ORLA/OLAA: Orthogonal Coexistence of LAA and WiFi in Unlicensed Spectrum

Andres Garcia-Saavedra, Paul Patras, Victor Valls, Xavier Costa-Perez, and Douglas J. Leith

Abstract—Future mobile networks will exploit unlicensed spectrum to boost capacity and meet growing user demands cost-effectively. The 3rd Generation Partnership Project (3GPP) has recently defined a License Assisted Access (LAA) scheme to enable global Unlicensed LTE (U-LTE) deployment, aiming at (i) ensuring fair coexistence with incumbent WiFi networks, i.e., impacting on their performance no more than another WiFi device, and (ii) achieving superior airtime efficiency as compared to WiFi. In this paper we show the standardized LAA fails to simultaneously fulfill these objectives, and design an alternative orthogonal (collision-free) listen-before-talk coexistence paradigm that provides a substantial improvement in performance, yet imposes no penalty on existing WiFi networks. We derive two LAA optimal transmission policies, ORLA and OLAA, that maximize LAA throughput in both asynchronous and synchronous (i.e., with alignment to licensed anchor frame boundaries) modes of operation, respectively. We present a comprehensive performance evaluation through which we demonstrate that, when aggregating packets, IEEE 802.11ac WiFi can be more efficient than LAA, whereas our proposals attain 100% higher throughput, without harming WiFi. We further show that long U-LTE frames incur up to 92% throughput losses on WiFi when using 3GPP LAA, whilst ORLA/OLAA sustain >200% gains at no cost, even in the presence of non-saturated WiFi and/or in multi-rate scenarios.

I. INTRODUCTION

5G Radio Access Network (RAN) architects actively seek to augment mobile systems with inexpensive spectrum in order to boost network capacity and meet growing user demand in a cost-effective manner. License-exempt 5 GHz U-NII channels, currently exploited almost exclusively by WiFi deployments, are of particular interest to the 3GPP community [1], [2], which is pursuing U-LTE standardization. However, the substantial differences between incumbent WiFi, which employs a Listen before Talk (LBT) contention-based multiplexing protocol, i.e. Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), and Long Term Evolution (LTE), which is inherently a scheduled paradigm, makes the design of U-LTE channel access protocols particularly challenging [3], [4]. The fundamental question facing U-LTE design is how to exploit this uncharted spectrum efficiently whilst playing fair with native technologies?

Initial U-LTE solutions employ a Carrier Sensing and Adaptive Transmission (CSAT) scheme based on channel selection and time-based duty cycling [5]. The simplicity of this approach has enabled a short time to market in some countries (USA, Korea, India), but it is incompatible with LBT regulations in regions such as Europe and Japan [6]. More recent 3GPP specifications put forward an LBT-based solution named LAA, to address this deficiency [2]. The LBT flavor of LAA is similar to WiFi CSMA/CA, thereby enabling global U-LTE deployment. The impact of different U-LTE approaches on the performance of WiFi technology native to unlicensed bands is continuously studied by the research community [7]–[11].

In this paper we argue that the 3GPP LAA scheme is unable to simultaneously meet the following key design criteria: (i) avoid harming the performance of preexisting WiFi Wireless Nodes (WNs), and (ii) provide superior Media Access Control (MAC) protocol efficiency as compared to WiFi. We illustrate this with an example in Fig. 1, where 1 LAA WN coexists with 5 backlogged IEEE 802.11ac WNs. We plot the MAC throughput gain/loss attained by LAA and WiFi relative to the performance that would have been achieved if an additional WiFi WN was added to the network, instead of the LAA WN using the same packet burst size as the remaining WiFi WNs (i.e. a 6-WiFi WNs scenario). Results are shown for a range of LAA frame durations (x-axis) and WiFi packet burst sizes (y-axis). Observe that LAA achieves its highest throughput gain when the background WiFi WNs are most penalized (bottom right corners). This is due to the fact that LAA transmits data for 10 ms upon winning a successful transmission opportunity whereas a WiFi WN only transmits for 62 µs (1 KB at 130 Mbits/s). This leads to WiFi having a high relative overhead (inter-frame spaces and idle backoff

1We refer the interested reader to §VIII for a review of related research.

2With data ready to transmit at all times, i.e., saturated.
periods). Further, observe that in the region where LAA is relatively harmless to WiFi (top left corners), it experiences barely any throughput performance benefit (if not a loss).

In essence, since the LAA access procedure fundamentally resembles that of WiFi’s CSMA/CA it does have the ability to provide fairness; however, increases in MAC efficiency are achieved at the cost of greatly penalizing WiFi devices. Clearly, despite these preliminary results (a thorough evaluation campaign is presented in §VI), U-LTE can provide enhanced PHY-layer efficiency (e.g., more robust error recovery mechanisms) and operational advantages (e.g., common radio resource management with licensed LTE), which may sustain the appeal of LAA irrespective of its (inefficient) MAC-layer coexistence mechanism. In this paper we argue, however, that U-LTE’s MAC-layer efficiency can (and should) be substantially improved without compromising incumbent WiFi. We therefore revisit U-LTE’s MAC coexistence design, making use of the observation that if the airtime used by WiFi and U-LTE were decoupled by eliminating inter-technology collisions, we could build U-LTE access schemes employing transmission strategies that are demonstrably harmless to WiFi and achieve higher data rates compared to CSMA/CA-based alternatives for cellular access to unlicensed spectrum. These include LTE-WLAN Aggregation (LWA/LTE-H) [13] and 3GPP LAA. In summary, we make the following contributions:

- We design an orthogonal (collision-free) LBT-based access protocol that coexists with WiFi truly seamlessly;
- We construct optimal LBT transmission policies that suit a variety of network conditions, including heterogeneous traffic loads and link rates;
- We devise optimal policies for synchronous LBT technologies, constrained to transmit during fixed intervals;
- We undertake a thorough simulation campaign to evaluate the proposed policies and demonstrate LBT throughput gains of over 200% with no negative impact on WiFi;
- We give practical guidelines to aid the implementation of the mechanisms we propose on off-the-shelf hardware.

The rest of the paper is organized as follows. §II reviews fundamental technical concepts underpinning the design of the proposed U-LTE coexistence scheme and in §III we present their operation principles. In §IV and §V we introduce optimal LBT transmission policies that simultaneously accomplish fair resource sharing with WiFi WNs and throughput maximization. We evaluate via simulations our access procedures and transmission policies in §VI, comparing against 3GPP’s LAA scheme. We discuss practical details that can facilitate the implementation of our solutions in §VII. Finally, §VIII reviews related work and §IX draws concluding remarks.

II. PRELIMINARIES

We begin by reviewing technical aspects fundamental to the design of our coexistence mechanism we propose, namely

(i) LBT regulatory constraints in 5 GHz bands, (ii) the Distributed Coordination Function (DCF) that governs communication in IEEE 802.11 WiFi, and (iii) the 3GPP LAA mechanism. To stress that our approach extends to any technology that seeks operation on 5 GHz channels, hereafter we refer to an LTE-LAA cell as an LBT WN.

A. Listen Before Talk (LBT)

We focus on European Telecommunications Standards Institute (ETSI) regulation 301 893 [6], as it is the most restrictive. ETSI specifies that Load Based Equipment (LBE) shall implement LBT following Clear Channel Assessment (CCA) using energy detection. Energy detection requires to observe the channel for a pre-defined duration and determine whether the energy level sensed exceeds a sensitivity threshold. In such case, the channel is regarded busy and transmission is deferred. ETSI mandates that CCA assertion can be performed in accordance with the IEEE 802.11 standard’s provisions [12]. Alternatively, minimum requirements should be met, as defined by two channel access options stipulated in [6].

In our design, we consider LBT WNs that comply with clause 18 of IEEE 802.11 [12], which requires Orthogonal Frequency Division Multiplexing (OFDM) transmitters to identify a busy channel within 4 µs. This also requires a WiFi station operating in the 5 GHz band to observe the channel idle for at least DCF Interframe Space (DIFS) time, $DIFS = 34 \mu s$, before attempting to transmit. We further note that 802.11 data frames and Acknowledgments (ACKs) are separated by a Short Interframe Space (SIFS), $SIFS = 16 \mu s$. We propose allowing the LBT WN to attempt transmission following an 802.11 frame exchange, immediately after the channel is sensed idle for an LTE Interframe Space (LIFS) of 20 µs that we introduce. This ensures an 802.11 frame exchange will not be interrupted, while avoiding potential collisions with co-existing 802.11 stations (which may transmit immediately after DIFS, if initializing random back-off counters with zero). We summarize the 802.11 channel access procedure next.

B. IEEE 802.11 MAC protocol (WiFi)

WiFi medium access is regulated by the IEEE 802.11 DCF, which performs CSMA/CA with Binary Exponential Backoff (BEB). While a detailed description is given in [14], we include here a brief description here for completeness. An IEEE 802.11 network divides time into MAC slots and a station transmits after observing $S_m$ idle slots, where $S_m$ is a random variable selected uniformly at random from $\{0, 1, \ldots, 2^m CW_{\text{min}} - 1\}$ where $CW_{\text{min}} \in \mathbb{N}$ is the minimum Contention Window (CW) and $m = 0, 1, 2, \ldots$ is the number of successive collisions experienced by the station. After a successful transmission $m$ is set to 0. IEEE 802.11 defines a parameter $CW_{\text{max}}$ that limits the expected number of idle slots a station has to wait after $m$ successive collisions, i.e., $2^m CW_{\text{min}} = CW_{\text{max}}$ for $m \geq \bar{m}$.

Two key features of WiFi systems relevant to our work are: (i) each 802.11 packet includes in the header information regarding the duration of the transmission, i.e., upon correct reception of a packet, a station knows the duration for which the channel will be busy; (ii) after a successful transmission all stations in the network wait for an Arbitration Interframe Space (AIFS) time of at least 34 µs. That is, importantly, after each successful transmission there will be at least 34 µs during which the channel is free of WiFi transmissions.

3LBE does not follow a fixed TX/RX pattern, but is driven by demand [6]. In contrast, Frame Based Equipment (FBE) transmits at fixed intervals.

4AIFS is equivalent to DIFS and lasts at least 34 µs in 5 GHz bands.
C. 3GPP U-LTE specification (LAA)

The 3GPP introduced LAA in release 13 specification, as a mechanism for boosting downlink capacity [2]. This is achieved by aggregating unlicensed spectrum in the 5GHz band with licensed spectrum available to the operator. To ensure seamless rollout, the licensed spectrum is regarded as an anchor that transports signaling and control messages.

Seeking fair coexistence with other technologies native to the 5GHz band, including WiFi, LAA adopts the LBT strategy specified by ETSI [6]. A CSMA/CA approach is implemented on the unlicensed carrier, similar to that of WiFi. Specifically, if the channel is sensed idle for an initial CCA period, a transmission is triggered. Otherwise, the base station draws a random counter withing a contention window range, which is doubled upon a collision and reduced to a minimum value after a successful transmission, as inferred through Hybrid Automatic Repeat reQuest (HARQ) ACK. Every time the channel is sensed idle for an extended CCA period, the counter is decreased. A transmission is triggered when the counter reaches zero and consists of a Physical Downlink Shared Channel (PDSCH) burst. Depending on channel activity and priority class, LAA can keep the channel occupied for at least 1 ms and up to 10 ms.

III. ORTHOGONAL AIRTIME COEXISTENCE

Our first objective is to design a MAC-layer protocol for a non-WiFi LBT-based WN (e.g., LTE-LAA cells) that fulfills the following criteria: (i) pose no harm to preexisting WiFi wireless networks, and (ii) improve MAC efficiency over WiFi WNs. To formalize these objectives, we first introducing the following definitions.

**Definition 1.** We let “WiFi WN” define a wireless transmitter operating in unlicensed bands following the IEEE 802.11 specification.

By the above definition, WiFi WNs observe a slotted channel, where slot duration varies as the channel can be idle, occupied by a transmission (whose length depends on the payload size and bit rate), or contain a collision between two or more simultaneous WiFi WN transmissions.

**Definition 2.** We let “LBT WN” define a non-WiFi wireless transmitter operating in unlicensed bands using a LBT access mechanism. “Synchronous LBT WN” refers to LBT WNs constrained to transmit in predefined fixed timeslots, and “asynchronous LBT WN” those which are not.

We note that LTE-LAA may require synchronization with Time Division Multiple Access (TDMA) LTE framing in anchor licensed bands (synchronous LTE-LAA) or not, as in the case of MulteFire [15] (or asynchronous LTE-LAA). Importantly, we let “WiFi/LBT WN” refer to any transmitter using unlicensed wireless channels including the WiFi AP in downlink scenarios and any WiFi station in uplink communication so in general we will avoid making explicit reference to uplink or downlink.

Our coexistence approach builds upon the key observation that the minimum duration of an DIFS/AIFS (34μs) is longer than the CCA minimum time (20μs) specified by ETSI. More specifically, to avoid inter-technology collisions, we assume that an LBT WN can acquire the channel if the medium is sensed idle for a LIFS = 20μs (LBT inter-frame space) duration, a timing constant we introduce. Note that SIFS < LIFS < PIFS (< AIFS/DIFS), which means LBT transmissions take priority, but cannot interrupt ongoing data–ACK exchanges (which are separated by a SIFS) in concurrent WiFi transmissions. Thus, allowing idle channel acquisition after LIFS enables an LBT system to opportunistically exploit orthogonal collision-free airtime in unlicensed spectrum.

To support synchronous LBT WNs that need to postpone data transmissions until the beginning of a frame, we employ the 802.11 compliant Clear To Send (CTS)-to-self mechanism. This enables the synchronous LBT WN to reserve the channel and so ensure this frame alignment. Specifically, if the synchronous LBT WN senses the channel idle for LIFS, it can either send a CTS-to-self to reserve the medium until the next frame, or it can defer access until the next opportunity. In either case, when an asynchronous/synchronous LBT WN decides to acquire the channel, it will hold the channel for a fixed duration $T_{LBT}$, as permitted by regulations for both FBE or LBE [6]. We exemplify the proposed modus operandi in Fig. 2.

Note that with the above scheme, transmissions by LBT and WiFi WNs will never collide, hence an LBT WN does not affect the transmission attempt probability of the stations in a WLAN. Therefore, the shared channel can be regarded as divided into two orthogonal airtimes. In addition, since an LBT WN will always sense the channel idle immediately after a WiFi transmission, the policies specified for FBE and LBE as to how to perform another CCA when the channel is sensed busy are irrelevant to this work. Note however that collisions amongst LBT WNs could occur. In the rest of the paper we will assume there is only one LBT WN, which is in line with

---

3 Point Coordination Function Interframe Space (PIFS) of 25 μs is used by WiFi access points that can support contention free periods.

6 This is true under certain assumptions which we discuss next and in §VII.
3GPP [2] and related work (e.g. [17]–[20]), and corresponds to the case of downlink offloading with cellular operators using different channels to circumvent interference. We discuss the case with multiple LBT WN sharing the same channel in §VII.

Making use of the LBT access protocol introduced above, our second objective it to find optimal transmission policies for both synchronous and asynchronous LBT. To this end, we finalize this section by formalizing our design criteria.

**Definition 3 (LBT transmission policy).** Index the LIFS transmission opportunities (the times when a station using our LBT access protocol could potentially transmit) by \( \omega = 1, 2, \ldots \) and gather these indices into set \( \Omega \subseteq \mathbb{N} \). Then \( \Pi \subseteq 2^\Omega \) defines an LBT transmission policy, where \( 2^\Omega \) denotes the set of all possible subsets (the superset of \( \Omega \)). An LBT station using policy \( \Pi \) transmits one burst of duration \( T_{LBT} \) at each LIFS transmission opportunity \( \omega \in \Pi \).

**Definition 4 (Optimal policies).** An airtime-optimal policy \( \Pi^* \) and a throughput-optimal policy \( \Pi^{\text{del}} \), respectively, the channel time held by an LBT WN and the throughput of an LBT WN, while preserving (at least) the same amount of channel time each coexisting Wi-Fi WN would have if the LBT WN implemented a Wi-Fi access protocol.

In what follows, we construct optimal LBT transmission policies for a range of scenarios.

**IV. ORLA: ORTHOGONAL RANDOM LBT UNLICENSED ACCESS STRATEGY**

In this section we devise optimal LBT transmission policies for asynchronous LBT WNs coexisting with Wi-Fi. We first address airtime maximization under circumstances where Wi-Fi contenders are backlogged (saturated) and operate with the same PHY bit rate (homogeneous), then focus on maximizing airtime usage when Wi-Fi is lightly loaded (non-saturated) and WNs employ dissimilar rates (heterogeneous links).

**A. Homogeneous and Saturated Conditions**

The scope of our work is MAC layer enhancement, therefore we assume ideal physical layer conditions, i.e LBT and Wi-Fi WNs are within carrier sensing range of each other (no hidden terminals), no capture effect, and perfect PHY rate control mitigates channel errors. Thus losses due to fading are negligible, whilst we consider practical multi-rate Wi-Fi operation in §IV-B. These assumptions not only ensure mathematical tractability but help us keep a focused analysis. We assume LBT WNs equipped with an off-the-shelf WiFi interface for channel sensing and medium access reservation purposes, in addition to the native LTE modem. We note this is common practice in the related literature (see e.g. [17]) and we discuss in detail the above practical implications in §VII.

Consider a scenario with \( n \) saturated WiFi WNs, i.e., each WN always has data ready for transmission. It is well known that under these conditions the transmission attempt probability \( \tau_i^{(n)} \) of WN \( i \) in random MAC slot can be related to the conditional collision probability \( p_i^{(n)} \). Given the homogeneous load assumption, \( \tau_i^{(n)} = \tau_i^{(1)} \) for all \( i \in \{1, \ldots, n\} \) (we relax this assumption later), \( \tau_i^{(n)} \) can be computed by solving the following system of non-linear equations [22]:

\[
\begin{align*}
\tau_i^{(n)} &= \frac{2(1-2p_i^{(n)})}{(1-2p_i^{(n)})C_{\text{channel}})+(1-2\sigma^{(n)}C_{\text{channel}})} \cdot \left(1 - (1 - \tau_i^{(n)})^{n-1}\right),
\end{align*}
\]

where recall \( C_{\text{channel}} \) is the minimum contention window and \( m \) is the backoff stage. Then, the probability that a MAC slot is idle is given by the probability that none of the WNs transmits, i.e., \( P_{\text{idle}}^{(n)} = (1 - \tau_i^{(n)})^n \); the probability that a slot is occupied by a successful transmission is \( P_{\text{success}}^{(n)} = nP_{\text{success}}^{(1)} \), where \( P_{\text{success}}^{(n)} = \tau_i^{(n)}(1 - \tau_i^{(n)})^{n-1} \) is the probability that a single WN transmits in a MAC slot. Finally, the probability that a slot is occupied by a collision is given by \( P_{\text{coll}}^{(n)} = 1 - P_{\text{idle}}^{(n)} - P_{\text{success}}^{(n)} \) and the probability of a slot being busy is \( P_{\text{tx}}^{(n)} = P_{\text{coll}}^{(n)} + P_{\text{success}}^{(n)} \). The throughput of a WiFi WN is given by

\[
s^{(n)} = \frac{P_{\text{success}}^{(n)}B}{P_{\text{idle}}^{(n)}} \cdot \sigma + (1 - P_{\text{idle}}^{(n)}) T,
\]

where \( \sigma, B, \text{and } T \) are the duration of an (idle) MAC slot, the expected number of bits in a transmission, and the duration of a transmission (successful or collision) respectively, which is equal to

\[ T = T_{\text{PLCP}} + f_{\text{agg}} \left( L_{\text{del}} + L_{\text{mac-oh}} + L_{\text{pad}} \right) + B + \text{SIFS} + T_{\text{ACK}} + \text{DIFS}. \]

SIFS, DIFS, \( T_{\text{PLCP}} \), \( L_{\text{del}} \), and \( L_{\text{pad}} \) are PHY layer constants (inter-frame spacing, delimiters, padding), \( L_{\text{mac-oh}} \) is the MAC layer overhead (header and FCS), \( f_{\text{agg}} \) is the number of packets aggregated in a transmission, \( B \) the expected number of data bits transmitted in the burst (payload), and \( C \) the PHY bit rate.

The duration of an acknowledgement is \( T_{\text{ACK}} = T_{\text{PLCP}} + \frac{L_{\text{ctrl}}}{C_{\text{ctrl}}} \), where \( C_{\text{ctrl}} \) is the bitrate used for control messages.

We aim to obtain the maximum fraction of orthogonal airtime that an LBT WN can use such that the average throughput experienced by a WiFi WN is not degraded more than if another WiFi WN were added to the network. Since LBT transmissions following the access procedure proposed in §III are orthogonal to WiFi transmissions, an LBT WN can be regarded (in terms of airtime) as a WiFi WN that transmits in MAC slots that otherwise would be idle. Then, the LBT airtime can be expressed as

\[
A_{\text{LBT}} = \rho P_{\text{idle}}^{(1)} \left( T' - \sigma \right),
\]

where \( \rho \in [0, 1] \) is the fraction of idle slots that would change to busy slots, and \( (T' - \sigma) = T_{LBT} > 0 \) is the duration of an LBT WN transmission, which depends on the LBT mode used (FBE or LBE). Note that the quantity \( \rho P_{\text{idle}} \) is the fraction of orthogonal LBT transmissions. With (2) we can write the throughput experienced by a WiFi WN when an LBT WN uses \( A_{\text{LBT}} \) airtime, as follows:

\[
s^{(n+1)} = \rho P_{\text{success}}^{(n)} B \cdot \frac{1}{P_{\text{idle}}^{(n)}} \left( P_{\text{tx}}^{(n)} + P_{\text{coll}}^{(n)} \right) T + \rho P_{\text{idle}}^{(n)} \left( T' - \sigma \right). \]
1, 2, ..., we have that
\[ s^{(n+1)} = \frac{p^{(n+1)} P^{(n)}_s}{p^{(n+1)} \sigma + p^{(n+1)} T} \leq s^{(n+1)\text{LBT}} \] (4)

will always hold, provided \( \rho \) in (3) is sufficiently small. We are interested in finding the value of \( \rho \) that makes (4) tight, i.e., maximizes the LBT airtime. We give the following lemma.

**Lemma 1.** Consider \( n \) homogeneous saturated WNs. Suppose \( T, T' > \sigma \). Then, (4) holds for every \( \rho \in [0, \bar{\rho}] \) with
\[ \bar{\rho} := \left( \frac{T}{T' - \sigma} \right) \min \left\{ 1, \frac{p^{(n)}}{p^{(n+1)}_s} \rho - \frac{p^{(n)}_s}{p^{(n+1)}_s} \right\} \] (5)

Proof: Rearranging terms in (4) with \( P_{\text{idle}} = (1 - P_{\text{idle}}) \) and \( A = \rho P_{\text{idle}}^{(n)} (T' - \sigma) \) we have that
\[ \rho P_{\text{idle}}^{(n)} (T' - \sigma) \leq \frac{p^{(n)}_s}{p^{(n+1)}_s} \left( p^{(n+1)}_s (\sigma - T) + T + \rho P_{\text{idle}}^{(n+1)} (T' - \sigma) \right), \] (6)

Further rearranging we obtain that
\[ \rho P_{\text{idle}}^{(n)} (T' - \sigma) \leq p^{(n)}_s \left( 1 + \frac{p^{(n)}_s}{p^{(n+1)}_s} \right) (T' - \sigma), \]
and dividing by \( P_{\text{idle}}^{(n)} (T' - \sigma) \) yields
\[ \rho \leq \frac{p^{(n)}_s}{p^{(n+1)}_s} \left( 1 + \frac{p^{(n)}_s}{p^{(n+1)}_s} \right). \]

Now fix \( T' = T \) and see that since \( T/\sigma > 1 \) we have
\[ \rho \leq \frac{1}{p^{(n+1)}_s} \left( 1 + \frac{p^{(n)}_s}{p^{(n+1)}_s} \right), \]
\[ \leq \frac{p^{(n+1)}_s}{p^{(n+1)}_s} \left( 1 + \frac{p^{(n)}_s}{p^{(n+1)}_s} \right), \] (7)

where in (7) we have used the fact that \( 1 - P_{\text{idle}} = P_{\text{idle}} \) and \( \rho \leq 1 \). Finally, when \( T' \neq T \), since all that matters is the total airtime \( A_{\text{LBT}} \) given in (2), if we multiply (7) by \( (T' - \sigma) \) the stated result follows.

With Lemma 1 we can obtain the fraction of orthogonal/successful LBT transmissions (\( \rho P_{\text{idle}}^{(n)} \)) of expected duration \( T_{\text{LBT}} = T' - \sigma \) that can be accommodated in order to be compliant with our coexistence criterion. Importantly, the bound in (5) depends on \( P_{\text{idle}}^{(n)} \) and \( P_{\text{idle}}^{(n+1)} \), however, in saturation conditions a very good approximation of these values can be easily obtained [14].

Observe that we can write
\[ P_{\text{idle}}^{(n)} (1 - P_{\text{idle}}^{(n)} T + \rho P_{\text{idle}}^{(n)} (T' - \sigma) = P_{\text{idle}}^{(n)} (1 - P_{\text{idle}}^{(n)} (T + \rho P_{\text{idle}}^{(n)} (T' - \sigma)), \]
where
\[ \pi(\rho) := \min \left\{ 1, \frac{p^{(n)}_s}{1 - P_{\text{idle}}^{(n+1)}} \right\}. \] (8)

That is, an LBT WN will be compliant with our coexistence criterion as long as it takes a fraction \( \pi(\rho) \) of collision-free LIFS opportunities after a busy slot (successful or collision), where \( \rho \) is given in Lemma 1. In this way, we can establish the following optimal transmission policy, hereafter referred to as “Orthogonal Random LBT Unlicensed Access (ORLA)”:

**Theorem 1 (ORLA transmission policy).** Consider a policy \( \Pi_{\text{ORLA}} \) by which an (LBT) WN initiates transmissions for a fixed duration \( T_{\text{LBT}} \) after a LIFS opportunity \( \omega \in \Omega \) with probability \( \pi(\rho) \):
\[ \Pi_{\text{ORLA}}(\rho) := \{ \omega \in \Omega | U_\omega(\rho) = 1 \}, \]
where \( U_\omega(\rho), \omega \in \Omega \) are random variables taking values 0 or 1 such that \( \Pr(U_\omega(\rho) = 1) = \pi(\rho) \). In an homogeneous scenario where all WiFi WNs are saturated and have the same channel access configuration, \( \Pi_{\text{ORLA}}(\rho) \) is an airtime-optimal policy \( \Pi^* \) and a throughput-optimal transmission policy \( \Pi^t \) for synchronous LBT and asynchronous LBT, respectively.

Proof. The proof follows directly from Lemma 1.

---

**B. Non-saturation and Heterogeneous Conditions**

Next we generalize the results above to heterogeneous conditions in terms of WiFi packet arrival rates and link qualities. Recall that since cellular deployments work permanently on licensed frequencies, we only investigate the performance of supplemental downlink services exploiting the unlicensed band for best effort data transfers. As such, we still consider backlogged LBT WN newcomers (i.e., always having data to transmit) and study the performance of the proposed system with (i) practical multi-rate WiFi operation, thus focusing on airtime instead of throughput fairness, while (ii) utilizing additional airtime released by WiFi WNs with finite loads.

We extend the ORLA transmission policy to guarantee that the aggregate channel time of \( n \) (non-saturated) WiFi WNs when we add a saturated WiFi WN, \( A^{(n+1)} \), remains constant or larger than the aggregate channel time of \( n \) saturated WiFi WNs when we add the LBT WN, \( A^{(n+1)\text{LBT}} \). Formally,
\[ A^{(n+1)} = \sum_{i=1}^{n} P_{\text{coll},i}^{(n+1)} T_{\text{coll},i}^{(n+1)} \] (9)

where
\[ T_{\text{coll}}^{(n+1)} = P_{\text{idle}}^{(n+1)} \sigma + \sum_{i=1}^{n+1} P_{\text{coll},i}^{(n+1)} T_{\text{coll},i}^{(n+1)} + P_{\text{coll}}^{(n+1)} T_{\text{coll}}^{(n+1)} \]

with \( P_{\text{coll},i} \) being the probability that WN \( i \) transmits successfully in a MAC slot, \( T_{\text{coll},i} = T_{\text{LCP}} + [f_{\text{agg},i}(L_{\text{del}} + L_{\text{mac-oh}} + L_{\text{pad}}) + B_i] / C_i + SIFS + T_{\text{ACK}} + DIFS \) the duration of a slot when WN \( i \) transmits successfully, and \( T_{\text{coll}} \) the time the channel remains busy during a collision. Note that the numerator in (9) sums over the \( n \) (non-saturated) WiFi WNs and not over all the WNs in the system.
To compute the WiFi WNs’ transmission attempt rates \( \tau_i, i = 1, \ldots, n \), we first rewrite the conditional collision probability \( p_i^{(n)} \) that the frames transmitted by WN \( i \) experience:

\[
p_i^{(n)} = 1 - \prod_{k=1,k\neq i}^{n} \left( 1 - \tau_k^{(n)} \right).
\]

We use a renewal-reward approach to model the WiFi BEB scheme in the presence of different packet arrival rates. To avoid notation clutter, we drop the \( i \) and \( (n) \) indexes when there is no scope for confusion. The transmission attempt rate of a WiFi contender can be thus expressed as

\[
\tau = \frac{E[A]}{E[S]},
\]

where \( E[A] \) is the expected number of attempts to transmit a packet burst and \( E[S] \) is the expected number of slots used during back-off, which we compute as follows:

\[
E[A] = 1 + p + p^2 + \ldots + p^M,
E[S] = t_{idle} + b_0 + pb_1 + p^2b_2 + \ldots + p^Mb_M,
\]

where \( M \) is the maximum number of retries (which we assume equal to the maximum back-off stage \( m \)) and \( b_m \) is the mean length of back-off stage \( m \) expressed in slots. \( t_{idle} \) is the mean idle time that a contender waits for new content after a transmission. Thus, we can express the transmission attempt rate of a WiFi transmitter as

\[
\tau = \frac{E[A]}{E[S]} = \frac{1 + p + p^2 + \ldots + p^M}{t_i + b_0 + pb_1 + p^2b_2 + \ldots + p^Mb_M},
\]

where \( \Pi \) is defined in Theorem 1 and \( \rho \) is the fraction of busy slots experienced by the channel with an LBT WN, \( \rho \) holds for every \( \rho \in [0, \bar{\rho}] \), where \( \bar{\rho} \) is computed as:

\[
\bar{\rho} := \frac{1}{\rho F_{idle}(T' - \sigma)} \left[ \sum_{i=1}^{\text{sat}} p^{(n) \ast}_{\text{succ},i} T_{s,i}^{(n) \ast} - T_{\text{slot}}^{(n) \ast} - t_{\text{slot}}^{(n) \ast} \right].
\]

Proof: Rearranging the terms in (14) yields

\[
\sum_{i=1}^{\text{sat}} p^{(n) \ast}_{\text{succ},i} T_{s,i}^{(n) \ast} \leq \frac{T_{\text{slot}}^{(n) \ast} + \rho F_{idle}(T' - \sigma)}{\rho F_{idle}(T' - \sigma)}.
\]

Then, the result in (15) can be verified by proceeding similarly to the proof of Lemma 1.

By applying Lemma 2 and estimating the packet arrival rates at the WiFi WNs, the LBT WN can maximize channel utilization, leading to the following extension to Theorem 1.

Corollary 1 (ORLA transmission policy). Given the ORLA transmission policy \( \Pi_{\text{ORLA}}(\rho) \) defined in Theorem 1 and a heterogeneous scenario where WiFi WNs transmit at different rates and have different loads, \( \Pi_{\text{ORLA}}(\rho) \) is an airtime-optimal policy \( \Pi \) and a throughput-optimal transmission policy \( \Pi^\ast \) for synchronous LBT and asynchronous LBT, respectively, where \( \bar{\rho} \) is derived with Lemma 2.

Proof. The proof follows directly from Lemma 2. \( \square \)

V. OLAA: OPTIMAL ORTHOGONAL LAA ACCESS STRATEGY

The ORLA transmission strategy introduced above maximizes the channel airtime an LBT WN can use while preserving the fairness constraint in Definition 4. Since the amount of airtime held by an LBT WN when transmitting is fixed and equal to \( T_{\text{LBT}} \), the ORLA policy also maximizes the throughput performance of asynchronous LBT WNs. However, this is not true for synchronous LBT WNs, since the amount of airtime between the start of a transmission and the frame boundary this must synchronize with is effectively overhead, i.e., cannot be exploited for actual data communication, as illustrated in Fig. 2. Indeed, ORLA may prove highly inefficient in terms of throughput for synchronous LBT WNs, even providing less throughput than if following the WiFi specification, e.g., when the time until the next frame boundary is comparable to \( T_{\text{LBT}} \).

Recall that an LBT WN has to make a choice about whether to transmit or not every LIFS opportunity. If
decides to transmit at LIFS opportunity \( \omega \), a synchronous LBT 
WN will send useful data for an amount of time equal to 
\( Y(\omega) = T_{LBT} - T_{res} \), where \( T_{res} \) is the (random) time it 
takes between a LIFS opportunity and the closest frame boundary 
(see Fig. 2). Otherwise, it skips a round and waits for the next 
LIFS opportunity \( \omega + 1 \) (or round). We can regard this as 
an investment of time on each LIFS opportunity \( \omega \) yielding a 
gross gain equal to \( Y(\omega) \) of useful channel time. Since this 
process is repeated over time, our goal is to design a policy \( \Pi \) 
that maximizes the expected long-term \textit{rate of return}. Clearly, 
a naive policy that \textit{takes all} LIFS opportunities to transmit will 
not necessarily maximize the rate of return of a synchronous 
LBT WN, since the amount of time \textit{wasted} on waiting for frame 
boundaries may exceed the time invested in skipping LIFS 
opportunities. This resembles well-known problems in 
optimal stopping theory [23].

Suppose now the above process is repeated \( K \) times. Let 
\( \{\omega_1, \cdots, \omega_K\} \) denote LIFS opportunities that have been 
taken (referred to as \textit{stopping times}), \( Y(\omega_i) \) the useful channel 
time (reward) obtained at opportunity \( \omega_i \), and \( \psi(\omega_i) \) the 
time invested by doing so. Then, by the law of large numbers,

\[
\frac{\sum_{i=1}^{K} Y(\omega_i)}{\sum_{i=1}^{K} \psi(\omega_i)} \to E[Y_{\Pi}] / E[\psi_{\Pi}] \text{ a.s.}
\]

Then, we can cast the problem of maximizing the long-term 
average goodput of the synchronous LBT WN as a maximal-
rate-of-return problem and use optimal stopping theory to 
solve it [23]. In this way, we need to characterize an optimal 
stopping rule \( \Pi^\dagger \) as

\[
\Pi^\dagger := \arg \max_{\Pi \in \Pi} E[Y_{\Pi}] / E[\psi_{\Pi}]
\]

and the optimal LBT WN goodput as

\[
\lambda^\dagger := \sup_{\Pi \in \Pi} E[Y_{\Pi}] / E[\psi_{\Pi}].
\]

Following [23, Ch. 4, Th. 1], we transform our maximal-rate-
of-return problem into the following ordinary stopping rule 
problem:

\[
\max E \left[ Y_{\Pi} - \lambda \sum_{i=1}^{\omega \in \Pi} T_{slot} K^{(i)} + T_{LBT} \right],
\]  

where \( K_i \) is the number of WiFi slots between LIFS 
opportunities \( i \) and \( i - 1 \). The intuition behind the above 
problem is that we invest \( c_{\Omega} := \lambda \sum_{i=1}^{\Pi} T_{slot} K^{(i)} \) and gain 
\( X_{\Pi} := Y_{\Pi} - \lambda T_{LBT} \) in return, when we use stopping rule 
\( \omega \in \Pi \). Then, according to [23, Ch. 6, Th. 1], the optimal rule 
\( \Pi^\dagger \) and the optimal throughput \( \lambda^\dagger \) is such that 

\[
V^\dagger(\lambda^\dagger) := \sup_{\Pi \in \Pi} E[Y_{\Pi} - \lambda \psi_{\Pi}] = 0.
\]

It can be shown then that our stopping rule is 

\[
\Pi^\dagger = \min \{ \omega \geq 1 \mid X_{\Pi} \geq V^\dagger \}
\]

and that \( V^\dagger \) satisfies the optimality equation 

\[
V^\dagger = E \left[ \max \{ X^{(1)}, V^\dagger \} \right] - \phi(1).
\]

Given that \( V^\dagger(\lambda^\dagger) = 0 \) and that \( Y_{\Pi} \) is i.i.d, the above becomes

\[
E \left[ \max \left\{ Y - \lambda^\dagger \left( \frac{T_{slot}}{1 - P_{idle}} + T_{LBT} \right), -\lambda^\dagger \left( \frac{T_{slot}}{1 - P_{idle}} \right) \right\} \right] = 0,
\]

since \( E[K] = 1/(1 - P_{idle}) \), and hence

\[
E \left[ Y - \lambda^\dagger T_{LBT} \right] = \lambda^\dagger \frac{T_{slot}}{1 - P_{idle}}.
\]  

This is a fixed point equation that can be solved with iterative 
methods and the optimal rule is computed as

\[
\Pi^\dagger = \min \{ \omega \geq 1 \mid Y_{\Pi} \geq \lambda^\dagger T_{LBT} \} = \min \{ \omega \geq 1 \mid T_{res} < T_{LBT} (1 - \lambda^\dagger) \}.
\]

In addition to finding the stopping policy that maximizes the 
rate of return, we must guarantee the conditions established in 
Lemma 1 or 2, i.e., the ratio of used LIFS opportunities must not exceed

\[
\pi(\bar{\rho}) := \bar{\rho} \frac{P_{idle}}{1 - P_{idle}}.
\]

In order to accommodate such constraint, let us first introduce 
the following lemma.

**Lemma 3.** \( T_{res} \) is uniformly distributed between \( 0 \) and \( T_{LBT} \).

\[
\text{Proof:} \text{ We have a time-slotted system } t = 1, 2, \cdots, \text{ each slot \textit{containing} an idle, collision or successful WiFi event. Let } X := \{X_1, X_2, \cdots\} \text{ be a sequence of slots where } X_i \text{ is a Bernoulli trial with probability } \phi = 1 - P_{idle}, \text{i.e., a successful trial is a LIFS opportunity. This corresponds to a Bernoulli process } S(t) := \sum_{i=1}^{t} X_i \text{ such the probability of having } K \text{ LIFS opportunities in } t \text{ slots follows a binomial probability distribution, i.e., } P(S(t) = K) = B(t, K), \text{ and the number of slots between two LIFS opportunities follows a geometric distribution. Let now } \mathcal{L} := \langle \delta, T_{LBT} \rangle \text{ be the sequence of slots containing an LTE-like frame boundary for } \delta = 1, 2, \cdots \text{ and denote } T^{(\omega)} \text{ the (random) slot in which LIFS opportunity } \omega \text{ occurs. Then, since in order to be a LIFS opportunity, this must occur within the interval } [[(\delta - 1)T_{LBT}, \delta T_{LBT})], \delta = 1, 2, \cdots, \text{ we need to compute the conditional distribution of } T^{(\omega)}. \text{ As}
\]

\[
P \left( T^{(\omega)} \leq t \mid B(T_{LBT}, 1) \right) = \frac{P \left( T^{(\omega)} \leq t \right) B(T_{LBT}, 1)}{B(T_{LBT}, 1)} = \frac{t \phi(1 - \phi)^{t-1} (1 - \phi)T_{LBT} - t}{T_{LBT} \phi(1 - \phi)^{T_{LBT} - t}}.
\]

is the CDF of a uniform distribution, \( T_{res} = \mathcal{L}^{(\omega)} - T^{(\omega)} \) 
(note \( \mathcal{L} \) is not a random process) is also uniformly distributed 
between \( 0 \) and \( T_{LBT} \). 

It is therefore sufficient to simply consider those LIFS 
opportunities closer than \( \pi(\bar{\rho})T_{LBT} \) slots to a frame boundary, 
that is,

\[
T_{res} < \pi(\bar{\rho})T_{LBT}.
\]

This leads to the following theorem describing an optimal 
policy for synchronous LBT, which we name “Optimal 
Orthogonal LAA (OLAA)”:

\[
\sum_{t=1}^{\infty} \sum_{i=1}^{\nu} T_{slot} \phi(1 - \phi)^{t-1} (1 - \phi)T_{LBT} - t
\]
Theorem 2 (OLAA transmission policy). Consider a policy $\Pi_{\text{OLAA}}$ a synchronous LBT follows to initiate transmission according to the rule:

$$\Pi_{\text{OLAA}}(\lambda, \rho) := \left\{ \Pi \geq 1 | T_{\text{res}} < \min \left( T_{\text{LBT}}(1-\lambda), \pi(\rho) T_{\text{LBT}} \right) \right\},$$

where $T_{\text{res}}$ is the (random) time between a LIFS opportunity and the nearest frame boundary. Then, $\Pi_{\text{OLAA}}(\lambda, \rho)$ is a throughput-optimal transmission policy $\Pi^*$ for synchronous LBT WNs, where $\lambda$ is computed by solving (18), while $\pi(\rho)$ and $\bar{\rho}$ are given by Lemma 2 and (8), respectively.

Proof. The proof follows directly from Lemma 3. $\blacksquare$

VI. SIMULATION RESULTS

In this section we undertake a comprehensive performance evaluation of the orthogonal LBT transmission policies we propose, i.e. ORLA and OLAA, by means of event-driven simulation. We will demonstrate our schemes attain superior throughput as compared to the de facto 3GPP LAA, while being substantially more fair to incumbent WiFi. We consider coexistence with WiFi WNs that implement the IEEE 802.11ac specification [12], with the parameters summarized in Table I. We examine the performance of both synchronous and asynchronous WNs for a wide range of frame aggregation, contention window parameters and modulation schemes.\footnote{Numerical values from our model and simulation results are remarkably close to each other. Therefore, the figures show only simulation results for the sake of readability.}

The latter employ the 3GPP’s LAA protocol with the same contention parameters as WiFi WNs or our mechanism with ORLA policy. Unless otherwise stated, $C_W = 16$, $\overline{r} = 4$, $T_{\text{LBT}} = 1$ms (i.e. LTE’s Transmission Time Interval or TTI), and 64-QAM modulation is employed by both technologies. We note that the values we select are common in household or office environments (see [16]).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slot Duration ($\sigma$)</td>
<td>9 $\mu$s</td>
</tr>
<tr>
<td>DIFS</td>
<td>36 $\mu$s</td>
</tr>
<tr>
<td>SIFS</td>
<td>16 $\mu$s</td>
</tr>
<tr>
<td>PLCP Preamble and Headers ($T_{\text{PLCP}}$)</td>
<td>40 $\mu$s</td>
</tr>
<tr>
<td>MPDU Delimiter Field ($T_{\text{ctrl}}$)</td>
<td>32 bits</td>
</tr>
<tr>
<td>MAC Overhead ($T_{\text{del}}$)</td>
<td>288 bits</td>
</tr>
<tr>
<td>ACK Length ($T_{\text{agg}}$)</td>
<td>256 bits</td>
</tr>
<tr>
<td>Data bit rate ($C_L$)</td>
<td>130 Megabits/s</td>
</tr>
<tr>
<td>Control bit rate ($C_{\text{ctrl}}$)</td>
<td>24 Megabits/s</td>
</tr>
</tbody>
</table>

TABLE I: IEEE 802.11ac [12] parameters used for simulations.

A. Variable Number of WiFi WNs

We first investigate the airtime and individual throughput performance of an LBT WN operating with our LBT access mechanism with ORLA and OLAA policies, the benchmark 3GPP LAA, and the legacy WiFi protocol, as we vary the number of (background) WiFi WNs sharing the channel. We consider backlogged transmitters, first working with $f_{agg} = 1$ and payload $B = 1500$B, and subsequently aggregating 10 packets (i.e., $f_{agg} = 10$) and sending an Media Access Control Protocol Data Unit (MPDU) $B = 15000$ B upon each attempt. In these experiments, the coexisting LTE WN works with $T_{\text{LBT}} = 1$ms. Our results are presented in Fig. 3.

It is important to observe first the behavior of 3GPP LAA and background WiFi WNs when the latter transmits bursts of $B = 1500$B (the red points Figs 3(a)–3(b)). Clearly LAA consumes almost 6× more airtime when the number of WiFi contenders is small, which results in up to a 2-fold reduction in the throughput of a WiFi WN (blue line vs. red crosses). This violates the first coexistence criterion, i.e., not harming the performance of incumbent WiFi. In contrast with our approach, both ORLA and OLAA, safeguard WiFi throughput (overlapping green crosses and blue line), while the LBT WN consumes additional airtime more wisely, almost doubling MAC throughput efficiency (which is the second coexistence criterion) without negatively impacting on WiFi (green circles). As such, 3GPP LAA attains more throughput as compared to ORLA though at the mentioned price
we are interested in quantifying the performance gains LBT may obtain under different configurations, and irrespectively any losses incurred onto WiFi. To this end, we consider both synchronous and asynchronous LBT operation, fixed LBT frame size (1ms), different WiFi burst sizes (1500 and 15000B), and 3 CW settings, namely (16,512), (32,1024), and (8,256), while we vary again the number of WiFi WNs. The results of these experiments are illustrated in Fig. 4, where we plot the individual throughput gains with respect to a scenario where the LBT WN is an additional WiFi WN instead.

We note that when the background WiFi stations transmit small payloads and the LBT WN operates with the 3GPP LAA scheme, the difference between contention configurations are subtle. However, the remarkable 3GPP LAA throughput gains are at the expense of WiFi losses (negative gain), as shown in Fig. 4(a). We note that in this case the relative throughput gain of incumbent WiFi WNs improves (from −60% to −20%) when the number of WiFi WNs grows. This is due by the fact that with more WiFi contenders there is also a higher rate of collisions; therefore, the overall channel efficiency is higher when there is an LAA transmitter instead of another WiFi WN, which holds the channel longer at each transmission opportunity and so it causes less collisions per unit of time. Obviously, the absolute throughput of each WN worsens when there are more nodes, as indicated by Fig. 3.

In contrast, our orthogonal coexistence approach imposes no penalty on WiFi (observe the constant lines at gain equal

![Fig. 4: Individual throughput gain different LBT approaches achieve with respect to legacy WiFi for a variable number of background WiFi WNs, different contention settings, and different MPDU and LBT frame sizes. Simulation results.](image-url)
to 0%), while we achieve throughput gains almost up to 200% with both ORLA and OLAA policies. These gains are slightly more prominent when WiFi contends aggressively (i.e. \( CW_{\text{min}} = 8 \)), which we attribute to more LIFS opportunities higher WiFi attempt rates create. Moreover, our previous findings about the questionable efficiency of 3GPP LAA in scenarios where WiFi WNs transmit 15000B payloads is further confirmed in Fig. 4(b), where we observe that LAA incurs no penalty to WiFi WNs (which virtually lie on a gain equal to 0% irrespective of the number of WNs) but 3GPP LAA has between 15 and 60% of throughput loss over the case where it followed the WiFi protocol of the background WNs (this become clearer at the zoomed area, at the bottom part of the figure). In contrast, ORLA and OLAA do not affect WiFi performance in neither asynchronous or synchronous settings. We do note that when the number of WiFi WNs is small (mostly less than 5), OLAA does not attain throughput gains because the LBT frame size is fixed and no aggregation is allowed (we will assess optimal ORLA/OLAA aggregation, as suggested in Eq.(16), later on). However, depending on contention setting, individual relative LBT performance grows to as much as 60% under high contention levels, with both orthogonal coexistence policies we propose.

Similar insights can be obtained for larger LBT frame sizes, shown in Fig. 4(c) and 4(d). From these results we can observe that, while OLAA/ORLA achieve similar gains as before, without harming the incumbent WiFi WNs, 3GPP LAA causes even more harm when using larger frame sizes.

C. Impact of LBT Frame and WiFi Burst Durations

Next, we provide further insight into the impact of different WiFi burst sizes and \( T_{\text{LBT}} \) settings, i.e., the duration for which an LBT holds the channel when transmitting. We are particularly interested in understanding the impact LBT may have on WiFi with large frame sizes, since by default LTE operates with 10ms frames, while ETSI allows transmission in the unlicensed band for up to 8ms. For this purpose, we consider a scenario with 5 backlogged WiFi WNs and the LBT WN operating with 3GPP LAA scheme and with the orthogonal coexistence mechanism we propose, respectively. We investigate again both asynchronous and synchronous operation, whereby our solution employs ORLA and OLAA policies, respectively. The results of this experiment are shown in Fig. 5, where we plot as heatmaps the throughput gains of both LBT (with each approach) and background WiFi.

Under asynchronous operation, the 3GPP LAA may attain as much as 983% throughput gains when working with 10ms frames. However, this cuts off WiFi transmissions almost completely (92%) throughput loss. Conversely, if WiFi performance is to be preserved, 3GPP LAA has minimal gains as compared to using legacy WiFi, and may even be 54% less efficient if the background WiFi WNs employ long bursts (Fig. 5(a)). This effect is further exacerbated in the case of synchronous LBTs, as shown by Fig. 5(b).

In contrast both ORLA and OLAA achieve more than 200% throughput gains when operating with long LBT frames, without negatively impacting on WiFi. Indeed the performance of the incumbent remains unaffected, irrespective of the LBT frame/WiFi burst settings. Further, ORLA achieves improvements even when sending 1ms frames.

D. Non-saturation Conditions

We now turn attention to circumstances where incumbent WiFi WNs have limited offered load (i.e., non-saturation). We expect LBT to effectively exploit the additional airtime available in light load regimes, without harming WiFi performance. To this end, we consider 5 WiFi WNs transmitting 1500B PDUs and increment their offer load, relative to the load that saturates the network, from 10% (light load) to 100% (saturation conditions). We study the performance of both LBT and background WiFi when the LBT frame duration is 1 and 10ms, respectively; the LBT WN operates again
with either 3GPP LAA or the proposed orthogonal coexistence scheme with ORLA (asynchronous) or OLAA (synchronous) policy, respectively. Note however, that aggregation level in this case is automatically adjusted for both ORLA and OLAA, as explained in Eq. (16).

The obtained results are plotted in Fig. 6. It can be seen that 3GPP LAA negatively impacts on WiFi even when lightly loaded. This is more obvious when the frame duration is long (10ms), while we note that when the $T_{LBT} = 1$ms, 3GPP LAA leaves WiFi unaffected but exhibits decreasing performance up to the point where the relative load is precisely 50% (observe the LBT minimum), following which the LBT gain grows at the expense of WiFi. The proposed transmission policies do not affect non-saturated WiFi WNs. ORLA provides steady throughput gains above 100% up to the point where the WLAN saturates, OLAA’s performance grows with WiFi activity level, again exceeding 100% improvements.


ey. Heterogeneous Bitrates

We conclude the evaluation by investigating the performance of the proposed scheme and policies in a multi-rate scenario where 5 backlogged WiFi WNs transmit at different bitrates in response to dissimilar channel conditions. Specifically, upon accessing the channel, each transmits 1500B at the following rates $\{156, 130, 78, 39, 13\}$ Mbits/s, respectively. This will yield longer slot durations whenever a slower station transmits, leaving less time available for both LBT and other WiFi contenders. In this scenario, the LBT WN operates with a frame of 1ms and transmits using 64-QAM (MCS level 6). We compare again against the 3GPP LAA benchmark in both asynchronous and synchronous settings.

As seen in Fig. 7, with asynchronous operation the 3GPP LAA attains remarkable gains as compared to legacy WiFi (500%), but has a negative impact on the WiFi contenders. In contrast, ORLA ensures harmless coexistence, while still providing nearly 175% throughput improvements. When the LBT transmissions align to frame boundaries (synchronous LBT), the 3GPP LAA’s gains are less impressive, while WiFi is more severely affected. The results obtained confirm the OLAA policy does not inflict penalties onto WiFi also in this case, while still achieving 150% performance gains.

VII. PRACTICAL CONSIDERATIONS

In this section we discuss the implications and validity of some of the assumptions made in our analysis, and give a set of technical guidelines for practical implementation of the ORLA and OLAA coexistence mechanisms. Recall that our objective is to provide superior MAC efficiency as compared to 3GPP’s LAA, while protecting native WiFi systems.

A. PHY layer aspects

Our focus is on the MAC layer and in our modeling we have made number of assumptions regarding the physical layer, primarily for tractability. We revisit these and explain how they could be captured more precisely in practice. In particular,

1) Channel errors: We assumed channel errors are completely handled by appropriate use of Modulation and Coding Schemes and as such packet loss occurs only due to collisions. Our analysis can however be readily extended to account for noise/fading induced channel losses, which directly impact the average slot duration. In [24] we showed that the average slot duration can be inferred by only measuring the duration of successfully received frames, which an LBT WN can easily perform via packet sniffing. Hence, accommodating channel errors makes use of available measurements and does not require explicitly quantifying the loss rate.

2) Hidden and exposed terminals: We note that hidden WNs could initiate a transmission even when another WN (WiFi or LBT) has already acquired the channel, therefore the attempt probabilities are coupled with the dynamics of the transmissions not directly overheard. While the goal of our work is not to address such WiFi-related issues, which are already well studied in the literature (see e.g. [25]), the coexistence mechanism we propose aims to avoid harming incumbent WiFi. For this purpose, a non-WiFi LBT WN can minimize the amount of additional hidden terminal issues by transmitting (at a sufficiently high power) a CTS-to-self every time a LIFS opportunity is taken. Likewise, exposed terminal issues are inherent to CSMA and thus would affect a LBT WN as much as they would any other WiFi station. To reduce susceptibility to this effect, an LBT WN can employ signature detection techniques, as recently proposed in [26]. On the other hand, we can ensure that a LBT WN does not create exposed WiFi terminals, by implementing adaptive
transmission power control algorithms similar to [27], to reduce LBT coverage.

(3) **Capture effect:** Our modeling assumes that concurrent transmissions cannot be decoded at the intended receivers, i.e., they result in collisions. In practice, however, a receiver could still decode a ‘strong’ packet, if the receive power difference between multiple transmissions is sufficiently large. This may cause unfairness among WiFi WNs [28], which our design does not aim to mitigate. However, the proposed coexistence mechanism neither alters the behavior of incumbent WiFi, nor induces additional unfairness due to capture effect issues. Precisely, WiFi remains agnostic about the orthogonal access procedure we introduce.

### B. MAC layer aspects

Regarding the MAC behavior modeling underpinning the design of our LBT scheme, we did not explicitly consider multi-operator scenarios, absence of WiFi activity, details concerning the computation of the average duration of slots containing collisions in multi-rate settings, and the impact on delay performance. We discuss these aspects next.

1. **Multiple non-WiFi LBT WNs:** Assuming a channel is mostly used by a single non-WiFi LBT WN, as in our model, is reasonable, since we expect that LTE operators have financial incentives to coordinate offline on exploiting spatial and spectral diversity. On the other hand, in case multiple non-WiFi LBT WNs are forced to contend for unlicensed spectrum, they can either coordinate transmissions via cross-carrier downlink control information (DCI) [29] or could distributively contend for non-WiFi LBT orthogonal channel time via CSMA methods. For instance, these can coordinate their access via a dynamic control channel established using rendezvous protocols [30], or employ a single-tone random access scheme as used in narrow band IoT over the Narrow band Physical Random Access Channel (NPRACH) [31].

2. **Absence of WiFi transmissions:** Our LBT design is built upon a seamless symbiosis between WiFi and non-WiFi LBT transmissions. Lack of WiFi activity in a network will effectively lead to no transmission opportunities for a LBT transmitter following our scheme. However, as we employ carrier sensing, such circumstances can be detected and we argue in favor of dynamic switching to 3GPP’s LAA in this corner case, especially since channel access would not be subject to collisions.

3. **Estimating collision durations:** The orthogonal transmission policies we propose depend on accurate knowledge of the average slot duration, and therefore the duration of slots containing collisions. Estimating this was long thought to be difficult, but recent results demonstrate analytically and practically that average collision durations can be accurately computed by simply inspecting correctly received packets [24].

4. **Impact on delay.** An LBT WN may impact on the delay experienced by background WiFi WNs when accessing the medium in two ways: (i) by causing collisions with WiFi transmissions, and (ii) by continuous airtime consumption. By design our coexistence mechanism and transmission policies, in contrast to LAA, incur minimal delay penalties on WiFi, since (i) our orthogonal access paradigm does not cause collisions, and (ii) the LBT frame size can be configured as equal (or smaller) to the amount of time taken by a WiFi transmission, with no cost in performance (see Fig. 5).

### C. Practical implementation

In the remainder we discuss how the coexistence mechanism and the optimal transmission policies we propose could be implemented with off-the-shelf wireless hardware. Recall our design relies on the assumption that LBT WNs are equipped with native LTE transceivers and additional IEEE 802.11 network cards. The latter are inexpensive, already provide access to a number of channel state parameters required by our policies (e.g. duration of idle/busy periods via open source firmware [32]) and have been used to develop custom features (e.g. precise control of TX engine timings [33]). For our design, key functionality can be implemented as follows.

1. **Estimating the number of WiFi WNs:** By employing passive packet sniffing on the 802.11 interface, the LBT WN can employ a range of practical methods to estimate the number of active WiFi contenders, including Kalman filtering [22] and Bayesian estimation [34].

2. **Estimating transmission probabilities (τ):** All the parameters required by our optimal transmission policies can be derived from the transmission attempt probabilities of WiFi WNs, τ, as we explain in §IV. Under saturation conditions, τ can be computed by solving a non-linear system of equations which depend on contention parameters CWmin and m, and conditional collision probabilities p (see §IV-A). The former can be directly extracted from overheard beacon frames sent periodically by APs. The latter can be estimated by inspecting the retry flag of overheard WiFi packets [35]. In case of non-saturation, an LBT WN also needs to estimate the probability q that a new packet arrives in a uniform slot time Tslot. If each WN managed sufficiently large number of heterogeneous flows, by the Palm–Khintchine theorem [36], the resulting arrival process will be Poisson and q can be related to the average load as λ = −log(1−q)/T slot. Although this assumption may not always hold, it is commonly used for testing purposes, including for LTE/WiFi coexistence evaluation [37].

3. **Detecting LIFS opportunities:** To avoid collisions, an LBT WN has to detect if a channel has been idle for LIFS duration. This can be achieved by inspecting channel state registers of 802.11 cards and signaling such event through the device’s shared memory [33], which the LTE protocol stack can periodically inspect. In case of synchronous LBT operation, the 802.11 interface can also be instructed by writing into the shared memory, to immediately generate a CTS-to-self frame that reserves the channel for the desired duration.
VIII. RELATED WORK

The coexistence between IEEE 802.11-based WiFi networks and other wireless technologies in unlicensed spectrum has received much interest from the research and industry communities. Coexistence mechanisms have been repeatedly proposed to handle joint operation of WiFi and Bluetooth [38], WiFi and Zigbee [39]–[41], WiFi and WiMax [42], and more recently between WiFi and LTE systems operating in unlicensed bands [4], [9]–[11], [17], [20], [43]–[52].

The work in [20] is one of the first to addressing LTE/WiFi coexistence, proposing a simple scheme based on Almost Blank Subframes (ABS) to support coexistence in a 900 MHz (TV) band. The performance of WiFi in the presence of LTE was analyzed numerically [45] and by means of simulation [43], [44], though the LTE access model in these studies is not based on LBT and therefore it is not compliant with ETSI regulation. Experimental results of CSAT (Carrier Sensing Adaptive Transmission) based access over software-defined platforms were reported in [10], the impact of U-LTE CSAT duty cycling patterns were quantified in [11], and WiFi signals were embedded into LTE in [52] to improve coexistence.

Ratasuk et al. studied an LBT-enabled LTE access mechanism very similar to what later was specified as Licensed-Assisted Access (LAA) [46]. 3GPP presented early simulation results with LBT-based LAA, claiming that in some scenarios an LAA node can be configured such that the performance of WiFi users would not be affected more than if another WiFi station were added to the network [1]. However, as we demonstrated through our evaluation, whether legacy LAA meets this fairness criteria while improving spectral efficiency (as our work does) remains questionable.

In [47], Sagari introduced an inter-network coordination architecture to solve the bandwidth and channel selection problem when LTE systems access WiFi bands. Coexistence was modeled as a classic opportunistic spectrum access problem in [48], whereby LTE and WiFi users are considered secondary primary users respectively, operating with cognitive full-duplex transceivers that follow a non ETSI-compliant access protocol.

Sadek et al. focused on a similar fairness criterion as the one we consider, i.e. in a network with a set of WiFi stations, if an arbitrary number of WiFi WNs are replaced with U-LTE nodes, the performance of the remaining WiFi stations should be comparable (or better) than in the all-WiFi scenario [49]. The authors undertook a simple analysis and simulation-based evaluation of a duty cycle-based approach (CSAT), which is not ETSI-compliant. More recently, Guan et al. analyzed the problem of channel selection and carrier aggregation when using U-LTE CSAT-like opportunistic spectrum access [9], while Liu et al proposed an LBT-based mechanism for LAA multi-carrier operation [51].

In [50], Yun et al. made a radically different proposal where both LTE and WiFi signals are transmitted together and a novel decoding method is designed to decode the overlapped transmissions. However, this scheme requires a redesign of the physical layer for both LTE and WiFi, which renders practical implementation cumbersome.

The work of Cano et al. in [17] and [4] proposed coexistence mechanisms based on both CSAT duty-cycling and LBT, which achieve a proportionally fair rate allocation across LTE/WiFi nodes. We note that our fairness criteria, in line with 3GPP’s, fundamentally differs from this approach, since WiFi systems do not naturally lean towards proportional fairness.

The interest in U-LTE shown by industry is unquestionable. While Nokia, Qualcomm, and Huawei released white papers showing promising results [53]–[55], the details of the access schemes and simulation models used are not public. Intel presented a performance analysis of the coexistence between both technologies from a system level perspective, using a large scale model based on stochastic geometry where both U-LTE and WiFi employ CSMA [56]. However, this work does not address the fundamental issue of designing an LTE-based access scheme that is fair to legacy WiFi. In fact, results show that WiFi performance is substantially degraded in the presence of the more spectrum-efficient U-LTE scheme.

To the best of our knowledge, the work we present in this paper and our preliminary results in [3] are the first attempt to explore orthogonal channel access in WiFi bands. Based on this paradigm, we devise optimal transmission policies that maximize spectral efficiency of LBT-based LAA, while remaining harmless to native WiFi. Here, we substantially extend our initial work in [3], as we specify optimal transmission policies in more general cases, including scenarios where transmitters use different modulations and/or loads. We further design an optimal transmission policy for synchronous LAA, which minimizes channel overhead when synchronizing to licensed LTE framing.

IX. CONCLUSIONS

To enable LTE deployment in unlicensed bands and seamless integration with existing cellular systems, 3GPP has recently specified an LTE based solution named LAA. Despite its potential to attain superior user multiplexing and robustness (e.g., via HARQ error recovery), in this paper we contended that the 3GPP LAA improves 802.11 MAC efficiency at the cost of penalizing incumbent WiFi networks, with 3GPP LAA configurations that are completely fair to WiFi achieving inferior MAC performance compared to 802.11. We present a radically different approach to coexistence in unlicensed bands, which overcomes the limitations of 3GPP LAA and is compliant with the listen-before-talk requirement of, e.g., ETSI’s EN 301 893 regulation. The proposed coexistence mechanism builds a symbiotic relationship between incumbent WiFi and U-LTE that creates orthogonal airtime blocks for each system, thereby avoiding collisions between them and substantially increasing the MAC layer efficiency of both technologies. Based on this orthogonal access procedure, we derive optimal transmission policies, namely ORLA and OLAA, for asynchronous and synchronous systems respectively, which maximize U-LTE throughput, yet cause no harm to background WiFi networks. Finally, by means of extensive system-level simulations, we demonstrated the proposed transmission policies attain LBT throughput gains above 200% with no negative impact on WiFi.
**GLOSSARY**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project.</td>
</tr>
<tr>
<td>ACK</td>
<td>Acknowledgment.</td>
</tr>
<tr>
<td>AIFS</td>
<td>Arbitration Interframe Space.</td>
</tr>
<tr>
<td>BEB</td>
<td>Binary Exponential Backoff.</td>
</tr>
<tr>
<td>CCA</td>
<td>Clear Channel Assessment.</td>
</tr>
<tr>
<td>CSAT</td>
<td>Carrier Sensing and Adaptive Transmission.</td>
</tr>
<tr>
<td>CSMA/CA</td>
<td>Carrier Sense Multiple Access with Collision Avoidance.</td>
</tr>
<tr>
<td>CTS</td>
<td>Clear To Send.</td>
</tr>
<tr>
<td>CW</td>
<td>Contention Window.</td>
</tr>
<tr>
<td>DCF</td>
<td>Distributed Coordination Function.</td>
</tr>
<tr>
<td>DIFS</td>
<td>DCF Interframe Space.</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute.</td>
</tr>
<tr>
<td>FBE</td>
<td>Frame Based Equipment.</td>
</tr>
<tr>
<td>HARQ</td>
<td>Hybrid Automatic Repeat reQuest.</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things.</td>
</tr>
<tr>
<td>LAA</td>
<td>License Assisted Access.</td>
</tr>
<tr>
<td>LBE</td>
<td>Load Based Equipment.</td>
</tr>
<tr>
<td>LBT</td>
<td>Listen before Talk.</td>
</tr>
<tr>
<td>LIFS</td>
<td>LTE Interframe Space.</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution.</td>
</tr>
<tr>
<td>LWA/LTE-H</td>
<td>LTE-WLAN Aggregation.</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control.</td>
</tr>
<tr>
<td>MPDU</td>
<td>Media Access Control Protocol Data Unit.</td>
</tr>
<tr>
<td>NPRACH</td>
<td>Narrow band Physical Random Access Channel.</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing.</td>
</tr>
<tr>
<td>OLAA</td>
<td>Optimal Orthogonal LAA.</td>
</tr>
<tr>
<td>ORLA</td>
<td>Orthogonal Random LBT Unlicensed Access.</td>
</tr>
<tr>
<td>PIFS</td>
<td>Point Coordination Function Interframe Space.</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network.</td>
</tr>
<tr>
<td>SIFS</td>
<td>Short Interframe Space.</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access.</td>
</tr>
<tr>
<td>U-LTE</td>
<td>Unlicensed LTE.</td>
</tr>
<tr>
<td>WN</td>
<td>Wireless Node.</td>
</tr>
</tbody>
</table>