IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 24, 2019, accepted May 13, 2019, date of publication May 20, 2019, date of current version June 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2917692

Protocol Design and Resource Allocation for LTE-U System Utilizing Licensed and Unlicensed Bands

WEIHUA WU[1], **QINGHAI YANG**[1], **RUNZI LIU**[2], **(Member, IEEE), AND KYUNG SUP KWAK**[3]

[1]State Key Laboratory of ISN, School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
[2]School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China
[3]Department of Geoinformatic Engineering, Inha University, Incheon 402-751, South Korea

Corresponding author: Runzi Liu (rzliu@xauat.edu.cn)

This work was supported in part by the NSF China under Grant 61701365, Grant 61801365, and Grant 61471287, in part by the National Research Foundation of Korea-Grant funded by the Korean Government (Ministry of Science, ICT and Future Planning)-NRF under Grant 2014K1A3A1A20034987, in part by the China Postdoctoral Science Foundation under Grant 2018M643581, in part by the National Natural Science Foundation of Shaanxi Province under Grant 2019JQ-152, in part by the Postdoctoral Foundation in Shaanxi Province of China, and in part by the Fundamental Research Funds for the Central Universities.

ABSTRACT LTE deployed with both licensed and unlicensed bands (LTE-U) is one of the promising approaches to meet the rapidly growing data demand in wireless networks. In this paper, both throughput and fairness for the LTE-U system are maximized by a multi-objective optimization problem. Then, a log-sum-exp approximation method is developed to convert the multi-objective optimization into a single objective optimization problem. At the same time, the tradeoff between throughput and fairness is mathematically depicted by a control parameter. To tackle the obtained single objective optimization problem, a Markov chain directed algorithm is developed to convert it into a coexistence protocol design subproblem at the MAC layer and a resource allocation subproblem at the physical layer, respectively. Then, we propose adaptive exponential backoff schemes for both the LTE-U devices and the incumbent devices on the unlicensed bands. After that, a low-complexity two-iterative optimization procedure is developed to jointly allocate the licensed and unlicensed resources of the LTE-U system. The simulation results show that our proposed coexistence protocol and resource allocation can achieve fair coexistence between the LTE-U devices and the incumbent devices on the unlicensed bands, moreover it can achieve higher throughput than the non-adaptive coexistence protocol in the unlicensed bands.

INDEX TERMS LTE-unlicensed bands, LBT, coexistence protocol.

I. INTRODUCTION

The amount of data generated by smartphones, tablets, PDAs and new mobile computing devices has doubled each year, and this trend will be likely to continue in the next decade [1]. As a result, the scarce licensed spectrum bands in cellular networks have been convinced difficult to carry these mobile data traffic. To overcome this difficult, the Federal Communications Committee (FCC) opened 295 MHz bandwidth in the unlicensed 5 GHz spectrum band to relieve the pressure on wireless networks [2]. Different from the licensed spectrum, unlicensed spectrum is a completely open spectrum resource to different radio access technologies (RAT) as long as they comply with the relevant regulations. Therefore, deploying LTE over unlicensed (LTE-U) spectrum has been proposed and is currently under the consideration of Third Generation Partnership Project (3GPP) in its future standards [3]. However, LTE utilizes centralized scheduling protocol and incumbent systems on unlicensed bands are deployed with the distributed coordination function (DCF), so that the coexistence of LTE-U and incumbent systems will cause increased interference to each other and the overall system performance will be degraded.

Currently, there are two major coexistence mechanisms for LTE-U system: (a) duty-cycle method [4], [5] and (b) listen-before-talk (LBT) [6], [7]. The basic idea of duty-cycle method is to rely on the LTE's centralized scheduling to periodically turn its signal off and on by using almost

The associate editor coordinating the review of this manuscript and approving it for publication was Khalil Afzal.

blank subframes (ABSs) so that the incumbent systems can have adequate access time. Obviously, LTE-U system dominates the operation of the whole system, and the incumbent systems' performance will be seriously affected. For the LBT mechanism, LTE-U should detect whether the channel is idle before transmission. If the channel is busy, it will wait for some time, otherwise it will occupy the unlicensed bands. Several LBT-based coexistence mechanisms have been proposed in these years. For example, in [8], static and uniformly distributed backoff scheme was developed for the equipments utilizing unlicensed band. Based on the LBT mechanism, the unlicensed band was shared in time among different users [7], [9], and the unlicensed-band usage was best characterized by the fraction of time that a device accesses the channel. However, the fairness and adaptability have not been thoughtfully investigated in these works.

Another challenge comes from the joint allocation of both licensed and unlicensed resource. In fact the resource allocation for LTE-U has been drawing attentions for a long time [10]–[13], while all of them neglected the fact that most practical systems adopt discrete power control with a limited number of power levels [14]. It's worth noting that the current 3GPP LTE standard of wireless networks only supports discrete power control in the downlink transmission via a user specific data-to-pilot-power offset parameter [15]. Compared with continuous power control, discrete power control not only can simplify the design of transmitter but also reduce the cost of the information exchange overhead within system. However, the discrete power control usually leads to a combinatorial optimization problem, which has been proved to be NP-hard. Despite that many suboptimal algorithms have been proposed in [16]–[18] for solving the discrete power control problem, these works only concentrated on licensed resource. It is difficult to directly applying these methods on the joint design of licensed and unlicensed resource allocation. That is because the interference from LTE-U system to incumbent devices on unlicensed bands should not be ignored and the fairness between different devices must be thoughtfully investigated.

In this paper, we consider the LTE-U deployed with both licensed and unlicensed bands. Based on the fact that 3GPP has identified LBT as a working assumption of the standardizing LTE-U systems [19], we focus on a similar LBT coexistence protocol design in our proposed system. Besides, we assume a more practical setting for the LTE-U that includes a limited number of power levels. The LTE-U shares the licensed bands by using Orthogonal Frequency Division Multiple Access (OFDMA) technique and the unlicensed bands by time division multiple access (TDMA) manner. For ensuring that each device has the opportunity to access the unlicensed bands, we define a fairness criterion by using the negative entropy function, restricted to the time occupancy fraction simplex. The problem of this paper is formulated as a multi-objective optimization to maximize the system throughput and as well the fairness of devices by joint power, licensed and unlicensed bands allocation, while guaranteeing the basic transmit rate requirements of devices. In brief, the main contributions of this paper can be summarized as follows:

- A markov chain directed strategy is developed to convert the multi-objective optimization into a coexistence protocol design subproblem at the MAC layer and a resource allocation subproblem at the physical layer. At the same time, the throughput and fairness tradeoff framework is mathematically depicted by a control parameter. Only by adjusting the control parameter, the proposed backoff scheme provides a controllable method to balance the tradeoff performance and let the network operate in a predefined state.
- The adaptive exponential backoff schemes are developed for both the LTE-U devices and the incumbent devices on unlicensed bands. It should be noted that the backoff window size can be adaptively adjusted based on the available device throughput utility.
- A low-complexity two-iterative optimization procedure is developed for the resource allocation subproblem, which is proved to be an mixed integer programming. First, with fixed transmit power, the optimal subchannel allocation solution is obtained. Then, a greedy algorithm is developed to control the transmit power on both licensed and unlicensed bands.

The organization of the rest of this paper is as follows: In Section II, we present the network model and problem formulation. The adaptive coexistence protocol are given in Section III. The resource optimization procedure is proposed in Section IV. In Section V, we present the simulation results to evaluate the proposed scheme. Finally, the conclusion is drawn in Section VI.

II. SYSTEM MODEL

We consider a LTE-U small base station (SBS), which can access both licensed bands and unlicensed bands concurrently. There exist $\mathcal{M} = \{1, 2, \cdots, M\}$ sDevices associated with the SBS. Conventionally, the SBS provides the basic service rate for the sDevices through the licensed bands, and utilizes unlicensed bands to cope with the large traffic data transmission service, such as eMBB in 5G. We assume that the sDevices coexist with many incumbent devices on the unlicensed bands, and these incumbent devices are represented as wDeivces. The set of $\mathcal{L} = \{1, 2, \cdots, L\}$ wDevices shair the unlicensed bands with sDeivces in time domain, so that the subchannels are not considered at the unlicensed bands [20]. In this paper, we only consider the downlink transmission for the ease of exposition.

A. sDevice TRANSMISSION RATE ON LICENSED BANDS

We assume that the licensed band is divided into K subchannels, each with a bandwidth of W. Then the transmission rate for sDevice m served by SBS on subchannel k is given as

$$r_m^k = W \log_2 \left(1 + \frac{p_m^k g_m^k}{N_0 W}\right), \quad (1)$$

where p_m^k and g_m^k are the transmit power and channel gain between SBS and sDevice m on sub-channel k, respectively, N_0 is the noise power density. Due to the fact that only a limited number of power levels are used in practical systems, we assume that the transmit power p_m^k can only be chosen from a discrete power level set $\mathcal{P}_L = \{px_m^k \cdot p_{max}^L/(\mathcal{P}_L - 1), px_m^k = 0, 1, 2, \cdots, \mathcal{P}_L - 1\}$, where px_m^k is the index of power level and p_{max}^L denotes the maximum transmission power over a sub-channel.

Then, the overall transmission rate of sDevice m on licensed bands can be computed as

$$R_m^1 = \sum_{k=1}^{K} x_m^k r_m^k, \tag{2}$$

where x_m^k is the subchannel allocation indicator, i.e. $x_{i,n}^k = 1$ if sDevice m is allocated with k-th subchannel, and $x_{i,n}^k = 0$ otherwise. Let R_m^{req} denote the basic data requirement of sDevice m, so the overall transmission rate of sDevice m should be larger than R_m^{req}, i.e. $R_m^1 \geq R_m^{req}$.

B. sDevice TRANSMISSION RATE ON UNLICENSED BANDS

In order to be compatible with wDevices, the SBS shares the unlicensed bands in a time division manner. The achieved data rate for sDevice m on unlicensed bands can be written as

$$R_m^2 = t_m B \log_2\left(1 + \frac{p_m^U h_m^U}{N_0 B}\right), \tag{3}$$

where t_m is the time fraction occupied by sDevice m on the unlicensed bands, p_m^U and h_m^U are the transmit power and channel gain between SBS and sDevice m, respectively. Similar to the transmit power at licensed bands, p_m^U can only be chosen from a discrete power level set $\mathcal{P}_U = \{pt_m \cdot p_{max}^U/(\mathcal{P}_U - 1), pt_m = 0, 1, 2, \cdots, \mathcal{P}_U - 1\}$, where pt_m is the index of power level and p_{max}^U denotes the maximum transmit power on unlicensed bands. In order to guarantee that each sDevice can access the unlicensed band, the time fraction t_m should be insured $t_m > 0, \forall m$.

Based on Eqs. (2) and (3), the achieved data rate for sDevice m is given as

$$R_m = R_m^1 + R_m^2. \tag{4}$$

C. wDevice TRANSMISSION RATE

For the wDeivces, distributed coordinated function (DCF) with exponential backoff scheme is employed to contend the unlicensed bands. Hence, at any instant only one wDevice is transmitting over the whole spectrum band except a collision happens. Let R_l denote the total transmit rate of wDevice l, then it can be computed as

$$R_l = t_l R_l^w, \tag{5}$$

where t_l is the time fraction occupied by wDevice l on the unlicensed bands, R_l^w is the instantaneous transmission rate according to the signal-to-interference-plus-noise ratio (SINR) and the PHY specifications of the IEEE 802.11 protocol [21].

D. FAIRNESS FOR COEXISTENCE

It should be noted that fairness is a key design consideration for the coexistence of different devices on unlicensed bands. Because that the unlicensed spectrum is a completely open spectrum resource to different devices as long as they comply with the relevant regulations, so it must be ensure that each device has the opportunity to access the unlicensed bands.

Entropy was introduced by Shannon [22] in communication system. Since it also reflects fairness aspects, so many researchers [23]–[25] employed it as a measure of fairness. Hence, we use the following entropy function to define the fairness criterion

$$F(\mathbf{t}) = -\sum_{m=1}^{M} t_m \log t_m - \sum_{l=1}^{L} t_l \log t_l, \tag{6}$$

where $\mathbf{t} = [t_i]_{M+L}$. Then, the following lemma can be obtained for the fairness criterion $F(\mathbf{t})$.

Lemma 1: The fairness criterion $F(\mathbf{t})$ is a concave function and the maximum can be achieved at the equal distribution of the unlicensed bands, i.e. $t_m, t_l = \frac{1}{M+L}, \forall m, l$.

Proof: See Appendix A. □

Therefore, maximizing the fairness criterion $F(\mathbf{t})$ can guarantee that each device has the equal opportunity to access the unlicensed bands. Moreover, it will found in Section III that this design of fairness criterion in fact opens a new design space for developing the coexistence protocol of the devices on unlicensed bands.

E. PROBLEM FORMULATION

In this paper, our objective is to maximize both the system throughput and the fairness with the constraint of satisfying each sDevice's basic data requirement. Thus, the optimization problem can be formulated as

$$\begin{aligned}
&\max w_1 \sum_{m=1}^{M} R_m + w_2 \sum_{l=1}^{L} R_l \\
&\quad \max -\sum_{m=1}^{M} t_m \log t_m - \sum_{l=1}^{L} t_l \log t_l \\
&\text{s.t. } C1: R_m^1 \geq R_m^{req}, \forall m, \\
&\quad C2: \sum_{m=1}^{M} x_m^k \leq 1, x_m^k \in \{0, 1\}, \\
&\quad C3: \sum_{m=1}^{M} t_m + \sum_{l=1}^{L} t_l \leq 1, \forall t_m > 0, \forall t_l > 0, \\
&\quad C4: \sum_{m=1}^{M} t_m \sum_{pt_m \in \mathcal{P}_U} pt_m \frac{p_{max}^U}{\mathcal{P}_U - 1} \leq p_{max}^U, \\
&\quad C5: \sum_{m=1}^{M} \left(\sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} \right. \\
&\qquad \left. + t_m \sum_{pt_m \in \mathcal{P}_U} pt_m \frac{p_{max}^U}{\mathcal{P}_U - 1} \right) \leq P_{max},
\end{aligned} \tag{7}$$

where $w_1 \in (0, 1)$, $w_2 \in (0, 1)$, and $w_1 + w_2 = 1$. Here, w_1 and w_2 represent the weights of throughput of sDevices and wDevices respectively. By choosing appropriate values, the operators can give higher preference to sDevices or wDevices. Constraint C1 specifies that the transmission rate of each sDevice should be larger than its basic rate requirement. C2 guarantees the orthogonality of channel allocation on licensed bands. C3 shows the total unlicensed band usage time should be less than 1. C4 is transmit power constraint on the unlicensed bands due to the corresponding regulations [26]. C5 is the transmit power constraint on both licensed and unlicensed bands.

In problem (7), both subchannel allocation variable x and transmit power allocation variables p_L and p_U are integer variables, while the time fractions t_m and t_l are continuous variables, these characteristics make the problem a combinatorial optimization, which is generally a NP-hard problem. Moreover, different from the existing works to guarantee the time fraction $t_m, t_l \geq 0$ [19], [26]–[28], problem (7) proposes $t_m, t_l > 0$ to promise that each device has the opportunity to access the unlicensed bands. It can prevent the case that some device will exclusively occupy the unlicensed bands greedily, i.e. $\exists t_m = 1$ or $\exists t_l = 1$. Besides, problem (7) is a multi-objective optimization problem and it is hard to derive a closed-form solution. In the following, we adopt a Markov approximation framework to solve problem (7) because of its ability of approximation and solving combinatorial problems [29], [30], which will be presented in the next section.

III. COEXISTENCE PROTOCOL DESIGN VIA MARKOV APPROXIMATION

Our proposed solution framework contains three steps: At the first step, we should analyze the throughput maximization problem and decompose it into a master problem and several subproblems; The second step is to convert the multi-objective optimization into a single objective optimization problem by using the log-sum-exp approximation and then give the tradeoff analysis between throughput and fairness. At the third step, we should derive the coexistence protocol for different devices using the unlicensed bands.

A. PROBLEM DECOMPOSITION

Due to the difficulties of solving problem (7), we first assume that the time fractions are given, and then analyze the throughput objective. To gain insight on the structure of problem (7), we write the throughput objective function as

$$\begin{aligned} &w_1 \sum_{m=1}^{M} R_m + w_2 \sum_{l=1}^{L} R_l \\ &= w_1 \sum_{m=1}^{M} \sum_{k=1}^{K} W \log_2 \left(1 + \frac{p_m^k g_m^k}{N_0 W}\right) \\ &\quad + w_1 \sum_{m=1}^{M} t_m B \log_2 \left(1 + \frac{p_m^U h_m^U}{N_0 B}\right) + w_2 \sum_{l=1}^{L} t_l R_l^w. \end{aligned} \tag{8}$$

For the case of convenience, we can use the same index i to denote the device that contends to the unlicensed bands. Then, let $i \in \mathcal{I}$ and $\mathcal{I} = \mathcal{M} + \mathcal{L}$, the transmit power p_i^U and transmit rate R_i of device i can be rewritten as

$$p_i^U = \begin{cases} \sum_{pt_i \in \mathcal{P}_U} pt_i \dfrac{p_{max}^U}{\mathcal{P}_U - 1}, & i \in \mathcal{M}, \\ 0, & i \in \mathcal{L}. \end{cases} \tag{9}$$

$$R_i = \begin{cases} w_1 B \log_2 \left(1 + \dfrac{p_i^U h_i^U}{N_0 B}\right), & i \in \mathcal{M}, \\ w_2 R_l^w, & i \in \mathcal{L}. \end{cases} \tag{10}$$

Replacing the corresponding terms by Eqs. (9) and (10), the objective function and C5 can be rewritten as

$$w_1 \sum_{m=1}^{M} \sum_{k=1}^{K} W \log_2 \left(1 + \frac{p_m^k g_m^k}{N_0 W}\right) + \sum_{i \in \mathcal{I}} t_i R_i, \tag{11}$$

$$\begin{aligned} &\sum_{m=1}^{M} \sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} \\ &\qquad + \sum_{i \in \mathcal{I}} t_i \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \leq P_{max}. \end{aligned} \tag{12}$$

It can be easily seen that the throughput objective function in problem (7) is an increasing function of t_i. In order to maximize the throughput, the equality must be achieved for constraint C3, that is,

$$\sum_{i \in \mathcal{I}} t_i = \sum_{m=1}^{M} t_m + \sum_{l=1}^{L} t_l = 1, \forall t_i > 0. \tag{13}$$

Then, Eq. (12) can be rewritten as

$$\begin{aligned} &\sum_{m=1}^{M} \sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} + \sum_{i \in \mathcal{I}} t_i \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \\ &= (\sum_{i \in \mathcal{I}} t_i) \sum_{m=1}^{M} \sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} + \sum_{i \in \mathcal{I}} t_i \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \\ &= \sum_{i \in \mathcal{I}} t_i \left(\sum_{m=1}^{M} \sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} + \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \right) \\ &\leq P_{max}. \end{aligned} \tag{14}$$

Similarly, Eq. (11) can be rewritten as

$$\sum_{i \in \mathcal{I}} t_i \left(w_1 \sum_{m=1}^{M} \sum_{k=1}^{K} W \log_2 \left(1 + \frac{p_m^k g_m^k}{N_0 W}\right) + R_i \right). \tag{15}$$

Let U_i be the maximum system throughput utility under the time fraction t_i. Problem (7) can be decomposed into one master problem and I subproblems. The master problem can be formulated as

$$\begin{aligned} \max \quad & \sum_{i \in \mathcal{I}} t_i U_i \\ \text{s.t.} \quad & \sum_{i \in \mathcal{I}} t_i = 1, \quad \forall t_i > 0. \end{aligned} \tag{16}$$

Given the time faction t_i, U_i can be calculated in two cases. When $i \in \mathcal{M}$, U_i is the optimal value of the following subproblem

$$\begin{aligned} U_i = \max w_1 \sum_{m=1}^{M}\sum_{k=1}^{K} W \log_2\left(1+\frac{p_m^k g_m^k}{N_0 W}\right) \\ +w_1 B \log_2\left(1+\frac{p_i^U h_i^U}{N_0 B}\right) \\ \text{s.t. C1, C2,} \\ \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \leq P_{max}^U, \\ \sum_{m=1}^{M}\sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} \\ + \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \leq P_{max}. \end{aligned} \quad (17)$$

Similarly, when $i \in \mathcal{L}$, the subproblem can be formulated as

$$\begin{aligned} U_i = \max w_1 \sum_{m=1}^{M}\sum_{k=1}^{K} W \log_2\left(1+\frac{p_m^k g_m^k}{N_0 W}\right) + w_2 R_i^w \\ \text{s.t. C1, C2,} \\ \sum_{m=1}^{M}\sum_{k=1}^{K} x_m^k \sum_{px_m^k \in \mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L - 1} \\ + \sum_{pt_i \in \mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U - 1} \leq P_{max}. \end{aligned} \quad (18)$$

Based on the decomposition method above, the throughput maximization in problem (7) has been decomposed into one master problem and several subproblems. However, there are still many difficulties for solving these problems. On the one hand, both subproblems (17) and (18) are integer problems, which are difficult to solve in practical networks, especially for the case with a large numbers of subchannels and devices. On the other hand, it seems that the upper bound of problem (16) can be given by computing the max function: $\max_{i \in \mathcal{I}} U_i$, and make the corresponding t_i equal to 1. While, both the fairness objective function and the constraint $\forall t_i > 0$ make this solution infeasible. Thus, finding a feasible solution for problem (16) is a challenge task. To overcome this challenge, we will propose the Markov approximation framework for searching the bounded near optimal solution of problem (16) in the following subsections.

B. LOG-SUM-EXP APPROXIMATION

We first define a differentiable log-sum-exp function to gain insights on the structure of problem (16)

$$g_\beta(\mathbf{U}) \triangleq \frac{1}{\beta} \log\left(\sum_{i \in \mathcal{I}} \exp(\beta U_i)\right), \quad (19)$$

where $\mathbf{U} = [U_1, U_2, \cdots]$ and β is a positive constant.

Because the convexity of log-sum-exp, $g_\beta(\mathbf{U})$ is a convex function. Besides, $g_\beta(\mathbf{U})$ is a closed function due to its closed set domain. Then, we use this function to approximate the upper bound of problem (16) [31]:

$$g_\beta(\mathbf{U}) \approx U_{max} = \max_{i \in \mathcal{I}} U_i. \quad (20)$$

For the log-sum-exp function $g_\beta(\mathbf{U})$, its conjugate function can be computed as [31]

$$g_\beta^*(\boldsymbol{t}) = \begin{cases} \frac{1}{\beta}\sum_{i \in \mathcal{I}} t_i \log t_i, & \text{if } t_i > 0 \text{ and } \mathbf{1}^T \boldsymbol{t} = 1 \\ \infty, & \text{otherwise} \end{cases} \quad (21)$$

Because the log-sum-exp function $g_\beta(\mathbf{U})$ is convex and closed function, the conjugate of its conjugate $g_\beta^*(\boldsymbol{t})$ is itself [29]. Based the definition of the conjugate function,[1] we obtain that $g_\beta(\mathbf{U})$ is equivalent to solving the following optimization problem

$$\begin{aligned} \max \sum_{i \in \mathcal{I}} t_i U_i - \frac{1}{\beta} \sum_{i \in \mathcal{I}} t_i \log t_i \\ \text{s.t.} \sum_{i \in \mathcal{I}} t_i = 1, t_i > 0, \quad \forall i. \end{aligned} \quad (22)$$

Here, we observe that the first term of the objective function happens to be the weighted throughput in problem (16), and that the negative entropy function at the second term of the objective function, restricted to the probability simplex, happens to be the fairness objective in problem (7). Therefore, the multi-objective optimization in (7) is converted into a single objective optimization problem by using the log-sum-exp approximation method.

Then, by using the KKT conditions, we obtain the time fraction probability distribution $\mathbf{t}^*$, which is given as

$$t_i^*(U_i) = \frac{\exp(\beta U_i)}{\sum_{i' \in \mathcal{I}} \exp(\beta U_{i'})}, \quad \forall 1 \leq i \leq I. \quad (23)$$

It can be easily seen that even U_i equals to zero, its corresponding time fraction probability $t_i^*(U_i)$ is larger than zero. Hence, Eq. (23) guarantees that each device has the opportunity to access the unlicensed bands. However, Eq. (23) requires completeness, i.e. complete information on $\mathcal{I}$ which can be difficult to find in a practical wireless network due to the large solution space. More specifically, for obtaining its time fraction probability t_i, each device i must solve all of the feasible problems on $\mathcal{I}$, which is computationally exhaustive. However, based on the time-sharing of the unlicensed band according to their portion $t_i^*(U_i)$, we can obtain a useful insight on designing the distributed coexistence protocol in the following subsection. Moreover, the throughput and fairness tradeoff performance can be mathematically depicted by the following theorem.

[1] The conjugate function of $g(\mathbf{y})$ is defined as $g^*(\mathbf{z}) = \sup_{\mathbf{y} \in dom_g}(\mathbf{z}^T\mathbf{y} - g(\mathbf{y}))$.

Theorem 1: The log-sum-exp approximation scheme can achieve the following tradeoff performance between throughput and fairness, i.e.

$$|U(\mathbf{t}^*) - U_{max}| \leq \frac{1}{\beta}\log|\mathcal{I}|,$$
$$|F(\mathbf{t}^*) - F_{max}| \leq \beta(U_{max} - U_{min}),$$

where U_{min} is the lower bound of the system throughput.

Proof: See Appendix B. □

It shows that the throughput and fairness performance are proportional and conversely with the value of β, respectively. As a result, we can balance the throughput-fairness performance with the control parameter β and let the network operate in a predefined state.

C. MARKOV CHAIN DIRECTED ALGORITHM (MCDA)

The product form (23) can be considered as the distribution of a time-reversible Markov chain, where each state represents a certain unlicensed bands allocation. Hence, the idea of this subsection is to design a Markov chain with the state space being $\mathcal{I}$ and the stationary distribution being $\mathbf{t}^*$. Let state i represent the unlicensed band allocated to specified device i. All of the feasible states are contained in set $\mathcal{I}$. Let $q_{i,i'}$ and $q_{i',i}$ be the transition rates from $i \to i'$ and $i' \to i$, respectively. When the Markov chain converges to stationary, its probability distribution will be time-shared according to Eq. (23). Driven by Eq. (23), the states with high utilities will achieve high probabilities. As a result, these states will be maintained at most of the time. Therefore, the tradeoff performance will be achieved at the approximated optimality gap in **Theorem 1**. It was proven in [29] that there exist more than one time reversible ergodic Markov chains whose probability distribution can be represented as the product form of Eq. (23).

In order to construct a time-reversible Markov chain with stationary distribution in Eq. (23), the following two sufficient conditions should be followed [29].

- Any two states should be reachable from each other.
- The following balanced equation should be satisfied for all of the states. i.e. $t_i^*(U_i)q_{i,i'} = t_{i'}^*(U_{i'})q_{i',i}$, i.e. $\exp(\beta U_i)q_{i,i'} = \exp(\beta U_{i'})q_{i',i}$.

Based on the two conditions above, the transition rates can be designed [29] as correlated to the target system performance $U_{i'}$ under state i

$$q_{i,i'} = \exp(\beta U_{i'}), \tag{24}$$

where $q_{i',i}$ is defined symmetrically.

For achieving the transition rate in Eq. (24), we propose a Markov Chain Directed Access (MCDA) mechanism. As shown in Fig. 1, device i waits for a backoff time with mean equal to H_i. During the count down process, if the sDevice senses another interfering device is in transmission, it will freeze the countdown process. When the transmission is over, the sDeivce restart the countdown process according to the residual back-off time. Based on the state transition of MCDA, the probability that the user will transmit its packet can be computed as [32]

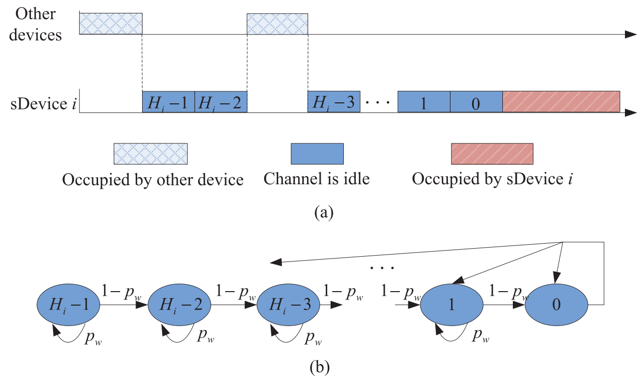

FIGURE 1. (a) Procedure of the MCDA. (b) State transition graph of MCDA.

$$\tau_i = \frac{2}{(H_i + 1)}. \tag{25}$$

Then, for making the MCDA converges to the stationary distribution in Eq. (23), we give the following theorem.

Theorem 2: If the backoff window satisfies

$$H_i = \frac{2}{(M+L)\exp(\beta U_i)} - 1$$

where $\beta \leq \frac{1}{U_{max}}\log\frac{2}{M+L}$, the MCDA mechanism can realize the transition rate in Eq. (24) and achieve a time-reversible Markov chain with stationary distribution given in Eq. (23).

IV. LOW-COMPLEXITY SUBOPTIMAL RESOURCE ALLOCATION

Recalled that the resource allocations in (17) and (18) are mixed-integer programmings, which are NP-hard problems. To tackle its computational complexity, we try to separate the problem into two subproblems and adopt a two-step-iterative approach (TSIA) to solve them. At the first step, under the assumption of equal power allocation on each subchannel, we should solve the subchannel allocation problem. Then, the derived subchannel allocation solution is used for searching the corresponding power control solution by **Algorithm 1** at the second step. Steps 1 and 2 are iteratively executed until the current subchannel and power control solutions are not much different from their values in the previous iteration. In other words, the iterative procedure can be given as

$$\underbrace{x[0] \to pt[0], px[0]}_{\text{initialization}} \to \cdots \to \underbrace{x[\iota] \to pt[\iota], px[\iota]}_{\text{Iteration t}} \to \cdots \to \underbrace{x^* \to pt^*, px^*}_{\text{Optimal solution}}, \tag{26}$$

where $\iota > 0$ is the iteration index. Obviously, this method can reduce the number of variables by half in each optimization subproblem and reduce the computational complexity of the original problems [33], [34]. However, it should be noted that both the subchannel allocation and power control problems are still integer programming, both of which still suffer high computational complexity. In the following, we will develop the relaxation and greedy algorithms for solving them respectively.

A. SUBCHANNEL ALLOCATION

The subchannel allocation problems are integer problems, which are difficult to solve in practical setting with a large set of sDevices and subchannels. In the following, we will relax the problem into continuous linear programming and employ the Lagrangian relaxation method for solving problem (17). Fortunately, the special structure of our optimization problem allows us to derive an optimal solution. Moveover, the similar method can be directly used to solve problem (18).

By relaxing the integer variable x into continuous interval [0, 1], we can rewritten problem (17) as

$$\begin{aligned}
\max w_1 & \sum_{m=1}^{M}\sum_{k=1}^{K} W x_m^k \log_2\left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right) \\
& + w_1 B\log_2\left(1+\sum_{pt_i\in\mathcal{P}_U} pt_i \frac{p_{max}^U}{\mathcal{P}_U-1}\frac{h_i^U}{N_0B}\right) \\
\text{s.t.} & \sum_{k=1}^{K} W x_m^k \log_2 \\
& \left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right)\geq R_m^{req}, \forall m, \\
& \sum_{m=1}^{M} x_m^k \leq 1, x_m^k\in[0,1], \sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}\leq P_{max}^U, \\
& \sum_{m=1}^{M}\sum_{k=1}^{K} x_m^k \sum_{px_m^k\in\mathcal{P}_L} px_m^k \frac{p_{max}}{\mathcal{P}_L-1} + \sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}\leq P_{max}.
\end{aligned} \tag{27}$$

Since all of the constraints are linear functions, the Slater's condition is always satisfied for the problem above, hence it leads to a zero Lagrange duality gap [31]. We relax the QoS constraint and the total transmit power constraint by introducing the dual variables λ_m and μ, obtaining the following Lagrangian:

$$\begin{aligned}
L(\mathbf{x},\boldsymbol{\lambda},\mu) = & \sum_{m=1}^{M}\sum_{k=1}^{K}\Delta_m^k x_m^k - \sum_{m=1}^{M}\lambda_m R_m^{req} \\
& -\mu\sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}+\mu P_{max},
\end{aligned} \tag{28}$$

where Δ_m^k is represented as

$$\begin{aligned}
&\Delta_m^k(\boldsymbol{\lambda},\mu) \\
&= -\mu\sum_{m=1}^{M}\sum_{k=1}^{K} x_m^k \sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1} \\
&\quad + w_1\sum_{m=1}^{M}\sum_{k=1}^{K} W\log_2\left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right) \\
&\quad + \sum_{m=1}^{M}\sum_{k=1}^{K} W\lambda_m\log_2\left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right).
\end{aligned} \tag{29}$$

Then, the corresponding Lagrange dual is

$$\begin{aligned}
d(\boldsymbol{\lambda},\mu) = \max & \sum_{m=1}^{M}\sum_{k=1}^{K}\Delta_m^k(\boldsymbol{\lambda},\mu)x_m^k - \sum_{m=1}^{M}\lambda_m R_m^{req} \\
& -\mu\sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}+\mu P_{max} \\
\text{s.t.} & \sum_{m=1}^{M} x_m^k\leq 1, x_m^k\in[0,1].
\end{aligned} \tag{30}$$

The dual problem of (27) is

$$\begin{aligned}
& \min d(\boldsymbol{\lambda},\mu) \\
& \text{s.t. } \mu>0, \quad \lambda_m>0, \ \forall m.
\end{aligned} \tag{31}$$

Due to its convexity, the optimal values μ^* and $\boldsymbol{\lambda}^*$ of dual problem (31) can be solved by using the gradient descent method. Given the optimal values μ^* and $\boldsymbol{\lambda}^*$, subchannel allocation solutions can be obtained from the classical linear assignment problem in (31). It's not difficult to understand that the optimal solution can be given as

$$x_m^k = \begin{cases} 1, & m = \arg\max_m\{\Delta_m^k(\boldsymbol{\lambda},\mu), \forall 1\leq m\leq M\}, \\ 0, & \text{otherwise.}\end{cases} \tag{32}$$

Eq. (32) shows that the optimal solution of the relaxed problem is still integer. That is to say, the optimal solution of the relaxed continuous linear programming problem in (28) is the same as that of the integer problem in (17).

B. SUBOPTIMAL POWER CONTROL

With the subchannel allocation results, the power control problem can be given as

$$\begin{aligned}
\max w_1 & \sum_{m=1}^{M}\sum_{k\in\mathcal{K}_m} W\log_2\left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right) \\
& + w_1 B\log_2\left(1+\sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}\frac{h_i^U}{N_0B}\right) \\
\text{s.t.} & \sum_{k\in\mathcal{K}_m} W\log_2\left(1+\sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1}\frac{g_m^k}{N_0W}\right)\geq R_m^{req}, \forall m, \\
& \sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}\leq P_{max}^U, \\
& \sum_{m=1}^{M}\sum_{k\in\mathcal{K}_m}\sum_{px_m^k\in\mathcal{P}_L} px_m^k\frac{p_{max}}{\mathcal{P}_L-1}+\sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p_{max}^U}{\mathcal{P}_U-1}\leq P_{max}.
\end{aligned} \tag{33}$$

Power allocation problem (33) with discrete power level is proved to be NP-hard [18], [35]. In the following, we will propose a greedy algorithm for solving this problem.

The lagrangian of objective function (33) associated with its QoS constraint can be written as

$$
\begin{aligned}
&L(\mathbf{px}, pt_i, \boldsymbol{\lambda})\\
&= w_1 B \log_2\left(1+\sum_{pt_i\in\mathcal{P}_U} pt_i \frac{p^U_{max}}{\mathcal{P}_U-1}\frac{h^U_i}{N_0 B}\right)\\
&\quad + w_1 \sum_{m=1}^{M}\sum_{k\in\mathcal{K}_m} W\log_2\left(1+\sum_{px^k_m\in\mathcal{P}_L} px^k_m \frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0 W}\right)\\
&\quad -\sum_{m=1}^{M}\lambda_m\left(R^{req}_m - \sum_{k\in\mathcal{K}_m} W\log_2\right.\\
&\qquad \left.\times\left(1+\sum_{px^k_m\in\mathcal{P}_L} px^k_m\frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0 W}\right)\right). \qquad (34)
\end{aligned}
$$

The derivative of Eq. (34) with respect to pt_i and px^k_m can be given as

$$
\frac{\partial L(\mathbf{px}, pt_i, \boldsymbol{\lambda})}{\partial pt_i} = \frac{w_1 \frac{p^U_{max}}{\mathcal{P}_U-1}\frac{h^U_i}{N_0}}{\left(1+\sum_{pt_i\in\mathcal{P}_U} pt_i\frac{p^U_{max}}{\mathcal{P}_U-1}\frac{h^U_i}{N_0B}\right)\ln 2}, \qquad (35)
$$

$$
\frac{\partial L(\mathbf{px}, pt_i, \boldsymbol{\lambda})}{\partial px^k_m} = \frac{(w_1+\lambda_m)\frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0}}{\left(1+\sum_{px^k_m\in\mathcal{P}_L} px^k_m\frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0W}\right)\ln 2}. \qquad (36)
$$

Eqs. (35) and (36) represent the influence of power increase on the system throughput. Thus, both of them can be regarded as the greedy rule in the greedy algorithm. Besides, from the *Karush-Kuhn-Tucker (KKT)* conditions, we also have

$$
\lambda_m\left(R^{req}_m - \sum_{k\in\mathcal{K}_m} W\log_2\left(1+\sum_{px^k_m\in\mathcal{P}_L} px^k_m\frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0W}\right)\right) = 0. \qquad (37)
$$

Thus, we can set $\lambda_m = 0$ if

$$
\sum_{k\in\mathcal{K}_m} W\log_2\left(1+\sum_{px^k_m\in\mathcal{P}_L} px^k_m\frac{p_{max}}{\mathcal{P}_L-1}\frac{g^k_m}{N_0W}\right) \le R^{req}_m. \qquad (38)
$$

On the other hand, if the transmission rate of sDevice can satisfy its basic QoS requirement, we will set $\lambda_m = -w_1$ to make the greedy rule in Eq. (36) become 0. As a result, the SBS will no longer allocate transmit power to this sDevice. However, if all of the sDevices QoS requirements are satisfied and that the total transmit power has not been run out, we will reset λ_m as 0 to restart the power allocation process for improving the system throughput.

Then, the greedy algorithm is summarized in **Algorithm 1**, where f^k_m and l are the indicating variables to denote whether or not the transmit power px^k_m and pt_i have achieved their corresponding maximum transmit power p_{max} and p^U_{max}.

Algorithm 1 Suboptimal Power Control (SPC)

1 Initialization: $P = P_{max}$; $\lambda_m = 0, \forall m$; $f^k_m = 1, \forall m, k$;
2 $l = 1$; $\Phi_{(M+1)\times K} = 0$;
3 Calculate $\frac{\partial L}{\partial pt_i}$ and $\frac{\partial L}{\partial px^k_m}$, $\forall m, k$;
4 Let $\Phi(m, k) = f^k_m \cdot \frac{\partial L}{\partial px^k_m}$, $\forall m, k$; $\Phi(M+1, 1) = l \cdot \frac{\partial L}{\partial pt_i}$;
5 **while** $P > 0$ **do**
6 $\quad (u^*, v^*) = \arg\max \Phi(m, k)$;
7 $\quad$ **if** $u^* \le M$ **then**
8 $\quad\quad px^{v^*}_{u^*} = px^{v^*}_{u^*} + \frac{p_{max}}{\mathcal{P}_L-1}$;
9 $\quad\quad$ **if** $px^{v^*}_{u^*} >= p_{max}$ **then**
10 $\quad\quad\quad f^k_m = 0$;
11 $\quad\quad$ **end**
12 $\quad\quad P = P - \frac{p_{max}}{\mathcal{P}_L-1}$;
13 $\quad$ **else**
14 $\quad\quad pt_i = pt_i + \frac{p^U_{max}}{\mathcal{P}_U-1}$;
15 $\quad\quad$ **if** $pt_i >= p^U_{max}$ **then**
16 $\quad\quad\quad l = 0$;
17 $\quad\quad$ **end**
18 $\quad\quad P = P - \frac{p^U_{max}}{\mathcal{P}_U-1}$;
19 $\quad$ **end**
20 $\quad$ Calculate R^1_m, $\forall m$;
21 $\quad$ **for** $m = 1$ *to* M **do**
22 $\quad\quad$ **if** $R^1_m >= R^{req}_m$ **then**
23 $\quad\quad\quad \lambda_m = -w1$;
24 $\quad\quad$ **end**
25 $\quad$ **end**
26 $\quad$ **if** $\lambda_m = -w1, \forall m$ **then**
27 $\quad\quad \lambda_m = 0$;
28 $\quad$ **end**
29 $\quad$ Calculate $\frac{\partial L}{\partial pt_i}$ and $\frac{\partial L}{\partial px^k_m}$, $\forall m, k$; Update Φ;
30 **end**

C. COMPUTATIONAL COMPLEXITY OF RESOURCE ALLOCATION

In the subchannel allocation problem, the subgradient method is utilized to solve dual problem (31). For achieving δ-optimality, i.e., $|d - d^*| < \delta$, the number of iterations is on the order of $O(\frac{1}{\delta^2})$ [36]. In each iteration, it is required to compute Eq. (32) for K subchannels. The computational complexity of Eq. (32) is $O(M)$. Consequently, the total computational complexity of subchannel allocation is $O(\frac{MK}{\delta^2})$. In suboptimal power control algorithm, Step 5 requires a complexity of $O(|\mathcal{P}_l|+|\mathcal{P}_u|)$ for the *while loop*. Step 6 needs a complexity of $O(MK)$ for finding the optimal (u^*, v^*). Step 21 requires a complexity of $O(M)$ to evaluate the *for* loop. Step 26 requires a complexity of $O(M)$ for *if* loop. Hence, the worst case computational complexity of SPC is calculated as $O((|\mathcal{P}_l| + |\mathcal{P}_u|)(MK + 2M))$. Then, the total complexity at each step of Eq. (26) is $O(\frac{MK}{\delta^2} + (|\mathcal{P}_l| + |\mathcal{P}_u|)(MK + 2M))$.

TABLE 1. Simulation Parameters.

Parameters	Values
Number of sDevice, M	10
Number of wDevice, L	4
Number of subchannel, K	20
Total transmit power on licensed bands, P_{max}	37dBm
Maximum transmit power for each subchannel, p_{max}	24dBm
Total transmit power on unlicensed bands, P^U_{max}	33dBm
The size of power level set on licensed bands, $\|\mathcal{P}_L - 1\|$	5
The size of power level set on unlicensed bands, $\|\mathcal{P}_U - 1\|$	10
Unlicensed bandwidth, B	20MHz
Licensed bandwidth, W	10MHz
sDevice basic QoS requirement, $R^{req}_m, \forall m$	10Kbps
AWGN noise power, N_0	-174dBm
Path loss model on licensed band (dB)	$-15.3 - 37.6\log_{10}(D)$
Path loss model on unlicensed band (dB)	$-15.3 - 50\log_{10}(D)$

Considering ε as the accuracy of the two step iterative algorithm in Eq. (26), solving the resource allocation will take $O\left(\log(\frac{1}{\varepsilon})\left(\frac{MK}{\delta^2} + (|\mathcal{P}_l| + |\mathcal{P}_u|)(MK + 2M)\right)\right)$.

V. SIMULATION RESULTS

In this section, the simulation results are presented to verify the proposed algorithm. We assume that there is an incumbent Wi-Fi network and a SBS located at (150m, 250m) and (200m, 200m) in a square with size 400m × 400m. The sDevices and wDevices are randomly distributed in the circles of SBS and Wi-Fi with radius of 150m and 100m, respectively. The major simulation parameters are initially set as in Table 1. The channel fading models in licensed and unlicensed bands are based on Refs [19] and [26], where PL is in decibels and D is in meters. We evaluate the proposed schemes from the following aspects. First, the convergence and throughput performance of two-step iterative approach (TSIA) in Seciton IV are compared with the GreHybrid algorithm [37], whose the subchannel allocation is converted into a tentative-cluster construction, and then the subchannel allocation solution is used for solving the power control problem. Second, we illustrate the effects of control parameter β on the tradeoff performance of system throughput and fairness. Finally, we compare our proposed MCDA algorithm with the non-adaptive channel access scheme (NAS), where all of the devices use the same backoff window size.

Fig. 2 illustrates the convergence of the suboptimal power control in **Algorithm 1**. In this experiment, we assume that both the power level sets on licensed and unlicensed bands are set as P_L. Moreover, we evaluate the performance of our proposed suboptimal power control through numerical comparisons with the continuous power allocation. With the increase of the number of power levels available, we observe that the performance of our proposed suboptimal power control increases and that the gap between the discrete power allocation and the continuous one becomes smaller. That is because the large number of power levels can improve the accuracy of power control, hence the performance can be improved. However, the figure also shows that the convergence rate of the suboptimal power control in **Algorithm 1** decreases with the increase of power levels. That is because at each iteration, the SBS can only allocate at most $p_{max}/(P_L - 1)$ or $p^U_{max}/(P_L - 1)$ transmit power, so it will need more iterations for allocating all of the increased transmit power.

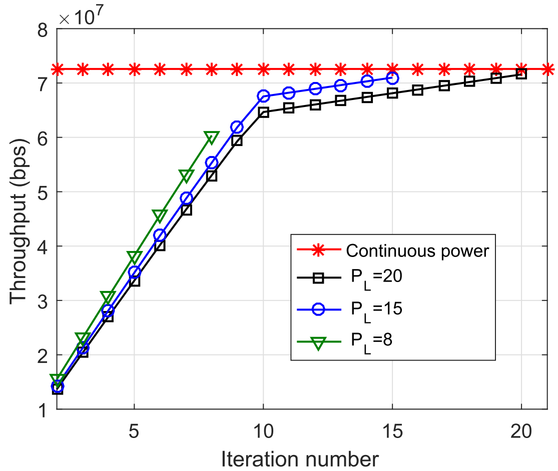

FIGURE 2. Convergence rate of suboptimal power control.

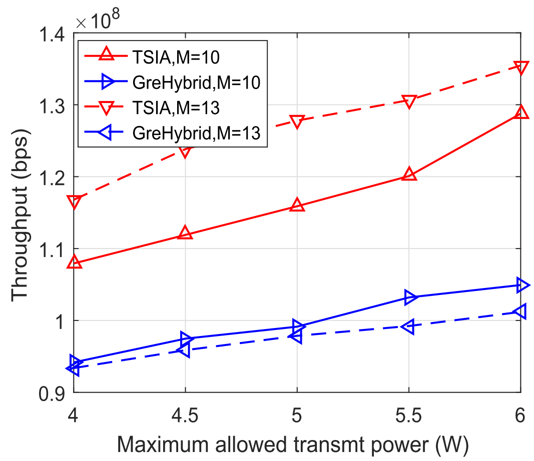

FIGURE 3. Throughput comparison of TSIA and GreHybrid algorithms.

Fig. 3 compares the proposed low-complexity suboptimal resource allocation with the GreHybrid algorithm. It shows that our proposed algorithm performs much better than the GreHybrid algorithm. We also observe that the throughput under both TSIA algorithm and GreHybrid algorithm increases with the increase of maximum allowed transmit power. However, with the increase of devices number, TSIA algorithm has a lager performance gain than GreHybrid algorithm. That is because the optimal subchannel allocation can be achieved at the first step of TSIA algorithm. While, the GreHybrid algorithm can only obtain the suboptimal subchannel allocation by constructing the tentative-cluster, so that its performance is affected by the device number.

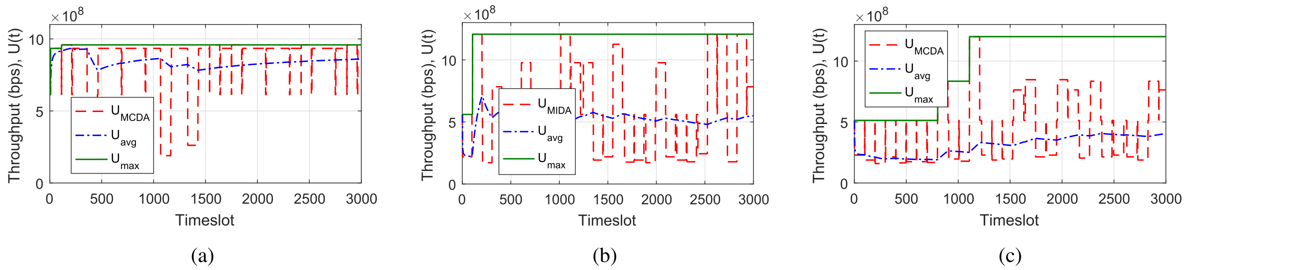

FIGURE 4. Changes in throughput value with elapsed time t. (a) Real-time throughput, $\beta = 10^{-8}$. (b) Real-time throughput, $\beta = 10^{-9}$. (c) Real-time throughput, $\beta = 10^{-10}$.

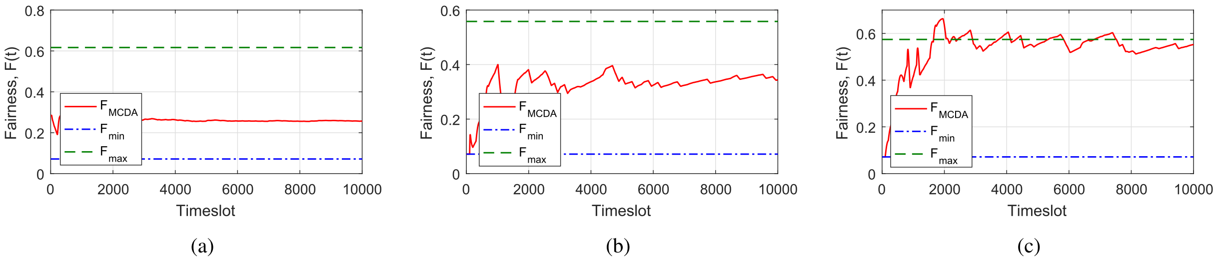

FIGURE 5. Changes in fairness value with elapsed time t. (a) Real-time fairness value, $\beta = 10^{-8}$. (b) Real-time fairness value, $\beta = 10^{-9}$. (c) Real-time fairness value, $\beta = 10^{-10}$.

The tradeoff performance between throughput and fairness is illustrated in Figs. 4 and 5. In these figures, U_{MCDA} refers to the instantaneous throughput value, U_{avg} is the average throughput value at each timeslot and $U_{max} = \max_{i \in \mathcal{I}} U_i$ is the upper bound of system throughput. F_{MCDA} is the average fairness value at each timeslot and F_{max} is defined based the optimal solution in **Lemma 1**. Fig. 4 shows the throughput value as the function of elapsed time t. It is easy to understand that each "sawtooth" of the U_{MCDA} curve corresponds to a situation that a specified device has accessed the unlicensed bands. The curve shows that different devices can competitively access the unlicensed bands along the operating of MCDA. We observe from these figures that the achieved average system throughput decreases with the decrease of the control parameter β. It shows that, when the control parameter β is large, the device with large throughput values are more likely to be selected. However, in this case, the system fairness will be reduced as shown in Fig. 5. When the control parameter is small, the fairness can achieve a very high value at a great sacrifice of the system throughput. Therefore, Figs. 4 and 5 show that our proposed algorithm achieves a tradeoff between system throughput and fairness. Therefore, we can obtain a significant rule for engineering design to flexibly balance the system throughput and fairness performance. We only need to adjust appropriately the control parameter β to let the network operate in a predefined state.

To explicitly show the tradeoff performance of sDevices and wDevices throughput versus the weight factor, $\alpha = w_2/w_1$, we demonstrate the corresponding throughput performance in Fig. 6. It shows that sDevices' throughput decreases with the increase of weight factor α, while wDevices throughput increases correspondingly. From these

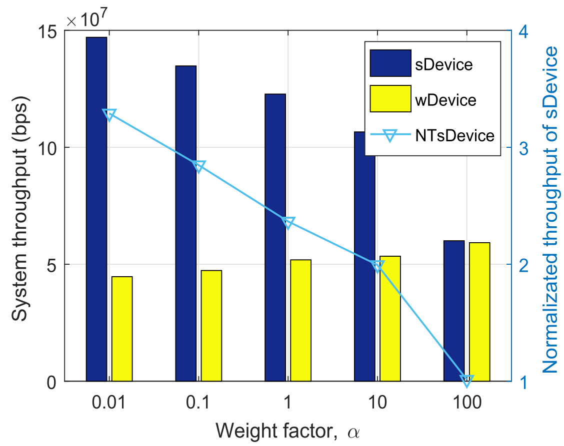

FIGURE 6. Tradeoff performance of sDevices and wDevices throughput versus the weight factor, $\alpha = w_2/w_1$.

observations, we can conclude that our proposed MCDA algorithm can achieve the tradeoff between sDevices and wDevices throughput through adjusting the weight factor α. If the throughput on sDevices is more important than it on wDevices, we can set higher w_1 than w_2. On the contrary, we set w_2 higher than w_1.

Finally, we compare our proposed MCDA algorithm with the non-adaptively channel access scheme (NAS) in Figs. 7-9. Fig. 7 shows the throughputs of sDevices under MCDA and NAS versus sDevice number. We observe that the throughputs of SBS under both schemes increase with the increase of sDevice number. On the contrary, as shown in Fig. 8, the throughput of wDevices decreases with the number of sDevice. That is because, with the increase of sDeivce number, the unlicensed bands utilization in the SBS increases,

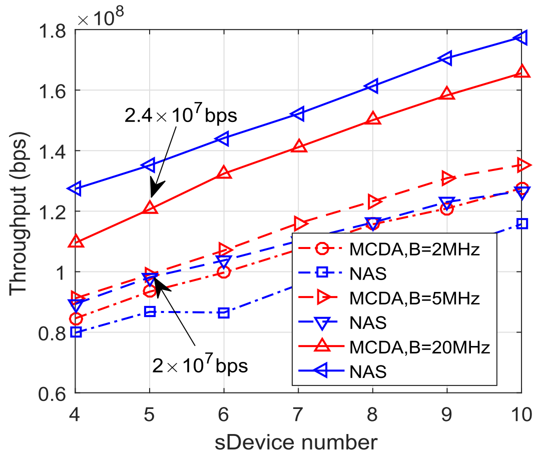

FIGURE 7. Throughput of sDevices under MCDA and NAS versus sDevice number.

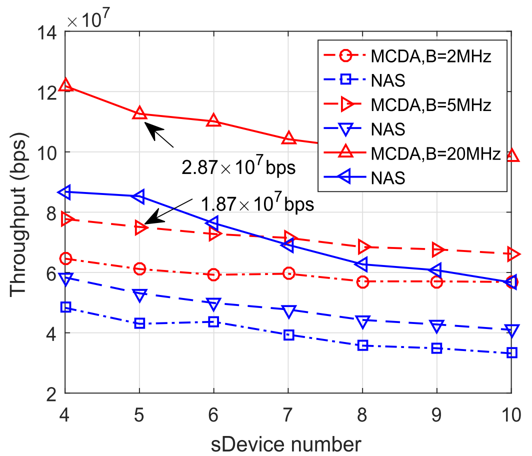

FIGURE 8. Throughput of wDevices under MCDA and NAS versus sDevice number.

so that it will bring more user diversity gain. However, due to the fact the increased sDeivces seize more unlicensed bands resource, so it will result in lower wDevices throughput as shown in Fig. 8. We also observe in Figs. 7 and 8 that MCDA algorithm achieves a larger throughput than NAS when the maximum allowed unlicensed bands is small. However, with the increase of unlicensed bands, Fig. 8 shows that NAS algorithm achieves a larger throughput than MCDA. That is because our proposed MCDA algorithm has an adaptive exponential backoff scheme, where the window size can be adaptively adjusted based on the available device throughput utility. When the maximum allowed unlicensed bands is small, per sDevice's throughput (e.g. 2×10^7bps in Fig. 7) is larger than per wDevice's throughput (e.g. 1.87×10^7bps in Fig. 8), hence the sDevices will seize more unlicensed bands, which results in larger throughput than NAS. When the maximum allowed unlicensed bands becomes larger, wDevice has a larger throughput (e.g. 2.87×10^7bps) than sDevice (e.g. 2.4×10^7bps). Therefore, wDevices will occupy the unlicensed bands for a longer time than the sDevices. On the contrary, under NAS algorithm, SBS greedily occupy more

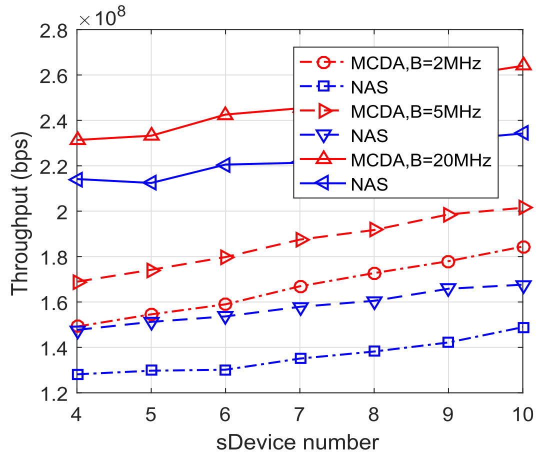

FIGURE 9. System throughput under MCDA and NAS versus sDevice number.

unlicensed bands than wDevices, which results a lower system throughput as shown in Fig. 9. It is also shown in Fig. 9 that, with any sDevice number and unlicensed bands, our proposed MCDA algorithm always achieves a larger throughput than NAS algorithm, which shows the effectiveness of our proposed algorithm.

VI. CONCLUSION

In this paper, we consider the coexistence protocol design and resource allocation of licensed and unlicensed bands for a LTE-U system, where the LTE-U SBS deployed with both licensed and unlicensed spectrum. We propose a fair coexistence mechanism, where the transmit power, licensed and unlicensed spectrum are jointly considered in the designing of channel access and resource allocation. Wherein, an adaptive exponential backoff scheme is developed for the channel access of the unlicensed spectrum. It should be noted that the backoff window size can be adaptively adjusted based on the available device throughput utility. Then, a low complexity two iterative optimization procedure is developed for the resource allocation subproblem. Finally, the tradeoff between network throughput and fairness is revealed. Only by adjusting a control parameter, the proposed backoff scheme provides a controllable method to balance the tradeoff performance and let the network operate in a predefined state.

APPENDIX A PROOF OF LEMMA1

Proof: For maximizing the system throughput, the following equality must be achieved

$$\sum_{i \in \mathcal{I}} t_i = \sum_{m=1}^{M} t_m + \sum_{l=1}^{L} t_l = 1, \quad \forall t_i > 0.$$

Then, we obtain the following optimization problem for maximizing the fairness

$$\max -\sum_{i \in \mathcal{I}} t_i \log t_i, \;\; \text{s.t.} \sum_{i \in \mathcal{I}} t_i = 1, \quad \forall t_i > 0.$$

By using the KKT conditions, we have

$$\sum_{i\in\mathcal{I}} t_i^* - 1 = 0; \log t_i^* + 1 + \lambda^* = 0, \quad \forall i \in \mathcal{I}.$$

Through solving the equations above, we obtain $t_i^* = \frac{1}{M+L}, \forall i \in \mathcal{I}$. □

APPENDIX B PROOF OF THEOREM 1

Proof: Followed from Section 3.1.5 in [31], we have

$$\max_{i\in\mathcal{I}} U_i \le \frac{1}{\beta}\log\left(\sum_{i\in\mathcal{I}}\exp(\beta U_i)\right) \le \max_{i\in\mathcal{I}} U_i + \frac{1}{\beta}\log \mathcal{I}.$$

Then, we can obtain

$$|g_\beta(\mathbf{U}) - U_{max}| = |U(\mathbf{t}^*) - U_{max}| \le \frac{1}{\beta}\log|\mathcal{I}|.$$

Lemma 1: has concluded that the maximum fairness criterion can be achieved at the equal distribution of the unlicensed bands, i.e. $t_m, t_l = \frac{1}{M+L}, \forall m, l$. By substituting Eq. (23) into the fairness criterion, we can obtain

$$\begin{aligned} &F(\mathbf{t}^*) - F_{max} \\ &= -\frac{1}{\beta}\sum_{i\in\mathcal{I}}\frac{\exp(\beta U_i)}{\sum_{i'\in\mathcal{I}}\exp(\beta U_{i'})}\log\frac{\exp(\beta U_i)}{\sum_{i'\in\mathcal{I}}\exp(\beta U_{i'})} \\ &\quad + \sum_{i\in\mathcal{I}}\frac{1}{M+L}\log\frac{1}{M+L} \\ &= \log\frac{1}{M+L} - \frac{\sum_{i\in\mathcal{I}}\beta U_i\exp(\beta U_i)}{\sum_{i\in\mathcal{I}}\exp(\beta U_i)} + \log\sum_{i\in\mathcal{I}}\exp(\beta U_i). \end{aligned}$$

Thus,

$$\begin{aligned} |F(\mathbf{t}^*) - F_{max}| &\le \log\frac{1}{M+L} - \frac{\sum_{i\in\mathcal{I}}\beta U_{min}\exp(\beta U_i)}{\sum_{i\in\mathcal{I}}\exp(\beta U_i)} \\ &\quad + \log((M+L)\exp(\beta U_{max})) \\ &= \beta(U_{max} - U_{min}). \end{aligned}$$

□

APPENDIX C PROOF OF THEOREM 2

Proof: By the two degrees of freedom of constructing the designed Markov chain, we see that all of the states can reach each other within a finite number transitions. Let $Pr_{i'\leftarrow i}$ denote the probability that the system comes from state i when it enter state i'. In the MCDA algorithm, the pervious state has equal probability to be any state $i \in \mathcal{I}$. Then, we have

$$Pr_{i'\leftarrow i} = \frac{1}{|\mathcal{I}|} = \frac{1}{M+L}.$$

Based on Wq. (25), device i' transmit its packet with $\tau_{i'}$. Therefore, we calculate the transition rate from i to i' as follows

$$q_{i,i'} = Pr_{i'\leftarrow i}\tau_{i'} = \frac{2}{(M+L)(W_{i'}+1)}.$$

Letting $W_{i'} = \frac{2\exp(-\beta U_{i'})}{M+L} - 1$, we can compute the above equation as

$$q_{i,i'} = \exp(\beta U_{i'})$$

Based on the balanced equation, we obtain

$$t_{i'} = t_i\frac{\exp(\beta U_{i'})}{\exp(\beta U_i)}.$$

Due to the fact that $\sum_{i'\in\mathcal{I}} t_{i'} = 1$, we obtain

$$1 = \sum_{i'\in\mathcal{I}} t_i\frac{\exp(\beta U_{i'})}{\exp(\beta U_i)}.$$

Thus

$$t_i = \frac{\exp(\beta U_i)}{\sum_{i'\in\mathcal{I}}\exp(\beta U_{i'})}, \quad \forall 1 \le i \le I.$$

Finally, for making the backoff window positive, the control parameter should satisfies

$$\beta \le \frac{1}{U_{max}}\log\frac{2}{M+L}.$$

□

REFERENCES

[1] J. Xiao and J. Zheng, "An adaptive channel access mechanism for LTE-U and WiFi coexistence in an unlicensed spectrum," in *Proc. IEEE ICC*, May 2016, pp. 1–6.
[2] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen, and L.-L. Xie, "LTE-unlicensed: The future of spectrum aggregation for cellular networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 150–159, Jun. 2015.
[3] F. M. Abinader *et al.*, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 54–61, Nov. 2014.
[4] Y. Su, X. Du, L. Huang, Z. Gao, and M. Guizani, "LTE-U and Wi-Fi coexistence algorithm based on Q-learning in multi-channel," *IEEE Access*, vol. 6, pp. 13644–13652, 2018.
[5] X. Ying, R. Poovendran, and S. Roy, "Detecting LTE-U duty cycling misbehavior for fair sharing with Wi-Fi in shared bands," in *Proc. IEEE PIMRC*, Oct. 2017, pp. 1–7.
[6] A. D. Shoaei, M. Derakhshani, T. Le-Ngoc, and M. Salem, "Efficient LTE/WiFi coexistence in unlicensed spectrum using virtual network entity," in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1–6.
[7] F. Liu, E. Erkip, M. C. Beluri, R. Yang, and E. Bala, "Dual-band femtocell traffic balancing over licensed and unlicensed bands," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 6809–6814.
[8] *Broadband Radio Access Networks (BRAN); 5 GHz High Performance RLAN; Harmonized EN Covering the Essential Requirements of Article 3.2 of the RTTE Directive*, document ETSI EN 301 893 V1.7.2, Jun. 2012.
[9] F. Liu, E. Bala, E. Erkip, and R. Yang, "A framework for femtocells to access both licensed and unlicensed bands," in *Proc. Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw.*, May 2011, pp. 407–411.
[10] M. Ali, S. Qaisar, M. Naeem, and S. Mumtaz, "Joint user association and power allocation for licensed and unlicensed spectrum in 5G networks," in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1–6.
[11] H. Zhang, B. Wang, K. Long, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation in heterogeneous small cell networks with WiFi spectrum sharing," in *Proc. GLOBECOM*, Dec. 2017, pp. 1–5.
[12] Q. Wei, L. Wang, Z. Feng, and Z. Ding, "Wireless resource management in LTE-U driven heterogeneous V2X communication networks," *IEEE Trans. Veh. Technol*, vol. 67, no. 8, pp. 7508–7522, Aug. 2018.
[13] S. Sagari, S. Baysting, D. Saha, I. Seskar, W. Trappe, and D. Raychaudhuri, "Coordinated dynamic spectrum management of LTE-U and Wi-Fi networks," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Sep. 2015, pp. 209–220.
[14] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution*. Hoboken, NJ, USA: Wiley, 2009.

[15] S. Li, Z. Shao, and J. Huang, "ARM: Anonymous rating mechanism for discrete power control," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 326–340, Feb. 2017.
[16] F. Paganini, M. Zubeldía, and A. Ferragut, "Reciprocity and efficiency in peer exchange of wireless nodes through convex optimization," *IEEE Trans. Netw. Sci. Eng.*, vol. 3, no. 4, pp. 257–270, Oct./Dec. 2016.
[17] T. Z. Oo, N. H. Tran, W. Saad, D. Niyato, Z. Han, and C. S. Hong, "Offloading in HetNet: A coordination of interference mitigation, user association, and resource allocation," *IEEE Trans. Mobile Comput.*, vol. 16, no. 8, pp. 2276–2291, Aug. 2017.
[18] H. Zhang, L. Venturino, N. Prasad, P. Li, S. Rangarajan, and X. Wang, "Weighted sum-rate maximization in multi-cell networks via coordinated scheduling and discrete power control," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1214–1224, Jun. 2011.
[19] 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, Oct. 2016.
[20] C. Pan, C. Yin, N. C. Beaulieu, and J. Yu, "Distributed resource allocation in SDCN-based heterogeneous networks utilizing licensed and unlicensed bands," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 711–721, Feb. 2018.
[21] *IEEE Standard for Information Technology–Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks–Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11-2016 (Revision IEEE Standard 802.11-2012), Dec. 2016, pp. 1–3534.
[22] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.
[23] H. Shi, R. V. Prasad, E. Onur, and I. G. M. M. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 5–24, 1st Quart., 2014.
[24] M. Uchida and J. Kurose, "An information-theoretic characterization of weighted alpha-proportional fairness," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 1053–1061.
[25] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *J. Stat. Phys.*, vol. 52, nos. 1–2, pp. 479–487, 1988.
[26] R. Yin, G. Yu, A. Maaref, and G. Y. Li, "A framework for co-channel interference and collision probability tradeoff in LTE licensed-assisted access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6078–6090, Sep. 2016.
[27] U. Challita, L. Dong, and W. Saad, "Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4674–4689, Jul. 2018.
[28] H. He, H. Shan, A. Huang, L. X. Cai, and T. Q. S. Quek, "Proportional fairness-based resource allocation for LTE-U coexisting with Wi-Fi," *IEEE Access*, vol. 5, pp. 4720–4731, 2017.
[29] M. Chen, S. C. Liew, Z. Shao, and C. Kai, "Markov approximation for combinatorial network optimization," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6301–6327, Oct. 2013.
[30] S. M. A. Kazmi *et al.*, "Mode selection and resource allocation in device-to-device communications: A matching game approach," *IEEE Trans. Mobile Comput.*, vol. 16, no. 11, pp. 3126–3141, Nov. 2017.
[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[32] 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.
[33] D. T. Ngo, S. Khakurel, and T. Le-Ngoc, "Joint subchannel assignment and power allocation for OFDMA femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 342–355, Jan. 2014.
[34] S. Parsaeefard, R. Dawadi, M. Derakhshani, and T. Le-Ngoc, "Joint user-association and resource-allocation in virtualized wireless networks," *IEEE Access*, vol. 4, no, pp. 2738–2750, 2016.
[35] X. Zhen, H. Shan, G. Yu, Y. Cheng, L. X. Cai, and A. Huang, "Joint resource allocation for LTE over licensed and unlicensed spectrum," in *Proc. IEEE VTC-Fall*, Sep. 2017, pp. 1–6.
[36] N. Mokari, M. R. Javan, and K. Navaie, "Cross-layer resource allocation in OFDMA systems for heterogeneous traffic with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 1011–1017, Feb. 2010.
[37] X. Li, F. Jin, R. Zhang, and L. Hanzo, "Joint cluster formation and user association under delay guarantees in visible-light networks," in *Proc. IEEE GLOBECOM*, Dec. 2016, pp. 1–6.

WEIHUA WU received the B.S. and M.E. degrees in telecommunications engineering and the Ph.D. degree in communication and information systems from Xidian University, China, in 2011, 2014, and 2017, respectively. From 2016 to 2017, he was a Visiting Student with the University of Avignon, France. He currently holds a postdoctoral position with the School of Telecommunications Engineering, Xidian University. Since 2019, he has been with the Singapore University of Technology and Design as a Postdoctoral Researcher sponsored by the China Scholarship Council (CSC). His research interests include wireless resource allocation, and stochastic network optimization and their applications in LTE-U networks.

QINGHAI YANG received the B.S. degree in communication engineering from the Shandong University of Technology, China, in 1998, the M.S. degree in information and communication systems from Xidian University, China, in 2001, and the Ph.D. degree in communication engineering from Inha University, South Korea, in 2007, with the University-President Award. From 2007 to 2008, he was a Research Fellow with UWB-ITRC, South Korea. Since 2008, he has been with Xidian University, China. His current research interests include autonomic communication, content delivery networks, and LTE-A techniques.

RUNZI LIU (M'16) received the B.S. degree (Hons.) in telecommunications engineering and the Ph.D. degree in communication and information systems from Xidian University, Xi'an, China, in 2011 and 2016, respectively. From 2016 to 2019, she held a postdoctoral position with the State Key Lab of ISN, Xidian University. In 2019, she joined the Xi'an University of Architecture and Technology, where she is currently an Associate Professor. Her research interests include resource managements and performance analysis in wireless networks and space networks.

KYUNG SUP KWAK received the Ph.D. degree from the University of California at San Diego. He was with Hughes Network Systems and the IBM Network Analysis Center, USA. Since then, he has been with the School of Information and Communication Engineering, Inha University, South Korea, as a Professor, and served as the Dean of the Graduate School of Information Technology and Telecommunications and the Director of the UWB Wireless Communications Research Center, an IT Research Center, South Korea, since 2003. In 2006, he served as the President of the Korean Institute of Communication Sciences (KICS), and in 2009, the President of the Korea Institute of Intelligent Transport Systems (KITS). In 1993, he received the Engineering College Achievement Award from Inha University, a Service Award from the Institute of Electronics Engineers of Korea (IEEK), and in 1996 and 1999 he received distinguished service awards from the KICS. He received the LG Paper Award, in 1998 and Motorola Paper Award, in 2000. He received official commendations for UWB Radio Technology Research and Development from the Ministry of Information & Communication, Prime Minister, and President of Korea, in 2005, 2006, and 2009, respectively. In 2007, he received the Haedong Paper Award and, in 2009, Haedong Scientific Award of research achievement. In 2008, he was elected for Inha Fellow Professor (IFP) and now for Inha Hanlim Fellow Professor. He published more than 200 peer reviewed journal papers and served as a TPC and Track chairs/organizing chairs for several IEEE related conferences. His research interests include multiple access communication systems, mobile &UWB radio systems, the future IoT, wireless body area network: nano networks, and molecular communications. He is a member of the IEICE, KICS, and KIEE.