Integrating Full-duplex Capabilities in Heterogeneous Spectrum Sharing

Wessam Afifi, Marwan Krunz, and Mohammed Hirzallah
Department of Electrical and Computer Engineering, University of Arizona, AZ, USA
{wessamafifi, krunz, hirzallah}@email.arizona.edu

***Abstract*—Driven by a persistent increase in wireless demand, spectrum sharing between heterogeneous systems has recently been the focus of extensive research. A key common concern is how to address coexistence-related interference, especially when the systems involved exhibit different protocols. In this paper, we outline a framework for exploiting full-duplex (FD) capabilities to support coexistence in a spectrum sharing environment. Besides the traditional simultaneous transmission/reception (STAR), we advocate a simultaneous transmission and sensing (STAS) mode that allows for rapid interference detection and migration to other channels. Specific examples of Wi-Fi/LTE-U coexistence and opportunistic systems are used to demonstrate the idea.**

***Index Terms*—Self-interference cancellation, full-duplex, LTE-U/Wi-Fi, spectrum awareness/efficiency tradeoff.**

I. Introduction

To face the exponential increase in the mobile data traffic, various solutions were proposed such as building more small cells and repurposing bands from their existing usage to mobile broadband either via auctions or for unlicensed usage (e.g., citizens broadband radio service or CBRS band, unlicensed national information infrastructure or U-NII band, mmW band, etc.). Another promising solution is to increase the spectrum efficiency by investigating new technologies such as inband full-duplex (FD) communications (see [1] for a survey), massive MIMO, and spectrum sharing.

Spectrum sharing can be categorized into three categories: vertical, horizontal, and combined spectrum sharing. In vertical spectrum sharing, users are divided into tiers, each can access the spectrum with different privilege. In opportunistic spectrum access (OSA) systems, two classes of users exist, where the highest priority class (i.e., primary users or PUs) can access the spectrum at will any time. On the other hand, secondary users (SUs) can only access the spectrum in an opportunistic fashion using spectrum sensing [2]. Horizontal sharing denotes the scenario where all coexisting systems share the spectrum with equal priority such as the coexistence between Wi-Fi and LTE-unlicensed (LTE-U) in the 5 GHz band. Finally, combined spectrum sharing includes both vertical and horizontal sharing in the same paradigm such as CBRS band. In CBRS, there are three tiers of users which are ordered from the highest to the lowest priority (i.e., vertical sharing) as incumbents, priority access licenses (PAL), and general authorized access (GAA). However, different systems can coexist in the GAA level with equal priority.

In this paper, we incorporate FD capabilities in dynamic spectrum sharing systems. FD communication allows a given

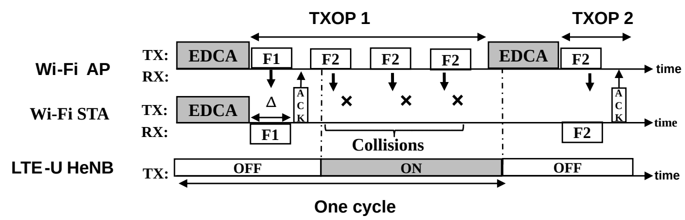

Fig. 1: Collision between LTE-U and Wi-Fi TXOP ('F1': Frame transmitted from AP).

radio to transmit and receive data simultaneously over the same channel which can increase the link's throughput significantly. FD communications are enabled by employing self-interference suppression (SIS) techniques which span the propagation, analog, and digital domains [1]. In addition to the traditional way of using SIS techniques for inband simultaneous transmission and reception (STAR), we investigate the exploitation of these techniques to enable simultaneous transmission and sensing (STAS). STAS mode enables inband channel monitoring which reduces the collision probability between heterogeneous systems sharing the same band. In this paper, we investigate two issues: spectrum sensing techniques for the STAS mode and FD-based coexistence framework for spectrum sharing environments.

We apply the proposed framework for two use cases: OSA networks and Wi-Fi/LTE-U coexistence. In the first use case, we investigate the tradeoff between spectrum awareness (i.e., STAS) and spectrum efficiency (i.e., STAR) in an OSA setting. Specifically, we propose to equip SUs with SIS capabilities to enhance the link throughput and/or collision probability. In the second use case (i.e., Wi-Fi/LTE-U coexistence), we further investigate the spectrum awareness/efficiency tradeoff to include the joint optimization of the transmission mode and rate. LTE-U is based on the same concept of carrier aggregation (CA) which has been used in LTE-A. However, instead of aggregating channels from licensed spectrum only, LTE-U enables CA between licensed and unlicensed channels. In an effort to reduce the impact of LTE-U on Wi-Fi, two approaches have been proposed: Carrier-sensing adaptive transmission (CSAT) [3] and licensed assisted access (LAA) [4].

LAA (standardized in 3GPP Rel. 13) targets countries that mandate using listen-before-talk (LBT) in the 5 GHz band (e.g., Europe and Japan). An HeNB senses the spectrum prior to each transmission and transmits if the measured signal is below -72 dBm. CSAT relies on channel selection and time-based duty cycle (see Figure 1). The HeNB measures the traffic density of neighboring Wi-Fi stations (STAs) during the OFF period of the LTE-U system and adapts its duty cycle accordingly. In the ON period, HeNB transmits DL frames without performing LBT. On the other hand, Wi-Fi STAs can access the spectrum using the enhanced distributed channel access (EDCA) scheme, which is an extension of the distributed coordination function (DCF). The successful STA can reserve the channel for a duration called a transmit opportunity (TXOP), which may last for 3.008 ms. During a TXOP, a Wi-Fi access point (AP) or STA transmits several frames. After each frame, the AP/STA could wait for an ACK from its peer.

Deploying LTE-U small cells in unlicensed bands may lead to severe service degradation for Wi-Fi STAs. As shown in Figure 1, the AP detects a transmission failure (e.g., frame 'F2') via an ACK timeout. However, the AP cannot tell the reason for this transmission failure (e.g., channel fading and Wi-Fi/LTE-U interference). The AP may retransmit the corrupted frame several times. Several transmission failures may lead to performance degradation in terms of long delays, reduced throughput, and power wastage. In the second use case, we consider Wi-Fi devices with SIS capabilities, which enable them to perform STAS to get real-time channel monitoring and interference detection. Increasing the spectrum awareness at the AP helps it optimize its actions to maintain connectivity with the STAs. Wi-Fi standards (e.g., IEEE 802.11n/ac) define multiple modulation and coding schemes (MCSs), which can be used by the AP to adapt to channel dynamics, interference, and contention. We leverage this degree of freedom to jointly optimize the MCS and transmission mode at the AP, taking into account the AP's belief about LTE-U interference.

II. FD-BASED COEXISTENCE FRAMEWORK

We consider an FD-capable wireless link exists in a dynamic spectrum sharing environment (e.g., an FD-capable SU link coexisting with a primary network or an FD-capable Wi-Fi link coexisting with an LTE-U small cell). Assume that every node in that link is capable of partial or complete SIS, enabling it to operate in the STAS or STAR modes. We use χ_i to represent the SIS capability of the ith node, $\chi_i \in [0,1]$. Specifically, χ_i is the ratio between the residual self-interference (RSI) signal and the original one. If $\chi_i = 0$, the node can completely suppress its self-interfering signal (i.e., perfect SIS); otherwise, it can only suppress a fraction $1-\chi_i$ of its self-interference (i.e., imperfect SIS). χ_i may differ from one node to another, depending on the employed SIS techniques.

Figure 2 shows an FD-based coexistence framework which consists of different modules. The first module is the learning or cognition module, which is responsible for building a profile for each of the monitored channels in terms of availability and channel quality. Note that these channels' profiles are reflected what the coexisting wireless link believes since the channels are partially observable. In other words, the FD-enabled link does not know the actual status (idle or busy) of

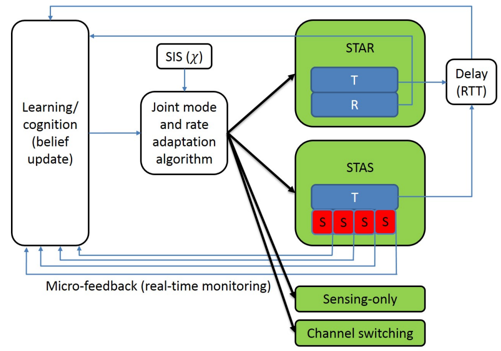

Fig. 2: FD-based coexistence framework for dynamic spectrum sharing systems.

the channels since observations are usually imperfect. After each action (e.g., STAR or STAS), the learning module takes the outcomes of that action and updates its belief about the channel availability.

The learning/cognition module then pass that belief value to the decision-maker module, where different algorithm could be implemented to adapt the operational mode and transmission rate (i.e., MCS). In addition to the belief about the channel availability, the decision-maker module takes as an input the SIS capability of the nodes of a given link. Having low SIS capabilities may affect the performance of different actions. For the STAR mode, high RSI reduces the SINR significantly, which in some cases may return lower "utility" than the STAS mode. On the other hand, and based on the employed spectrum sensing technique, having low SIS capability may return poor sensing performance. One way of implementing the decision-maker module is to build a policy function offline about the optimal action that needed to be taken for each belief and time instance. In the online phase, the function of the decision-maker is to check the policy function and take the corresponding action.

There are four possible modes for a given FD-capable link: STAS, STAR, sensing-only (SO), and channel switching (CS). Using SIS techniques, a node can carry out spectrum sensing while simultaneously transmitting its data (i.e., STAS) as shown in Figure 2. This sensing process may be done over multiple (consecutive) short periods instead of one long sensing period. Specifically, a given node may perform m sensing actions, each of duration T_{Sk}, $k = 1, 2, \ldots, m$, while transmitting data for a period of T seconds. The motivation behind this approach is to account for the tradeoff between sensing efficiency and timeliness in detecting other users' activities. Increasing T_{Si} improves the sensing accuracy (smaller false-alarm and mis-detection probabilities). However, such an increase implies delaying the time to make a decision regarding the signal presence. If at the end of any given sensing period $k, k = 1, 2, \ldots, m$, a signal is detected, communication is aborted. We use the term *FD sensing* to refer to the sensing process in the STAS mode.

In the STAR mode, the node transmits and receives data simultaneously over the same channel. Denote the transmission and reception durations by T and T_R, respectively. For simplicity, we assume $T_R = T$, which can be justified as follows. To operate in the TR mode, both nodes must have data to send to each other. If their packets are not of the same size, then the duration T can be set to the smaller of the two packets. Although operating in the STAR mode enhances the link's throughput, no online monitoring will be possible. Hence, the collision probability between heterogeneous systems will be higher than that of the STAS mode.

Besides the STAS and STAR modes, an FD-capable link may need to operate in an SO mode to improve the sensing accuracy (which would otherwise be impacted by imperfect SIS). Considering the availability of multiple channels, the link may decide to perform CS action if there is a high belief that the channel may get occupied soon.

A. FD Sensing Techniques

Extensive literature has been published on spectrum sensing techniques for traditional HD devices [2]. Energy-based sensing is quite general, as it requires no prior knowledge of the structure/waveform of the PU signal. On the other extreme, matched-filtering requires perfect knowledge of the features of the PU signal (i.e., modulation scheme, frame format, etc) to demodulate this signal. Waveform-based sensing can be utilized by SUs for detecting PU signals with known signal patterns, which is quite common, while cyclostationary-based sensing can be used in cases where enough cyclostationarity features exist in the PU signal.

Consider an FD SU device with an arbitrary SIS factor χ. Under FD sensing, the hypothesis test of whether the channel is occupied by a PU or not can be formulated as:

$$r(n) = \begin{cases} \chi\, s(n) + w(n) & \text{under } H_0 \text{ (PU is idle)} \\ l(n) + \chi\, s(n) + w(n) & \text{under } H_1 \text{ (PU is busy)} \end{cases}$$

where $r(n), s(n), l(n)$, and $w(n)$ are, respectively, the nth samples of the received signal, the self-interfering SU signal, the received PU signal, and the additive white Gaussian noise with variance σ_w^2. In the case of HD sensing, the hypothesis test reduces to $w(n)$ (under H_0) and $l(n) + w(n)$ (under H_1).

Among the aforementioned sensing techniques, energy detection is widely used in HD sensing because of its low computational and implementation complexities. However, a key disadvantage of this technique is its inability to differentiate between a PU signal and noise (specially under low signal-to-noise ratios). Furthermore, for FD sensing, energy detection cannot differentiate between RSI, PU signal, and noise. Better accuracy can be achieved with waveform-based sensing which can robustly differentiate between different signal types. In terms of complexity, waveform-based sensing is a little bit more complex than energy detection, because it requires the SU to know the PU pattern. Hence, it can be considered as a good candidate, with reasonable complexity, for the TS mode. The performance of any sensing technique is measured by the false-alarm probability (P_f) and the detection probability (P_d). P_f and P_d are defined as the probabilities that the SU declares the sensed channel to be busy given hypothesis H_0 and H_1, respectively.

Consider energy-based sensing. The main idea is to compute the average energy of N samples of the signal $r(n)$ and compare this average with a threshold γ_e to determine whether the PU is idle or not. The decision metrics for the energy detector and waveform-based sensing can be formulated respectively as $M_e \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^{N} |r(n)|^2$ and $M_w \overset{\text{def}}{=} \text{Re}\left[\sum_{n=1}^{N} r(n)\, l^*(n)\right]$, where $l^*(n)$ is the conjugate of the known part of the PU signal. The metric M_w correlates the received samples with the samples of a static part of the PU signal. The value of M_w is then compared to a threshold γ_w to determine the presence/absence of a PU signal. Let M be a generic random variable that refers to either M_e or M_w, depending on the context. Also, let γ be an arbitrary threshold.

In the FD case, P_f and P_d can be formulated, respectively, as $P_f = \Pr[M > \gamma/H_0] = 1 - F_{M/H_0}(\gamma)$ and $P_d = \Pr[M > \gamma/H_1] = 1 - F_{M/H_1}(\gamma)$, where F_{M/H_0} and F_{M/H_1} are the conditional CDFs of the random variable M given hypothesis H_0 and H_1, respectively. Using the central limit theorem, we can determine these two CDFs. Specifically, for a large N, the pdfs of M/H_0 and M/H_1 can be approximated by Gaussian distributions with means μ_{M/H_0} and μ_{M/H_1} and variances σ^2_{M/H_0} and σ^2_{M/H_1}, respectively. Hence, the false-alarm and detection probabilities can be written, respectively, as $P_f = Q\left(\frac{\gamma - \mu_{M/H_0}}{\sigma_{M/H_0}}\right)$ and $P_d = Q\left(\frac{\gamma - \mu_{M/H_1}}{\sigma_{M/H_1}}\right)$, where Q is the complementary CDF of a standard Gaussian random variable. Substituting for μ_{M/H_0}, μ_{M/H_1}, σ^2_{M/H_0}, and σ^2_{M/H_1} in the aforementioned equations, we get the false-alarm and detection probabilities for FD sensing under both energy-based and waveform-based sensing (see [5, 6] for details). As reported in [5, 6], waveform-based sensing expectedly outperforms energy-based sensing in reliability and convergence time. Despite its susceptibility to synchronization errors, waveform-based sensing results in low false-alarm and mis-detection probabilities with very short sensing times. To compensate for the RSI, SUs need to increase their sensing times if energy-based sensing is used.

III. Use cases

A. Opportunistic Spectrum Access Systems

We consider an OSA network, where SUs operate opportunistically over licensed PU channels. The PU activity is modeled as an alternating ON/OFF random process. Each SU is capable of partial or complete SIS, enabling it to operate in the STAS or STAR modes, along with the SO and CS modes. At any given time and over any given channel, we assume at most one SU link can be active in a given neighborhood (i.e., collision domain). Hence, different SU links do not interfere with each other. Various spectrum access protocols have been proposed to handle SU-SU interference.

One way of of optimize the mode selection for an SU link is to formulate the problem as a partially observable decision process. Let $\mathbb{S} = \{0, 1\}$ be the state space, which indicates the actual state (idle or busy) of the channel that is being observed by the SU. The action set at the SU is given by $\mathbb{A} = \{\mathit{STAR}, \mathit{STAS}, \mathit{SO}, \mathit{CS}\}$. While observing the PU channel, the SU has to choose an action from the set $\mathbb{A}$. The outcome/observation space for the SU depends on the action taken. Because a STAR action consists of two simultaneous processes (transmission and reception), there are two outcomes for each of these processes. Specifically, for the reception part, the SU may observe the outcome $\{D\}$, which means that the MD was able to decode the received message, or the outcome $\{U\}$, which stands for an undecoded message. For the transmission part of the STAR mode, the SU may get an ACK or NACK from its peer, which are denoted by $\{A\}$ and $\{N\}$, respectively. Similarly, a STAS action consists of two simultaneous processes (transmission and sensing). The SU will also observe two possible outcomes for the sensing process: $\{F\}$ for free or $\{B\}$ for busy. The outcomes of the transmission part are similar to those of the STAR mode. Finally, the observed outcomes for the SO/CS actions are $\{F\}$ or $\{B\}$. Altogether, these actions result in an observation space $\mathbb{O} = \{D, U, A, N, F, B\}$. A reward function is then defined which maps the state and action space to a reward value.

The goal of the SU is to choose actions sequentially in time so as to maximize the expected reward over some random but finite horizon. This can be done via stochastic dynamic programming. First, note that a sufficient statistics for choosing the optimal action at any time t is the belief [7], which is defined as the a posteriori probability $p_t \in [0, 1]$ that the PU is idle at time t given the observation history. The time index t is defined here as the time elapsed since the PU switched from ON to OFF [8]. Hence, $t = 0$ is the start of the PU idle period, which is assumed to be known to the SU, and therefore $p_0 = 1 - P_f$. Starting from $t = 0$, the SU keeps applying the optimal mode selection policy until switching to a new channel. At that point, the SU resets the algorithm and searches for a new channel to operate on.

After any given action $a \in \mathbb{A}$ and depending on the observation $o \in \mathbb{O}$, the SU updates its belief p_t and computes the corresponding reward. Let π_t be the policy that maps the SU's belief p_t to the action space $a \in \mathbb{A}$ at time t. Define the value function $U(p_t, t)$ as the maximum expected total reward at time t when the current belief is p_t. This function specifies the performance of the optimal policy, denoted by π^*, starting from belief p_t. Based on Bellman equation [9], we have the following:

$$U(p_t, t) = \max \{U_{TR}(p_t, t), U_{TS}(p_t, t), U_{SO}(p_t, t), U_{CS}(p_t, t)\}$$

where $U_{TR}(p_t, t), U_{TS}(p_t, t), U_{SO}(p_t, t)$, and $U_{CS}(p_t, t)$ are the SU's expected total rewards if the SU decides to operate in the STAR, STAS, SO, and CS modes, respectively, at time t and then follows the optimal policy π^* after that. The SU utility for a given action can be formulated as an addition of two terms: the myopic and long-term reward. The myopic reward takes into account the instantaneous benefit that results from operating under a specific mode such as throughput enhancement, collision probability reduction, etc. On the other hand, the long-term reward (which is multiplied by a discount factor) takes into account the utility achieved in the future based on the expectations of getting different observations.

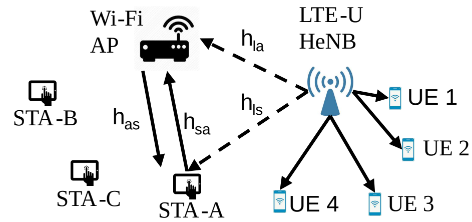

Fig. 3: System model of LTE-U/Wi-Fi coexistence (dashed lines represent interference from HeNB to Wi-Fi AP and STA).

The discount factor determines the importance of the long-term reward compared to the myopic reward.

To find the optimal policy function that maximizes the SU utility while having a constraint on the collision probability, backward induction can be used. In [6], it was found that the optimal policy has a threshold-based structure. The policy recommends that the SU should exploit its high belief that the PU is idle and operate in the STAR mode if the belief is larger than a given threshold. In that case, the SU will dramatically increase its throughput by transmitting and receiving data simultaneously over the same channel. As the belief decreases, the SU should monitor the channel while transmitting (i.e., operate in the STAS mode), as the probability that the PU returns is now relatively high. In that case, the SU achieves lower throughput than in the TR mode, but also a lower collision probability. As the belief decreases beyond a specific threshold, the SU should stop transmitting and either carry out HD sensing (i.e., SO mode) or switch to new channel. For relatively low belief values, the probability that the PU returns to the channel is quite high and the PU collision constraint will not be satisfied. Hence, more accurate sensing and a temporary pause in the transmission are required. At very low belief values, where the PU is most likely to return to the channel, the SU should take the CS action. This happens when the probability that the PU is idle in a new channel (where no information is available) is higher than the current belief.

B. Wi-Fi/LTE-U Coexistence

We consider an LTE-U small cell that coexists with a Wi-Fi network in the unlicensed band (see Figure 3). The LTE-U small cell consists of an HeNB that communicates with a number of UEs over an aggregation of licensed and unlicensed channels. Without loss of generality, we focus on the LTE-U DL. The Wi-Fi system consists of one FD-enabled AP that communicates with a number of FD-enabled STAs. A Wi-Fi network implements an exclusive channel occupancy policy among its STAs. Specifically, a channel is allocated to only a single Wi-Fi. Contention is resolved using CSMA/CA, where neighboring STAs defer from accessing the channel by setting their network allocation vector (NAV) after decoding the duration field in the MAC header.

In LTE-U, the HeNB must search for a free channel to use. If no idle channel is found, HeNB shares the spectrum with the Wi-Fi system according to an adaptive duty cycle. During the OFF period, the HeNB measures the traffic intensity of neighboring Wi-Fi STAs (e.g., by recording the MAC addresses of overheard transmissions) and adapts its duty cycle accordingly.

We now propose a modified TXOP scheme for FD-enabled Wi-Fi systems. We divide the TXOP into N_p time slots of equal duration, during which AP and STA can exchange UL and DL frames. We consider two FD modes: The simultaneous Transmit-Receive (TR) mode and the simultaneous Transmit-Sense (TS) mode, as shown in Figure 4. Wi-Fi AP switches between these modes to mitigate the interference caused by LTE-U transmission. We assume that the AP is the session "master". It instructs the STA about the recommended mode of operation (e.g., TR or TS) and the associated MCS indices that the STA has to use by embedding this information in the DL frame's optional header field (e.g., filed 'H' in Figure). This information requires a few bits, and hence represents small overhead. When LTE-U interference is relatively high the Wi-Fi AP has an option of quitting the TXOP period early and switching to a new channel. We use CS to refer to this channel-switching mode. AP also has the option of backing off until LTE-U completes transmission and the channel becomes idle again.

In the TR mode, the transmitted DL and UL frames can have different MCS indices (e.g., k_D and k_U, respectively). The Wi-Fi STA first reads the 'H' field in the DL frame and extracts the mode/MCS indices. Next, STA initiates a simultaneous UL transmission with MCS index k_U. After transmitting DL and UL frames, the AP and STA have to exchange ACK frames in both directions, indicating successful reception. In the TS mode, the AP sends a DL frame with an MCS index k_D, and simultaneously senses for any LTE-U signal using the detection scheme. At the end of each time slot, AP updates its belief about LTE-U HeNB interference, and selects a new FD mode with suitable MCS indices for the next time slot.

An example of the proposed TXOP scheme is shown in Figure 5, where the AP sends five DL frames (e.g., $N_p = 5$), each of duration Δ. In this example, the AP starts in the TR mode with MCS indices $k_D = k_U$, whose modulation is 64QAM. AP sends in the DL direction frame 'Fa1'. STA reads the header field and starts transmitting the 'Fs1' frame in the UL direction using 64QAM modulation. HeNB ON cycle starts just after the start of 'Fa1' and 'Fs1' transmission, which causes collision. Both AP and STA are not able to decode their received frames, and hence no ACKs are transmitted. In this case, the AP updates its belief about HeNB interference and selects a new action. For instance, the next optimal action might be retransmitting the 'Fa1' frame in the TS mode with QPSK modulation. If STA is able to decode this frame, it will send back an ACK. Upon receiving an ACK for 'Fa1' and sensing an HeNB signal, the AP updates its belief about HeNB interference and may decide to raise the modulation to 16QAM with TS mode for the next transmitted frame (e.g., frame 'Fa2'). The process continues as shown in Figure 5. In order for the AP to select the optimal action, which maximizes the link utility in the TXOP period, it should be able to quantify the amount of LTE interference that the AP and

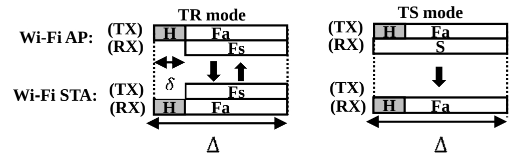

Fig. 4: FD modes: TR-TR and TS-RO modes ('S': Sense, 'Fa': AP frame, 'Fs': STA frame, and 'H': Header).

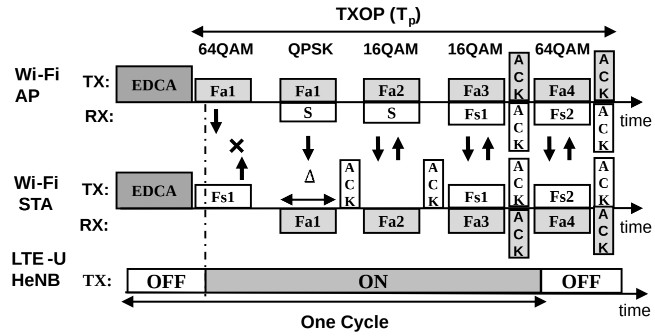

Fig. 5: Example of a modified TXOP operation mode ('S': Sense, 'Fa1', 'Fa2', 'Fa3', 'Fa4': Frames sent by AP, and 'Fs1', 'Fs2': Frames sent by STA).

STA receive. This interference is affected by the channel gains between AP/STA and HeNB. We model LTE activity and its interference using a FSMC model.

Wi-Fi AP mitigates the interference caused by LTE-U transmissions by jointly adapting FD modes and transmission rates during the TXOP period. This requires the knowledge of h_{la}, h_{ls}, h_{as}, and h_{sa} channel gains. The channel gains of h_{as} and h_{sa} can be implicitly and explicitly estimated. However, the HeNB cannot estimate the h_{la} and h_{ls} channel gains, because Wi-Fi and LTE-U uses different techologies. Wi-Fi AP can still obtain partial knowledge about these channel gains by monitoring the performance of Wi-Fi UL and DL links over time. For example, AP can indirectly deduce interference levels through monitoring ACKs and decoding received frames during TXOP period. Therefore, AP has to jointly control rates/modes in response to LTE-U hidden processes using this partial knowledge. This motivates the need for a HMM control scheme which can be formulated through a POMDP framework [10]. POMDP assigns a belief (probability) for each unknown parameter, and updates this belief sequentially over time during the execution based on the resultant outcomes. POMDP maximizes the Wi-Fi utility through mapping its belief about the interference caused by LTE-U to a set of actions, consisting of recommended joint rate/mode configurations during the TXOP period. This mapping function is known also as the policy of POMDP.

For simplicity, we assume that channels between Wi-Fi AP and STA (i.e., h_{as} and h_{sa}) are static, and focus on formulating the POMDP problem for the channels between LTE-U HeNB and Wi-Fi nodes (i.e., h_{la} and h_{ls}). First, we introduce the main components needed for formulating the POMDP problem. Then, we introduce the reward functions and explain the policy evaluation.

a) Time Horizon: POMDP will take place over a finite horizon equals to the duration of one TXOP period (i.e., T_p

second), where a total of $N_p = T_p/\Delta$ frames have to be exchanged each of Δ duration. In other word, there will be N_p time slots during each TXOP transmission. We denote each time slot as $\ell = \{1, \cdots, N_p\}$.

b) State Space: The state space represents the status of h_{la} and h_{ls} channel gains. We model the state space according to the FSMC model. We introduce a two dimensional finite state space $\mathcal{S}: M \times M$, where each state corresponds to h_{la} and h_{ls} channel gains. The number of states per each channel is $M = |\mathcal{K}| + 1$. We denote the $(h_{la}^{(i)}, h_{ls}^{(m)})$ state as $s^{(i,m)} \in \mathcal{S}$.

c) Action Space: At the start of each time slot, Wi-Fi AP has to take two decisions simultaneously; the FD mode (e.g., TR , TS, or CS) and the applicable transmission rates (i.e., the MCS indices k_U and k_D for the UL and DL transmissions, respectively). The channel switching CS mode is only selected when the transmission with the lowest MCS index is believed to be unsuccessful and the AP has enough knowledge about suitable channels for switching to. Instead, channel switching is replaced by a 'backoff' action. The action space is written as $\mathcal{A} = \{\mathrm{TR}\,(k_D, k_U), \mathrm{TS}\,(k_D), \mathrm{CS}\,\}$, and it has $|\mathcal{K}|^2 + |\mathcal{K}| + 1$ possible actions. We denote the action that the AP takes at the start of time slot ℓ as a_ℓ.

d) Observation Space: Wi-Fi AP takes an action $a_\ell \in \mathcal{A}$ at the start of time slot ℓ and waits for an observation at the end. This observation depends on the action that the AP takes and the true state of interference. The AP takes a TR action and receives four possible observations: Decode or Undecode $\{D, U\}$ for the UL frame and ACK or NACK $\{A, N\}$ for the DL frame. At the end of a TS action, Wi-Fi AP receives four possible outcomes: ACK or NACK for the DL frame and busy B or idle I for the sensing. The observation space is written as $\mathcal{O} = \{\{\bar{o}_{\mathrm{TR}}\}, \{\bar{o}_{\mathrm{TS}}\}\}$, where $\bar{o}_{\mathrm{TR}} \in \{(D, A), (D, N), (U, A), (U, N)\}$, and $\bar{o}_{\mathrm{TS}} \in \{(I, A), (I, N), (B, A), (B, N)\}$. Let $\bar{o}_\ell$ denotes the observation vector that AP receives at the end of time slot ℓ. Let $q_{a_\ell, \bar{o}_\ell}^{(i,m)}$ denotes the probability of receiving an observation vector $\bar{o}_\ell$ when the AP takes an action a_ℓ, while the channel states are $(h_{la}, h_{ls}) = (i, m)$:

IV. Performance Evaluation

Unless stated otherwise, we use the following parameters. $f_S = 6$MHz, $\sigma_s^2 = 5$, $m = 500$, $SNR^{(HD)} = -20$ dB, $\alpha = 1$, $p = 0.5$, T_{ON} and T_{OFF} are exponentially distributed random variables with means $\bar{T}_{ON} = \bar{T}_{OFF} = 5$, and $SNR_{TO} = 20$ dB.

We first evaluate the performance of waveform-based spectrum sensing for the FD TS mode and compare it with the energy-based sensing. Figures 6 and 7 depict P_f and P_d versus the sensing duration for different values of χ. Generally, the performance of any spectrum sensing technique expectedly improves (i.e., P_f decreases and P_d increases) with the sensing duration, as more samples are being used for PU detection. Also, as χ increases the performance of waveform-based sensing (and similarly for energy-based sensing) degrades due the increase in the RSI. At perfect SIS, P_f and P_d converge to the HD case. As shown in the Figures, SUs need about 20% longer sensing durations to achieve the same sensing accuracy of the HD mode with 20% RSI.

We use LabVIEW to simulate the proposed adaptive scheme, referred to as TSRA, and compare its performance with the LBT scheme. Some Figures are drawn with a 95% confidence interval. Figures 8 and 9 depict the average throughput and collision rate, respectively, as functions of the noise power for TSRA and the LBT scheme. In some cases, TSRA achieves up to 2x throughput gain over the standard LBT scheme, and reduces the average collision rate by more than 30%.

V. Conclusions

In the last few years, there has been lots of interest in developing powerful SIS techniques that enable radios to operate in inband FD mode to increase the link's throughput. In this paper, we investigated the incorporation of SIS techniques in dynamic spectrum sharing systems. In addition to the inband FD mode, an FD-capable radio can simultaneously transmit and sense (STAS) over the same channel which reduces the collision probability with coexisting systems. We investigated various sensing techniques for the STAS mode. We also proposed an FD-based coexisting framework and studied two uses cases related to opportunistic spectrum access systems and Wi-Fi/LTE-unlicensed coexistence.

References

[1] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[2] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access: Signal processing, networking, and regulatory policy," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[3] LTE-U Forum, "LTE-U CSAT procedure TS v1.0," , Oct. 2015.

[4] 3GPP, "Study on licensed-assisted access to unlicensed spectrum," 3GPP TR. 36.889 v13.0.0., Jun. 2015.

[5] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proc. of the IEEE INFOCOM'13 Conf.*, Turin, Italy, Apr. 2013, pp. 1258–1266.

[6] ——, "Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities," in *Proc. of the IEEE DySPAN'14 Conf.*, McLean, VA, Apr. 2014, pp. 149–160.

[7] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[8] S. Huang, X. Liu, and Z. Ding, "Optimal sensing-transmission structure for dynamic spectrum access," in *Proc. of the IEEE INFOCOM'09 Conf.*, April 2009, pp. 2295–2303.

[9] R. Bellman, "Dynamic programming and lagrange multipliers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 42, no. 10, p. 767, 1956.

[10] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden Markov model sensors," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1382–1397, Jun 2002.

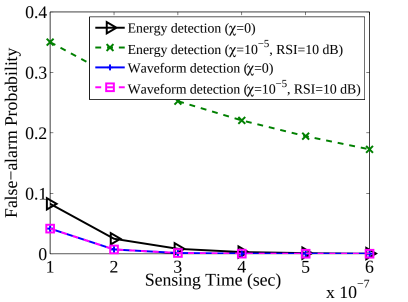

Fig. 6: False-alarm probability vs. sensing time in the FD STAS mode.

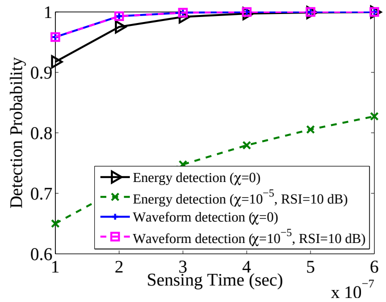

Fig. 7: Detection probability vs. sensing time in the FD STAS mode.

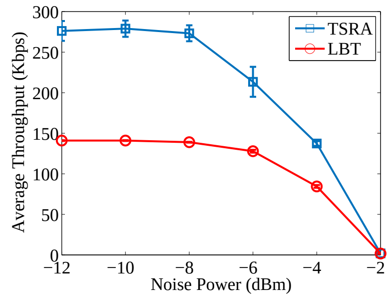

Fig. 8: Average SU throughput vs. noise power in dBm for TSRA and LBT schemes.

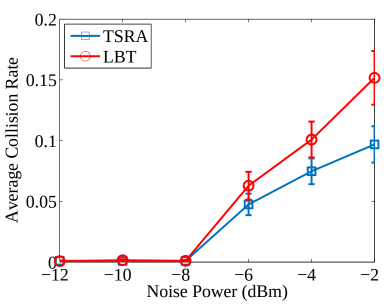

Fig. 9: Average SU collision rate vs. noise power in dBm for TSRA and LBT schemes.