

Modeling Cross-Technology Coexistence with Prioritized Traffic Classes

Mohammed Hirzallah¹, Yong Xiao¹, and Marwan Krunz^{1,2}

¹Department of Electrical and Computer Engineering, University of Arizona, AZ, USA

² School of Electrical and Data Engineering, University of Technology Sydney, NSW, Australia

Email: {hirzallah,yongxiao,krunz}@email.arizona.edu

Technical Report

TR-UA-ECE-2018-2

Last Update: Nov 26th, 2018

Abstract

The dramatic growth in demand for mobile data service has prompted mobile network operators (MNOs) to explore new spectrum resources in unlicensed bands. MNOs have been recently allowed to extend LTE-based service called LTE-LAA over 5 GHz U-NII bands, currently occupied by Wi-Fi. To support applications with diverse QoS requirements, both LTE and Wi-Fi technologies introduce multiple priority classes with different channel contention parameters for accessing unlicensed bands. How these different priority classes affect the interplay between coexisting LTE and Wi-Fi technologies is still relatively under explored. In this paper, we develop a simple and efficient framework that helps MNOs assess the fair coexistence between MNOs and Wi-Fi operators with prioritized channel access under multi-channel setting. We derive an approximated close-form solution for each MNO to pre-evaluate the probability of successful transmission (PST), average contention delay, and average throughput when adopting different priority classes to serve different traffics. MNOs and Wi-Fi operators can fit our model using measurements collected offline and/or online, and use it to further optimize their systems' throughput and latency. Our results reveal that PSTs computed with our approximated closed-form model approach those collected from system-level simulations with around 95% accuracy under scenarios of dense network deployment density and high traffic intensity.

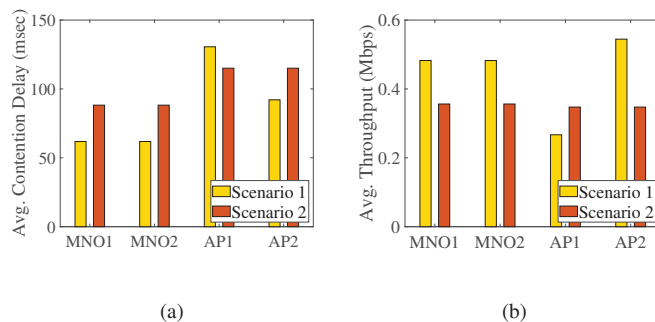


Fig. 1. (a) Average downlink contention delay of traffic with priority class 1, (b) Average downlink throughput of traffic with priority class 4.

I. INTRODUCTION

To cope with the fast growing demand for mobile data services, mobile network operators (MNOs) have taken steps to secure more spectrum resources. One solution promoted by FCC is to allow MNOs to extend their operations into unlicensed spectrum, including the unlicensed national-information-infrastructure (U-NII) bands at 5 GHz, currently used by Wi-Fi [1]. Licensed-Assisted Access (LAA), ‘enhanced LAA’ (eLAA), and ‘further enhanced LAA’ (feLAA) have been introduced in 3GPP Release-13 and Release-14 standards and the draft of Release-15 to let MNOs enable LTE services in unlicensed bands whereby MNOs supplement their licensed-spectrum downlink and uplink operations. These channel access schemes follow a similar listen-before-talk-based channel access mechanism adopted by Wi-Fi technology. Features and procedures introduced for LAA, eLAA, and feLAA are also expected to play key roles in the design of future 5G New Radio Unlicensed (NR-U) [2].

Next generation networks are expected to support a plethora of different services with diverse and often conflicting performance requirements [3]. For example, recent IoT services can support up to 10 years of battery life with high tolerance to service latency. Some emerging interactive services such as Augmented & Virtual Reality (AR/VR) require both extremely low latency (as low as 10 ms) and high throughput for data streaming. Other non-realtime applications, such as high-definition video downloading, are more tolerant to latency but have more stringent requirement on throughput. To meet the diverse service demands and requirements of newly emerging applications, 3GPP standards introduced four priority classes (PCs), labeled as P_1 , P_2 , P_3 , and P_4 , with different channel access parameters currently only enabled for downlink unlicensed operation [4]. 3GPP Release-14 and Release-15 also introduce an equivalent set of PCs for the eLAA to support uplink unlicensed

This research was supported in part by NSF (grants # IIP-1265960, IIP-1535573, CNS- 1731164, and CNS-1563655) and by the Broadband Wireless Access & Applications Center (BWAC). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF.

cellular operation [4]. In particular, for saturated traffics, among all the PCs, P_1 has the lowest inter-frame waiting time, which determines the latency between two consecutive channel access attempts and shortest channel occupancy time (COT), which specifies the time duration for which channel can be occupied and data transmission takes place. These characteristics of PC P_1 make it ideal for applications requiring low latency and low throughput such as interactive/streaming voice services. PCs P_3 and P_4 , on the other hand, have higher inter-frame waiting times with COTs. Therefore, these two PCs are more ideal for applications that require higher throughput but are more latency tolerant. Similarly, Wi-Fi standards, such as IEEE 802.11ac, also introduce the enhanced-distributed-channel access (EDCA) scheme employing similar idea as LAA that can offer multiple sets of contention parameters supporting differentiated services with different QoS requirements, referred to as the access categories (ACs) (labeled as A_1, A_2, A_3 , and A_4) [5]. Although LAA and Wi-Fi technologies follow a similar channel access scheme, they adopt different channel contention parameters, resulting in different channel access probabilities, contention delay, as well as throughput when they coexist on the same unlicensed channel.

The interplay between LTE PCs and Wi-Fi ACs over multi-channel setting further complicates the operation of these two technologies. To shed more light on this issue, let us consider the following example. Suppose two MNOs, labeled MNO1 and MNO2, sharing two unlicensed channels with two Wi-Fi access points (APs), labeled AP1 and AP2. Each MNO deploys one SBS and can have access to the two channels, and each Wi-Fi AP operates on one of these two channels. Suppose both MNO SBSs and Wi-Fi APs can offer two types of services to their users: voice and FTP data streaming. Suppose voice and FTP data streaming services have been allocated with PCs P_1 and P_4 by both MNOs, respectively, and with ACs A_1 and A_4 by Wi-Fi APs, respectively. We carried out a study using discrete-event simulations to investigate the impact of the channel assignment of these traffics on the performance of both coexisting systems. We present the simulation results for two scenarios in Figure 1. In Scenario 1, both MNOs assign the same type of service traffic to the same channel, i.e., both MNO SBSs send P_1 traffic to channel 1 and P_4 traffic to channel 2. In Scenario 2, each channel has been assigned with different types of services from each MNO, i.e., MNO1 P_1 and MNO2 P_4 traffics were assigned to channel 1, while MNO1 P_4 and MNO2 P_1 traffics were assigned to channel 2. We present the average contention delay and throughput for MNOs and Wi-Fi APs in Figure 1(a) and Figure 1(b), respectively. We observe that, compared to Scenario 2, Scenario 1 achieves lower contention delay and higher throughput for both MNOs. Scenario 1, however, results in lower throughput and higher contention delay for the Wi-Fi AP1. In other words, Scenario 1 provides better performance for LTE systems. Scenario 2, on the other hand, improves the the performance of the Wi-Fi AP1 but reduces the throughput with increased contention delay for MNOs.

The above example reveals the complexity for optimizing the performance and fairness between LTE and Wi-Fi when multiple

MNOs and Wi-Fi networks can adopt different priority classes to further improve QoS of their services. This motivates our work in which we develop a novel multi-channel performance evaluation framework, referred to as *Harmony*, for MNOs to pre-evaluate different QoS performance metrics, including the probability of successful transmission (PST), contention delay, as well as throughput for any possible traffic assignment decision. Our framework incorporates queuing and Markov-based models to evaluate the above metrics while considering multi-channel settings. We further simplify the outcomes of this model and derive closed-form PST expressions for LTE and Wi-Fi priority classes. We develop a discrete-event simulation environment using C++ and apply it to perform extensive fitting and evaluation to verify our closed-form expressions in various LTE/Wi-Fi coexistence scenarios. These expressions are flexible and can be leveraged with offline and/or online measurements, making them ideal for optimizing the resource allocation, traffic distribution, as well as fairness control for the coexistence of LTE and Wi-Fi technologies. Due to space limit, we move detailed proofs of this paper into an online technical report [6].

II. RELATED WORKS

Existing works on LTE/Wi-Fi coexistence mainly focus on optimizing LBT channel access [7], designing of collision detection schemes [8], [9], optimizing the resource allocation among MNOs [10]–[12], detecting misbehaving of LTE/Wi-Fi [13]. Detailed surveys are given in [14], [15], [16]. On the other side, several works focused on extending Wi-Fi-oriented Bianchi’s approach into modeling LTE/Wi-Fi coexistence [17]–[24]. Vallas et al. [17] modeled LTE/Wi-Fi coexistence using Markov chains and investigated maximizing the capacity of LTE in unlicensed bands by specifying the maximum airtime LTE could use. Zhang et al. [18] modeled the coexistence of LTE and Wi-Fi using a Markov-based framework, and derived the optimal size of contention window that maximizes LTE and Wi-Fi throughput. Lee et al. [19] investigated the problem of asymmetric hidden node in LTE/Wi-Fi coexistence, mainly focusing on mathematical modeling using a Markov-based framework. Yin et al. [20] introduced an adaptive LBT scheme that balances LTE operations by minimizing Wi-Fi collisions while ensuring a robust LTE SBS performance. Han et al. [21] introduced a MAC design for harmonious LTE operation in unlicensed bands. Abdelfattah and Malouch [22] presented a solid model based on random-walks for modeling the duty-cycle-based LTE-U operation. Sutton et al. [23] focused on analyzing delay of LAA. Mehrnoush et al. [24] modeled the impact of energy detection in LTE/Wi-Fi coexistence. Although previous works include interesting analysis and insightful results, they still limited because they focus on a single priority class or fail to include key parameters distinguishing the different LTE and Wi-Fi priority classes, such as the arbitration inter-frame space (AIFS). Bianchi et al. [25] coined the importance of including AIFS parameter while modeling performance of the EDCA scheme. The AIFS value used by Wi-Fi systems (a.k.a., ‘deferment period’ in LAA) decides who

can access the channel earlier. In contrast to these existing works, in this paper, we address the issues of the previous works and provide a closed-form solution for modeling the prioritized LTE/Wi-Fi coexistence.

III. OVERVIEW OF UNLICENSED CHANNEL ACCESS SCHEMES IN IEEE 802.11 AND LTE-LAA STANDARDS

A. IEEE 802.11

IEEE 802.11 standards support several channel access schemes in which the distributed coordination function (DCF) and EDCA are the most frequently used ones. The EDCA is an enhanced version of DCF that is introduced to improve QoS provisioning. EDCA defines four ACs ($A_1 - A_4$): Voice (AC_VO), video (AC_VI), best effort (AC_BE), and background (AC_BK) as shown in Table I. The duration of AIFS, T_{AIFS} , is computed as $T_{\text{AIFS}} = T_{\text{SIFS}} + d_i T_{\text{slotTime}}$ [5], where $T_{\text{SIFS}} = 16 \mu\text{sec}$ is the short inter-frame space and $T_{\text{slotTime}} = 9 \mu\text{sec}$ is the MAC time slot, and d_i is the AIFS number (AIFSN). In addition, EDCA scheme limits the transmission time T_i , a.k.a, transmission-opportunity (TXOP) period, for ACs A_1 and A_2 . The TXOP times for ACs A_3 and A_4 are not restricted. During each TXOP period, it is possible to send single or multiple data frames. The EDCA scheme works as follows. Each station (e.g., Wi-Fi AP or device trying to initiate channel access) first senses the channel for an AIFS period of time and can only start transmission if the channel is sensed idle during the AIFS. If the channel is sensed busy during the AIFS, a backoff mechanism will be triggered in which the transmitter randomly picks an integer k between 0 and $K - 1$ for

$$K \in \min\{W_j^{(i)}, W_{\max}^{(i)}\}, \quad (1)$$

where $W_j^{(i)} = 2^j W_{\min}^{(i)}$, j is the retransmission attempt, $W_{\min}^{(i)}$ is contention window minimum (CWmin) size, $W_{\max}^{(i)}$ is the contention window maximum (CWmax) size.

The station counts down for k successive time slots as long as the channel is idle. Whenever the channel is sensed to be busy during the count down, the station has to freeze its counter until it becomes idle again. Once the counter becomes zero, the station can then start its transmission for a duration that is less or equal its TXOP period. The receiver acknowledges the transmitter about the successful reception by sending back an acknowledgment (ACK) or a block acknowledgement (BA) frame. A station can infer a failure of frame transmission or collision by an acknowledgment timeout, i.e., the transmitter does not receive ACK framework within a certain period of time. Failed or collided data frames should be retransmitted for at most R times before being discarded, and for each retransmission the CWmin needs to be doubled which is characterized by the 2^j term in (1). The smaller the values of AIFS, as well as CWmin and CWmax sizes the higher the probability for a station to successfully occupy the channel [25]. An example of EDCA operation over time is shown in Figure 2.

TABLE I

EDCA AND LBT CAT-4 CHANNEL ACCESS PARAMETERS FOR EACH AC AND PC, RESPECTIVELY [5] [4]

AC A_i (EDCA)	d_i/T_{AIFS}	$W_{\text{min}}^{(i)}$	$W_{\text{max}}^{(i)}$	Max TXOP T_i
A_1 :Voice (AC_VO)	2/34 μsec	4	8	1.504 ms
A_2 :Video (AC_VI)	2/34 μsec	8	16	3.008 ms
A_3 :Best effort (AC_BE)	3/43 μsec	16	1024	-
A_4 :Background (AC_BK)	7/79 μsec	16	1024	-
Legacy DCF	2/34 μsec	16	1024	-
PC P_i (CAT-4)	d_i/T_{df}	$W_{\text{min}}^{(i)}$	$W_{\text{max}}^{(i)}$	T_i
P_1	1/25 μsec	4	8	2 msec
P_2	1/25 μsec	8	16	3 msec
P_3	3/43 μsec	16	64	6, 8, or 10 msec
P_4	7/79 μsec	16	1024	6, 8, or 10 msec

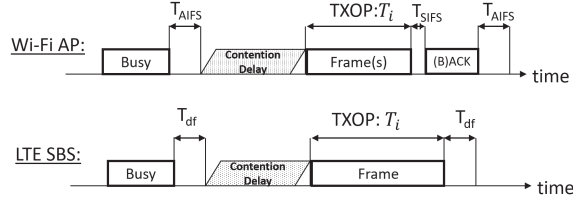


Fig. 2. Examples of operation over time for EDCA (top) and LAA LBT CAT-4 (down) channel access schemes.

B. LTE-LAA

The listen-before-talk (LBT) Category 4 (CAT-4) scheme, similar to DCF/EDCA in Wi-Fi, is adopted by the 3GPP Release-13 and Release-14 standards [4]. LAA defines four priority classes, $P_1 - P_4$ as shown in Table I, which, in some sense, can be considered as the equivalent to ACs A_1 - A_4 in Wi-Fi. The deferment period T_{df} in LAA is equivalent to AIFS in Wi-Fi, and therefore, in this paper, we generally use AIFS to refer for both LAA ‘deferment period’ and Wi-Fi AIFS. PC P_1 has the smallest AIFS as well as both CW_{max} and CW_{min} among all PCs that is equivalent to AC A_1 in EDCA scheme. In TXOP period, the SBS sends an OFDMA frame where it schedules different user equipments (UEs) to access different resource blocks (time and frequency). In LAA, SBS infers the failure of transmission by monitoring the HARQ-ACK feedback messages sent by UEs over the licensed channel [4]. By comparing the entries of EDCA and CAT-4 schemes in Table I, we notice that LAA supports smaller AIFS values and hence devices with LTE-LAA technology are expected to capture channels faster than those with Wi-Fi. An example of LAA LBT CAT-4 operation over time is shown in Figure 2.

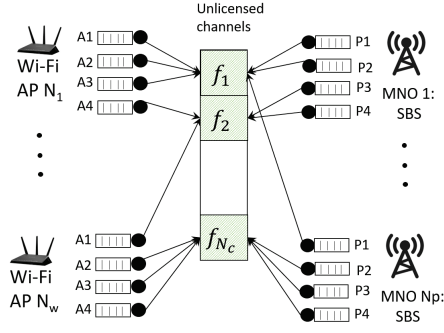


Fig. 3. Example of LTE/Wi-Fi coexistence with prioritized traffics over a set of N_c channels.

IV. NETWORK COEXISTENCE MODEL

We consider a set \mathcal{N} of N_p MNOs, each of which has deployed a set B of SBSs that can offer services in unlicensed bands using LAA technology. MNOs share a set \mathcal{F} of N_c unlicensed channels with another set \mathcal{M} of N_w Wi-Fi APs. To simplify our discussion, we focus on the downlink transmission for both LTE-LAA and Wi-Fi. Our model however can be directly extended into uplink scenarios. In this paper, we consider a general model in which each Wi-Fi AP can support a set of ACs denoted as $\mathcal{C} = \{A_1, \dots, A_{N_{ac}}\}$, where N_{ac} is the number of supported ACs (e.g., $N_{ac} = 4$ for IEEE 802.11ac Wi-Fi technology). Each LTE-LAA SBS can also support a set of PCs labeled as $\mathcal{L} = \{P_1, \dots, P_{N_{pc}}\}$ where N_{pc} is the number of total supported PCs (e.g., $N_{pc} = 4$ for LTE-LAA Release 13 and Release 14). Let $m_{k,\theta}$ be the number of Wi-Fi APs that assign their traffics of type AC A_θ to channel k . Let $n_{k,\eta}$ be the number LTE-LAA SBSs that assign traffic of type PC P_η to channel k .

We define an LTE/Wi-Fi channel sharing structure (CSS) for characterizing the assignment of LTE-LAA SBSs and Wi-Fi APs traffic types to the N_c unlicensed channels. Formally, a CSS \bar{cs} is defined as a tuple:

$$\bar{cs} = \langle \langle \mathbf{n}_1, \mathbf{m}_1 \rangle, \dots, \langle \mathbf{n}_{N_c}, \mathbf{m}_{N_c} \rangle \rangle, \quad (2)$$

where $\langle \mathbf{n}_k, \mathbf{m}_k \rangle$ specifies the traffics sent by LTE-LAA SBSs and Wi-Fi APs on channel k for $\mathbf{n}_k = \langle n_{k,1}, \dots, n_{k,N_{pc}} \rangle$ and $\mathbf{m}_k = \langle m_{k,1}, \dots, m_{k,N_{ac}} \rangle$ and $n_{k,i}$ and $m_{k,j}$ are the numbers of co-located LTE-LAA SBSs and Wi-Fi APs transmitting traffics in channel k with PC P_i and AC A_j , respectively.

Each MNO needs to pre-evaluate the possible performance when assigning service traffics into different channels. The expected PST is closely related to the channel assignment decisions as well as the assigned PCs and ACs by MNOs and Wi-Fi APs. We propose a Markov-based framework for each MNO to estimate their expected PST under each possible channel assignment.

V. MODELING PRIORITIZED LTE/WI-FI COEXISTENCE

To simplify our description, in this section, we focus on a single channel shared by a set of n_k LTE SBSs and m_k Wi-Fi APs. We first model traffics of LTE and Wi-Fi according to their priority classes, and introduce a Markov-based model to characterize EDCA and LBT CAT-4 channel access schemes, and explain how PST can be derived and computed. We then explain how PST computations can be simplified by introducing an approximated closed-form expression and discuss the implications of these expressions.

A. Queuing Model

We follow the same line as [25] and consider the traffic generated by each AC in IEEE 802.11 standards as an individual queue. In this way, traffics generated by all N_{ac} ACs can be considered as N_{ac} parallel queues. Similarly, LTE traffics generated by N_{pc} PCs can be modeled as N_{pc} parallel queues (see Figure 3 with $N_{ac} = N_{pc} = 4$). Frames arriving at each queue line are served on a First-Come-First-Served (FCFS) basis. We follow a commonly adopted setting and assume the data frame arrival process follows a Poisson distribution [26] (see Annex A.2.1.3.1 ‘Traffic models’). Our model can be directly extended into more general settings, e.g., frame arrival follows other more general distributions. Let λ_i be the arrival rate of frames that are associated with i th priority class for $i \in \mathcal{C} \cup \mathcal{L}$, measured in frames per MAC time slot. We use ‘priority class’ for referring for LTE PCs and Wi-Fi ACs and distinguish between them when necessary.

1) *Probability of Frame Arrival*: The probability of having k frames arriving in a δ -duration time slot can be written as $(\lambda_i \delta)^k \exp^{-\lambda_i \delta} / k!$. In our case, we set δ to be equal to the duration of MAC time slot. We define the *probability of frame arrival* g_i as the probability of having at least one frame arriving at the queue during a time slot δ , and this can be written as:

$$g_i = \Pr(k \geq 1) = 1 - \Pr(k = 0) = 1 - \exp^{-\lambda_i \delta}. \quad (3)$$

2) *Average Contention Delay*: Data frame at the head of each queue can experience a contention delay $D_{k,i}$ before being sent over-the-air. Contention delay is caused by the random time a transmitter should wait before accessing the channel and it constitutes of AIFS period plus the time for which the channel k becomes busy during the countdown process, as illustrated in Section III. We can prove the following result regarding the contention delay $D_{k,i}$:

Proposition 1. *The average contention delay for the i th priority class is given by:*

$$D_{k,i} = 1/p_{s,k,i}, \quad (4)$$

where the $p_{s,k,i}$ is the PST which will be defined later in Section V-B4.

Proof: See Appendix B. ■

Note that $D_{k,i}$ is the number of MAC time slots spent on the contention. Delivering a frame over-the-air requires a time duration that is equivalent to TXOP period T_i . The sum of contention delay $D_{k,i}$ and TXOP period T_i constitutes the *service time* of the queue; i.e., the time required for a data frame reaching queue head to get successfully delivered to its destination, which is given by $S_i = D_{k,i} + T_i$. The service rate of the queue can then be written as $\mu_i = \frac{1}{S_i}$.

3) *Probability of Saturation:* We define the *probability of saturation* q_i as the probability of having a non-empty queue, i.e., $q_i = 1$ means that there are always data frames available in the queue waiting to be served. The number of frames in each priority class queue increases and decreases according to the frame arrival rate λ_i and queue service rate μ_i . Therefore, it becomes natural to model the queue dynamics of each priority class using a birth-and-death process. To compute the saturation probability, we consider two queue states ('Idle' and 'Occupied') with birth rate of λ_i and death rate of μ_i . The probability of saturation is equivalent to the long-term probability of being at the state of 'Occupied', which is given by [27]:

$$q_i = \lim_{t \rightarrow \infty} (1 - P_{\text{Idle}}(t)) = \frac{\lambda_i}{\mu_i + \lambda_i} = \frac{\lambda_i(D_{k,i} + T_i)}{1 + \lambda_i(D_{k,i} + T_i)}. \quad (5)$$

B. Markov-based Model

We model the LBT CAT-4 for LTE SBSs (or EDCA schemes for Wi-Fi APs) as a Markov chain consisting of four states: The state of idle-queue (I), contention (C), successful transmission (S), and failed transmission (F), as illustrated in Figure 4. Please see Appendix A in [6] for a more detailed description of discrete-time Markov chain (DTMC) for this Markov process. A transmitter could spend a random number of time slots in each state based on the contention parameters and class types adopted by itself as well as other transmitter in proximity. Possible transitions between states can happen in the following scenarios:

- *Transition from I to C, $T_{I \rightarrow C}$:* Initially, a transmitter will be at state I when its queue is empty. Once a new frame arrives at the queue, the transmitter will be in state C. It can be observed that the transition probability from state I to state C is equivalent to the probability of frame arrival g_i defined in (3), i.e., we have $T_{I \rightarrow C} = g_i$.
- *Self-Transition at C, $T_{C \rightarrow C}$:* A transmitter resides in state C until the channel is clear and the data frame is free to be sent, otherwise it remains at the state of contention. We define the *probability of channel-access* (PCA), denoted as τ_i , as the probability that a transmitter accesses the channel in an arbitrary time slot after contention. Therefore, we define the self-transition probability of state C as $T_{C \rightarrow C} = 1 - \tau_i$.
- *Transition from C to F, $T_{C \rightarrow F}$:* Let p_i be the *probability of collision* and C_i be the time spent in collision. The transition from C to F happens with a probability $T_{C \rightarrow F} = \tau_i p_i$, meaning that the transmitter tries to access the channel but its

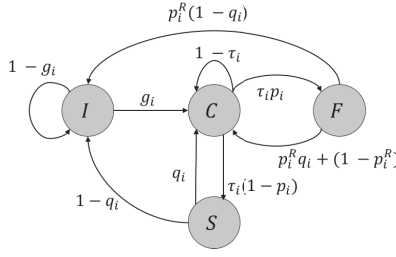


Fig. 4. Markov chain that models different states of LTE/Wi-Fi transmitter with traffic type i (I: State of idle-queue; C: State of contention; S: State of successful-transmission; and F: State of failed transmission).

transmission is collided. A transmitter fails to deliver a frame because of collisions with other simultaneous transmissions or because of bad channel conditions.

- *Transition from F to C, $T_{F \rightarrow C}$:* When collision happens, the transmitter should backoff and re-contend again for a new channel access, and this happens with a probability $1 - p_i^R$. A frame will be dropped after R consecutive re-transmission failures, which happens with a probability $p_i^{R_i}$. Once transmitter fails to deliver the frame it starts serving a new one and goes again into the state of contention, provided queue already has frames to be transmitted. The transition into state C after R collided retransmissions happens with a probability $p_i^{R_i} q_i$, where q_i is the queue-saturation probability as defined in (5). Therefore, the transition between F and C happens with probability $T_{F \rightarrow C} = p_i^R q_i + (1 - p_i^R)$.
- *Transition from F to I, $T_{F \rightarrow I}$:* If queue is idle and collision happens for R consecutive retransmission attempts, a transmitter drops frame and goes back into state I, awaiting for a new frame to arrive. The event of having R consecutive failed retransmissions happens with probability p_i^R , and the probability of having an idle queue is $1 - q_i$. Therefore, the transition from state F to I happens with probability $T_{F \rightarrow I} = p_i^R(1 - q_i)$.
- *Transition from C to S, $T_{C \rightarrow S}$:* The transition from the state C to the state of S happens when a transmitter accesses the channel and no collision happens. The transition from state C to S happens with probability $T_{C \rightarrow S} = \tau_i(1 - p_i)$.
- *Transition from S to C and I, $T_{S \rightarrow C}$ and $T_{S \rightarrow I}$:* The transmitter spends T_i consecutive time slots for a successful transmission, afterward it jumps to the state C to serve a new data frame with a probability of $T_{S \rightarrow C} = q_i$, provided that the queue already has frames to be served, otherwise, it jumps to state I with a probability $T_{S \rightarrow I} = 1 - q_i$.

1) *Probability of Collision:* Collision happens when two or more transmitters start accessing the same channel simultaneously. By including the PCAs of all priority classes, we write the probabilities of collision p_η and p_θ for all LTE and Wi-Fi

priority classes, respectively, as follow:

$$p_\eta = 1 - (1 - \tau_\eta)^{n_{k,\eta}-1} \prod_{j \in \mathcal{L}, j \neq \eta} (1 - \tau_j)^{n_{k,j}} \prod_{\theta \in \mathcal{C}} (1 - \tau_\theta)^{m_{k,\theta}}, \quad (6)$$

$$p_\theta = 1 - (1 - \tau_\theta)^{m_{k,\theta}-1} \prod_{i \in \mathcal{C}, i \neq \theta} (1 - \tau_i)^{m_{k,i}} \prod_{\eta \in \mathcal{L}} (1 - \tau_\eta)^{n_{k,\eta}}. \quad (7)$$

Equations (6) and (7) are equivalent to the event of having two or more transmitters access the channel simultaneously.

2) *Probability of Counter Freezing*: Channel becomes busy when at least one transmitter becomes active and starts transmission. In this case, any other LTE or Wi-Fi transmitter who can successfully detect this transmission will need to freeze its counter. Suppose there is no hidden nodes. The probability of counter-freezing b_i for the i th priority class can be written as:

$$b_i = 1 - \prod_{\eta \in \mathcal{L}} (1 - \tau_\eta)^{n_{k,\eta}} \prod_{\theta \in \mathcal{C}} (1 - \tau_\theta)^{m_{k,\theta}}. \quad (8)$$

From the above equation, we can observe that the probability of counter-freezing is related to transmitter's ability to detect transmissions started by other coexisting devices. For simplicity, we assume a dense network model in which all coexisting devices from different technologies are in close proximity, and hence their signals have transmit powers that are sufficiently high to be detected by others. Under this assumption, the dissimilarity in between the energy detection thresholds adopted by the different coexisting technologies can be ignored. Therefore, the counter-freezing probability can be written as the probability of the the event of having one or more transmitters accessing channel simultaneously in the same time slot, as shown in (8).

3) *Probability of Channel Access*: Ability of an LTE/Wi-Fi transmitter to access a channel depends on the parameters listed in Table I. Therefore, PCA τ_i for the i th LTE or Wi-Fi priority class depends on the following parameters: AIFS number d_i , CWmin $W_{\min}^{(i)}$, CWmax $W_{\max}^{(i)}$, and the maximum re-transmission limit R . It also depends on the time period spent in both successful T_i and failed transmission C_i , as well as the intensity of its traffic characterized by g_i and q_i as in (3) and (5). We can prove the following result regarding PCA τ_i :

Proposition 2. *Probability of channel access for the i th LTE and/or Wi-Fi priority class is:*

$$\begin{aligned} \tau_i = & \frac{1 - p_i^{R+1}}{1 - p_i} \cdot \left[\overbrace{\frac{1 + b_i B}{b_i} \cdot \frac{1 - (1 - b_i)^{d_i+1}}{(1 - b_i)^{d_i+1}}}^{(a)} \right. \\ & + \overbrace{T_i (1 - p_i^{R+1})}^{(b)} + \overbrace{(1 + p_i C_i) \frac{1 - p_i^{R+1}}{1 - p_i}}^{(c)} \\ & \left. + \overbrace{\frac{(1 - q_i)(1 + p_i^R - p_i^{R+1})}{g_i}}^{(d)} + \overbrace{\frac{1 + b_i B}{2(1 - b_i)^{d_i}} \sum_{j=0}^R W_j^{(i)} p_i^j}^{(e)} \right]^{-1}, \quad (9) \end{aligned}$$

where B is the continuous time duration for which the channel becomes busy because other transmitters use it (i.e., the time period for one continuous counter freezing).

Proof: See Appendix A. ■

B is random and depends on TXOP periods adopted by all coexisting transmitters and the potential of having two or more transmitters having a simultaneous or overlapping over-the-air transmission. It is of a low probability that the channel remains busy continuously for a time period that exceeds the sum of TXOP periods of all coexisting transmitters. Therefore, we consider an intermediate case and set the busy time to $B \approx \max_i (T_i), i \in \mathcal{C} \cup \mathcal{L}$, i.e., the maximum duration of TXOP period of all LTE PCs and Wi-Fi ACs. It can be observed that terms (a) - (e) in the channel access probability equation (9) correspond to the different states illustrated in Figure 4:

- Terms (a) and (e) correspond to two potential scenarios of the state of contention C: Term (a) corresponds to the scenario of freezing the backoff counter when the channel becomes busy, while term (e) accounts for having multiple backoffs due to frame retransmission. Notice that terms (a) and (e) include the key parameters that control the backoff process for the i th priority class (i.e., size of contention window $W_j^{(i)}$ and AIFS number d_i).
- Term (b) corresponds to the state of successful transmission S , and it includes the time duration needed to achieve a successful transmission T_i . This term models a scenario in which a transmitter successfully sends a frame over the air.
- Term (c), on the other hand, corresponds to the state of collision C , and it models the scenario in which the transmitter fails to deliver its frame due to collisions with transmissions started by other coexisting transmitters.
- Term (d) corresponds to the state of idle-queue I , and models the scenario on which the transmitter does not have frames in its queues. In particular, this term includes parameters that characterize the traffic intensity of the transmitter and its queue dynamics represented in the probabilities g_i and q_i as illustrated in equations (3) and (5), respectively.

4) *Probability of Successful Transmission:* We can express PSTs for LTE and Wi-Fi priority classes by computing the event of having collision-free channel access, which can simply be written as the product of τ_i and $1 - p_i$. By substituting the expression of collision probabilities in $p_{s,k,\eta}^{(l)} = \tau_\eta(1 - p_\eta)$ and $p_{s,k,\theta}^{(w)} = \tau_\theta(1 - p_\theta)$, the PSTs for LTE and Wi-Fi priority classes can be expressed, respectively, as:

$$p_{s,k,\eta}^{(l)} = \tau_\eta(1 - \tau_\eta)^{n_{k,\eta}-1} \prod_{j \in \mathcal{L}, j \neq \eta} (1 - \tau_j)^{n_{k,j}} \prod_{\theta \in \mathcal{C}} (1 - \tau_\theta)^{m_{k,\theta}}, \quad (10)$$

$$p_{s,k,\theta}^{(w)} = \tau_\theta(1 - \tau_\theta)^{m_{k,\theta}-1} \prod_{j \in \mathcal{C}, j \neq \theta} (1 - \tau_j)^{m_{k,j}} \prod_{\eta \in \mathcal{L}} (1 - \tau_\eta)^{n_{k,\eta}}. \quad (11)$$

In the absence of LTE technology, it should be observed that (11) reduces to the probability of successful transmission of Wi-Fi in [28].

To be able to compute and/or optimize PSTs, we need to put PSTs in a closed form expression, which includes the key parameters governing the different LTE and Wi-Fi priority classes. Computing PSTs in (10) and (11) requires computing the probabilities of channel access τ_i of the different LTE and Wi-Fi priority classes. We can easily observe from (9) that τ_i also depends on the probability of collision p_i as well as the probabilities governing queue dynamics (i.e., g_i and q_i). However, the formulation of g_i , q_i , and p_i are also dependent on τ_i , as can be observed in (3), (5), (6), and (7). This inter-dependency in between the formulations of τ_i , p_i , g_i , and q_i limits us from putting PST in a ‘clean’ closed-form expression that can be easily manipulated and optimized. Putting PST formula in an easy to manipulate expression is extremely important to apply further studies and optimization to the problem of prioritized LTE and Wi-Fi coexistence. Therefore, we next introduce a closed-form expression of PSTs.

C. Closed-form Modeling of Key Performance Measures

Our target is to simplify the expression of PST by avoiding the complicated expressions linking τ_i with g_i , q_i , and p_i . We developed discrete-event-based simulator to study and monitor changes happening in PST values as a function of network settings. We observed that under saturated network conditions (i.e., network with dense deployment and high traffic intensity) the PST values tend to decrease exponentially with the number of coexisting devices. The rate of reduction was also different for the different LTE and Wi-Fi priority classes. This led us to investigate approximations for PST where we put PST in a closed-form expression that incorporates key parameters distinguishing the different LTE/Wi-Fi priority classes such as the number of LTE and Wi-Fi transmitters, i.e., $\mathbf{n}_k = \langle n_{k,1}, \dots, n_{k,N_{pc}} \rangle$ and $\mathbf{m}_k = \langle m_{k,1}, \dots, m_{k,N_{ac}} \rangle$ as well as key contention parameters of each priority class (i.e., $W_j^{(i)}$, d_i , and T_i). We prove the following key result:

Theorem 1. *The average PSTs of the i th priority class can be characterized using the following model, assuming saturated traffic:*

$$\begin{aligned}
 p_{s,k,i}^{(t)} &\approx c_{i,0} + \\
 &\sum_{l \in \mathcal{L}} c_{i,l} (n_{k,l} + 1) \log \left(\beta_{i,l} (W_0^{(l)} d_l + T_l) n_{k,l} + \epsilon_{i,l} \right) \\
 &+ \sum_{j \in \mathcal{C}} h_{i,j} (m_{k,j} + 1) \log \left(\gamma_{i,l} (W_0^{(j)} d_j + T_j) m_{k,j} + \epsilon_{i,j} \right),
 \end{aligned} \tag{12}$$

where $t \in \{l, w\}$, the constants $c_{i,0}$, $c_{i,l}$, $\beta_{i,l}$, $e_{i,l}$, $h_{i,j}$, $\gamma_{i,j}$, $\epsilon_{i,j}$ are obtained by fitting this closed-form expression to PST samples collected from real implementations or system-level simulations. The fitting and verification processes are discussed next in Section V-C1.

Proof: See Appendix C. ■

The closed form expression in Theorem 1 can also be used to model the PSTs when only one technology occupy an unlicensed channel by setting the corresponding constants of the other technology to zero.

1) *Fitting PSTs Closed-Form Expressions:* We perform extensive simulations using our developed discrete-event simulator and collect traces of frames sent by LTE and Wi-Fi transmitters while contending with parameters of various priority classes, including different scenarios with various number of LTE/Wi-Fi transmitters (see Section VI for more details on simulation setup and simulator capabilities). We measured PST for each frame by taking the inverse of the number of MAC time slots spent in contention. We scramble the measured PSTs and divide them into two disjoint sample sets. The first set is used for fitting the model in (12) using the ‘curve_fit’ tool in Python [29], while the second set is used to test our model. Figure 5 shows samples of measured PSTs for a large number of transmitted frames. We also plot the sample mean of these measured PSTs and the PST values obtained using the approximate model in (12) for LTE PCs P_1 and P_2 and Wi-Fi ACs A_1 and A_2 . The stairs in these plots correspond to different scenarios (i.e., network with fixed number of LTE/Wi-Fi transmitters and traffic setting). We report the average absolute testing error for these in Figure 6. The closed-form approximate model in (12) computes the expected PST with high accuracy.

2) *Implications of PST Closed-form Expression:* The PST closed-form expression in (12) has many exciting implications:

- The proportionality constants in (12) can be found by fitting this closed-form expression to PST samples collected through offline and/or online measurements collected from real implementations or system-level simulations. This makes the expression in (12) ideal for applying in online learning solutions.
- By solving for the PSTs, we could evaluate the average throughput $\tilde{r}_{k,\eta}^{(l)}$ and $\tilde{r}_{k,\theta}^{(w)}$ of the η th and θ th LTE and Wi-Fi priority classes, respectively, when they operate at the k th channel by computing the following:

$$\tilde{r}_{k,i}^{(t)} = \frac{p_{s,k,i}^{(t)} \mathbb{E}[Q_i]}{D_{k,i} + p_{s,k,i}^{(t)} T_i + (1 - p_{s,k,i}^{(t)}) C_i} \quad (13)$$

where $t \in \{l, w\}$, Q_i is the payload size (in bits) of the data frame that belongs to the i th priority class, and $D_{k,i}$ is the corresponding average contention delay, as shown in (4).

- PST is both proportional to the average throughput in (13) and inversely to the average contention delay in (4). Although the

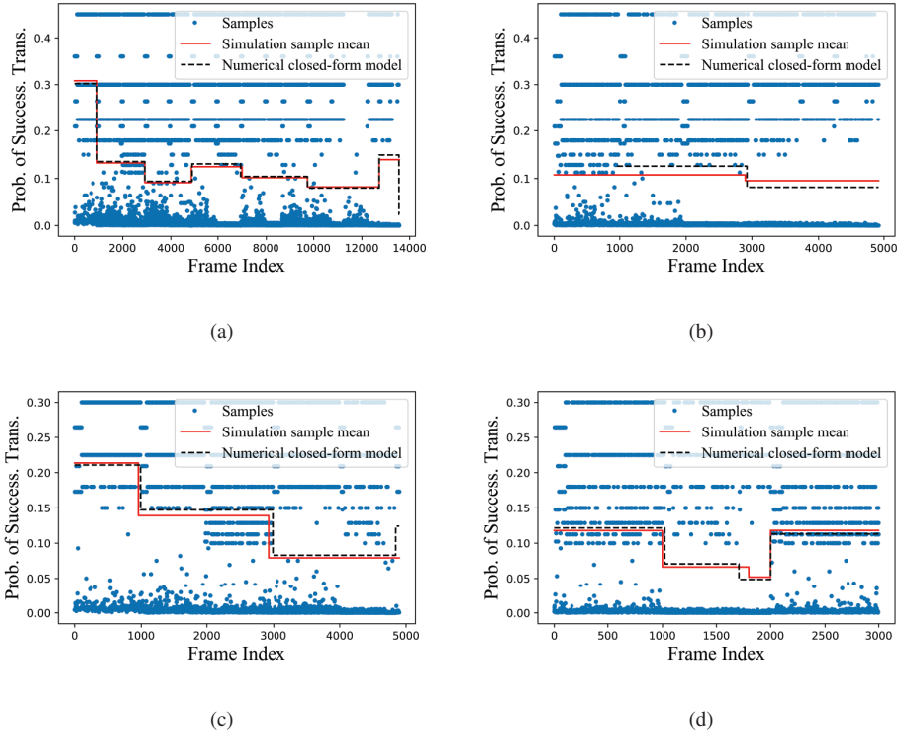


Fig. 5. Probability of successful transmission vs. frame index (a) LTE-LAA PC P_1 , (b) LTE-LAA PC P_2 , (c) Wi-Fi AC A_1 , and (d) Wi-Fi AC A_2 .

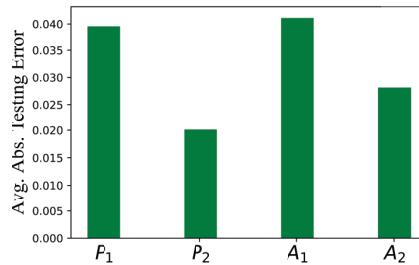


Fig. 6. Average absolute testing error for the closed-form approximate model.

different LTE/Wi-Fi priority classes have different performance measures, the PST can be linked to their key performance QoS metrics. Therefore, optimizing PST can be sufficient to optimize and balance QoS conflicting objectives.

VI. DISCUSSION AND SIMULATION RESULTS

A. Simulation Methodology

To capture the interplay between LTE and Wi-Fi technologies with prioritized channel access, we need a tool that provides independent and simultaneously active processes for modeling LTE SBSs/UEs Wi-Fi APs/STAs and their corresponding traffic streams with the four priority classes. Therefore, we develop a discrete-event simulation framework using CSIM, a C++ library that supports discrete-event-based and process-oriented simulations [30]. CSIM provides functions for creating parallel

TABLE II
SIMULATION PARAMETERS

Parameters	Value
Center frequency	channel 36 at 5.18 GHz
Channel bandwidth	20MHz
PHY rate	6.5 Mbps
Path Loss Model	$43.3 \log(d) + 11.5 + 20 \log(f_c)$
Antenna gain	5 dBi
Transmit power	23 dBm
LTE (Wi-Fi) ED-CCA thresholds	-62(-72) dBm
LTE (Wi-Fi) noise floor	-100 (-90)dBm
Small cell radius	30 meter
LTE(LAA)/Wi-Fi MAC time slot	$9\mu\text{sec}$
Wi-Fi SIFS	$16\mu\text{ sec}$

processes, and capabilities for enabling control and signaling between them. We build on CSIM functionalities and implement the most recent LBT CAT-4 and EDCA channel access schemes adopted in LTE-LAA and IEEE 802.11ac standards [4] [5] with all channel access parameters as described in Table I. We used our developed simulator to carry out various studies and collect measurements for fitting the PST approximate expression in (12). We set our simulator to operate on a granularity of one microsecond, whereby we capture the exact timing for AIFS, CWmin, CWmax, and TXOP period of all LTE and Wi-Fi priority classes.

We run the simulator for 10 seconds and collect traces and logs from all LTE and Wi-Fi transmitters, including time stamps for frame arrival to MAC queue, time spent in queue, time spent during contention, time spent during transmission. Each LTE/Wi-Fi transmitter serves four priority queues and contend using parameters in Table I. We adopt a Poisson frame arrival rate of $\lambda = 1000$ frames per second for all LTE and Wi-Fi traffics. We consider several scenarios, where in each scenario we vary the number of MNOs and Wi-Fi systems according to our investigation objective. The rest of PHY- and MAC-layer simulation parameters are summarized in Table II. We consider a centric topology where all SBSs/UEs and APs/STAs locations are generated uniformly over an area of 1600 square meter.

B. Number of LTE/Wi-Fi Transmitters

The increase in the number of LTE and Wi-Fi transmitters add more congestion to unlicensed channels. We evaluate the probabilities of successful transmission and collision probabilities as well as the average throughput, and average contention

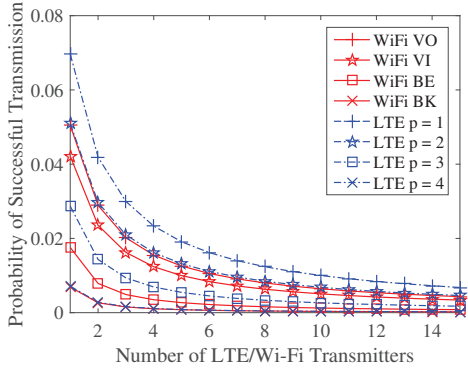


Fig. 7. PSTs vs. number of LTE/Wi-Fi transmitters ($N_w = N_p$).

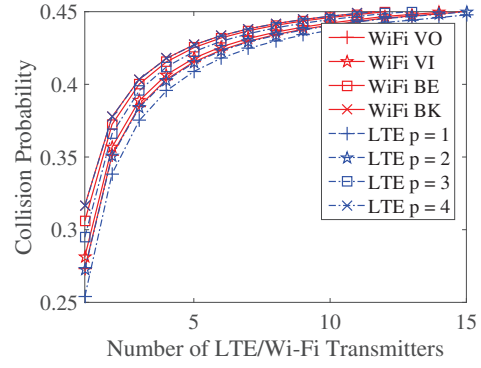


Fig. 8. Collision probability vs. number of LTE/Wi-Fi transmitters ($N_w = N_p$).

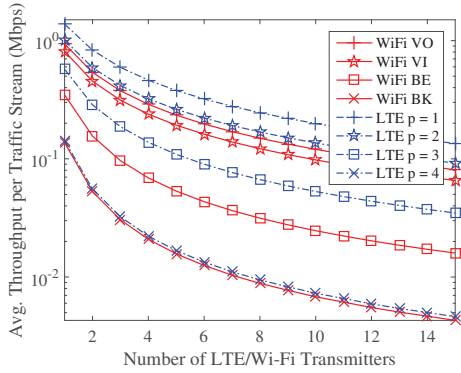


Fig. 9. Average throughput vs. number of LTE/Wi-Fi transmitters ($N_w = N_p$).

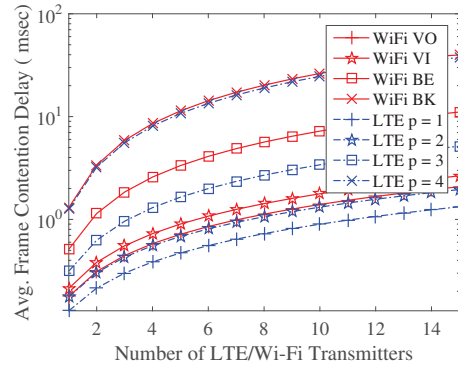


Fig. 10. Average frame contention delay vs. number of LTE/Wi-Fi transmitters ($N_w = N_p$).

delay versus the number of transmitters, as shown in Figures 7, 8, 9, and 10 (the number of LTE and Wi-Fi transmitters are set equal $N_w = N_p$). As the number of LTE and Wi-Fi transmitters increase, the achieved performance degrades and the difference in performance between the different priority classes becomes negligible. The priority classes P_1 and A_1 in both LTE and Wi-Fi systems achieve higher PSTs and average throughput, as well as lower collision probabilities and average contention delays when compared to other supported priority classes. However, we notice that LTE priority classes achieve higher PST and average throughput values with lower collision probability and average contention delay when compared to those achieved by Wi-Fi ACs. This happens due to the fact that LTE PCs P_1 & P_2 adopt smaller AIFS duration and longer TXOP period than those adopted by Wi-Fi ACs A_1 & A_2 .

C. Size of Contention Window and Fairness Tradeoffs

We investigate how changing LTE-LAA PC P_1 's CWmin size affects the performance of MNOs and Wi-Fi systems. We evaluate the average throughput and delay achieved by a transmitter by taking the mean of throughput and delay achieved

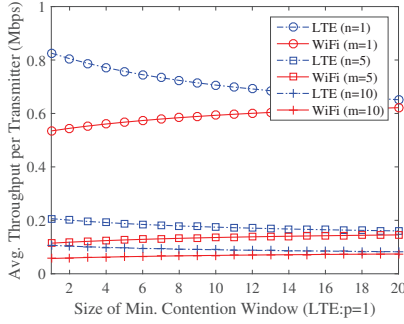


Fig. 11. Average throughput per traffic stream vs. size of minimum contention window of LTE PC P_1 .

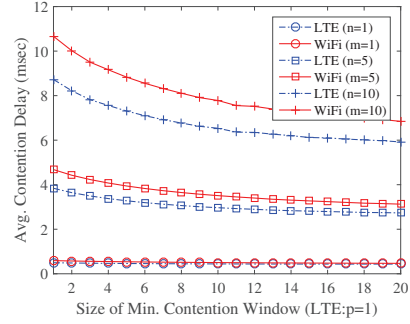


Fig. 12. Average contention delay per transmitter vs. size of minimum contention window of LTE PC P_1 .

by the four priority classes it supports, respectively. We plot the average throughput and average contention delay per each transmitter versus PC P_1 's CWmin size, as shown in Figures 11 & 12. Although increasing the size of PC P_1 's CWmin improves the fairness between LTE and Wi-Fi systems, this improvement becomes negligible when the number of LTE and Wi-Fi transmitters becomes relatively high (see LTE and Wi-Fi plots with $m = n = 10$).

D. Arbitration Inter-frame Space and Fairness Tradeoffs

In Figures 13 & 14, we investigate how fairness between LTE and Wi-Fi systems can be improved by changing the size of AIFS number of LTE-LAA PC P_1 . We notice that changing the size of AIFS number has a higher impact on LTE and Wi-Fi performance when compared to changing the size of CWmin. Another interesting result relates to the average contention delay. The size of AIFS duration impacts the contention delay more than CWmin does. The increase in AIFS value reduces the average contention delay of LTE PCs up to a certain limit, but afterward the delay increases significantly, as shown in Figure 14. This happens because beyond this limit a transmitter contending with PC P_1 has to wait for a longer AIFS duration before resuming the counting, and during this duration other transmitters will be more likely to occupy the channel and start transmission. This forces transmitters with PC P_1 to freeze their counters for longer time, causing tremendous contention delay and starvation. Another important observation is that the fairness issue becomes less stringent when network grows (see the $m = n = 10$ plots in Figures 13 & 14).

VII. CONCLUSION

We presented a framework for allowing MNOs to assess the performance of assigning their prioritized traffics to channels in the unlicensed bands. We introduced a novel approximate closed-form expressions for computing the probability of successful transmission for LTE/Wi-Fi priority classes, and performed extensive simulations using discrete-event-based simulations to verify our numerical closed-form expressions and study the fairness of the prioritized LTE/Wi-Fi coexistence. Our simulation

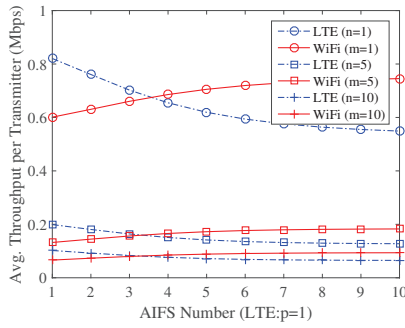


Fig. 13. Average throughput per transmitter vs. AIFS number of LTE priority class P_1 .

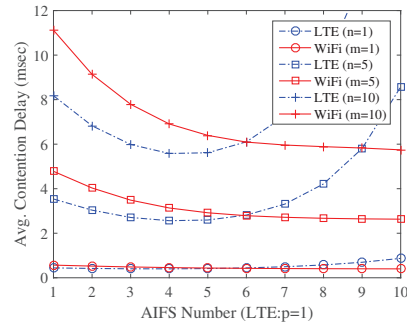


Fig. 14. Average contention delay per transmitter vs. AIFS number of LTE priority class P_1 .

results reveal that our closed-form model estimates the probability of successful transmission with high degree of accuracy, especially for networks with dense deployment and high traffic intensity.

REFERENCES

- [1] 3GPP, "Study on licensed-assisted access to unlicensed spectrum," 3GPP TR. 36.889 v13.0.0., Jun. 2015.
- [2] S. Verma and S. Adhikari, "3GPP RAN1 status on LAA and NR-Unlicensed," IEEE 802.11-18/0542r0, Mar. 2018.
- [3] "IMT vision - "framework and overall objectives for the future development of IMT for 2020 and beyond"," Recommendation ITU-R M.2083, Sep. 2015.
- [4] 3GPP, "Physical layer procedures," 3GPP TR. 36.213 v15.0.0., April. 2018.
- [5] "IEEE-part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications-amendment 4," <http://ieeexplore.ieee.org/servlet/opac?punumber=6687185>, 2013.
- [6] M. Hirzallah, Y. Xiao, and M. Krunz, "Modeling cross-technology coexistence with prioritized traffic classes," University of Arizona, Department of ECE, TR-UA-ECE-2018-2, [Available online] <http://wireless.ece.arizona.edu/sites/default/files/techrep-ua-ece-2018-2-mohammed.pdf>.
- [7] C. Hasan, M. K. Marina, and U. Challita, "On LTE-WiFi coexistence and inter-operator spectrum sharing in unlicensed bands: Altruism, cooperation and fairness," in *Proc. of MobiHoc'16 Conference*, 2016, pp. 111–120.
- [8] M. Hirzallah, W. Afifi, and M. Krunz, "Full-duplex-based rate/mode adaptation strategies for Wi-Fi/LTE-U coexistence: A POMDP approach," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 20–29, Jan 2017.
- [9] X. Cao, Z. Song, B. Yang, and Z. Han, "Full-duplex MAC protocol for Wi-Fi/LTE-U coexistence networks," in *Proc. of the IEEE ICC'18 Workshop*, May 2018, pp. 1–6.
- [10] F. Teng, D. Guo, and M. L. Honig, "Sharing of unlicensed spectrum by strategic operators," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 3, pp. 668–679, Mar. 2017.
- [11] Z. Guan and T. Melodia, "CU-LTE: Spectrally-efficient and fair coexistence between LTE and Wi-Fi in unlicensed bands," in *Proc. of IEEE INFOCOM'16 Conference*, April 2016, pp. 1–9.
- [12] Y. Xiao, M. Hirzallah, and M. Krunz, "Optimizing inter-operator network slicing over licensed and unlicensed bands," in *Proc. of the IEEE SECON'18 Conf.*, June 2018, pp. 1–9.
- [13] I. Samy, L. Lazos, Y. Xiao, M. Li, and M. Krunz, "LTE misbehavior detection in Wi-Fi/LTE coexistence under the LAA-LTE standard," in *Proc. of ACM WiSec'18 Conference*. ACM, 2018, pp. 87–98.

- [14] 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, Firstquarter 2017.
- [15] J. Zhang, M. Wang, M. Hua, T. Xia, W. Yang, and X. You, "LTE on license-exempt spectrum," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 647–673, Firstquarter 2018.
- [16] A. M. Voicu, L. Simić, and M. Petrova, "Survey of spectrum sharing for inter-technology coexistence," *arXiv preprint arXiv:1712.08589*, 2017.
- [17] V. Valls, A. Garcia-Saavedra, X. Costa, and D. J. Leith, "Maximizing LTE capacity in unlicensed bands (LTE-U/LAA) while fairly coexisting with 802.11 WLANs," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1219–1222, June 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 Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2968–2980, Nov 2016.
- [19] H. R. Lee, H. Kim, H. J. Yang, J. T. Kim, and S. K. Baek, "Performance analysis of license assisted access LTE with asymmetric hidden terminals," *CoRR*, vol. abs/1612.04244, 2016.
- [20] R. Yin, G. Yu, A. Maaref, and G. Y. Li, "LBT-based adaptive channel access for LTE-U systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6585–6597, Oct 2016.
- [21] S. Han, Y. C. Liang, Q. Chen, and B. H. Soong, "Licensed-assisted access for LTE in unlicensed spectrum: A MAC protocol design," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2550–2561, Oct 2016.
- [22] A. Abdelfattah and N. Malouch, "Modeling and performance analysis of Wi-Fi networks coexisting with LTE-U," in *In Proc. of IEEE INFOCOM 2017*, May 2017, pp. 1–9.
- [23] G. J. Sutton, R. P. Liu, and Y. J. Guo, "Delay and reliability of load-based listen-before-talk in LAA," *IEEE Access*, vol. 6, pp. 6171–6182, 2018.
- [24] M. Mehrmouh, V. Sathya, S. Roy, and M. Ghosh, "Analytical modeling of Wi-Fi and LTE-LAA coexistence: Throughput and impact of energy detection threshold," *arXiv preprint arXiv:1803.02444*, 2018.
- [25] G. Bianchi, I. Tinnirello, and L. Scalia, "Understanding 802.11e contention-based prioritization mechanisms and their coexistence with legacy 802.11 stations," *IEEE Network*, vol. 19, no. 4, pp. 28–34, July 2005.
- [26] "Further advancements for E-UTRA physical layer aspects," 3GPP TS. 36.814 V9.2.0, Mar. 2017.
- [27] L. Kleinrock, *Queueing systems, volume 2: Computer applications*. Wiley New York, 1976, vol. 66.
- [28] Z. ning Kong, D. H. K. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 10, pp. 2095–2106, Dec. 2004.
- [29] "Scipy optimization and root finding," [Available online] https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html.
- [30] "CSIM 20 - development toolkit for simulation & modeling," [Available online] <http://www.mesquite.com>.
- [31] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [32] S. M. Ross, *Stochastic processes*, 2nd ed. John Wiley & Sons New York, 1996.

APPENDIX A

DISCRETE TIME MARKOV CHAIN CHANNEL ACCESS MODEL

Channel access schemes described in Section III can be modeled using a three dimensional Markov chain by following a similar line as in [31] [28]. We define a three dimensional stochastic process $\xi^{(i)}(t) = \{(J^{(i)}(t), K^{(i)}(t), L^{(i)}(t))\}$, composed

of three stochastic processes, $J^{(i)}(t)$, $K^{(i)}(t)$, and $L^{(i)}(t)$. These three processes describe the channel access and contention behavior for a transmitter serving traffics of the i th priority class:

- The $J^{(i)}(t)$ process models all transmission attempts for a frame, including retransmissions. When collision happens the contending transmitter should backoff and recontend for a new channel access. Therefore, $J^{(i)}(t)$ takes an integer value in $[0, R_i]$, where R_i is the maximum retransmission limit of the i th priority class. When $J^{(i)}(t) = j$, we say the transmitter is at the j th *backoff stage* at time t . We also use $J^{(i)}(t)$ to model some special cases. For instance, $J^{(i)}(t) = -1$ indicates that the transmitter at time t has already finished frame transmission, and is ready to start serving a new frame, $J^{(i)}(t) = -2$ indicates that the transmitter at time t is transmitting a frame successfully, while $J^{(i)}(t) = -3$ denotes the transmitter at time t has no frames to be transmitted, i.e., the transmitter has an empty queue.
- The process $K^{(i)}(t)$ models the length of the backoff period at time t , taking an integer value in $[0, W_j^{(i)} - 1]$, where $W_j^{(i)}$ is the minimum contention window in (1) and j is the retransmission attempt. $K^{(i)}(t)$ process helps tracking the value of backoff counter, when $K^{(i)}(t) = k$ this indicates that at time t transmitter's backoff counter equals k .
- The $L^{(i)}(t)$ process models the time remaining before a channel becomes idle after a busy period. When channel becomes busy, the transmitter should freeze its counting and wait for channel to become idle again. The transmitter deems the channel as idle by waiting for d_i idle time slots, i.e., AIFS duration. $L^{(i)}(t) = l$ indicates that the transmitter at time t should wait for l time slots for the channel to be considered idle. The channel busy duration is random, and depends on the TXOP lengths adopted by coexisting transmitters. It is of a low probability that the channel remains busy continuously for a time period that exceeds the sum of TXOP periods of all coexisting transmitters. Therefore, we consider an intermediate case and set the busy time to $B = \max_i (T_i), i \in \mathcal{C} \cup \mathcal{L}$, i.e., the maximum duration of TXOP period of all LTE PCs and Wi-Fi ACs. We also utilize $L^{(i)}(t)$ process to model other special cases. For instance, $L^{(i)}(t)$ takes a value in $[1, T_i]$ to model the time spent in successful transmission, and a value in $[1, C_i]$ to model the time spent in collided transmission, where C_i is the number of time slots consumed in a collided transmission.

Channel access mechanisms adopted by LTE and Wi-Fi technologies consider slotted time structure, where each time slot corresponds to a MAC time slot, i.e., $T_{\text{slot}} = 9$ microseconds [4] [5]. Therefore, we introduce a discrete version of $\xi^{(i)}(t)$ denoted as $\xi^{(i)}(n) = \{(j, k, l) : j \in J^{(i)}(n), k \in K^{(i)}(n), l \in L^{(i)}(n)\}$. We are interested in modeling the behavior of LTE and Wi-Fi channel access mechanisms at the steady state, i.e., $\lim_{n \rightarrow \infty} \xi_n^{(i)} = \{(J^{(i)}, K^{(i)}, L^{(i)})\}$, therefore, we drop n for the rest of this section. We introduce a detailed Markov chain wherein each state takes a value in the combined stochastic process $\xi^{(i)} = \{(j, k, l)\}$, as shown in Figure 15. For simplicity of illustration, we categorize $\xi^{(i)}$ into the following categories:

- *Contention-with-Idle-Channel*: This category includes the set of states wherein the channel is idle and the transmitter backs off for k idle time slots, where k is the value of the backoff counter, i.e., $\xi^{(i)} = \{(j, k, 0), k \in [1, W_j^{(i)} - 1]\}$.
- *Contention-with-Busy-Channel*: This category includes the set of states wherein the channel is busy and the transmitter freezes its counter, waiting for the channel to become idle again, i.e., $\xi^{(i)} = \{(j, k, l), j \in [0, R_i], k \in [1, W_j^{(i)} - 1], l \in [1, d_i + B]\}$. A more detailed illustration for these states can be seen in Figure 16, where we show these states at the j th backoff stage. Transmitter detects a busy channel with a probability b_i , and the channel might become busy for several consecutive times. A transmitter deems a channel to be idle by waiting for d_i time slots after a busy period.
- *Channel-Access-Attempt*: This category includes the set of states for which the transmitter can access the channel and start transmission, i.e., $\xi^{(i)} = \{(j, 0, 0), j \in [0, R_i]\}$. The sum of the stationary probabilities of these states constitutes the probability of channel access (PCA), i.e., τ_i .
- *Transmission-with-Collision*: This category includes the set of states wherein transmission is failed due to collision or bad channel condition, i.e., $\xi^{(i)} = \{(j, 0, l), l \in [1, C_i]\}$. When collision happens, the transmitter doubles its minimum contention window, as in (1), and initiates the backoff counter with a new integer value in $[0, W_j^{(i)} - 1]$. This triggers a transition probability between state $(j - 1, 0, C_i)$ and one of the states $\xi^{(i)} = \{(j, k, 0), k \in [0, W_j^{(i)} - 1]\}$. Collision happens with a probability p_i .
- *Transmission-with-Success*: This category includes the set of states wherein the transmitter is involved in a successful transmission, i.e., $\xi^{(i)} = \{(-2, 0, l), l \in [1, T_i]\}$. When the transmitter reaches state $(j, 0, 0)$, it is free to start transmission over the air. Successful transmission of frames with i th priority class happens with a probability $p_{s,i} = \tau_i(1 - p_i)$.
- *Idle-Queue*: After finishing transmission, the transmitter checks its queue to serve a new frame. When queue is empty the transmitter waits for a new frame to arrive. This category models the state of the transmitter when its queue is empty, i.e., state $(-3, 0, 0)$. The probability of having an empty queue is $1 - q_i$, where q_i is computed in (5). A transmitter will be pending in this state until a new frame arrives. New frames arrive with probability g_i .
- *Post-Transmission*: The channel might become busy directly after either a successful or collided transmission, therefore, the transmitter should wait for the channel to become idle again before start contending for a new frame. This category models this post-transmission phase, i.e., $\xi^{(i)} = \{(-1, 0, l), l \in [1, d_i + B]\}$.

States in the set $X^{(i)} = \{(j, 0, 0) : 0 \leq j \leq R_i\}$, in Figure 15, represent all attempts to access the channel, and the sum of

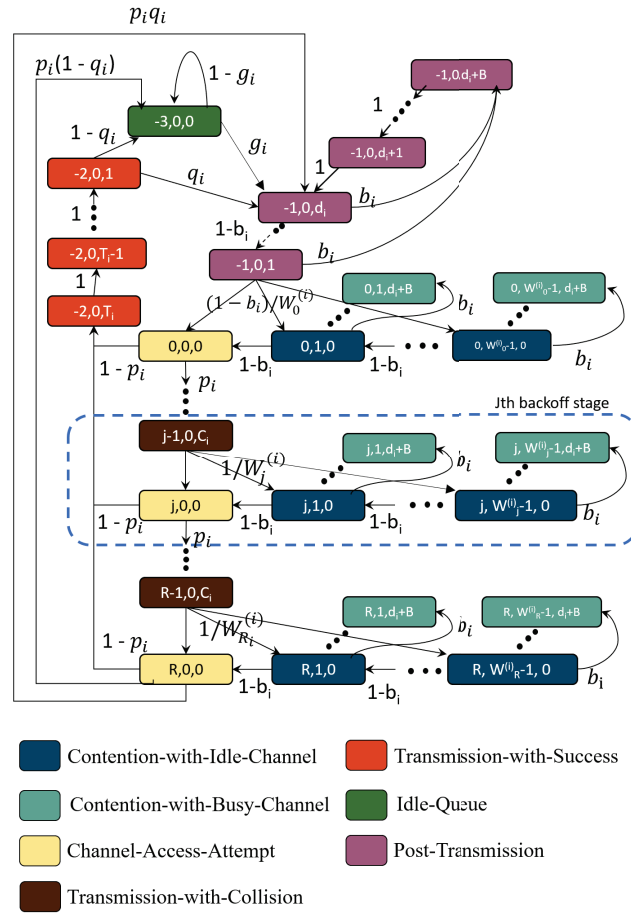


Fig. 15. Markov chain depicting channel access and contention behavior of the i th priority class belonging to EDCA/CAT-4 channel access mechanisms ($i \in \{C, L\}$).

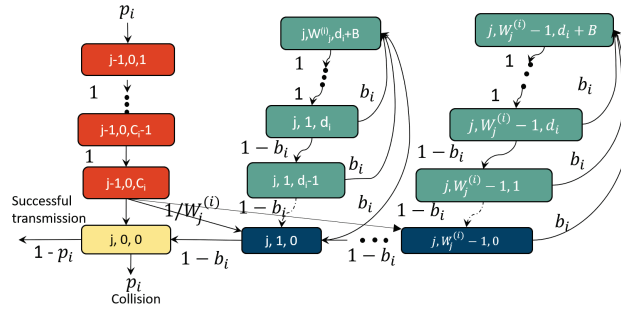


Fig. 16. Detailed illustration of the sub-chain corresponding to the j th backoff stage shown in Figure 15.

their steady state probabilities, i.e., $\pi_{(j,0,0)}^{(i)}$, gives the probability of channel access:

$$\tau_i = \sum_{j=0}^{R_i} \pi_{(j,0,0)}^{(i)} \quad (14)$$

To find $\pi_{(j,0,0)}^{(i)}$, we have to solve for the steady state probabilities for all states in Figure 15. Markov chain $\xi^{(i)}$ is a discrete time irreducible and aperiodic chain, and we can solve for its steady state probabilities by exploiting chain regularity and adding

all steady states to one. Let $x, y \in \xi^{(i)}$, then by chain regularity $\pi_x = \sum_y \pi_y Pr(x|y)$, where π_x and π_y are the steady-state probabilities of states x and y , respectively, and $Pr(x|y)$ is the transition probability from state y to state x . The steady state probabilities for all states can be expressed in terms of the steady state probability of state $(0, 0, 0)$, i.e., $\pi_{(0,0,0)}^{(i)}$. For instance, $\pi_{(j,0,0)}^{(i)}$ can be expressed as follows:

$$\pi_{(j,0,0)}^{(i)} = p_i^j \pi_{(0,0,0)}^{(i)} \quad (15)$$

By substituting (15) in (14), we get:

$$\tau_i = \frac{1 - p_i^{R_i+1}}{1 - p_i} \pi_{(0,0,0)}^{(i)} \quad (16)$$

The steady state probabilities for all other states can be expressed in a similar manner:

$$\pi_{(j,0,l)}^{(i)} = p_i \pi_{(j,0,0)}^{(i)}, j \in [0, R_i], l \in [1, C_i] \quad (17)$$

$$\pi_{(j,k,0)}^{(i)} = \frac{W_j^{(i)} - k}{W_j^{(i)}} \pi_{(j,0,0)}^{(i)},$$

$$j \in [0, R_i], k \in [1, W_j^{(i)} - 1] \quad (18)$$

$$\pi_{(j,k,l)}^{(i)} = \frac{b_i}{(1 - b_i)^{d_i}} \pi_{(j,k,0)}^{(i)},$$

$$j \in [0, R_i], k \in [1, W_j^{(i)} - 1], l \in [d_i, d_i + B] \quad (19)$$

$$\pi_{(j,k,l)}^{(i)} = \frac{b_i}{(1 - b_i)^l} \pi_{(j,k,0)}^{(i)},$$

$$j \in [0, R_i], k \in [1, W_j^{(i)} - 1], l \in [1, d_i - 1] \quad (20)$$

$$\pi_{(-1,0,l)}^{(i)} = \frac{1}{(1 - b_i)^l} \pi_{(0,0,0)}^{(i)}, l \in [1, d_i] \quad (21)$$

$$\pi_{(-1,0,l)}^{(i)} = \frac{1 - (1 - b_i)^{d_i}}{(1 - b_i)^{d_i}} \pi_{(0,0,0)}^{(i)},$$

$$l \in [d_i + 1, d_i + B] \quad (22)$$

$$\pi_{(-2,0,l)}^{(i)} = (1 - p_i^{R_i+1}) \pi_{(0,0,0)}^{(i)}, l \in [1, T_i], \quad (23)$$

and finally for the idle-queue state:

$$\begin{aligned} \pi_{(-3,0,0)}^{(i)} &= \pi_{(R_i,0,0)}^{(i)} p_i (1 - q_i) + \\ &\quad \pi_{(-2,0,1)}^{(i)} (1 - q_i) + \pi_{(-3,0,0)}^{(i)} (1 - g_i) \\ &= \frac{(1 - q_i)(1 + p_i^{R_i} - p_i^{R_i+1})}{g_i} \pi_{(0,0,0)}^{(i)} \end{aligned} \quad (24)$$

We sum the steady state probabilities of all states in (15) and (17)-(24) to one [32], and solve for the steady state probability $\pi_{0,0,0}^{(i)}$ of state $(0, 0, 0)$:

$$1 = \pi_{(-3,0,0)}^{(i)} + \sum_{d=1}^{T_i} \pi_{(-2,0,d)}^{(i)} + \sum_{l=1}^{d_i+B} \pi_{(-1,0,l)}^{(i)} + \sum_{j=0}^{R_i} \sum_{l=1}^{C_i} \pi_{(j,0,l)}^{(i)} + \sum_{j=0}^{R_i} \sum_{k=1}^{W_j^{(i)}-1} \sum_{l=0}^{d_i+B} \pi_{(j,k,l)}^{(i)}, \quad (25)$$

$$\pi_{0,0,0}^{(i)} = \left[\frac{1+b_i B}{b_i} \frac{1-(1-b_i)^{d_i}}{(1-b_i)^{d_i}} + T_i(1-p_i^{R_i+1}) + (1+p_i C_i) \frac{1-p_i^{R_i+1}}{1-p_i} + \frac{(1-q_i)(1+p_i^{R_i}-p_i^{R_i+1})}{g_i} \frac{1+b_i B}{2(1-b_i)^{d_i}} \sum_{j=0}^{R_i} W_j^{(i)} p_i^j \right]^{-1}, \quad (26)$$

and finally substituting (26) in (16) results in (9).

APPENDIX B

PROOF OF PROPOSITION 1

Proof: Our proof is based on results obtained from Section V. The average contention delay can be taken by considering the expected ‘return time’ to the states of successful transmission in Figure 15 (i.e., $\xi^{(i)} = \{(-2, 0, l) \text{ for } 1 \leq l \leq T_i\}$). The expected return time to a state of discrete time Markov chain (DTMC) can be obtained by taking the inverse of its steady state probability. Therefore, the average contention delay for frames of the i th priority class can be obtained as:

$$D_i = \frac{1}{\pi_{(-2,0,l)}^{(i)}} \quad (27)$$

Notice that the steady state probability $\pi_{(-2,0,l)}^{(i)}$ can be expressed using chain regularity in terms of the steady state probability of the channel access states (i.e., $\pi_{(j,0,0)}^{(i)}$):

$$\pi_{(-2,0,l)}^{(i)} = \sum_{j=0}^{R_i} (1-p_i) \pi_{(j,0,l)}^{(i)} = (1-p_i) \sum_{j=0}^{R_i} \pi_{(j,0,0)}^{(i)} \quad (28)$$

By noticing that $\sum_{j=0}^{R_i} \pi_{(j,0,0)}^{(i)} = \tau_i$, we observe that $\pi_{(-2,0,l)}^{(i)} = (1-p_i)\tau_i = p_{s,i}$. ■

APPENDIX C

PROOF OF THEOREM 1

Proof: Probability is proportional to its logarithm. Although it is not a direct proportionality, we still can make use of the property to simplify PST formulation. To simplify illustration, we focus on the PST expression in (10), where similar

manipulation can be applied to the PST in (11). We take the logarithm of PST expression in (10):

$$\begin{aligned} \log(p_{s,i}^{(l)}) &= \log(\tau_i) + (n_i - 1) \log(1 - \tau_i) \\ &+ \sum_{j \in \mathcal{L}, j \neq i} n_j \log(1 - \tau_j) \quad + \sum_{\theta \in \mathcal{C}} m_\theta \log(1 - \tau_\theta) \end{aligned} \quad (29)$$

Enlightened by the expression in (29) and by noticing that the logarithm is monotonically increasing with probability, we consider the following approximate formulation for $p_{s,i}^{(l)}$:

$$\begin{aligned} p_{s,i}^{(l)} &\approx c_{i,0} + \sum_{j \in \mathcal{L}} c_{i,j} (n_j + 1) \log(1 - \tau_j) \\ &+ \sum_{\theta \in \mathcal{C}} h_{i,\theta} (m_\theta + 1) \log(1 - \tau_\theta) \end{aligned} \quad (30)$$

where $c_{i,0}$, $c_{i,j}$, and $h_{i,\theta}$ are constant and proportionality coefficients that we can obtain by fitting the formula in (30) with PST measurements collected from system-level simulations and/or real-life networks. The one terms in $(n_j + 1)$ and $(m_\theta + 1)$ are added as artificial regulation constants to ensure that the fitting process is well-behaved and does not diverge for scenarios in which n_j or m_θ could be zeros.

Our formulation in (30) still requires computing $1 - \tau_j$ and $1 - \tau_\theta$. We focus on a general $1 - \tau_i$, and later apply the same logic to them. By looking at τ_i formulation in (9), we notice that it depends on channel access parameters, i.e., size of contention window $W_j^{(i)}$, AIFSN d_i , and TXOP T_i , as well as the probability of collision p_i and counter freezing b_i . As can be seen from (6), (7), and (8), p_i and b_i also depend on the number of coexisting nodes, i.e., $\mathbf{n} = \langle n_1, \dots, n_{N_{pc}} \rangle$ and $\mathbf{m} = \langle m_1, \dots, m_{N_{ac}} \rangle$, and their channel access probabilities, i.e., $\bar{\tau}^{(l)} = \langle \tau_{P_1}, \dots, \tau_{P_{N_{pc}}} \rangle$ and $\bar{\tau}^{(w)} = \langle \tau_{A_1}, \dots, \tau_{A_{N_{ac}}} \rangle$. The final closed-form expression of τ_i can be put as $\tau_i = f_i(\bar{\tau}^{(l)}, \bar{\tau}^{(w)})$, where $f_i(\cdot)$ is a non-linear complicated function. Therefore, simplifying τ_i by starting at (9) is not straightforward, and this is due to the complicated dependency on other τ 's.

To avoid complications, we follow an engineering-oriented approach to simplify $1 - \tau_i$. Based on extensive results that we obtained by solving for τ_i in (9) numerically, as well as carrying system-level simulations, we notice that the increase in the number of coexisting nodes reduces the probability of channel access, i.e., $n_i \uparrow$ leads to $\tau_i \downarrow$. Similarly, increasing the size of the minimum contention window, AIFSN, and TXOP, also reduces the probability of channel access, i.e., $W_0^{(i)} \uparrow, d_i \uparrow, T_i \uparrow$

leads to $\tau_i \downarrow$. Therefore, we can lead to the following observations:

$$\begin{array}{lll}
 n_i \uparrow, & \tau_i \downarrow, & 1 - \tau_i \uparrow \\
 W_0^{(i)} \uparrow, & \tau_i \downarrow, & 1 - \tau_i \uparrow \\
 d_i \uparrow, & \tau_i \downarrow, & 1 - \tau_i \uparrow \\
 T_i \uparrow, & \tau_i \downarrow, & 1 - \tau_i \uparrow
 \end{array}$$

Therefore, we propose the following expression to approximate $1 - \tau_i$:

$$1 - \tau_i \approx \beta_i(W_0^{(i)}d_i + T_i)n_i + e_i, \quad (31)$$

where β_i and e_i are proportionality coefficient and constant that can be obtained by fitting. Based on our numerical as well as simulation result, we have noticed that AIFSN affects τ_i more than the size of the minimum contention window. A similar observation has also been captured by author in [25]. Our choice of multiplying d_i and $W_j^{(i)}$ as well as multiplying $W_0^{(i)}d_i + T_i$ and n_i in (31) is based on making a better distinction between LTE and Wi-Fi priority classes relevant to the number of their serving transmitters. By substituting (31) in (30), we obtain the approximate expression in (12). ■