Joint Mode and Rate Adaptation for Asymmetric Full-duplex Communications in WLANs

Mohammed Hirzallah[1], Wessam Afifi[2], and Marwan Krunz[1]
[1]Department of Electrical and Computer Engineering, University of Arizona, AZ, USA
[2]Mavenir Systems, Richardson, TX, USA
{hirzallah, krunz}@email.arizona.edu

***Abstract*—We consider the application of asymmetric full-duplex (AFD) communications in wireless local area networks (WLANs), exemplified by a Wi-Fi system. A full-duplex (FD)-enabled Wi-Fi access point communicates simultaneously uplink (UL) and downlink (DL) with a pair of half-duplex (HD) Wi-Fi stations (STAs). AFD enhances spectrum efficiency and reduces latency for STAs; however, it faces challenges related to node selection, switching between AFD and HD modes, and rate control. In this paper, we propose the *asymmetric FD-mode and rate adaptation* (AFRA) scheme. AFRA relies on partially observable Markov decision process (POMDP) to adapt the UL and DL transmission rates as well as the transmission mode (i.e., AFD or HD; UL-only or DL-only), considering the future expected channel and interference conditions. Our simulation results indicate that AFRA outperforms classical rate-adaptation schemes and achieves up to 95% of the optimal performance of AFD communications.**

***Index Terms*—Asymmetric full-duplex, TXOP sharing, joint rate and duplex-mode control.**

I. Introduction

Classical bidirectional communication is achieved by separating the forward and reverse links in time (TDD) or frequency (FDD). Simultaneous transmission and reception on the same frequency channel, i.e., full-duplex (FD) mode, is challenging due to the existence of strong self-interference. The infeasibility of FD communications was challenged by several studies (see [1] for a survey), which successfully demonstrated the possibility of FD communications using self-interference suppression (SIS) techniques.

Symmetric FD communications require the two ends of a wireless link to be capable of SIS. Although implementing SIS techniques in Wi-Fi access points (APs) and relatively large communication devices (e.g., laptops, TVs, large tablets, etc.) is foreseeable, it is currently quite challenging to do that in small-factor devices (e.g., smart phones). Hence, in this paper we consider half-duplex (HD) Wi-Fi stations (STAs) but FD-enabled APs. Under this assumption, an AP can operate in an asymmetric full-duplex (AFD) mode, transmitting downlink (DL) frames to an STA (e.g., STA-D) while receiving uplink (UL) frames simultaneously from another STA (e.g., STA-U) on the same frequency channel, as shown in Figure 1.

Wi-Fi systems use the concept of the *transmit opportunity* (TXOP), which involves the exchange of several data frames after one successful channel contention. Fading and the bursty interference behavior create a complicated wireless channel with a dynamic coherence time, which makes assigning one transmission rate for each TXOP inefficient from a spectral utilization point of view. Furthermore, assigning rates to the UL/DL channels is not trivial as the two receivers (STA-D and AP) face different channel and interference conditions.

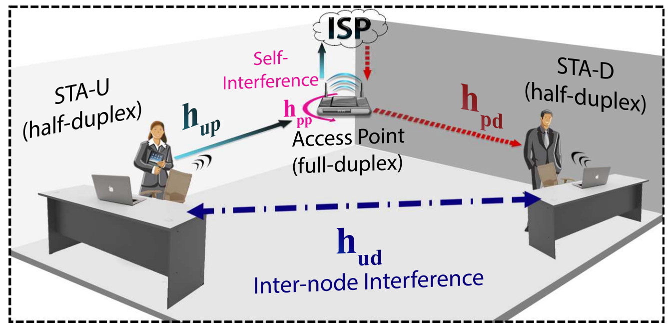

Fig. 1: System model of a Wi-Fi network, consisting of one FD AP and two HD STAs (STA-U and STA-D).

In the AFD mode, simultaneous UL and DL frames are subject to fading, AP's self-interference, inter-node interference, and external interference from other coexisting systems (e.g., unlicensed LTE [2]). This motivates the idea of adapting the modulation and coding schemes (MCSs) of transmitted frames. Because a TXOP period consists of multiple data frames, we consider MCS adaptation on a per-frame basis, using the TXOP duration as our optimization horizon. The objective of such adaptation is to enhance spectrum efficiency. Furthermore, in some scenarios with excessive external and/or self-interference (due to limited SIS capabilities) or fading, and in dense networks, it is more beneficial for the AP to operate in the HD mode [3] (i.e., UL or DL but not both).

Transmission-rate adaptation in WLANs has been addressed in the literature (see [4] for a survey). Proposed schemes include the auto-rate fallback (ARF), Onoe, SampleRate, and receiver-based auto-rate (RBAR). Most of these schemes adapt the rate in an ad-hoc fashion, based on either long-term statistics, packet successful rate, random channel probing, or SNIR. A common feature of these schemes is their relatively long reaction time, which can range from hundreds to thousands of milliseconds. For example, ARF adjusts the rate for a sequence of frames, Onoe adjusts the rate once every 100 msec, and

This research was supported in part by NSF (grants # IIP-1265960, IIP-1535573, CNS-1409172, and CNS-1563655) and by the Broadband Wireless Access & Applications Center (BWAC). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF.

SampleRate requires a monitoring window of 10 seconds to facilitate its adaptation.

In contrast, rate selection approaches based on stochastic decision theory, such as partially-observable-Markov-decision processes (POMDPs), can provide relatively faster adaptation responses [5], [6], [7]. In this paper, we introduce the *asymmetric FD-mode and rate adaptation* (AFRA) scheme. AFRA is a POMDP-based adaptation scheme that jointly adapts the rate and mode over a "*shared-TXOP* period" between an AP and two Wi-Fi STAs. In shared-TXOP, several links could be active simultaneously in one TXOP period. The Wi-Fi AP can switch between AFD, UL-only, and DL-only modes, while adapting the MCS index on a per-frame basis (see Figure 2). We characterize the SINR at UL and DL connections using a two-dimensional finite-state Markov chain model, and utilize this chain in POMDP design. We also introduce payoff functions that reward the AP for successful UL and DL transmissions, and penalize it when collision or outage occurs. We address the required changes in the protocol design to facilitate such adaptation where by introducing a modified TXOP sharing design. To the best of our knowledge, this is the first work that addresses the joint mode and rate adaptation in AFD communications. Our adaptation framework achieves up to 95% of the optimal performance and outperforms other classical approaches. In [7], we addressed the problem of joint rate and mode adaptation for symmetric FD communications, where AP and STA are both equipped with SIS capabilities.

The rest of the paper is organized as follows. We present the system model in Section II. We discuss the proposed AFRA adaptation scheme and the POMDP framework in Section III, and discuss some related issues in Section IV. Finally, we evaluate the performance of the proposed scheme and conclude the paper in Sections V and VI, respectively.

II. System Model

We consider an AFD-based WLAN that consists of one FD-enabled AP and a number of HD-enabled stations. AP/STAs perform CSMA/CA prior to their transmission based on the *enhanced distributed channel access* (EDCA) scheme. In EDCA, Wi-Fi AP/STAs contend for channel access and the winning device has the opportunity to transmit several frames continuously for a duration equal to the TXOP period. During the TXOP, the AP could communicate with two STAs simultaneously, one at the uplink (e.g., STA-U) and the other at the downlink (e.g., STA-D). During the TXOP, the AP has four modes of communications: *uplink-only* (UL-only) with STA-U; *downlink-only* (DL-only) with STA-D; simultaneous UL/DL with STA-U/STA-D, which we call as *asymmetric FD* mode (AFD); and *Backoff* (BO) mode. The BO mode is selected when neither UL nor DL connections are successful.

The AP adapts the communication mode and the transmission rate for its contained frames using AFRA scheme; as we introduce in Section III-B. We divide the TXOP period into time slots of equal length where the AP can switch between the four modes at the slot boundaries. Each time slot is divided into two phases, *data phase* and *control phase*; as shown in Figure 2. The data phase is used to exchange UL and DL data frames, while the control phase is used to exchange control

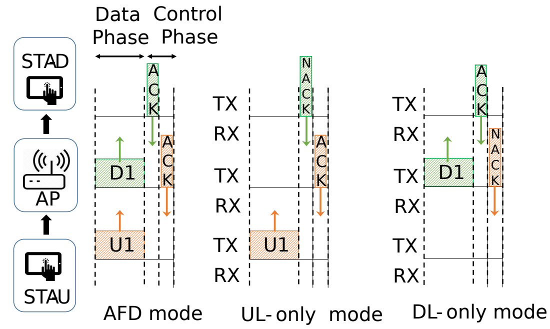

Fig. 2: Communication modes: AFD, UL-only, and DL-only. STA-U is the uplink and STA-D is the downlink (D1: Downlink frame 1; U1: Uplink frame 1).

frames such as *acknowledgment* (ACK), *Proceed* (PRC), and *negative-ACK* (NACK). The NACK frames are used for synchronization and '*keep-alive*' purposes and to keep other Wi-Fi STAs silent during the TXOP period. PRC frame is not shown and will be discussed in more detail in Section IV-B.

Let h_{up}, h_{pd}, and h_{ud} be the channel gains between STA-U and AP; AP and STA-D; and STA-U and STA-D, respectively. Let h_{pp} be the channel gain of the self-interference channel at the AP, modeling the medium between its transmit and receive chains. Let χ_p be the SIS capability at the AP (perfect SIS occurs at $\chi_p = 0$). The UL and DL received signals depends on the communication mode $a \in \{\text{AFD}, \text{UL-only}, \text{DL-only}\}$, and they are expressed, respectively, as:

$$y_p^{(a)} = h_{up} s_u + h_{\text{pp}} \chi_{\text{p}}\, s_p \mathbf{1}_a + w_p$$

$$y_d^{(a)} = h_{pd} s_p + h_{ud} s_u \mathbf{1}_a + w_d$$

where s_u and s_p are STA-U and AP transmitted signals, respectively, and w_p and w_d are the additive-white-Gaussian noise (AWGN) signals at AP and STA-D receivers, respectively, and $\mathbf{1}_a = \{1 : a = \text{AFD}\}$. The signal-to-interference-and-noise ratios (SINRs) for both UL (i.e., SINR_p^a) and DL (i.e., SINR_d^a) connections, respectively, are functions of the communication mode a, and are written as:

$$\text{SINR}_p^a = \frac{|h_{up}|^2 P_{\text{st}}}{\underbrace{|h_{\text{pp}}|^2 \chi_p^2 P_{\text{ap}} \mathbf{1}_a}_{\text{self-interference}} + N_p}, \quad \text{SINR}_d^a = \frac{|h_{pd}|^2 P_{\text{ap}}}{\underbrace{|h_{ud}|^2 P_{\text{st}} \mathbf{1}_a}_{\text{inter-node interference}} + N_d} \tag{1}$$

where P_{st} and P_{ap} are STA-U and AP transmit powers, respectively. N_a and N_d are AWGN powers at AP and STA-D receivers, respectively. Selecting the AFD mode causes a *self-interference* at AP and *inter-node interference* at STA-D.

III. Asymmetric FD-Mode and Rate Adaptation (AFRA) Scheme

Characterizing real-time changes in SINR values is crucial for selecting a proper joint communication mode and its respective transmission rate. There are two sources of these variations: Channel fading and shadowing; and the self-interference and inter-node interference caused by switching between the different communication modes, as stated in (1). In this section, we discuss how SINR variations at UL and

DL links can be characterized in a probabilistic structure using finite-state Markov chain (FSMC).

A. Finite-State Markov Chain-based SINR Model

FSMC is used in literature to characterize the instantaneous variation in SINR overtime (due to channel's shadowing and fading) [8]. Let γ and $\bar{\gamma}$ be the instantaneous and mean values of SINR. Let ν_i be the probability that γ is at the ith interval, i.e., $\gamma \in [g_i, g_{i+1})$, where g_i and g_{i+1} are two arbitrary SINR thresholds. Assuming a Rayleigh fading distribution, then $\nu_i = \Pr(g_i \leq \gamma < g_{i+1}) = \int_{g_i}^{g_{i+1}} p(\gamma) d\gamma$ where $p(\gamma) = \frac{\gamma}{\bar{\gamma}^2} e^{-\gamma^2/2\bar{\gamma}^2}$. The *level-crossing rate* L_i defines how often SINR passes a certain threshold g_i, and this rate depends on user's mobility, expressed in Doppler frequency f_d, as follows: $L_i = \sqrt{\frac{2\pi g_i}{\bar{\gamma}}} f_d e^{-g_i/\bar{\gamma}}$ [9].

We define the states in the FSMC based on the values that γ would take. For example, we say that the SINR is in the ith state when $\gamma \in [g_i, g_{i+1})$. We construct the transition probabilities between these states based on channel fading statistics. Let T be the time for which SINR remains static, a.k.a, *average-fade duration*, then the nonzero transition probabilities are expressed as:

$$\tilde{q}_{i,i-1} = \frac{L_i T}{\nu_i}, \quad \tilde{q}_{i,i+1} = \frac{L_{i+1} T}{\nu_i}, \quad \tilde{q}_{i,i} = 1 - \tilde{q}_{i,i+1} - \tilde{q}_{i,i-1}. \quad (2)$$

As explained before, switching between the different modes changes the SINRs at UL and DL receivers; see (1). For example, switching from AFD to UL-only or DL-only modes improves SINRs, while switching from the UL-only or DL-only to AFD mode does the opposite. Therefore, any switching to/from the AFD mode modulates the transition probabilities in (2) as follows. Let $\ell_p^{a,a'}$ and $\ell_d^{a,a'}$ be the respective change in the state of SINRs at AP and STA-D receivers, respectively, due to the switching from mode a to a', then:

$$(\ell_p^{a,d}, \ell_d^{a,d}) = \begin{cases} (-\delta_p, -\delta_d), & a \in \{\text{UL-only, DL-only}\}, a' = \text{AFD} \\ (\delta_p, \delta_d) & , a = \text{AFD}, a' \in \{\text{UL-only, DL-only}\} \\ (0,0) & , \text{otherwise} \end{cases}$$

where δ_p and δ_d represent the change in SINRs due to the self- and inter-node interference, respectively. The nonzero transition probabilities of FSMC become:

$$q^{a,a'}_{i,i+\ell_c^{a,a'}-1} = \frac{L_{i+\ell_c^{a,a'}} T}{\nu_{i+\ell_c^{a,a'}}}, \qquad q^{a,a'}_{i,i+\ell_c^{a,a'}+1} = \frac{L_{i+\ell_c^{a,a'}+1} T}{\nu_{i+\ell_c^{a,a'}}},$$
$$q^{a,a'}_{i,i+\ell_c^{a,a'}} = 1 - q^{a,a'}_{i,i+\ell_c^{a,a'}+1} - q^{a,a'}_{i,i+\ell_c^{a,a'}-1} \quad (3)$$

where $c \in \{p, d\}$. When $\ell_c^{a,a'} = 0$, these probabilities reduce to those in (2).

To account for joint variations in SINRs at AP and STA-D receivers, we extend the FSMC into a two-dimensional one. Let (im) be the joint state where i and m are the states of SINRs at AP and STA-D receivers, respectively. The transition probability between the joint states (im) and (jn) are written as:

$$p^{a,a'}_{im,jn} = q^{a,a'}_{i,j} \, q^{a,a'}_{m,n}. \quad (4)$$

We specify the boundaries between states based on their supported transmission rates. The IEEE 802.11ac standard specifies the *error-vector magnitude* (EVM) thresholds $\mathcal{V} = \{v_1, \cdots, v_M\}$ for the supported *modulation-coding schemes* (MCSs) [10]. The EVM threshold specifies the maximum allowable displacement in constellation points, which happens due to noise and interference, of each MCS. The EVM thresholds can be translated into SINR thresholds using the the approximate relation $g_i \approx 1/v_i^2$ [11]. Let $\mathcal{K} = \{1, \cdots, K_m\}$ be the set of supported MCSs, and $\mathcal{M} = \{1, \cdots, M_s\}$ be the set of supported MCSs. FSMC have $N = M + 1$ possible states, where each state corresponds to one of the MCSs defined in the standard. The first state corresponds to the case when no MCS would be supported. We define the *outage-indicator function* to specify when an MCS is not supported, i.e., $\gamma < g_i$. A transmitter should avoid using an MCS when the outage-indicator function is one. Let ρ_i^k be the outage-indicator function when the transmitter chooses the kth MCS while the SINR is at the ith state for any $k \in \mathcal{K}$ and $i \in \mathcal{M}$, then:

$$\rho_i^k = \begin{cases} 1 & , \quad i < k \\ 0 & , \quad i \geq k. \end{cases} \quad (5)$$

B. AFRA POMDP-based Design

Adapting the communication mode and rates in asymmetric FD communication requires knowledge about the channel gains and SINR states at both UL and DL receivers. Although such knowledge is hidden to the AP, it could still infer these states by decoding UL frames sent by STA-U and monitoring the ACKs sent by STA-D. SINRs at UL and DL receivers vary in a Markov-based fashion and AP has a partial knowledge about them, therefore AFRA utilizes the *partially-observable-Markov-decision process* to adapt the communication modes and their transmission rates in the TXOP period.

In POMDP, we define *actions* to be the set of the communication modes and their possible transmission rates, and assign each action a *reward* value. We solve for the optimal sequence of actions to be taken during the TXOP period by considering all possible actions, *observations*, and *beliefs* about SINR states. We assign beliefs (probabilities) about the SINR states because its not possible to fully observe them. The optimal solution becomes a function of AP's observations and beliefs, and is defined using a *policy* that tells what action the AP should take based on its beliefs and the observations it receives. The process of computing the policy takes place offline. Once the AP occupies the channel and starts the TXOP, it initiates beliefs $\bar{\pi}_0$ about SINRs at UL and DL receivers; see Figure 3. AP consults with the policy for an action a_1 to be taken, and at the end of the first time slot it receives observations o_1. AP uses these observations to update its beliefs $\bar{\pi}_1$, and consults a gain with the policy and takes a new action a_2 for the next time slot. The same process repeats again until the end of the TXOP period.

Next, we introduce the main POMDP elements, including state, action, and observation spaces. We also introduce the reward and value function formulations, and explain how to obtain the optimal policy.

1) POMDP Elements: We consider a discrete *time horizon* $\mathcal{T} = \{0, 1, \cdots, L\}$ that corresponds to the TXOP period with L time slots. The *action space* $\mathcal{A}$ includes the three possible communication modes: AFD (i.e., simultaneous UL

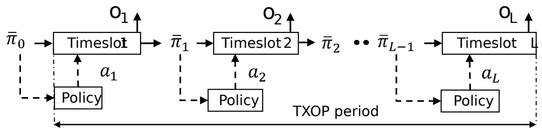

Fig. 3: POMDP operation during the TXOP period (a_i: Action at the ith time slot; o_i: Observation at the ith time slot; and $\bar{\pi}_i$: Belief vector at the end of the ith time slot).

and DL), UL-only, and DL-only modes, with their corresponding rates. When the SINR becomes too low for both UL and DL links, it would be better to quit TXOP early, hence, we introduce another action, backoff (BO), to account for such extreme condition. We define the action space as $\mathcal{A} = \{\text{AFD}_{k_u,k_d}, \text{UL}_{k_u}, \text{DL}_{k_d}, \text{BO} : \forall k_u, k_d \in \mathcal{K}\}$, where k_u and k_d are the MCS assigned for UL and DL links, respectively. Let $a_t \in \mathcal{A}$ be the action taken at time slot t for $t \in \mathcal{T}$.

We define the *state space* $\mathcal{S}$ to be the possible SINR values in (1) quantized according to the two-dimensional FSMC chain presented in Section III-A. Let $(im) \in \mathcal{S}$ be the joint state for which the SINRs at AP and STA-D receivers are at the ith state and the mth state, respectively. The transition probabilities $p_{s',s}^{a_{t-1},a_t}$, for any $s, s' \in \mathcal{S}$ and $a_{t-1}, a_t \in \mathcal{A}$, are as defined in (4).

The *observation space* $\mathcal{O}$ consists of all possible outcomes that the AP could receive after taking an action. Each action $a \in \mathcal{A}$ has specific possible observations. For actions with AFD mode, the AP either decodes (D) or fails to decode (F) the UL frame, and either receive an ACK or NACK from STA-D for the DL frame. BO action has no observations because the AP terminates the TXOP period. These observations constitute the observation space defined as $\mathcal{O} = \{(\text{D}, \text{ACK}), (\text{D}, \text{NACK}), (\text{F}, \text{ACK}), (\text{F}, \text{NACK}), (\text{F}), (\text{D}), (\text{ACK}), (\text{NACK})\}$. We define the *observation probability* as the probability of receiving an observation o_t when the AP takes an action a_t while the SINRs are at the sth joint state and denote it by $r_{s,o_t}^{a_t}$. The observation probabilities for various actions/observations are defined as follows:

$$r_{(im),o_t}^{\text{AFD}_{k_u,k_d}} = \begin{cases} (1-\rho_i^{k_u})(1-\rho_m^{k_d}) & \text{, for } o_t = (\text{D}, \text{ACK}) \\ (1-\rho_i^{k_u})\,\rho_m^{k_d} & \text{, for } o_t = (\text{D}, \text{NACK}) \\ \rho_i^{k_u}\,(1-\rho_m^{k_d}) & \text{, for } o_t = (\text{F}, \text{ACK}) \\ \rho_i^{k_u}\rho_m^{k_d} & \text{, for } o_t = (\text{F}, \text{NACK}) \end{cases} \quad (6)$$

$$r_{(jn),o_t}^{\text{UL}_{k_u}} = \begin{cases} (1-\rho_j^{k_u}) & \text{, for } o_t = (\text{D}) \\ \rho_j^{k_u} & \text{, for } o_t = (\text{F}) \end{cases} \quad (7)$$

$$r_{(k\ell),o_t}^{\text{DL}_{k_d}} = \begin{cases} (1-\rho_\ell^{k_d}) & \text{, for } o_t = (\text{ACK}) \\ \rho_\ell^{k_d} & \text{, for } o_t = (\text{NACK}) \end{cases} \quad (8)$$

$$r_{(im),o_t}^{BO} = \frac{1}{|\mathcal{O}|}, \forall\,(im) \in \mathcal{S}, \forall o_t \in \mathcal{O} \quad (9)$$

where $\rho_i^{k_u}$ is the outage-indicator function defined in (5).

As discussed before, the AP cannot monitor the true state of SINR values, and hence it assigns beliefs about them. These beliefs are simply the probabilities for potential SINR states. Let Ω be the probability space $\Omega = \{\omega : \omega \in [0,1]\}$. We define the *state-belief space* as $\mathcal{B} : \mathcal{S} \times \Omega$. At time slot t, we assign each state in $\mathcal{S}$ a belief value $\pi_{s,t} \in \mathcal{B}$. We define the *belief vector* at the tth time slot $\bar{\pi}_t = \langle \pi_{1,t}, \cdots, \pi_{|\mathcal{S}|,t} \rangle$. The sum of beliefs sums to one. After taking an action $a_t \in \mathcal{A}$ at time t and getting an observation $o_t \in \mathcal{O}$, the AP updates its belief about each SINR state using the following Bayes rule:

$$\pi_{s,t} = f(\bar{\pi}_{t-1}, a_t, o_t) = \frac{\sum_{s' \in \mathcal{S}} \pi_{s',t-1} p_{s',s}^{a_{t-1},a_t} r_{s,o_t}^{a_t}}{\sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \pi_{s',t-1} p_{s',s}^{a_{t-1},a_t} r_{s,o_t}^{a_t}}. \quad (10)$$

The belief vector is a sufficient statistic that helps AP trace the state of the environment without need to keep record for all previous actions and their corresponding observations [12].

2) Immediate Reward Formulations: We define the reward that the AP receives at each time slot to be the amount of data communicated successfully minus the cost represented in the associated power losses. Wi-Fi frames are OFDM modulated and the amount of data we can accommodate in one time slot is $R_k = N_o N_c b_k c_k$, where N_o is the number of OFDM symbols that fit in one time slot, N_c is the number of subcarriers, b_k is the modulation order, and c_k is the coding rate c_k of the kth MCS. Let $W_{o_t}^{a_t}$ be the reward that the AP receives after taking an action a_t and receiving an observation o_t:

$$W_{o_t}^{\text{AFD}_{k_u,k_d}} = \begin{cases} R_{k_u} + R_{k_d} - \eta(P_{\text{ap}} + P_{\text{st}}) & \text{, for } o_t = (\text{D}, \text{ACK}) \\ R_{k_u} - \eta(P_{\text{ap}} + P_{\text{st}}) & \text{, for } o_t = (\text{D}, \text{NACK}) \\ R_{k_d} - \eta(P_{\text{ap}} + P_{\text{st}}) & \text{, for } o_t = (\text{F}, \text{ACK}) \\ -\eta(P_{\text{ap}} + P_{\text{st}}) & \text{, for } o_t = (\text{F}, \text{NACK}) \end{cases} \quad (11)$$

$$W_{o_t}^{\text{UL}_{k_u}} = \begin{cases} R_{k_u} - \eta P_{\text{st}} & \text{, for } o_t = (\text{D}) \\ -\eta P_{\text{st}} & \text{, for } o_t = (\text{F}) \end{cases} \quad (12)$$

$$W_{o_t}^{\text{DL}_{k_d}} = \begin{cases} R_{k_d} - \eta P_{\text{ap}} & \text{, for } o_t = (\text{ACK}) \\ -\eta P_{\text{ap}} & \text{, for } o_t = (\text{NACK}) \end{cases} \quad (13)$$

$$W_{o_t}^{\text{BO}} = \eta(P_{\text{ap}} + P_{\text{st}}), \text{ for } \forall o_t \in \mathcal{O} \quad (14)$$

where η is a scaling coefficient that we use to match data and power terms, P_{ap} and P_{st} are the transmit powers of the AP and STAs, respectively. We include the power as a cost to penalize the AP when communication becomes unsuccessful due to outage conditions. When the AP takes an action at the start of the tth time slot it does not know whether this action results in successful transmission. Therefore, we define the *expected immediate reward* as the average reward over all possible outcomes and beliefs. Let D^{a_t} be the expected immediate reward of the a_t action:

$$D^{a_t} = \text{E}[W_{o_t}^{a_t}] = \sum_{o_t \in \mathcal{O}} \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \pi_{s',t-1} p_{s',s}^{a_{t-1},a_t} r_{s,o_t}^{a_t} W_{o_t}^{a_t} \quad (15)$$

3) Value Function Formulation: Our goal is to maximize the accumulated reward that the AP receives along all time slots during the TXOP period. The actions that the AP takes at the start of the TXOP affects its subsequent belief updates, and hence this impacts the actions to be taken subsequently. Therefor, we have to account for the expected immediate reward and the *expected long-term reward*. We define the *value function* to account for immediate and future rewards. The

Algorithm 1 Selection Algorithm of Downlink STA

AP receives a request from STA-U to start a TXOP
AP filters the STAs based on their inter-node interference with STA-U
if No DL STA with buffered data was found in the filtered set **then**
    AP continues the TXOP period in UL mode
else
    AP filters the STAs again based on their fairness and AC status
    if a DL STA, say STA-D, with the same AC is found **then**
        AP selects STA-D as a DL STA
    else if the STA-U AC is VI and STA-D is VO **then**
        AP selects a STA-D for DL, otherwise it proceeds with UL mode
    else
        AP proceeds with UL-only mode
    end if
end if

value function at the tth time slot can be written in a recursive fashion as:

$$V_t(\bar{\pi}_{t-1}) = \max_{a_t \in \mathcal{A}} \Big[D^{a_t} + \kappa \sum_{o_t \in \mathcal{O}} \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \pi_{s,t-1} p_{s,s'}^{a_{t-1},a_t} r_{s',o_t}^{a_t} V_{t+1}[f(\bar{\pi}_{t-1}, a_t, o_t)] \Big] \quad (16)$$

where $V_{t+1}[f(\bar{\pi}_{t-1}, a_t, o_t)]$ is the value function at the $(t+1)$th time slot, and κ is known as the discount factor. The discount factor characterizes how much future rewards are important. The value function defined in (16) has been proved to be piece-wise linear convex function [12]. The optimal policy maps the beliefs $\bar{\pi}_t$ about the SINR states to the optimal actions that maximizes the value function in (16) (i.e., $\mu_t^* : \mathcal{B} \mapsto \mathcal{A}$).

4) Solving POMDP: To determine the optimal policy μ^*, we need to find the set of actions that satisfy (16). This optimization can be solved through dynamic programming. However, the large number of states makes solving such a problem daunting and obtaining one policy may require days. Many algorithms were proposed in literature to solve such a program in exact form, while others followed approximate and heuristic approaches. A comparison between all of these approaches and their relative computational complexities can be found in [13]. We solved our problem using an approximate point-based POMDP solver called SARSOP [14]. SARSOP improves the computational efficiency for solving (16) by sampling a few initial values of the Belief space $\mathcal{B}$, and checking for the optimal solutions reachable from these initials. Point-based algorithms have a polynomial time complexity, and are efficient when the problem have tens of states.

IV. Discussion

A. Nodes Selection Algorithm

In this section, we propose an algorithm for selecting the downlink STA, e.g., STA-D, that shares the TXOP with STA-U. First, the AP builds and maintains an interference graph for its associated STAs. To enable this, each STA includes in its transmitted frame the *association identities* (AIDs) of the neighboring STAs that it can overhear. AP updates the interference graph frequently, where it disassociates any STA that remains silent for a certain period of time.

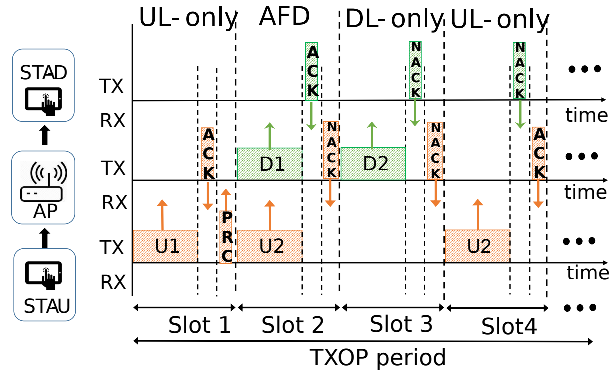

Fig. 4: Example of shared-TXOP operation in asymmetric full-duplex communications, frame 'U2' is lost in the second time slot, and frame 'D2' is lost in the third one (U1 and U2: UL frames; D1 and D2: DL frames; PRC: Proceed; ACK: Acknowledgment; NACK: Negative Acknowledgement).

Algorithm 1 shows the procedure for selecting the downlink STA, e.g., STA-D, where the AP utilizes the interference graph to filter the STAs that are not impaired by STA-U interference. When the filtered set contains no candidate downlink STA with a buffered data, the AP continues with the TXOP period in a half-duplex transmission. Otherwise, it sorts the STAs based on their access categories (ACs) as well as on how fair they have been treated. EDCA scheme supports several priority categories and assigns each category specific channel access parameters and TXOP period. For example, voice (VO) and video (VI) ACs have TXOP periods of 1.5 and 3.008 milliseconds, respectively. For fair treatment, the AP monitors the number of successful frame transmissions in the last T_A seconds for all candidate STAs and excludes the greedy ones (i.e., STAs exceeding a certain threshold value). In the selected set, the AP searches for a candidate downlink STA with the same AC of STA-U (e.g., VO, VI, background: BK, best-effort: BE). If the AC of STA-U is VO and no candidate downlink station has the same AC, the AP proceed with an HD-based transmission. When the AC of STA-U is VI and no downlink STA has the same AC, AP selects any existing STA with AC VO and treats it as VI, otherwise, it proceeds in an HD transmission.

B. TXOP Sharing Protocol

Assume that STA-U has already contended for a TXOP period, and AP has already selected STA-D for downlink as described in Algorithm 1. In the first TXOP time slot, STA-U initiates the TXOP by sending the frame 'U1' to AP, and AP acknowledges with an ACK frame, as shown in the example of Figure 4. An ACK timeout during the first time slot forces STA-U to terminate the TXOP period. The ACK frame sent by AP contains information about the timing structure of the TXOP and the MCS that STA-U has to use over the second time slot. Upon the successful reception of AP's ACK, STA-U sends a *proceed* (PRC) control frame, informing the AP to proceed with leading the operation in TXOP period. The PRC

see [15] for more details on IEEE 802.11e amendment.

frame is used only once at the start of TXOP period. For all subsequent time slots, AP has to send NACK frame when it cannot decode STA-U's frames. This NACK helps synchronize STA-U and update it about further actions for the next coming time slots.

At the start of the second time slot, AP initiates the adaptation process where it simultaneously sends frame 'D1' to STA-D and receives frame 'U2' from STA-U. The header of 'D1' contains information needed for decoding, and also also includes information about the timing structure of TXOP period. AP waits for an ACK from STA-D. After receiving STA-D's ACK, AP updates its belief vector as in (10) and consults with the policy μ^* for the best action to take on the third time slot. If AP cannot decode frame 'U2', it sends a NACK to STA-U, updating it about the next action that STA-U has to take and informing it that frame 'U2' was not delivered successfully. At the third time slot, STA-D cannot decode frame 'D2', and hence STA-D sends a NACK back to AP. AP utilizes these observations, updating its beliefs vector and taking new actions The same process continues until the end of TXOP period. The NACK frames are sent to synchronize the whole process and to ensure that channel will not be seized by other Wi-Fi nodes.

V. Performance Evaluation

A. Simulation Methodology

We consider a full-duplex-enabled Wi-Fi AP that communicates with two half-duplex stations, STA-U at the uplink and STA-D at the downlink, as shown in Figure 1. All Wi-Fi devices operate on the same channel with center frequency of 5.5 GHz. We adopt a path-loss channel model with pathloss exponent of 4, and a Rayleigh fading with mean value of 10 dB. In our simulation, we assume an imperfect self-interference suppression, and set the residual self-interference to be 5 dB above the AP noise floor. We consider eight MCS indices $k \in \{0, \cdots, 7\}$ with modulation index $b_k \in \{1, 1, 2, 2, 4, 4, 6, 6\}$ and coding rate $c_k \in \{0.5, 0.75, 0.5, 0.75, 0.5, 0.75, 0.666, 0.75\}$. The EVM thresholds for these MCSs are $\mathcal{V} \in \{-5, -8, -10, -13, -16, -19, -22, -25\}$ dB [10].

We assume that the Wi-Fi devices have already contended for a TXOP channel access, and UL/DL stations are already selected. We vary the locaiton of STA-U and STA-D and repeat the simulation for 100 runs. On each run, we evaluate the minimum average fade duration, and compute the number of frames (i.e., the number of time slots) that can be exchanged over a one TXOP period of 3 msec. The discount factor κ in (16) and the cost scaling coefficient η in (11)-(14) are usually selected empirically. In our simulations, we set $\kappa = 0.95$. We compare our scheme against the following:

1) The AP has a full knowledge about SINR states at both uplink and downlink receivers. The AP selects the maximum supported rate and the best communication mode based one the SINR values. We call this as the *Optimal* Scheme. This is an ideal scheme and its performance represents the upper bound that ever can be achieved.

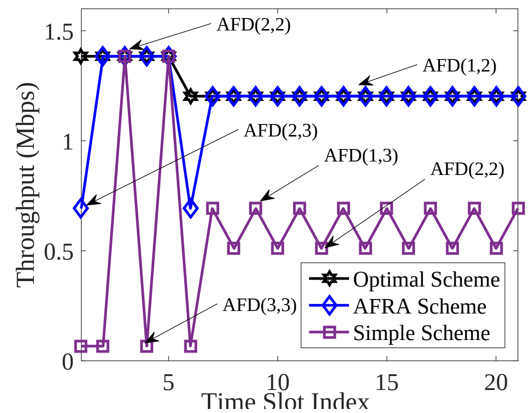

Fig. 5: Network throughput vs. time.

2) AP always operates in the AFD mode. After each successful frame transmission, it increases the rate gradually while it does the opposite after each failed transmission. This process is applied to UL and DL transmissions separately. We call this as the *Simple* Scheme. This scheme mimics the behavior of the classical Ad-hoc rate control schemes mentioned in Section I.
3) The AP always operates in the AFD mode, but with a fixed rate.

B. Sample of Shared TXOP Simulation

We show an example on how our scheme behaves where we plot the throughput sum for uplink and downlink communications versus time during one TXOP period, as shown in Figure 5. During this TXOP the SINR changes, and the 'Optimal' plot shows how an optimal scheme should adapt in response. Our scheme approaches the optimal cases. The true SINR values require an AFD mode with $k_u = k_d = 2$ (i.e., AFD$(2, 2)$) for the first five time slots and AFD$(1, 2)$ for the rest of the TXOP period. Our scheme starts with AFD$(2, 3)$ mode and adapts the MCSs until it converges to the optimal action. On contrast, the 'Simple' scheme starts with AFD$(4, 4)$ that fails and triggers the AP to reduce its rate to AFD$(3, 3)$ and then AFD$(2, 2)$. Although the latest action is successful, the AP should not raise its rate for the next time slot because channel conditions are still not improving. The AP falsely reacts by raising its rate, resulting in an oscillating behavior where it keeps jumping between AFD$(2, 2)$ and AFD$(3, 3)$ actions. This oscillating behavior also happens after the seventh time slot. Our scheme avoids such oscillating behavior because it is designed to react according to the expected future channel conditions, while the 'Simple' Scheme reacts instantly.

C. Performance Comparisons

First, we compare the performance of our AFRA scheme with others operating with AFD mode and fixed rates. In Figure 6, we plot the average Wi-Fi throughput (i.e., sum of UL and DL) versus SINR at the AP receiver (we set the SINRs for UL and DL links to be similar). Fixed rate and mode schemes use always the same MCS index and so they are agnostic to SINR changes. These schemes are optimal when

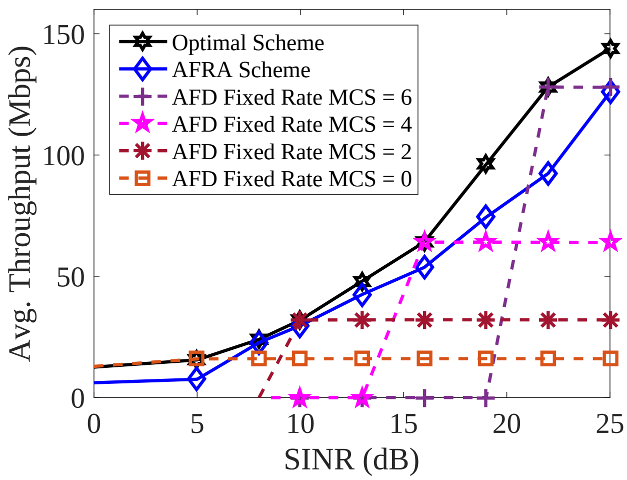

Fig. 6: Average network throughput vs. SINR.

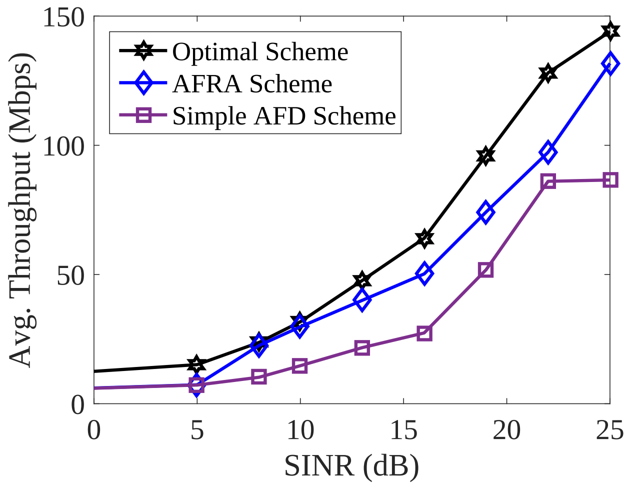

Fig. 7: Average network throughput vs. SINR.

SINR keeps constant, but fails to react to real-time changes. Our scheme adapts to channel changes, and approaches the optimal performance.

We compare our AFRA Scheme against the 'Simple' Scheme in Figure 7. The 'Simple' Scheme starts initially with $k_d = k_u = 4$ MCS indices. Our scheme outperforms the 'Simple' scheme due to the fact that AFRA is based on POMDP, providing it more awareness about the predicted channel changes that might happen in the Future. The 'Simple' Scheme adapts the rate in an Ad-Hoc fashion where it reacts instantly to channel changes, and this might be too conservative or too aggressive in some situations. We report the percentage throughput of the 'Simple' and AFRA Schemes when compared to the 'Optimal' Scheme in Table I. The AFRA Scheme achieves up to 95% of the 'Optimal' Scheme performance, while the 'Simple' Scheme achieves at the best about 67% of it.

TABLE I: Avg. throughput percentage when compared to the 'Optimal' scheme

SINR (dB)	5	8	10	13	16	19	22	25
AFRA (%)	48	95	93	83	78	77	76	92
Simple (%)	47	43	46	45	42	54	67	60

VI. Conclusions

Incorporating SIS techniques into WLANs is a promising solution for boosting the spectrum efficiency, however, it dictates a new perspective for MAC-layer and rate control design. In this paper, we considered asymmetric FD (AFD) communications, where an FD-enabled AP handles simultaneous UL and DL communications with a pair of HD-enabled STAs. We introduced a framework for jointly adapting the communication mode and its transmission rates. Our scheme, AFRA, relies on POMDP and accounts for the future expected environmental changes. AFRA achieves up to 95% of the optimal performance and outperforms classical rate-adaptation schemes. We also introduced a protocol design that facilitates the operation of AFRA and described a mechanism for node-selection that minimizes the inter-node interference between UL and DL links.

References

[1] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 4, pp. 2017–2046, 2015.

[2] LTE-U Forum, "LTE-U SDL coexistence specifications v1.3," , Oct 2015.

[3] W. Afifi, M. J. Abdel-Rahman, M. Krunz, and A. B. MacKenzie, "Full-duplex or half-duplex: A Bayesian game for wireless networks with heterogeneous self-interference cancellation capabilities," *IEEE Transactions on Mobile Computing*, 2017.

[4] S. Biaz and S. Wu, "Rate adaptation algorithms for IEEE 802.11 networks: A survey and comparison," in *Proc. of IEEE ISCC '08 Symposium.*, July 2008, pp. 130–136.

[5] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, "POMDP-based coding rate adaptation for type-I hybrid ARQ systems over fading channels with memory," *IEEE Trans. Wireless Comm.*, vol. 5, no. 12, pp. 3512–3523, December 2006.

[6] I. Koutsopoulos and L. Tassiulas, "Optimal transmission rate control policies in a wireless link under partial state information," *IEEE Trans. on Aut. Cont.*, vol. 55, no. 1, pp. 127–131, Jan 2010.

[7] M. Hirzallah, W. Afifi, and M. Krunz, "Full-duplex-based rate/mode adaptation strategies for Wi-Fi/LTE-U coexistence: A POMDP approach," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 20–29, Jan 2017.

[8] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state markov modeling of fading channels - a survey of principles and applications," *IEEE Sig. Proc. Mag.*, vol. 25, no. 5, pp. 57–80, Sep 2008.

[9] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[10] IEEE, "IEEE–part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications–amendment 4," http://ieeexplore.ieee.org/servlet/opac?punumber=6687185, 2013.

[11] A. Georgiadis, "Gain, phase imbalance, and phase noise effects on error vector magnitude," *IEEE Trans. on Veh. Tech.*, vol. 53, no. 2, pp. 443–449, Mar. 2004.

[12] 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.

[13] G. Shani, J. Pineau, and R. Kaplow, "A survey of point-based POMDP solvers," *Auton Agent Multi-Agent Syst*, pp. 1–51, 2013.

[14] H. Kurniawati, D. Hsu, and W. Lee, "SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces," in *Proc. Robotics: Science and Systems*, 2008.

[15] IEEE, "IEEE–part 11: Wireless LAN MAC and PHY layer specifications - amendment 8: MAC quality of service enhancements," http://tinyurl.com/gs9g3qh, pp. 1–212, Nov 2005.