, we normalize G to 1 (i.e., $G = 1$) in the following discussions. If the received signal level falls below the receiver threshold, the receiver cannot successfully decode the signal denoting an *outage* event. Thus, the system performance is influenced by the outage probability defined as the probability that the received SNR is less than the receiver threshold v,

$$\psi_n = \Pr\left(\frac{P_t h_n d_n^{-\xi}}{\sigma^2} \leq v\right) = 1 - e^{-\frac{\sigma^2 d_n^{-\xi} v}{P_t}}. \quad (2)$$

C. Analytical Model for p-persistent CSMA

The WiFi system operates on the CSMA protocol. Using a p-persistent CSMA, a WiFi user with a non-empty queue performs the channel sensing. If the channel is detected busy, the WiFi user n_w with a non-empty queue waits until channel becomes idle, and then transmits the packet with probability p_{n_w}. In this subsection, we model the throughput of p-persistent CSMA protocol in an unsaturated network.

In order to calculate the throughput, we define P_{idle}, which indicates the probability that channel is idle in a backoff unit. This probability is calculated as

$$P_{idle} = \prod_{n_w \in \mathcal{N}_W} (1 - \theta_{n_w} p_{n_w}), \quad (3)$$

where $\theta_{n_w} p_{n_w}$ represents the transmission probability of user n_w. Furthermore, we calculate the probability of *successful* transmission which happens if only one user transmits on the channel. This probability for transmission initiated by user n_w (denoted by $P_{succ}^{n_w}$) can be obtained as

$$P_{succ}^{n_w} = (1 - \psi_{n_w}) \theta_{n_w} p_{n_w} \prod_{n'_w \in \mathcal{N}_W, n'_w \neq n_w} (1 - \theta_{n'_w} p_{n'_w}). \quad (4)$$

As introduced in [25], we define the normalized throughput of user n_w (denoted by ρ_{n_w}) as the fraction of time the channel is used for its successful transmission. Accordingly, ρ_{n_w} can

be written as

$$\rho_{n_w} = \frac{P_{\text{succ}}^{n_w} T_{\text{succ}}}{P_{\text{idle}}\delta + (1 - P_{\text{idle}})T_{\text{succ}}}, \tag{5}$$

where δ is the duration of a backoff unit and T_{succ} is the duration of a successful transmission, which includes the data transmission for a fixed time, inter-frame spaces, and signaling overheads. Since signaling and inter-frame spaces are relatively small (in the order of μs) compared with data transmission length (in the order of ms), we approximately assume that both collided and successful transmissions are of the same size (i.e., T_{succ}). Furthermore, we assume that each transmission occupies 1 time-slot, i.e., $T_{\text{succ}} = T_s$. Consequently, the denominator in (5) represents the expected length of a general time-slot.

To simplify (5), we define a new variable z_{n_w} as,

$$z_{n_w} = \frac{\theta_{n_w} p_{n_w}}{1 - \theta_{n_w} p_{n_w}}. \tag{6}$$

By applying the substitution introduced in (6) in (3) and (4), P_{idle} and $P_{\text{succ}}^{n_w}$ can be written as

$$P_{\text{idle}} = \frac{1}{\prod_{n_w \in \mathcal{N}_w}(1 + z_{n_w})} \tag{7}$$

$$P_{\text{succ}}^{n_w} = \frac{z_{n_w}}{\prod_{n_w \in \mathcal{N}_w}(1 + z_{n_w})} = z_{n_w} P_{\text{idle}}. \tag{8}$$

Finally, ρ_{n_w} can be obtained in terms of z_{n_w} as

$$\rho_{n_w} = \frac{z_{n_w}}{\prod_{n_w \in \mathcal{N}_w}(1 + z_{n_w}) - t'}, \tag{9}$$

where $t' = \frac{T_s - \delta}{T_s}$.

III. HYBRID TDMA-CSMA SCHEDULING VIA CGP FOR WIFI AND LTE COEXISTENCE

A. Problem Formulation

Consider a system consisting of an LTE BS and a WiFi AP, which are connected to a virtual network entity. To facilitate the coexistence between WiFi and LTE and increase the spectral efficiency, the central controller dynamically divides each duty cycle between two slices. This would be done in a way that the overall throughput is maximized, while the throughput of WiFi users does not degrade significantly compared to the case in which the band is not shared with LTE users. In particular, this problem can be formulated as follows [2],

$$\max_{\boldsymbol{X},\boldsymbol{Z}} \quad S_{\text{td}}(t) + S_{\text{cs}}(t), \quad \text{subject to,} \tag{10}$$

C10.1: $\quad S_{\text{cs}}(t) \geq \eta$

C10.2: $\quad \sum_{n_l \in \mathcal{N}_L} x_{n_l} \geq C_{\min}.$

where $S_{\text{td}}(t)$ and $S_{\text{cs}}(t)$ represent LTE-U and WiFi throughput at duty cycle t, respectively. $\boldsymbol{X} = [x_{n_l}]$ is the vector indicating

[2]The results can be easily extended for downlink scenario of WiFi as well since the AP acts as a station in downlink and accesses the channel in the same manner

the time-slot allocation for LTE users in one duty cycle. In particular, $x_{n_l} \in \{0, 1\}$, where $x_{n_l} = 1$ if a time-slot is allocated to the LTE user n_l in the TDMA phase and $x_{n_l} = 0$ otherwise. Furthermore, $\boldsymbol{Z} = [z_{n_w}]$ is the vector, where the element z_{n_w} is defined in (6). In this optimization problem, the objective function represents the total throughput of network in both TDMA and CSMA phases in the duty cycle t, where the throughput in the TDMA and CSMA phases represents the throughput of the U-LTE and WiFi systems, respectively. In addition, the first constraint is to guarantee that WiFi throughput does not fall below a required threshold (denoted by η). Furthermore, to keep LTE users satisfied, we add the second constraint to reserve at least $C_{\min}$ time-slots for LTE users. It should be noted that the values of η and $C_{\min}$ are fixed dictated based on the service level agreements of WiFi and LTE-U systems.

Considering the throughput of each WiFi user in (9), the optimization problem in (10) can be expanded as

$$\max_{\boldsymbol{X},\boldsymbol{Z}} \sum_{n_l \in \mathcal{N}_L}(1 - \psi_{n_l})\theta_{n_l} x_{n_l} + $$
$$\sum_{n_w \in \mathcal{N}_W} \frac{(1 - \psi_{n_w})z_{n_w}(T - T_s \sum_{n_l \in \mathcal{N}_L} x_{n_l})}{\prod_{n_w \in \mathcal{N}_W}(1 + z_{n_w}) - t'} \tag{11}$$

subject to:

C11.1: $\sum_{n_w \in \mathcal{N}_W} \dfrac{(1 - \psi_{n_w})z_{n_w}(T - T_s \sum_{n_l \in \mathcal{N}_L} x_{n_l})}{\prod_{n_w \in \mathcal{N}_W}(1 + z_{n_w}) - t'} \geq \eta$

C11.2: $\sum_{n_l \in \mathcal{N}_L} x_{n_l} \geq C_{\min}$

In the objective function of (11), the first term represents the expected throughput associated with U-LTE users in the TDMA phase which is $S_{\text{td}} = \sum_{n_l \in \mathcal{N}_L} \theta_{n_l} x_{n_l}$. Furthermore, the second term denotes the entire WiFi network throughput, i.e., $S_{\text{cs}} = T_{\text{cs}} \sum_{n_w \in \mathcal{N}_W} \rho_{n_w}$, where T_{cs} denotes the duration of CSMA phase in a duty cycle.

Before solving this optimization problem, we provide a discussion on its feasibility. To determine the feasibility of this optimization problem, we obtain the feasible region for C, which indicates the length of TDMA phase in a duty cycle and is the common parameter in the conflicting constraints of C11.1 and C11.2. According to C11.2, we constrain $C \geq C_{\min}$. Furthermore, from C11.1, we can obtain the maximum value that C can take, while the WiFi throughput requirement is met. This value is obtained when $S_{\text{cs}} = \eta$, however as S_{cs} depends on $\boldsymbol{Z}$ which is the optimization variable of (11), derivation of C leading to $S_{\text{cs}} = \eta$ is not trivial. Thus,

$$C_{\min} \leq C \leq T - f(\eta).$$

This means that as long as $T - f(\eta)$ is larger than $C_{\min}$, the optimization problem in (11) is feasible. It should be noted that as far as $C_{\min}$ is sufficiently smaller than $T - \eta$, the problem would be feasible.

Now that we derive the feasibility region of (11), we discuss

on solving this problem. The optimization problem in (11) has a non-convex objective function and a non-convex constraint with the combination of continuous and binary variables. Consequently, (11) is a non-convex mixed-integer, NP-hard optimization problem. Therefore, an efficient algorithm with reasonable computational complexity is needed to solve this scheduling problem.

With a closer look, the optimization problem in (11) potentially looks like an extension of geometric programming (GP). A GP is an optimization problem of the form

$$\min_{\boldsymbol{x}} \quad f_0(\boldsymbol{x}) \tag{12}$$
$$\text{s.t.:} \quad f_i(\boldsymbol{x}) \leq 1, \ i = 1, 2, ..., I$$
$$g_j(\boldsymbol{x}) = 1, \ j = 1, 2, ..., J,$$

where $\boldsymbol{x} = [x_1,,...,x_N]$ is a non-negative vector of optimization variables, $g_j(\boldsymbol{x}) = c_i \prod_{n=1}^{N} x_n^{b_{i,n}}$ for all j are monomial functions, and $f_i(\boldsymbol{x}) = \sum_{k=1}^{K_j} c_{j,k} \prod_{n=1}^{N} x_n^{b_{j,k,n}}$ are posynomial functions for $i = 0, ..., I$, where the multiplicative constants are positive (i.e., $c_i, c_{j,k} > 0$) and $b_{i,n}, b_{j,k,n} \in \boldsymbol{R}$.

The ultimate goal is to transform the optimization problem in (11) to a GP problem. To this end, we first maximize the objective function by minimizing its negative. However, in GP the objective function should be positive. This can be handled by adding sufficiently large constant H to the objective function. Furthermore, we employ two auxiliary variables $T_{cs} = T - T_s \sum_{n_l \in \mathcal{N}_L} x_{n_l}$ and $d = \prod_{n_w \in \mathcal{N}_W}(1 + z_{n_w}) - t'$. By replacing of these auxiliary variables with their corresponding terms and applying transformations in the objective function, (11) becomes

$$\min_{\boldsymbol{X},\boldsymbol{Z},T_{cs},d} H -$$
$$\sum_{n_l \in \mathcal{N}_L} (1 - \psi_{n_l})\theta_{n_l} x_{n_l} - \sum_{n_w \in \mathcal{N}_w} (1 - \psi_{n_w}) T_{cs} z_{n_w} d^{-1} \tag{13}$$

subject to:

C13.1: $\sum_{n_w \in \mathcal{N}_W} (1 - \psi_{n_w}) T_{cs} z_{n_w} d^{-1} \geq \eta$,

C13.2: $\sum_{n_l \in \mathcal{N}_L} x_{n_l} \geq C_{\min}$,

C13.3: $d = \prod_{n_w \in \mathcal{N}_W}(1 + z_{n_w}) - t'$,

C13.4: $T_{cs} = T - T_s \sum_{n_l \in \mathcal{N}_L} x_{n_l}$.

In (13), the objective function is not a posynomial because of the negative in the second term. This can be handled by introducing and minimizing a new auxiliary variable x_0 in addition to guaranteeing the constraint C14.4. Finally, we reach to the following optimization problem

$$\min_{\boldsymbol{X},\boldsymbol{Z},T_{cs},d,x_0} \quad x_0, \quad \text{subject to:} \tag{14}$$

C14.1: $\dfrac{\eta}{\sum_{n_w \in \mathcal{N}_W} (1 - \psi_{n_w}) T_{cs} z_{n_w} d^{-1}} \leq 1$,

C14.2: $\dfrac{\prod_{n_w \in \mathcal{N}_W}(1 + z_{n_w})}{t' + d} = 1$,

C14.3: $\dfrac{T}{T_{cs} + T_s \sum_{n_w \in \mathcal{N}_W} x_{n_w}} = 1$,

C14.4:

$$\dfrac{H}{x_0 + \sum\limits_{n_l \in \mathcal{N}_L} (1 - \psi_{n_l})\theta_{n_l} x_{n_l} + \sum\limits_{n_w \in \mathcal{N}_W} (1 - \psi_{n_w}) T_{cs} z_{n_w} d^{-1}} \leq 1,$$

C14.5: $\dfrac{C_{\min}}{\sum\limits_{n_l \in \mathcal{N}_L} x_{n_l}} \leq 1$.

In this optimization problem [3], all upper-bound inequality constraints are in the form of a ratio between two posynomials and equality constraints are in the form of a ratio between a monomial and a posynomial. This problem belongs to the class of complementary geometric programming, which potentially looks like an extension of GP. In particular, a CGP can be presented as

$$\min_{\boldsymbol{x}} \quad P_0(\boldsymbol{x}) \tag{15}$$
$$\text{s.t.:} \quad P_i(\boldsymbol{x}) \leq 1, \ i = 1, ..., I,$$
$$Q_j(\boldsymbol{x}) = 1, \ j = 1, ..., J,$$

where $P_0(\boldsymbol{x})$ is a posynomial and $P_i(\boldsymbol{x}) = \frac{p_i(\boldsymbol{x})}{p_i^+(\boldsymbol{x})}$ for all $i = 1, ..., I$, in which $p_i(\boldsymbol{x})$ and $p_i^+(\boldsymbol{x})$ are posynomial functions. Moreover, $Q_j(\boldsymbol{x}) = \frac{q_j(\boldsymbol{x})}{q_j^+(\boldsymbol{x})}$ for all j, in which $q_j(\boldsymbol{x})$ are monomial and $q_j^+(\boldsymbol{x})$ are posynomial functions.

B. Algorithm

In this section, we focus on the solution of the optimization problem and present a computationally efficient algorithm which provides a locally optimal solution. In this algorithm, by applying successive transformations, we can convert the CGP problem into a sequence of ordinary GP problems [26], [27].

By approximating the posynomials in the denominator of constraints in (15) with monomials, a complementary GP can be turned into a standard form of GP. The arithmetic-geometric mean inequality can be used to approximate a posynomial with a monomial. In the following, we describe the monomial approximation, which is useful to transform the problem into GP. Let $p_i^+(\boldsymbol{x}) = \sum_{k=1}^{K_i} m_{i,k}^p(\boldsymbol{x})$ and $q_j^+(\boldsymbol{x}) = \sum_{k=1}^{K_j} m_{j,k}^q(\boldsymbol{x})$, where $m_{i,k}^p$ and $m_{j,k}^q$ are monomials. Using the arithmetic-geometric mean inequality, at iteration l, $p_i^+(\boldsymbol{x})$ and $q_j^+(\boldsymbol{x})$ can be approximated as

$$\widetilde{p_i^+}(\boldsymbol{x}(l)) = \prod_{k=1}^{K_i} \left(\dfrac{m_{i,k}^p(\boldsymbol{x}(l))}{\alpha_{i,k}(\boldsymbol{x}(l))} \right)^{\alpha_{i,k}(\boldsymbol{x}(l))}, \tag{16}$$

---
[3] In addition to constraints explained in (14), $\boldsymbol{X} \leq \boldsymbol{1}$ and $\boldsymbol{P} \leq \boldsymbol{1}$ should also be considered to derive the solution of (14).

$$\widetilde{q}_j^+(\boldsymbol{x}(l)) = \prod_{k=1}^{K_j} \left(\frac{m_{j,k}^q(\boldsymbol{x}(l))}{\zeta_{j,k}(\boldsymbol{x}(l))} \right)^{\zeta_{j,k}(\boldsymbol{x}(l))}. \tag{17}$$

The parameters $\alpha_{i,k}(\boldsymbol{x}(l))$ and $\zeta_{j,k}(\boldsymbol{x}(l))$ can be computed as

$$\alpha_{i,k}(\boldsymbol{x}(l)) = \frac{m_{i,k}^p((\boldsymbol{x}(l-1))}{p_i^+((\boldsymbol{x}(l-1))}, \qquad \forall i, \ k \tag{18}$$

$$\zeta_{j,k}(\boldsymbol{x}(l)) = \frac{m_{j,k}^q((\boldsymbol{x}(l-1))}{q_j^+((\boldsymbol{x}(l-1))}, \qquad \forall k, \ j \tag{19}$$

where $\boldsymbol{x}(l-1)$ is the value of last-round solution of the optimization problem. It is proved that arithmetic-geometric mean approximation gives the best local monomial approximation for a posynomial function [28].

Using the approximations in (16) and (17), in each iteration, the optimization problem in (15) would be in the form of a standard GP problem. Consequently, the optimal solution can be achieved by iteratively applying monomial approximations and solving a series of GPs [26]. Using such iterative approach, we propose an algorithm to solve the optimization problem in (14). The details of this approach are summarized in Algorithm 1.

It should be noted using the algorithm proposed in [27], the $\boldsymbol{P}$ value obtained from Algorithm 1 can be implemented in CSMA/CA by configuring the MAC parameters such as minimum contention window, arbitrary inter frame space (AIFS) and retry transmission limit.

C. Computational Complexity of the Proposed Algorithm

The computational complexity of the Algorithm 1 consists of two parts:

1) Converting the CGP problem to a GP problem by applying AGMA approximations
2) Solving the GP problem.

For the first part, AGMA approximations cost is $2N_L + 3N_W + 4$ operations and this means the computational complexity caused by approximations is $C_{\text{APP}} = \mathcal{O}(N_L + N_W)$.

For the second part, the GP problem is solved by CVX using the interior point method. According to [28], this iterative method requires $\log(n_c/t^0\varrho)\log\varepsilon$ number of iterations (Newton steps) to solve the GP problem, where n_c is the total number of constraints, t^0 is the initial point used by the solver in CVX to apply the interior point method, $0 < \varrho \ll 1$ is the stopping criterion and ε is used for updating the accuracy of the method. In the optimization problem , the number of constraints is $n_c = 5 + N_L + N_W$. Thus, the required number of iterations is

$$\mathcal{O}(\log N_L + N_W). \tag{20}$$

Each iteration (Newton step), the Newton step of the interior point method costs $\mathcal{O}(n_c n_v^2)$ operations, where n_c represents the number of constraints and n_v is the number of variables. For the optimization problem (14), $n_c = 5 + N_L + N_W$ and

$n_v = N_L + N_W + 3$. Thus, the computational complexity for solving the GP problem is

$$\mathcal{O}((N_L + N_W)^3). \tag{21}$$

Having the complexity of each iteration (Newton step) and the required number of iterations to solve the GP problem, we can derive the total computational complexity for solving each GP problem which is

$$C_{\text{GP}} = \mathcal{O}(\log(N_L + N_W)(N_L + N_W)^3) \tag{22}$$

Finally, the proposed algorithm is iterative solving successive GP problems until convergence. Since the complexity order of converting the CGP to the GP (C_{APP}) is less than solving the GP problem (CGP), the order of computational complexity for each outer iteration is equal to C_{GP}. The required number of outer iterations to achieve convergence is studied via simulation results.

D. Signaling Aspect of the Proposed Structure

The signaling needed for running Algorithm 1 by the central controller is dependent on the amount of information required to update $\theta_n(t)$ at each duty cycle t. In order to compute $\theta_n(t)$, according to (1), the controller needs the knowledge of the packet arrival probabilities of users as well as piggybacked bit information of users. It is assumed that packet arrival probabilities of users are fixed and can be forwarded once, thus the only information that should be received by the controller from the BS and the AP at each duty cycle is whether the users transmitted at the previous duty cycle have more packets for transmission or not. This information can be conveyed by 1 bit for LTE users who were assigned a time-slot in the previous duty cycle and 2 bits for WiFi users indicating three possible states: the user has a backlogged packet, no backlogged packet and no packet was received from the user at that duty cycle. Thus, the amount of information that should be received by the controller at each duty cycle is $N_L + 2N_W$ bits which can be considered as affordable.

IV. ADMISSION CONTROL FOR U-LTE USERS

The proposed hybrid TDMA-CSMA scheduling algorithm in Section III can satisfy a minimum requirement for WiFi throughput at each duty cycle. However, in addition to WiFi system, the U-LTE users may have quality assurance constraints. Therefore, it is important to admit a precise number of LTE users such that their QoS requirements such as delay constraints can be met. Thus, our objective is to study the number of LTE users that can be admitted to operate in unlicensed spectrum, while their delay requirements can be satisfied. To this end, we need to study the average packet delay for an LTE user.

A. Assumptions

To be able to study the average packet delay and derive admission control rules, we consider a homogeneous U-LTE network in which all users have the same packet arrival probabilities, i.e., a_n. Thus, we omit the subscript n from a_n

Algorithm 1 CGP-based U-LTE & WiFi scheduling
___
Input: Θ, Ψ, η, $C_{\min}$
Initialization: Set initial value to $(\boldsymbol{X}, \boldsymbol{Z}, T_{cs}, d, x_0)$
repeat
  Step 1: Monomial approximation
  1) Compute $\boldsymbol{\alpha}$ for denominators of C14.1, C14.4 and 14.5
  2) Use (16) to approximate the posynomials
  3) Compute $\boldsymbol{\zeta}$ for denominators of C14.2 and C14.3
  4) Use (17) to approximate the posynomials
  Step 2: Solve the transformed GP problem
  1) $gp \leftarrow$ replace denominators of (14) with obtained monomial terms in **Step 1**
  2) $(\boldsymbol{X'}, \boldsymbol{Z'}, T'_{\text{CS}}, d', x'_0) \leftarrow \text{CVX}(gp)$
until $|x_0 - x'_0| < \varsigma$
$p_{n_w} \leftarrow \frac{z_{n_w}}{\theta n_w (1 + z_{n_w})}$
Set $x_{n_l} = 1$ if it is the sum(X) largest elements of X, otherwise set $x_{n_l} = 0$
Output: $\boldsymbol{X}$, $\boldsymbol{P}$
___

in the rest of the paper. Furthermore, we assume that $\psi_n = 0$ for all users.

To calculate the average packet delay, the probability density function (pdf) for TDMA length is needed. This is because the duration of TDMA phase varies at each duty cycle in the proposed scheduling algorithm, depending on the probability that users have packets to transmit. Due to the algorithm dynamics, the derivation of pdf for TDMA length is not tractable. Therefore, we focus on the *minimum* number of users that can be supported by computing the minimum time that is assigned to the LTE system.

Algorithm 1 is run at each duty cycle t where at each duty cycle, $\boldsymbol{\theta}$ is computed based on the Equation (1). Let assume C_s denotes the minimum value of $C(t)$ over all t. In other words, C_s is defined as

$$C_s = \min_{\{t=1,2,...,T_1\}} C(t), \tag{23}$$

if Algorithm 1 runs for $t = 1 : T_1$, assuming that we have a set of LTE users with a specific set of traffic parameters, a set of WiFi users with specific traffic parameters, a given WiFi threshold, and fixed $C_{\min}$. The following theorem describes the condition for $\boldsymbol{\theta}$ which leads to $C(t) = C_s$.

Theorem 1. *For a saturated WiFi, C_s time-slots are assigned to the U-LTE network, if no LTE user has piggybacked a request packet in the previous duty cycle. In other words, having $\boldsymbol{\theta}_L = \boldsymbol{a}_L$, the scheduling algorithm always assigns C_s time-slots to the LTE users. While for the rest of duty cycles, we have $\boldsymbol{\theta}_L \geq \boldsymbol{a}_L$, which leads to $C(t) \geq C_s$.*

Proof. See Appendix A   □

Assuming a case with an unsaturated WiFi, the achievable minimum value of $C(t)$ can be obtained when LTE users have the lowest $\boldsymbol{\theta}$, while WiFi users have the highest $\boldsymbol{\theta}$. However, since it is not easy to derive the highest $\boldsymbol{\theta}$ for WiFi users, we use $C_{\min}$ for the delay analysis when we have unsaturated

WiFi. The reason is that in the optimization problem (11), the second constraint guarantees that the duration allocated for LTE users is larger than $C_{\min}$ time-slots, therefore it can be concluded that $C_s \geq C_{\min}$.

B. Modeling the U-LTE Scheduling Algorithm

As mentioned above, we assume that at each duty cycle, $C_{\min}$ time-slots are dedicated to LTE users to derive an upper bound on the delay. From the proposed scheduling algorithm, it is clear that these time-slots are assigned to users who sent a piggybacked request in the previous duty cycle. Since these users have packets for transmission, by assigning time-slots to them the network throughput can be maximized. However, if the number of requests is less than $C_{\min}$, according to Theorem 2 the rest of time-slots are allocated to users who have the highest θ_n among others.

Theorem 2. *At each duty cycle, time-slots are assigned to users with the highest θ_n's.*

Proof. See Appendix B.   □

In other words, a user will obtain a time-slot either by *demand*, i.e., sending a piggyback packet (if it is currently scheduled) or as *free-assignment*, i.e., waiting for its turn to be served (if it has no assigned time-slot in the current duty cycle). Consequently, time-slot assignment to the user can be modeled by intermittently attending two servers: called server and free server. The server is designated as the called server if it is available to the user as a result of demand assignment. Otherwise, it is called free server for the case of a free-assigned time-slot.

We now discuss the procedure that a user obtains a free time-slot. According to (1), θ_n depends on the a_n and $v_n(t)$. However, since we assume that users are homogeneous and have equal a_n, $\theta_n(t)$ is only dependent on $v_n(t)$, which indicates the last time that the user has been assigned a time-slot. The larger $v_n(t)$ results in the larger $\theta_n(t)$. Thus, among all users that have not been served by the called server, the time-slot goes to the one that has not received a time-slot for the longest time in the most recent duty cycles. On the other hand, those users that have been assigned time-slots more recently, have the lowest $\theta_n(t)$. In other words, the order that users would be served by the free server can be modeled by a queuing system, in which the user is shifted in the end of queue whenever it is served either by the called server or the free server.

C. Delay Analysis

Since we study a homogeneous LTE system, all users experience a same level of performance. Therefore, our derivations will be from the view point of a single user. To this end, we first define two events that are useful to perform the delay analysis. Event E_f occurs if the forthcoming server is a free server (the next transmission of the user happens when it is her turn and there is a free time-slot), while in event of E_c the forthcoming server is a called server.

The type of the forthcoming server depends on whether the queue of the user was empty or nonempty when the precedent server departed from the user. If the user queue was empty at the departure instant of the precedent server, then the user did not request for time-slot allocation. Therefore, the forthcoming server will be a free server and used by at most one packet at the front of the queue.

Packet arrivals happen either at event E_f or E_c. Therefore, for a specific packet, the pmf of the packet delay (denoted by d) can be computed as

$$\mathbb{P}(d = n) = \mathbb{P}(d = n|E_f)\mathbb{P}(E_f) + \mathbb{P}(d = n|E_c)\mathbb{P}(E_c), \quad (24)$$

where $\mathbb{P}(d = n|E_f)$ and $\mathbb{P}(d = n|E_c)$ are conditional probabilities of delay given E_f and E_c [29].

To calculate $\mathbb{P}(d = n)$ based on (24), in the following, we first explain how to obtain p_0 (the probability of empty user queue), which facilitates computing $\mathbb{P}(d = n|E_f)$ and $\mathbb{P}(d = n|E_c)$, and finally we derive $\mathbb{P}(E_f)$ and $\mathbb{P}(E_c)$.

1) Derivation of p_0: In order to derive p_0, we define r_k that represents the number of remaining packets in the queue at the departure instant of k (i.e. the time that the user leaves the server).

First, let assume that the user will be served by a free server. This happens when user queue was empty at the the departure instant of the precedent server (i.e., $r_{k-1} = 0$). Let define the random variable y as the number of packet arrivals between the departure instants of $k - 1$ and k. Therefore, in this case, $r_k = \max\{y - 1, 0\}$. Subsequently, r_k will be zero if at most one packet arrives between the departure instants of $k - 1$ and k. This is because only one packet can be served by the free server. Thus, $\mathbb{P}(r_k = 0|E_f)$ can be computed as

$$\mathbb{P}(r_k = 0|E_f) = \sum_{b=b_{\min}}^{+\infty} \left[\mathbb{P}(y = 0) + \mathbb{P}(y = 1) \right] B_f(b, N_L) \tag{25}$$

where $B_f(b, N_L)$ represents $\mathbb{P}(\mathbf{b} = b|\mathbf{n} = N_L)$, i.e., the probability that it takes b time-slots for the user to be served by a free server after the last departure from the precedent server, given that N_L users exist in the system. Moreover, $b_{\min}$ represents the feasible lower bound of b that can be computed as

$$b_{\min} = \begin{cases} \frac{N_L}{C_{\min}}T + N_L \bmod C_{\min} & \text{if} \quad N_L \bmod C_{\min} = 0 \\ \frac{N_L - C_{\min}}{C_{\min}}T + C_{\min} & \text{otherwise.} \end{cases} \tag{26}$$

where $x \bmod y$ represents the remainder of x divided by y. This value is realized if each user has at most one packet for transmission at the time it is allocated a time-slot. We can derive $\mathbb{P}(y = 0)$ and $\mathbb{P}(y = 1)$ as

$$\mathbb{P}(y = 0) = \sum_{b=b_{\min}}^{\infty} (1 - a)^{\frac{b}{T}} B_f(b, N_L) \tag{27}$$

$$\mathbb{P}(y = 1) = \sum_{b=b_{\min}}^{\infty} \frac{b}{T} a(1 - a)^{\frac{b}{T} - 1} B_f(b, N_L) \tag{28}$$

Thus, $\mathbb{P}(r_k = 0|E_f)$ can be obtained as

$$\mathbb{P}(r_k = 0, E_f) = \tag{29}$$
$$\sum_{b=b_{\min}}^{\infty} \left[(1 - a)^{\frac{b}{T}} + \frac{b}{T} a(1 - a)^{\frac{b}{T} - 1} \right] B_f(b, N_L)$$

Now, let assume that the next server is a called server. Then, $\mathbb{P}(r_k = 0, E_c)$ can be computed as

$$\mathbb{P}(r_k = 0|E_c) = \sum_{b=b_{\min}}^{+\infty} \sum_{k=2}^{+\infty} \sum_{i=1}^{b-T} \mathbb{P}(y = k)B_f(b, N_L) \tag{30}$$

where

$$\mathbb{P}(y = k) = \frac{b - i}{b \times T} \binom{[\frac{b-i}{T}] + k - 2}{k - 1} a^k (1 - a)^{\frac{b}{T}} \tag{31}$$

In order to derive the probability that the user leaves the called server (i.e., $\mathbb{P}(r_k = 0|E_c)$), we take a summation over the all possibilities of y. For $y = k$, the user leaves the called server after k duty cycles from being served by the free server. In other words, it means that k packets arrived at the user's queue during the time that the user was waiting for the free server and the time it spent in the called server. Furthermore, the last packet should arrive at the last duty cycle before the departure, otherwise the user will depart from the server after $k - 1$ duty cycles.

Suppose $\mathbb{P}(r_k = 0)$ is denoted by p_0. Then, p_0 can be found as

$$p_0 = \mathbb{P}(r_k = 0|E_f)\mathbb{P}(E_f) + \mathbb{P}(r_k = 0|E_c)\mathbb{P}(E_c) \tag{32}$$

2) B_f and B_c: In the proposed scheme, after a user departs the free server, it will be served by the free server again if during this interval, all the other users have left either free server or called server. With the assumption that at each duty cycle $C_{\min}$ users are served, those users that are currently allocated a time-slot will be served in the subsequent duty cycles until their queues become empty. Thus, B_f for $b > T$ can be calculated as

$$B_f(b, N_L) = \sum_{i=0}^{M} \binom{M}{i} p_0^i (1 - p_0)^{M-i} B_f(b - T, N_L - i) \tag{33}$$

where $M = \min(N_L, C_{\min})$. Furthermore, for $b < T$ we have

$$B_f(b, n) = \begin{cases} p_0^n & \text{if} \quad b = n \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

The recursive equation (33) indicates that at each duty cycle, i from M users leave the server. Thus, in order to have $\mathbf{b} = b$, the remaining number of users (which is $N_L - i$) should leave the server in $b - T$ time-slots. It is obvious that for $b < T$, only if the number of time-slots is equal to the number of users with probability of p_0^n, all users depart the server. Otherwise, i.e., $b \neq n$, it means either users left the server earlier than b

or later. In both cases, the probability is equal to zero.

Now that B_f is derived, we can calculate p_0. To this end, we use the equation (32) in which all terms are expressed in terms of p_0. To solve this one variable equation, numerical methods can be applied.

3) $\mathbb{P}(E_f)$ and $\mathbb{P}(E_c)$: To derive $\mathbb{P}(E_f)$ and $\mathbb{P}(E_c)$, we need to compute $\bar{M}_f$ and $\bar{M}_c$ which represent the average length of events E_f and E_c, respectively.

$$\bar{M}_f = \sum_{b=b_{\min}}^{\infty} b \times B_f(b, N_L) \tag{35}$$

and

$$\bar{M}_c = \sum_{k=2}^{\infty} \sum_{b=b_{\min}}^{\infty} \sum_{i=1}^{b-T} k \frac{b-i}{b} \times \binom{\left[\frac{b-i}{T}\right] + k - 2}{k - 1} a^k (1-a)^{\frac{b}{T}} B_f(b, N_L) \tag{36}$$

Based on [30], $\mathbb{P}(E_f)$ and $\mathbb{P}(E_c)$ can be derived as

$$\mathbb{P}(E_f) = \frac{p_0 \bar{M}_f}{p_0 \bar{M}_f + (1-p_0)\bar{M}_c} \tag{37}$$

$$\mathbb{P}(E_c) = \frac{(1-p_0)\bar{M}_c}{p_0 \bar{M}_f + (1-p_0)\bar{M}_c} \tag{38}$$

4) Average Packet Delay: Here, we first derive the conditional probability of delay given E_f (i.e., $\mathbb{P}(d = n|E_f)$). Assuming that the length of an interval between two departures is j and the arrival time of the first packet is at i, the packet delay can be computed for two cases as follows

- **First packet case**: The delay is equal to $(j - i)$ if the packet is the first one to arrive in the queue in the interval. We call this event F_1.
- **Non-first packet case**: In this case, denoted by F_2, there are some packets already in the queue. Therefore, the packet will be transmitted, if all packets ahead of it in the queue are scheduled first. In this situation, the first packet of queue is served by the free server. Since the queue is not empty, the piggyback bit will be set to 1. Thus, the remaining packets will be served by the called server. Consequently, the packet delay will be the summation of two parts: $(j - i)$ that is the waiting time required for the first packet to be served by the free server and the second part is the waiting time needed for serving the rest of packets by the called server.

Based on these two cases, the probability of packet delay can be computed as

$$\mathbb{P}(d = n|E_f) = \sum_{b=b_{\min}}^{+\infty} \mathbb{P}(d = n|\mathbf{b} = b, E_f) B_f(b, N_L) \tag{39}$$

where

$$\mathbb{P}(d = n|\mathbf{b} = b, E_f) = \tag{40}$$
$$\mathbb{P}(d = n, F_1|\mathbf{b} = b, E_f) + \mathbb{P}(d = n, F_2|\mathbf{b} = b, E_f)$$

Subsequently, we derive

$$\mathbb{P}(d = n, F_1|\mathbf{b} = b, E_f)$$
$$= \sum_{i=1}^{I} \frac{1}{b} \times \mathbb{P}(\text{no arrival in } i \text{ slots}) \tag{41}$$

and

$$\mathbb{P}(d = n, F_2|\mathbf{b} = b, E_f)$$
$$= \sum_{i=1}^{b} \sum_{h=2}^{\left[\frac{b-i}{T}\right]} \frac{1}{b} \times \mathbb{P}(h \text{ arrivals in } [b-i-n+hT] \text{ slots})$$
$$\times \mathbb{P}(\text{no arrival in } i \text{ slots})$$
$$= \sum_{i=1}^{b} \sum_{h=2}^{\left[\frac{b-i}{T}\right]} \frac{1}{b} \times \binom{\frac{b-i-n}{T} + h}{h} \times a^h \times (1-a)^{(b-n)/T}. \tag{42}$$

The packet can also arrive at Event E_c, in this case the delay can be computed as

$$\mathbb{P}(d = n|E_c) = \mathbb{P}(r_k = n/T)$$
$$= \sum_{i=1}^{b} \frac{1}{b} \times \mathbb{P}(n/T - 1 \text{ arrivals in } [b - i - n/T - 2] \text{ slots})$$
$$\times \mathbb{P}(\text{no arrival in } b \text{ slots})$$
$$= \sum_{i=1}^{b} \frac{1}{b} \times \binom{\frac{b-i}{T} + n/T - 2}{n/T - 1} \times a^{n/T} \times (1-a)^{b/T} \tag{43}$$

which means that in order to have a delay equals to $d = k \times T$, the packet must be arrived when there are k packets at the user queue.

V. ILLUSTRATIVE RESULTS

For performance evaluation, we consider a system with one LTE BS and one WiFi AP, both operating on the same channel. We obtain the results in MATLAB environment and we use CVX to derive the solution of GP problems [31]. We assume that $T = 100$ ms, each time-slot is equal to $T_s = 6$ ms and the backoff unit duration is 10 μs. In the following, we present the considered scenarios for performance evaluation along with their results.

A. Effect of Increasing N_L

We first investigate how increasing the number of users in the U-LTE system can affect the WiFi throughput and the overall network throughput. For this scenario, we assume that the WiFi throughput threshold η is equal to 4 time-slots. We consider a case where $N_W = 14$ and packet arrival probabilities are $\mathbf{A}_L = \{[0.8]_4, [0.5]_{N_L-4}\}$ and $\mathbf{A}_W = \{[0.8]_4, [0.5]_{10}\}$. More specifically, this means that the WiFi network serves 4 users with $a_n = 0.8$ and 10 users with $a_n = 0.5$. Similarly, U-LTE has 4 users with $a_n = 0.8$ and the rest of its users are the ones with $a_n = 0.5$. Furthermore, we assume that users of both networks are randomly located in a circular region with radios of 5m, and channel parameters are as following:
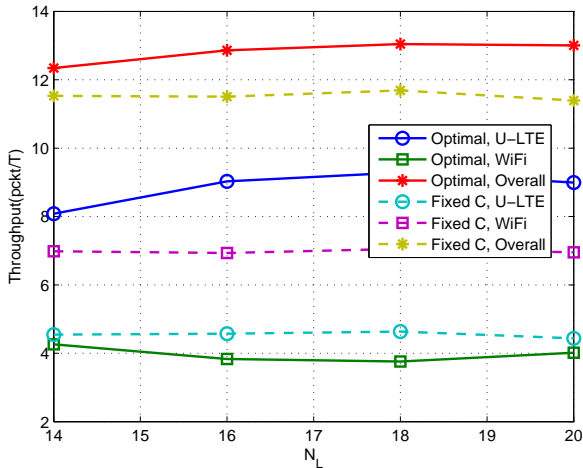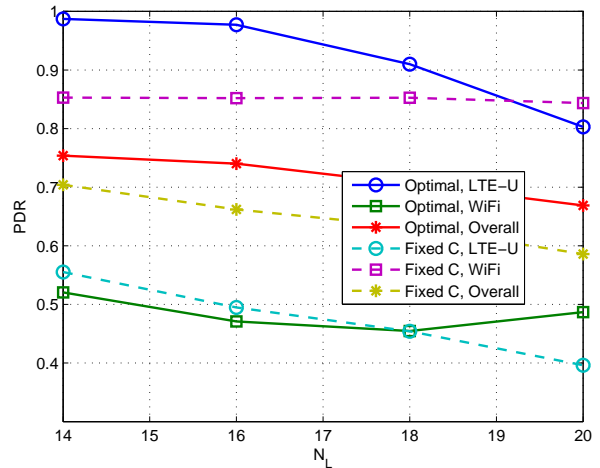
Fig. 2: Throughput vs N_L

Fig. 3: PDR vs N_L

path loss exponent $\xi = 2$, receiver threshold $\upsilon = 0$ dB and $\frac{P_t}{\sigma^2} = 20$ dB.

We compared the performance of the proposed algorithm with Fixed C algorithm as a benchmark to verify the effectiveness of our approach. In the Fixed C algorithm, at each duty cycle a fixed number of time-slots is assigned to the LTE users with highest θ, where θ is the probability that the user has a packet for transmission. Furthermore, in this algorithm, WiFi users compete with each other using fixed p parameters. Here, we assume that $C = 5$ and p is 0.05 for all WiFi users.

As can be seen in Fig. 2, by increasing the number of U-LTE users, i.e., N_L, greater throughput can be achieved for both U-LTE system and the overall network. The reason is that more packets are generated, therefore with higher probability time-slots are allocated to the users who have packets for transmission. On the other hand, WiFi throughput decreases because greater throughput can be achieved by assigning more time-slots to the U-LTE system. However, due to the WiFi throughput constraint, WiFi throughput never falls below the targeted threshold. On the other hand, for the Fixed C algorithm, the LTE-U and overall throughput remains the same, the reason is that number of time-slots assigned to these users is fixed and since these time-slots are allocated to the LTE users with highest θ, LTE-U throughput does not change by increasing N_L.

Furthermore, the effect of increasing N_L on the packet delivery ratio (PDR) are demonstrated in Fig. 3. PRD is defined as the ratio of number of transmitted packets to the number of generated packets. As observed, the proposed algorithm outperforms the Fixed C algorithm in terms of the U-LTE and overall PDR. Furthermore, for $N_L = \{14, 16\}$, U-LTE PDR is so close to 1, however by increasing N_L it starts to drop, which means that U-LTE users' quality of service requirements would be affected. In order to avoid this situation, the number of U-LTE users should be controlled otherwise U-LTE users will suffer from a performance degradation.

B. Effect of Increasing N_W

In another scenario, we have the results for increasing the number of users in the WiFi network, where $\boldsymbol{A}_L = \{[0.8]_4, [0.5]_{10}\}$ and $\boldsymbol{A}_W = \{[0.8]_4, [0.5]_{N_w-4}\}$ and the rest of the parameters are the same as subsection V.A. In Fig. 4, it is evident that in both U-LTE and WiFi, throughput remains unchanged for the optimal algorithm. In fact, when U-LTE system has high-traffic users, higher throughput can be achieved by allocating more time to them, since for high-traffic users TDMA achieves better performance compared to the CSMA. The reason is that in CSMA, due to the backoffs and collisions, time would be wasted. Therefore, in Fig. 4 where U-LTE users have high traffic, the following statement is true. The higher throughput could be achieved if more time slots were allocated to LTE-U. Consequently, the algorithm is reluctant to add more time slots for WiFi system since its throughput requirement is already satisfied, even if the number of WiFi users is increasing. The other point is that increasing N_W may lead to a larger number of collisions, therefore, to meet the WiFi throughput threshold, lower p probabilities are assigned by the optimal algorithm to the users. However, the WiFi throughput of the Fixed C algorithm drops by increasing N_W, as p values are fixed.

Furthermore, as shown in Fig. 5 the WiFi PDR of the optimal algorithm decreases by growing number of WiFi users. The reason is that increasing N_W leads to the larger number of packets, while the number of transmitted packets remains the same.

C. Homogeneous U-LTE Network

Here, we obtain the results for homogeneous LTE network. For this scenario, we set the duty cycle period T equal to 16 time-slots, and the WiFi throughput threshold η equal to 4 time-slots. We derive the results for $C_{\min} = 5$ and 7. Moreover, we consider the packet arrival probabilities as $\boldsymbol{A}_L = \{[0.3]_{N_L}\}$ and $\boldsymbol{A}_W = \{[1]_8\}$. As can be observed from
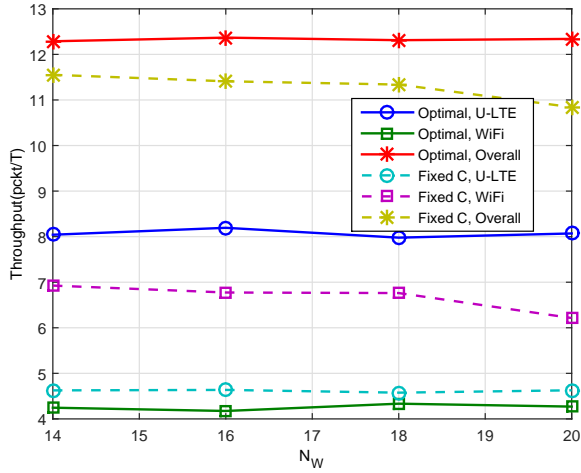
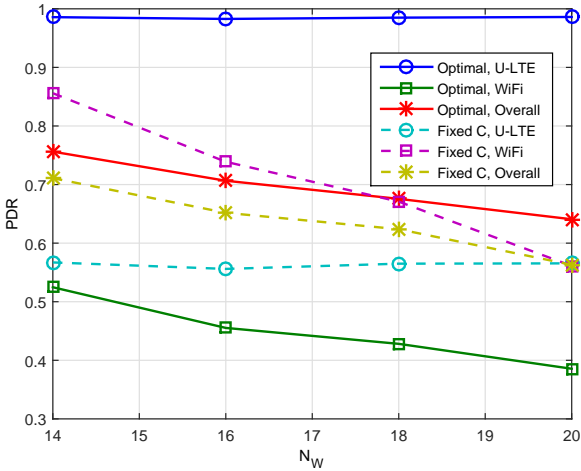Fig. 4: Throughput vs N_W
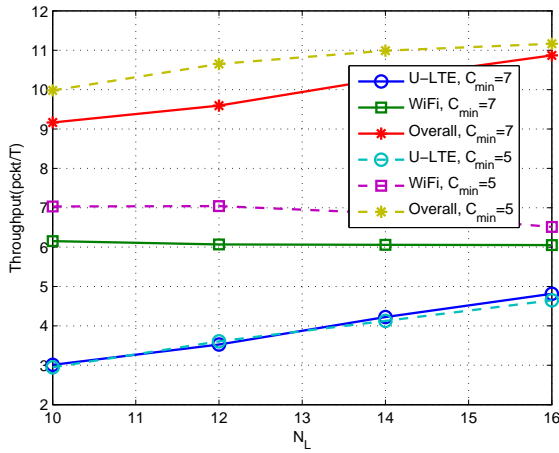

Fig. 5: PDR vs N_W
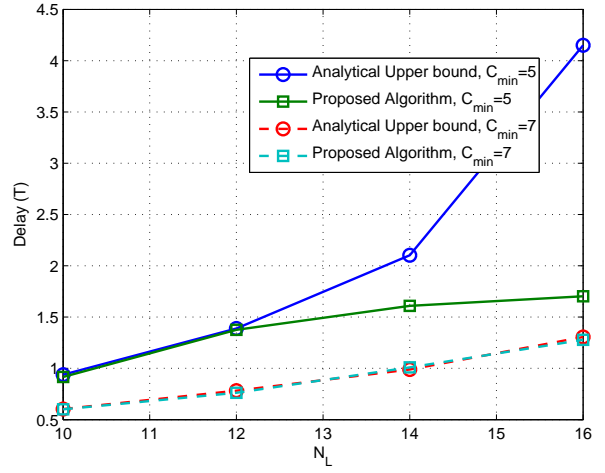

Fig. 6: Throughput vs N_L

Fig. 7: delay vs N_L

Fig. 6, by increasing N_L, LTE throughput increases as well. The reason is that since more number of packets are generated, less time-slots are left idle. Moreover, although more number of time-slots are assigned to the LTE network for $C_{\min} = 7$ but same throughput is achieved for two values of $C_{\min}$. The reason is that LTE network is underutilized, therefore adding more time-slots does not lead to greater throughput. However, for $C_{\min} = 5$ more time-slots are left for WiFi and since WiFi users are saturated, the increased time-share leads to larger throughput.

Furthermore, we obtain delay results for this scenario shown in Fig.7. The results are compared with the analytical upper-bound which is derived for saturated WiFi by assigning $C_{\min}$ time-slots to LTE users at each duty cycle. As can be observed, the gap between the upper-bound and the proposed scheduling algorithm increases as N_L grows. The reason is that by increasing N_L, more traffic is generated in the LTE network. Therefore, greater throughput can be achieved by assigning more time-slots to LTE users. Thus, the probability that $C(t) > C_{\min}$ increases which makes the gap between the proposed scheme and the upper-bound becomes larger.

Furthermore, for $C_{\min} = 7$ the upper bound delay and simulation results are close, while for $C_{\min} = 5$, the gap is larger. The reason is that with $C_{\min} = 7$, in most of the duty cycles, the derived $C(t)$ is equal to $C_{\min}$. However, $C_{\min} = 5$ leads to $C(t) > C_{\min}$ more frequently.

Moreover, as can be observed, $C_{\min} = 7$ achieves smaller delay compared to the $C_{\min} = 5$. Since larger $C_{\min}$ imposes to assign more time-slots to the LTE network. Therefore, at each duty cycle, more number of packets are served which leads to shorter delays for U-LTE users. However, for case of $N_L = 16$, the same amount of delay is achieved for both values of $C_{\min}$. Since for this case, LTE network is more loaded, therefore in most of the duty cycles, we have $C(t) > C_{\min}$, meaning that the output of the scheduling is not dependent on $C_{\min}$.
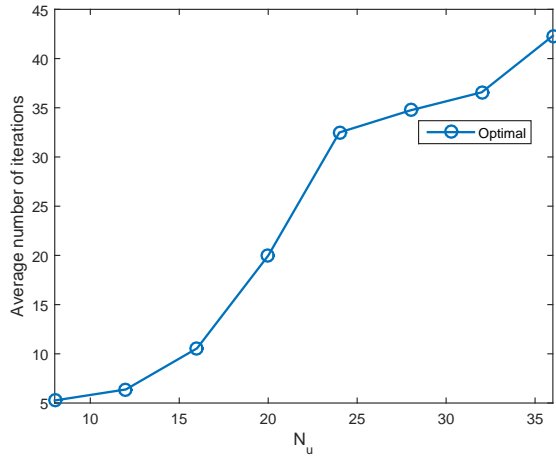
Fig. 8: Average number of iterations vs N_u

D. Computational Complexity

Here, we investigate the average number of iterations required for Algorithm 1 to converge. The results shown in Fig. 8 are obtained for $\boldsymbol{A}_{\mathrm{L}} = \boldsymbol{A}_{\mathrm{W}} = \{[0.8]_4, [0.5]_{N_{\mathrm{u}}/2-4}\}$, where $N_{\mathrm{u}} = N_{\mathrm{L}} + N_{\mathrm{W}}$. Furthermore, for this scenario, we set $\eta = 2$ time-slots, the convergence parameter $\varsigma = 0.05$ and the same setting for channel parameters as subsection V.A. As observed, the average number of iterations grows only linearly with N_{u}.

VI. CONCLUSION

In order to satisfy the increasing demand for mobile traffic, LTE operation over unlicensed bands has been proposed. In this paper, we consider the scenario that both LTE and WiFi systems share the same unlicensed band. In such a setting, the main challenge for U-LTE deployment is that the performance of WiFi system should not degrade significantly. In order to address this issue, we consider a coordinated approach in which both systems are connected to a central network entity. This entity manages the channel access between these two systems such that the overall spectrum efficiency is improved, while the WiFi performance does not fall below a certain level. In order to reach this goal, a duty-cycle-based approach is used, in which the time is divided into duty cycles and the exclusive share of each system is dynamically optimized by the network entity. It is shown that the developed algorithm can ensure a minimum throughput requirement for WiFi while maximizing the total throughput. Furthermore, we obtain an upper-bound for average delay of LTE users. Using this analysis, we can derive minimum number of LTE users that can be admitted by the network while their delay requirements are met.

REFERENCES

[1] F. Chaves, A. Cavalcante, E. Almeida, F. Abinader Jr, R. Vieira, S. Choudhury, and K. Doppler, "LTE/Wi-Fi coexistence: Challenges and mechanisms," *XXXI SIMPOSIO BRASILEIRO DE TELECOMUNICA-COES*, 2013.

[2] S. Sagari, I. Seskar, and D. Raychaudhuri, "Modeling the coexistence of LTE and WiFi heterogeneous networks in dense deployment scenarios," in *IEEE Intl. Conf. on Commun. Workshop (ICCW)*, 2015, pp. 2301–2306.

[3] F. M. Abinader, E. P. Almeida, F. S. Chaves, A. M. Cavalcante, R. D. Vieira, R. C. Paiva, A. M. Sobrinho, S. Choudhury, E. Tuomaala, K. Doppler *et al.*, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 54–61, 2014.

[4] B. Chen, J. Chen, Y. Gao, and J. Zhang, "Coexistence of LTE-LAA and Wi-Fi on 5 GHz with corresponding deployment scenarios: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 7–32, 2017.

[5] A. Mukherjee, J.-F. Cheng, S. Falahati, H. Koorapaty, R. Karaki, L. Falconetti, D. Larsson *et al.*, "Licensed-assisted access LTE: coexistence with IEEE 802.11 and the evolution toward 5G," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 50–57, 2016.

[6] H. Cui, V. C. Leung, S. Li, and X. Wang, "LTE in the unlicensed band: Overview, challenges, and opportunities," *IEEE Wireless Communications*, 2017.

[7] S. Hajmohammad and H. Elbiaze, "Unlicensed spectrum splitting between femtocell and WiFi," in *Proc. IEEE Intl. Conf. Commun. (ICC)*. IEEE, 2013, pp. 1883–1888.

[8] 3GPP Study Item RP-141397, "Study on licensed-assisted access using LTE," Tech. Rep., 2014.

[9] S. Dama, A. Kumar, and K. Kuchi, "Performance evaluation of LAA-LBT Based LTE and WLAN's co-existence in unlicensed spectrum," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2015, pp. 1–6.

[10] H. Ko, J. Lee, and S. Pack, "A fair listen-before-talk algorithm for coexistence of LTE-U and WLAN," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 116–10 120, 2016.

[11] C. Cano and D. J. Leith, "Coexistence of WiFi and LTE in unlicensed bands: A proportional fair allocation scheme," in *IEEE Intl. Conf. on Commun. Workshop (ICCW)*. IEEE, 2015, pp. 2288–2293.

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

[13] 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, 2015.

[14] 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, 2016.

[15] X. Wang, T. Q. Quek, M. Sheng, and J. Li, "Throughput and fairness analysis of Wi-Fi and LTE-U in unlicensed band," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 63–78, 2017.

[16] S. Han, Y.-C. Liang, Q. Chen, and B.-H. Soong, "Licensed-assisted access for LTE in unlicensed spectrum: A MAC protocol design," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2550–2561, 2016.

[17] Y. Song, K. W. Sung, and Y. Han, "Impact of packet arrivals on Wi-Fi and cellular system sharing unlicensed spectrum," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 204–10 208, 2016.

[18] Q. Zhang, Q. Wang, Z. Feng, and T. Yang, "Design and performance analysis of a fairness-based license-assisted access and resource scheduling scheme," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2968–2980, 2016.

[19] Q. Chen, G. Yu, H. Shan, A. Maaref, G. Y. Li, and A. Huang, "Cellular meets WiFi: Traffic offloading or resource sharing?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3354–3367, 2016.

[20] S.-Y. Lien, J. Lee, and Y.-C. Liang, "Random access or scheduling: Optimum LTE licensed-assisted access to unlicensed spectrum," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 590–593, 2016.

[21] M. Zhang, X. Zhang, Y. Chang, and D. Yang, "Dynamic uplink radio access selection of LTE licensed-assisted access to unlicensed spectrum: An optimization game," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2510–2513, 2016.

[22] A. Dalili Shoaei, M. Derakhshani, and T. Le-Ngoc, "Efficient LTE/WiFi coexistence in unlicensed spectrum using virtual network entity," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2017.

[23] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, and E. Sousa, "5G communications race: pursuit of more capacity triggers LTE in unlicensed band," *IEEE Trans. Veh. Technol.*, vol. 10, no. 1, pp. 43–51, 2015.

[24] Q. Chen, G. Yu, H. M. Elmaghraby, J. Hamalainen, and Z. Ding, "Embedding LTE-U within Wi-Fi bands for spectrum efficiency improvement," *IEEE Netw.*, vol. 31, no. 2, pp. 72–79, 2017.

[25] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[26] G. Xu, "Global optimization of signomial geometric programming problems," *European Journal of Operational Research*, vol. 233, no. 3, pp. 500–510, 2014.

[27] M. Derakhshani, X. Wang, T. Le-Ngoc, and A. Leon-Garcia, "Virtualization of multi-cell 802.11 networks: Association and airtime control," *arXiv preprint arXiv:1508.03554*, 2015.

[28] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[29] T. Le-Ngoc and I. M. Jahangir, "Performance analysis of CFDAMA-PB protocol for packet satellite communications," *IEEE Trans. Commun.*, vol. 46, no. 9, pp. 1206–1214, 1998.

[30] D. R. Cox, *Renewal theory*. Methuen, 1962, vol. 58.

[31] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," http://cvxr.com/cvx, Aug. 2012.

APPENDIX

A. Proof of Theorem 1

Assuming that $(\boldsymbol{X'}, \boldsymbol{Z'})$ is the optimal solution for $\boldsymbol{\Theta'}_{\mathcal{N}_L} = \boldsymbol{a}_{\mathcal{N}_L}$, we have

$$\boldsymbol{X'}\boldsymbol{\Theta'}_{\mathcal{N}_L} < \boldsymbol{X'}\boldsymbol{\Theta''}_{\mathcal{N}_L} \qquad (44)$$

where $\boldsymbol{\Theta''}_{\mathcal{N}_L} \geq \boldsymbol{a}_{\mathcal{N}_L}$. Following that, we can conclude that the total throughput obtains for $\boldsymbol{\Theta''}_{\mathcal{N}_L}$ at $(\boldsymbol{X'}, \boldsymbol{Z'})$ is greater or equal to the total throughput of $\boldsymbol{\Theta'}_{\mathcal{N}_L}$ at $(\boldsymbol{X'}, \boldsymbol{Z'})$ Thus, for $\boldsymbol{\Theta''}_{\mathcal{N}_L}$, the optimal solution $\boldsymbol{X''}$ is

$$\boldsymbol{X''} \geq \boldsymbol{X'} \qquad (45)$$

Which means that for $\boldsymbol{X''} > \boldsymbol{X'}$, we have $\boldsymbol{X''}\mathbf{1} > \boldsymbol{X'}\mathbf{1}$. Furthermore, $\boldsymbol{X''}\mathbf{1} = D''$ and $\boldsymbol{X'}\mathbf{1} = D'$, which leads to $D'' > D'$ or equivalently, $D'' > C_{\min}$.

B. Proof of Theorem 2

Proof by contradiction: Let assume that

$$\exists n' : x_{n'}^* = 1 \text{ and } \exists n'' : x_{n''}^* = 0 \text{ and } \theta_{n'} < \theta_{n''} \qquad (46)$$

Also, let $\boldsymbol{x'}$ be a suboptimal solution for the optimization problem in which $x_n' = x_n^*$, $\forall n \in \mathcal{N}_L$ except for $n \in \{n', n''\}$, in which $x_{n'}' = 0$ and $x_{n''}' = 1$. Consequently, we have

$$\boldsymbol{X^*}\boldsymbol{\Theta} < \boldsymbol{X'}\boldsymbol{\Theta}, \qquad (47)$$

and since $\boldsymbol{X^*}\mathbf{1} = \boldsymbol{X'}\mathbf{1}$, therefore we derive same WiFi parameter for both $\boldsymbol{X^*}$ and $\boldsymbol{X'}$; $\boldsymbol{Z'} = \boldsymbol{Z^*}$. This means that the total throughput of $(\boldsymbol{X'}, \boldsymbol{Z'})$ is larger than $(\boldsymbol{X}^*, \boldsymbol{Z}^*)$, which contradicts the fact that $(\boldsymbol{X}^*, \boldsymbol{Z}^*)$ is the optimal solution.

Atoosa Dalili Shoaei received her B.Sc. degree in information technology engineering from Isfahan University of Technology, Isfahan, Iran, in 2009 and her M.Sc. degree in information technology engineering form Amirkabir University of Technology, Tehran, Iran, in 2012. Since 2014, she has been with the Department of Electrical and Computer Engineering, McGill University, Canada, where she is currently working toward her Ph.D. degree. Her main areas of research interest are media access control techniques, Internet of Things (IoT) and wireless virtualization.

Mahsa Derakhshani (S10-M13) is a Lecturer (Assistant Professor) in Digital Communications with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, UK. Prior to joining Lboro, she was an Honorary NSERC Postdoctoral Fellow with Department of Electrical and Electronic Engineering, Imperial College London (2015 to 2016), a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada and a Research Assistant with the Department of Electrical and Computer Engineering, McGill University (2013 to 2015). She received her Ph.D. in electrical engineering degree from McGill University, Montral, Canada, in 2013. Her research interests include radio resource management for wireless networks, software-defined wireless networking, applications of convex optimization and game theory for communication systems, and spectrum sensing techniques in cognitive radio networks. Dr. Derakhshani has received several awards and fellowships including the John Bonsall Porter Prize, the McGill Engineering Doctoral Award, the Fonds de Recherche du QuebecNature et Technologies (FRQNT) and Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowships. Currently, she serves as an Associate Editor of the IET Signal Processing Journal.

Tho Le-Ngoc (F97) received the B.Eng. degree (with distinction) in electrical engineering and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 1976 and 1978, respectively, and the Ph.D. degree in digital communications from the University of Ottawa, Ottawa, ON, Canada, in 1983. During 19771982, he was with Spar Aerospace Limited, Sainte-Anne-de-Bellevue, QC, where he was involved in the development and design of satellite communications systems. During 19821985, he was Engineering Manager of the Radio Group, Department of Development Engineering, SRTelecom Inc., Saint-Laurent, QC, where he developed the new point-to-multipoint demand-assigned time-division multiple-access/time-division multiplexing (DA-TDMA/TDM) Subscriber Radio System SR500. During 19852000, he was a Professor with the Department of Electrical and Computer Engineering, Concordia University, Montreal. Since 2000, he has been with the Department of Electrical and Computer Engineering, McGill University. His research interest is in the area of broadband digital communications. Dr. Le-Ngoc is a Fellow of the Engineering Institute of Canada, the Canadian Academy of Engineering, and the Royal Society of Canada. He received the 2004 Canadian Award in Telecommunications Research and the 2005 IEEE Canada FessendenAward.He holds a Canada Research Chair (Tier I) on Broadband Access Communications.