Adaptive Preamble Embedding with MIMO to Support User-defined Functionalities in WLANs

Zhengguang Zhang, *Student Member, IEEE,* Hanif Rahbari, *Member, IEEE,* Marwan Krunz, *Fellow, IEEE*

Abstract—As the Wi-Fi technology transitions into its sixth generation (Wi-Fi 6), there is a growing consensus on the need to support security and coordination functions at the Physical (PHY) layer. In contrast to the costly approach of extending the PHY-layer header to support new functions (e.g., Spatial Reuse field in the Wi-Fi 6 frame), we propose to turn specific parts of the frame preamble into a reliable data field while maintaining its primary functions. Specifically, in this paper, we develop a scheme called *extensible preamble modulation (eP-Mod)* for 802.11n/ac/ax protocols that are built on multiple-input-multiple-output (MIMO) and orthogonal frequency-division multiplexing (OFDM). For each frame, *eP-Mod* can embed up to 144 user bits into the 802.11ac preamble of an 8×8 MIMO 40 MHz transmission. The proposed scheme is adaptive to channel conditions and enables several promising PHY-layer services, such as PHY-layer encryption and channel/device authentication, and PHY-layer signaling. At the same time, it allows legacy (*eP-Mod*-unaware) devices to continue to process the received preamble as normal by guaranteeing that the proposed preamble waveforms satisfy the structural properties of a standardized preamble. Through numerical analysis, extensive simulations, and hardware experiments, we validate the practicality and reliability of *eP-Mod*.

Index Terms—Preamble embedding, OFDM, MIMO, IEEE 802.11ax, USRP experiments.

1 Introduction

In Wi-Fi protocols, the primary objective of the Physical (PHY) layer is to assist the receiver (Rx) in decoding incoming frames. The Rx may use PHY-layer fields to synchronize with the transmitter (Tx); estimate the Tx-Rx channel state information (CSI); and learn frame duration, rate, and number of parallel spatial streams, among others. However, it has become increasingly evident that the PHY layer needs to support additional functions that facilitate protocol adaption and security. For example, the Tx may need to adjust its spatial reuse factor [1] or adapt its full-duplex mode [2]. It may also need to obfuscate the transmission attributes of a communication to achieve higher secrecy [3], detect a fraudulent channel switching to prevent subsequent man-in-the-middle attacks (e.g., [4]), or identify the Tx in case of PHY-layer encryption [5]. Presently, the rather rigid PHY-layer frame structure prevents Wi-Fi protocols from communicating information needed to support these new functions.

The need for additional PHY-layer functions has been accentuated in part by the surge in wireless attacks. For example, the eye-opening Key Reinstallation Attacks (KRACKs) [6] against WPA2 leverages a channel-based man-in-the-middle [4]. Its underlying channel hijacking attack, along with other similar ones, underscores the insufficiency of existing security measures, and the need for devices to directly authenticate and coordinate with each other at the PHY layer. At the same time, and from an operational perspective, Wi-Fi systems can significantly benefit from a signaling mechanism at the PHY layer. The new features that can be supported with such a mechanism include, but are not limited to, frequency resource allocation for multi-user MIMO (MU-MIMO), coloring of overlapping basic service set (OBSS) for adaptive collision avoidance in Wi-Fi 6 [1], Target Wake Time (TWT) signaling for power saving in resource-constrained (e.g., IoT) devices [7], operation mode advertisement for full-duplex devices (e.g., transmit/receive vs. transmit/sense), and so on.

On the *security* front, existing PHY-layer security measures include friendly jamming [8], [9] for confidentiality and radio frequency (RF) fingerprinting for device authentication. The effectiveness of friendly jamming as a security solution depends on the relative locations of eavesdropping devices or the availability of accurate CSI [10]. RF-based fingerprinting is sensitive to channel impairments and measurement errors [11], [12], making it prone to high false alarm and mis-detection rates. Thus, a more reliable and flexible identifier/signature/nonce exchange mechanism is needed for encryption and channel/device authentication at the PHY layer. On the *operational* front, 802.11ah and 802.11ax currently support TWT scheduling and negotiation through a special trigger frame, which aims at significantly increasing the lifetime of an energy-constrained device by awakening it only when communication is necessary [7]. One can reduce the overhead of transmitting such a trigger frame if the same information in the trigger frame can be embedded/encoded within the PHY-layer fields of a data frame; a philosophy that we advocate in this paper.

In the PHY frame of 802.11 protocols (see Fig. 1), the preamble consists of a few training fields and several signal (SIG) fields, which could potentially be used for signaling. In fact, the IEEE Task Group AX (TGax) has moved part

- *Zhengguang Zhang and Marwan Krunz are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721. E-mail: {zhengguangzhang,krunz}@email.arizona.edu*
- *Hanif Rahbari is with Department of Computing Security, Rochester Institute of Technology, Rochester, NY, 14623. E-mail: rahbari@mail.rit.edu*

An abridged version of this paper appeared in the Proc. of the IEEE INFOCOM 2020 Conference, July 2020.
Manuscript received 11 Nov. 2020; revised 21 May 2021; accepted 25 May 2021.

Fig. 1. Simplified PHY frame format in 802.11a/n/ac/ax standards.

of MAC signaling to the SIG fields of the PHY frame because they are always transmitted with the most robust modulation scheme and can be decoded before the payload is fully received [1]. However, introducing additional SIG fields to support new functions is a costly approach because of the rigid structure of the PHY frame and the fact that SIG fields are transmitted at the lowest rate, so each new SIG field extends the frame duration by a nonnegligible amount.

1.1 Proposed Approach

In this paper, we explore a novel approach that allows devices to embed user-defined bits within (a subset of) the training fields of the preamble waveform. The broadcast nature of wireless channels makes the unencrypted training fields heard by all neighboring devices, which is ideal for signaling. The embedded bits can be used to support new features/functions at the PHY layer. We focus on the latest MIMO-based 802.11 protocols (i.e., 802.11n/ac/ax) and make use of mandatory training fields of the frame preamble that are common across all MIMO-based Wi-Fi devices. The information embedded in the preamble can be decoded in a timely manner before decoding the payload, so it is not impacted by an encrypted or corrupted payload. By repurposing the selected training fields into a generic data field that can be used for a wide range of signalling purposes, we avoid the overhead associated with introducing new SIG fields and, more importantly, maintain backward compatibility with existing Wi-Fi standards.

Our proposed MIMO-based preamble embedding technique is called *extensible preamble modulation* (*eP-Mod*). At the Tx side, *eP-Mod* leverages multiple antennas to transmit several jointly-designed variants of the preamble waveform. These variants satisfy several requirements related to standardized preamble's functions, including frame detection, time and frequency synchronization, and CSI estimation. On the Rx side, a reverse mechanism, called *eP-Demod*, treats the training fields of the preamble as a special data field and efficiently combines the multiple received waveforms to reconstruct the particular variant of the preamble waveform that carries user-specified information. Our *eP-Mod* design provides flexibility to support new services at the PHY layer and offers *extensibility* to different MIMO schemes, channel widths and MIMO-based 802.11 protocols.

1.2 Example Applications of *eP-Mod*

eP-Mod is intended to support PHY-layer signaling and security applications in wireless local area networks (WLANs). Fig. 2 illustrates three example applications of *eP-Mod*. In these examples, the embedded bits in the preamble are split into the *type* bits that indicate what application it is used for, and *value* bits that provides the supportive information for that application.

First, we focus on its application for OBSS Coloring. Consider a dense WLAN, where two OBSSs are centered at access points (AP) AP_1 and AP_2. The two APs operate on the same frequency due to limited spectrum. This often

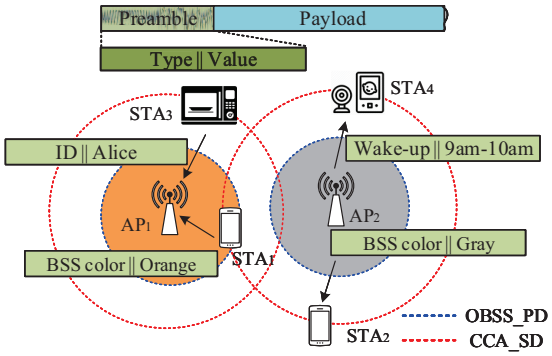


Fig. 2. Example applications of *eP-Mod*.

causes network congestion and inefficiency (e.g., exposed terminal problem). To address this issue, the 802.11ax standard proposes to identify these two OBSSs by coloring, such that a station (STA) can ignore signals from neighbors with different BSS colors, whenever possible, although they operate on the same channel. The BSS color is a 6-bits identifier that is currently indicated in the HE-SIGA field of the 802.11ax preamble. Signaling the BSS color via *eP-Mod* is advantageous over HE-SIGA field as discussed in Section 1.1.

Under *eP-Mod*, APs and STAs indicate their BSS color in the preamble. For detected OBSS transmissions, the clear channel assessment signal detection (CCA_SD) threshold is increased from -82 dBm to OBSS_PD (in the range of $(-82, -62]$ dBm). The STA sets channel to busy if and only if the detected BSS color matches its own BSS color and the received signal strength (RSS) is greater than OBSS_PD. When AP_2 tries to send a packet to STA_2, STA_1 hears the ongoing transmission as the RSS is beyond the CCA_SD. Then, STA_1 runs preamble demodulation and the decoded BSS color is gray, which does not match its own BSS color. Since the RSS is below OBSS_PD, STA_1 will ignore this packet and go ahead to send its own packet to AP_1.

We also briefly describe two other applications of *eP-Mod*: (1) To mitigate contention between STAs and save the power of a baby monitor (STA_4), AP_2 embeds wake-up schedule for the baby monitor; (2) *eP-Mod* can also help with authentication by indicating a device's cryptographic ID in the preamble, as shown for STA_3.

1.3 Contributions

The main contributions of this paper are as follows:

1) We redesign the PHY frame preamble of MIMO-based 802.11n/ac/ax systems, while maintaining backward-compatibility with these systems. The proposed design allows reliable embedding of user-defined bits to support important PHY-layer applications;
2) Considering various MIMO channel conditions, we show our proposed *eP-Mod* is able to adapt its reliability (via diversity gain techniques) and capacity (via multiplexing gain techniques) to channel conditions. Specifically, we extend space-time diversity and space-time multiplexing techniques to further take advantage of two separate training fields of the preamble and achieve an *embedding throughput* of 15 bits per frame at low signal-to-noise ratio (SNR) (≤ 5 dB) and 40 bits per frame with high SNR (≥ 30 dB), respectively, in the

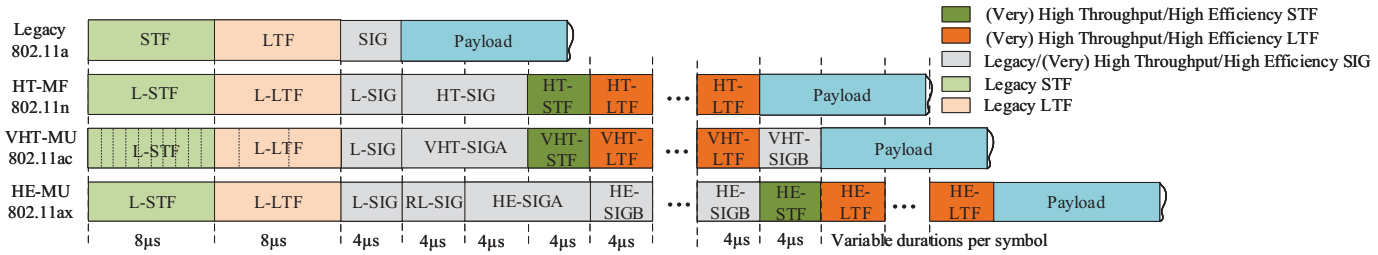


Fig. 3. Preambles in the PHY frame of 802.11a/n/ac/ax standards.

2×2 MIMO scenarios over a 40 MHz channel. In an 8×8 MIMO 40 MHz transmission, *eP-Mod* further achieves an embedding capacity of up to 144 bits per frame.

3) Besides theoretical analysis and simulation-based evaluation of *eP-Mod* on an 802.11ac MIMO system, we implement *eP-Mod* for 1×2, 2×1, and 2×2 MIMO scenarios on a USRP testbed. Our results demonstrate the feasibility, extensibility, and practicality of *eP-Mod* in achieving a high preamble decode ratio (PDR) ≥ 90 percent with $16\sim40$ embedded bits per frame, while maintaining comparable frame detection, synchronization, and payload bit error rate (BER) performance as the standardized preamble.

Paper Organization—We go over the preamble structure, generation, and functions in Section 2. The system architecture with integrated *eP-Mod* modules and its core idea are presented in Section 3. In Sections 4 and 5, we explore and analyze applying various MIMO techniques to *eP-Mod* along with its adaptability to channel status. Extension of *eP-Mod* to higher bandwidths as well as its computational complexity, are discussed in Section 6. We evaluates the performance of *eP-Mod* in Section 7 via simulations and experiments. We review related work in Section 8, before concluding in Section 9 with a discussion on future work.

2 Frame Preamble – A Primer

In WLANs, the preamble is the first part of the PHY frame. The preambles of MIMO-OFDM-based 802.11n/ac/ax systems (see Fig. 3) start with a copy of the legacy 802.11a preamble for backward compatibility, followed by additional short training fields (STFs), long training fields (LTFs), and SIG. In here, we take 802.11ac as the representative example of MIMO-OFDM-based 802.11 systems, and study its preamble structure, generation, and functions[1].

Important Notations—Table. 1 summarizes the most important notations.

2.1 IEEE 802.11ac Preamble

As shown in Fig. 3, the legacy part of the 802.11ac preamble consists of a legacy STF (L-STF) and a legacy LTF (L-LTF) field. The L-STF waveform is constructed by transmitting only on subcarriers that are four subcarrier-spacings apart from each other, resulting in a periodic OFDM waveform, with each period called the short training signal (STS). So within the 8 μs duration of the L-STF, there are 10 repetitive STSs. Yet the L-LTF is periodic with two OFDM symbols.

1. To simplify the exposition, in this paper, we refer to the preamble as the training fields.

The legacy STS is repeated five times in the very high throughput (VHT) preamble to build the VHT-STF. While the VHT-LTF waveform is slightly different from L-LTF and is mapped to multiple VHT-LTFs by an orthogonal matrix. For different MIMO transmit chains, different cyclic shifts are applied to the legacy and VHT preambles to create *cyclic shift diversity* (CSD), which prevents unintentional beamforming [13]. For channels wider than 20 MHz (40, 80, and 160 MHz), their preambles are essentially replicas of that of the base 20 MHz channel but with a subcarrier phase rotation on each additional 20 MHz channel. This rotation is necessary to reduce the peak-to-average power ratio (PAPR) of OFDM symbols [14], [15].

As an illustrating example, consider an IEEE 802.11ac frame with a 40 MHz channel. The L-STF occupies 24 out of 128 available subcarriers, indexed by $4k$, where $k\in\Lambda=\{-14,-13,\ldots,-9,-7,-6,\ldots,-2,2,3,\ldots,7,9,10,\ldots,14\}$. Let $\mathcal{S}=\{S_k\}_{k\in\Lambda}$ be the sequence of symbols carried in these 24 subcarriers. The L-STF signal for the ith Tx antenna at time t, $0\leq t<8$ μs, is:

$$s^{(i)}(t)=\sqrt{\frac{128}{24\,N_t}}\sum_{k\in\Lambda}\gamma_{4k}S_k\exp\left(j2\pi4k\Delta_F(t-T^i_{lcs})\right), \tag{1}$$

where N_t is the number of Tx antennas, γ_{4k} represents the subcarrier phase rotation at subcarrier $4k$, $\Delta_F=312.5$ kHz is the subcarrier spacing, and T^i_{lcs} is the legacy cyclic shift on the ith Tx antenna [13, Eq.(22-20)]. As mentioned before, The VHT-STF signal for the ith Tx antenna is almost the same as $s^{(i)}(t)$, but with a VHT cyclic shift T^i_{hcs}.

2.2 Basic Preamble Functions in 802.11ac Systems

The training fields assist the Rx in performing the following functions, which should be protected under any modification to the preamble:

TABLE 1
Important notations.

$\mathcal{S}$	STF symbol sequence for 40 MHz channel.
S_k	STF symbol on $4k$th subcarrier, i.e., an element of $\mathcal{S}$.
$S_{n,k}$	STF symbol on $4k$th subcarrier of nth Tx antenna.
$\tilde{S}_l$	Symbol sequence of lth received STS.
θ_i	Wrapped phase differences of the successive S_k's.
Θ	*Dependency pattern* of S_k's, i.e., a sequence of θ_i.
t_{cs}	Cyclic time shift of STF imposed by *eP-Mod*.
ν	Change of θ_i caused by cyclic time shift.
$\Delta\varphi$	Phase shift of S_k imposed by *eP-Mod*.
$\Theta^{(\nu)}$	*Child pattern* of S_k's for *eP-Mod* STF with cyclic time shift.
Q	Order of Q-DPSK symbol to embed *Q-Seq*.
M	Order of MPSK symbol to embed *M-Seq*.

2.2.1 Frame Detection and Time Synchronization

The repetitive structure of the L-STF (or the superposition of multiple received L-STFs in the case of MIMO) is used by the Rx to detect the start of a frame and time-synchronize with the Tx. Because the channel coherence time in WLANs (discussed in Section 2.2.3) is larger than the preamble duration, the channel response does not change during the transmission of the preamble, meaning that the Rx can still detect the repetitive structure of the L-STF using autocorrelation methods. Note that this method does not require knowledge of the transmitted L-STF waveform.

For fine-tuned time synchronization, however, the Rx must calculate the cross-correlation between the transmitted L-LTF and the received L-LTF. This method requires full knowledge of the L-LTF signal, but is usually more accurate.

2.2.2 Frequency Synchronization

Similar to time synchronization, carrier frequency offset (CFO) estimation can also be performed using the periodic L-STF and L-LTF based on an autocorrelation method. Assuming perfect symbol timing and a stationary channel, a CFO of δ_f results in a phase difference of $2\pi t\delta_f$ between two repetitions of an STS that are transmitted t seconds apart. CFO can be indeed estimated by measuring such phase differences, averaged over a period of time. Since the period of an L-STF is shorter than that of the L-LTF, the Rx can achieve higher accuracy by using the L-LTF to counter noise effects. Again, the Rx does not require knowledge of the transmitted L-STF waveform for CFO estimation.

2.2.3 CSI Estimation

L-LTF is also used for single-input-single-output (SISO) channel estimation, whereas VHT-LTFs are utilized for MIMO channel estimation and sounding. In the frequency domain, SISO channel estimation on each subcarrier is done by simply dividing every subcarrier of the received L-LTF or VHT-LTF by its counterpart in the transmitted signals. However, CSI estimation in MIMO must leverage the orthogonal mapping matrix $\mathbf{P}$ specified in 802.11ac [13, Eq. 22-43] since the uth VHT-LTF on the ith Tx antenna, denoted as $r^{(i,u)}(t)$, is pre-multiplied by the (i,u) element of $\mathbf{P}$, denoted by $P_{i,u}$. More specifically, $\mathbf{P}$ and the symbol sequence $\mathcal{V} = \{V_k\}, k \in \mathbb{Z}, -58 \leq k \leq 58$ [13, Eq. 22-37] are used for VHT-LTF generation shown in (2). The Rx takes advantage of the orthogonality of $\mathbf{P}$ and multiplies its Hermitian transpose by the received VHT-LTF signals before dividing $\mathcal{V}$ to obtain the CSI estimation. Here, the cyclic shift T^i_{hcs} and the subcarrier phase rotation γ_k are left with the actual frequency domain channel. So, the estimated CSI incorporates the impact of T^i_{hcs} and γ_k. Since the same T^i_{hcs} and γ_k are also applied to the data symbols, this approach automatically removes their effect when equalizing the data with the estimated channel.

$$r^{(i,u)}(t) = \sqrt{\frac{64}{57Nt}} \sum_k \gamma_k P_{i,u} V_k \exp\left(j2\pi k \Delta_F (t - T^i_{hcs})\right). \tag{2}$$

We further note that because of mobility and Doppler spread, channel may be time-varying. To check the channel coherence time T_c, we use the following approximation:

$$T_c = \frac{1}{16\pi a_0 f_c / c}, \tag{3}$$

where c is the speed of light, f_c is the operating frequency, and a_0 is the device speed. We note that device mobility in 802.11ac WLANs almost never exceeds $a_0 = 108$ km/h$= 30$ m/s. The resulting coherence time at this speed is $83\,\mu$s and $40\,\mu$s for 2.4 GHz and 5 GHz bands, respectively. It means that the preamble is almost always within the coherence time and the same CSI estimate can be safely used throughout the preamble duration.

2.3 Preamble Design Criteria

Above all, we can redesign only the STF waveforms for embedding bits, because knowledge of the transmitted STF is not necessary at the Rx. On the other hand, LTFs must be known for CSI estimation, and hence cannot be used for embedding user-defined bits. Given the criticality of the aforementioned preamble functions for correct frame decoding, any change in the waveform of the STFs should maintain the properties required to support these functions. As such, the following restrictions must be imposed on our design: (1) The repetitive structure with the exact same period and duration of the standardized STFs must be maintained to facilitate time and frequency synchronization; (2) different cyclic shifts at different antennas must be used to prevent unintentional beamforming; (3) the same (or better) dynamic range (DR) of the IEEE 802.11 standards (7.01 dB) must be achieved for fast automatic gain control (AGC) locking and A/D conversion without overflow/underflow; and (4) low PAPR (≤ 2.24 dB) must be ensured to accommodate the nonlinearities of typical power amplifiers [16].

3 *eP-Mod* Core Design

We are now ready to present the core design of *eP-Mod* by showing how it modulates and demodulates a given bit sequence through *a* redesigned STF waveform, which can be either L-STF or VHT-STF as the generation of these two waveforms is almost the same. For readability, we refer to them as STF, unless mentioned otherwise, but note that *eP-Mod* is applied to both of them simultaneously.

3.1 System Architecture

We consider a MIMO system where multiple Tx antennas modulate *eP-Mod* bits, either independently or collaboratively, in their STFs. For each transceiver antenna, *eP-Mod* and *eP-Demod* are two key modules that can be optionally added to the traditional transmit and receive chains, respectively, as shown in Fig. 4. In the *eP-Mod* module, the user bit sequence for each STF is used to intelligently generate $\mathcal{S}$, which then goes through IFFT, CSD, and CP concatenation to obtain one *eP-Mod* STF. Substituting default STFs, dynamically generated *eP-Mod* STFs are assembled into the frame preamble together with the default LTFs (L-LTF and VHT-LTFs). The Rx has to treat the synchronized STFs as special data fields, run FFT, and equalize them using the estimated CSI before performing *eP-Demod* (see Section 3.3). Note that the default LTFs are kept intact to ensure accurate CSI estimation. As long as the waveform that carries the modulated bits satisfies the target properties discussed in Section 2.3, the primary preamble functions needed for normal data transmission are maintained.

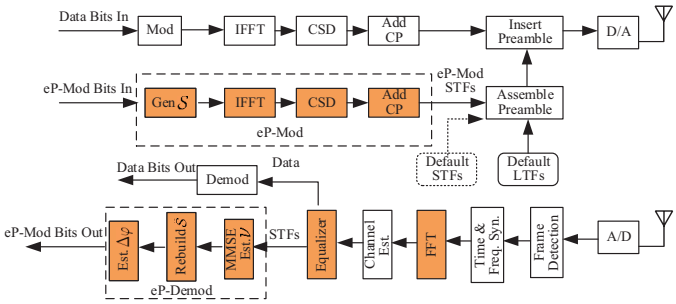


Fig. 4. Tx and Rx system architecture with integrated *eP-Mod* modules (colored blocks indicate new or modified components).

3.2 Preamble Modulation for a Single STF

We start by explaining the generation of a single STF. We propose to generate variants of the STF waveform with distinctive characteristics, each for modulating a different bit sequence. Conventionally, a bit sequence is punctuated and modulated into a plurality of symbols. However, the strict requirements for STF design restrict us to treating the entire waveform as one single modulation symbol. Specifically, we need to maintain the periodicity, the target PAPR, and the DR of the STF. The most reliable way to do that is to maintain its amplitude through time and/or phase shift operations of the entire STF waveform in the time domain as a means to generate variants of this waveform and modulate input bit sequences. Based on this idea, we design a low complexity modulation scheme for a STF of the 40 MHz 802.11ac channel, which can also be readily applied to any bandwidth in any OFDM-based 802.11 system.

First, we consider the time shift operation. Suppose the STF waveform is to be cyclically shifted by t_{cs}. Plugging this shift into (1), we end up with a linear phase shift of $2\pi 4k\Delta_F t_{cs}$ along the S_k. Let $\theta_i, i = 1, \ldots, 23$, be the (wrapped) phase differences of the successive S_k's (for example, $\theta_1 = \measuredangle(S_{-13}) - \measuredangle(S_{-14})$) and define $\Theta = [\theta_1, \ldots, \theta_{23}]$ as the symbols *dependency pattern* for an STF waveform. Denote the dependency pattern for a standardized STF as the *parent pattern* $\Theta^{(0)} = \left[\theta_1^{(0)}, \ldots, \theta_{23}^{(0)}\right]$ (see Table 2). Under a linear phase shift, all θ_i's will change by $\nu = 2\pi 4\Delta_F t_{cs}, \nu \in [-\pi, \pi]$, from $\theta_i^{(0)}$'s. Thus, we generate other compliant variants of STFs by generating *child patterns* $\Theta^{(\nu)}$, whose elements are determined by:

$$\theta_i^{(\nu)} = \begin{cases} \theta_i^{(0)} + 4\nu, \text{if } i = 12 \\ \theta_i^{(0)} + 2\nu, \text{if } i = 6, 18 \\ \theta_i^{(0)} + \nu, \text{otherwise} \end{cases} \tag{4}$$

The changes by a multiple of ν in (4) are due to null tones between successive S_k's in $\mathcal{S}$. For instance, S_8 is a null tone, the linear phase shift causes additional 2ν phase differences between S_7 and S_9, so $\theta_{18}^{(\nu)} = \theta_{18}^{(0)} + 2\nu$.

The notion of dependency pattern allows us to alternatively represent $\mathcal{S}$ solely by its first symbol S_{-14} and the associated $\Theta^{(\nu)}$. The rest of the sequence $\mathcal{S}$ can be derived recursively: $S_{-13} = S_{-14}\exp(j\theta_1^{(\nu)})$, $S_{-12} = S_{-13}\exp(j\theta_2^{(\nu)})$, and so on. In general, we embed $\log_2 Q$ user bits, called a *Q-Seq*, by setting $\nu = 2\pi q/Q, q \in \{0, 1, \ldots, Q-1\}$ to get Q time-shifted versions of STF waveform. An example of using four different ν's, corresponding to four time shifts, to embed two bits $(b_2 b_1)$ is shown in Table 3. The bits are Gray-coded based on ν.

TABLE 2
Part of the dependency pattern $\Theta^{(0)}$ of the standard $\mathcal{S}$.

i	8	9	10	11	12	13	14	15	16
$\theta_i^{(0)}$	π	0	0	0	0	π	π	π	0

TABLE 3
Dependency patterns and user-embedded bits when $Q = 4$.

t_{cs} (µs)	0	0.2	0.4	0.6
ν	0	$\pi/2$	π	$-\pi/2$
$b_2 b_1$	00	01	11	10

The standard $\mathcal{S}$ uses the dependency pattern $\Theta^{(0)}$, i.e., $\nu = 0$ and $\Delta\varphi = 0$ when $S_{-14} = \sqrt{1/2 + j/2}$.

In addition to time-shift operation, we rotate the constellation of each S_k by a common phase offset $\Delta\varphi$. To do this, we leverage the propagation of S_{-14}'s phase change because of the dependency pattern. First, we select a M-PSK-modulated S_{-14} with an amplitude of $\sqrt{2}$ and a phase of $\Delta\varphi$. Then, following one of the patterns $\Theta^{(\nu)}$ with the chosen S_{-14}, the phase shift $\Delta\varphi$ is propagated to the rest of the symbols in $\mathcal{S}$. If we make full use of all M symbols of the M-PSK scheme, $\log_2 M$ Gray-coded bits, referred to as *M-Seq*, can be embedded into the STF waveform.

Embeding $\log_2 Q$ bits using different ν's under the same $\Delta\varphi$ is equivalent to a form of frequency-domain differential PSK (FD-DPSK), where ν is encoded into $\theta_i^{(\nu)}, i = 1, \ldots, 23$. In contrast, $\Delta\varphi$ is encoded into the phases shift of all 24 symbols in $\mathcal{S}$. Thus, the former approach is more robust to channel phasor and CFO estimation errors than the latter, especially at high order of M and Q. One can utilize one or a combination of these two techniques. When only ν is used to embed bits, we call this "*eP-Mod(Q*)*" to distinguish it from "*eP-Mod*" that combines the two techniques. In general, we can embed $L = \log_2 M + \log_2 Q$ bits using the STF waveform, where time and phase shift are represented by Q-DPSK and M-PSK symbols, respectively. An example of combined scheme for preamble modulation is depicted in Fig. 5.

$$\big[\underbrace{b_{\log_2 MQ}, \ldots, b_{1+\log_2 M}}_{\text{Q-Seq}}, \underbrace{b_{\log_2 M}, \ldots, b_1}_{\text{M-Seq}}\big]. \tag{5}$$

Now that we have the frequency-domain sequence for the STF, the STF waveform can be generated by (1). In Fig. 6, we show the amplitudes and phases of two different *eP-Mod* STFs of the same dependency pattern with parameter $t_{cs} = 0.2$ µs, but different phase shift $\Delta\varphi = \pi, \pi/2$ compared to the default STF. The 0.2 µs cyclic time shift is obvious in Fig. 6(a) and the same dependency pattern results in the same envelope. However, the phases of STFs are impacted both by the pattern and phase shift, and their relationship are not explicit as seen in Fig. 6(b).

The resulting STF waveform under *eP-Mod* maintains the same periodicity (0.8 µs), DR (7.01 dB), PAPR (2.24 dB), and other features of the standardized STF waveform including standard cyclic shift and subcarrier phase rotation. Therefore, *eP-Mod* is backward compatible and *eP-Mod*-enabled devices can interoperate with those that do not support it.

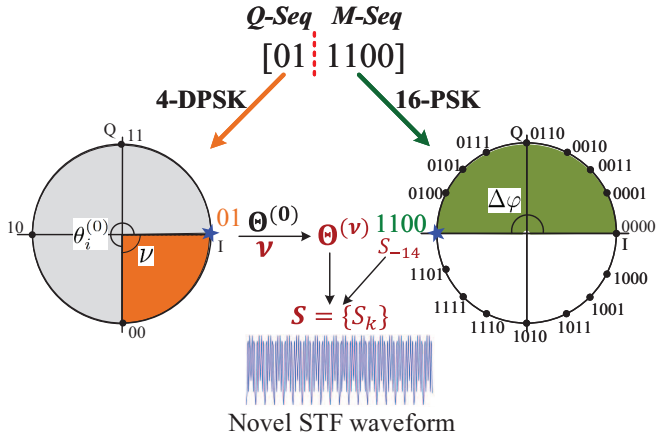


Fig. 5. Example of combined time and phase shift scheme for *eP-Mod*. First, *Q-Seq* 01 for $t_{cs} = 0.2\,\mu s$ is mapped to a 4-DPSK symbol which gives a phase difference of $\nu = \pi/2$ for the generation of dependency pattern $\Theta^{(\nu)}$ from $\theta_i^{(0)}$. *M-Seq* 1100 is mapped to a 16-PSK symbol which is set as the first symbol S_{-14}, whose phase corresponds to $\Delta\varphi = \pi$. Then, all S_k could be generated by S_{-14} with the dependency pattern $\Theta^{(\nu)}$. The resulting STF is equivalent to cyclically shifted standardized STF waveform by $0.2\,\mu s$ and phase-shifted by π.

3.3 Preamble Demodulation

The preamble demodulation process at the Rx side is basically the inverse of the modulation process.

Pattern Detection. As shown in Fig. 4, *eP-DeMod* starts by detecting the dependency pattern identified by ν. Let $\tilde{\mathcal{S}}_l = \{\tilde{S}_{l,k}\} : k \in \Lambda$, and $l \in \{1, \ldots, 9\}$ denote the symbol sequence of the lth received STS, whose dependency pattern is Θ_l, and let $\tilde{\nu}$ be the estimate of ν based on all the nine[2] $\tilde{\mathcal{S}}_l$'s. Minimum mean-square error (MMSE) is used to estimate ν with respect to the parent pattern $\Theta^{(0)}$.

$$\tilde{\nu} = \underset{\substack{\nu=2\pi q/Q,\\ q\in\{0,1,\ldots,Q-1\}}}{\arg\min} \sum_{l=1}^{9} \|\Theta_l - \Theta^{(\nu)}\|^2. \tag{6}$$

As the pattern reflects a time shift, any timing error will add to ν, so the estimation of ν requires accurate frame detection and timing.

Phase Shift Estimation. Next, a reference symbol sequence $\tilde{\mathcal{S}} = \{\tilde{S}_k\}_{k\in\Lambda}$ is constructed based on $\Theta^{(\tilde{\nu})}$ for the sake of estimating $\Delta\varphi$, say $\Delta\tilde{\varphi}$. Thereafter, the Rx calculates the phase shifts from $\tilde{S}_k$ to corresponding $\tilde{S}_{l,k}$ for all $k \in \Lambda$, and $l = 1, \ldots, 9$. Owing to the impact of noise and imperfect CSI and CFO estimation, the Rx sums $24l$ ratios, as shown below, to get a more accurate $\Delta\tilde{\varphi}$:

$$\Delta\tilde{\varphi} = \measuredangle \sum_{l=1}^{9} \sum_{k\in\Lambda} (\tilde{S}_{l,k}/\tilde{S}_k). \tag{7}$$

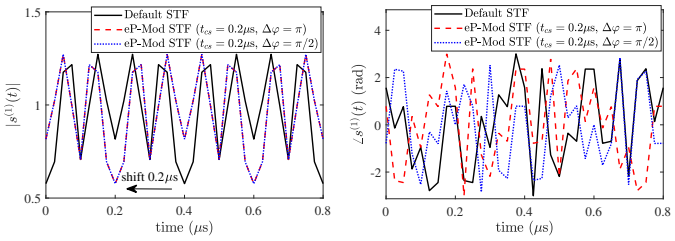
The estimation accuracy first depends on pattern detection and then CFO and CSI estimation accuracy (because it is sensitive to any phase errors).

Finally, the embedded *Q-Seq* and *M-Seq* are decoded from $\tilde{\nu}$ and $\Delta\tilde{\varphi}$, respectively.

4 *eP-Mod* WITH DIVERSITY GAIN

With the building block of *eP-Mod* for a single STF, we now explain our proposed *eP-Mod* for MIMO systems. In this section, we apply MIMO techniques to improve the reliability of *eP-Mod* by exploiting diversity gain. We first consider two special MIMO cases, 1×2 SIMO and 2×1 MISO, to illustrate how we adapt Rx-side maximal ratio combining (MRC) and equal gain combing (EGC), and also Tx-side Alamouti space-time coding to *eP-Mod*, as opposed to how they are applied to payload data. We will also discuss the trade-off between reliability and "capacity", defined as the number of embedded bits per frame. As we will explain later, reliability cannot be studied independently from capacity in *eP-Mod*. Finally, we show *eP-Mod* variants with high diversity gains in general MIMO cases.

2. The first STS $\tilde{\mathcal{S}}_1$ may be distorted by multipath effect, pulse shaping and other RF impairments, so we discard it during detection.



(a) Amplitude of the STFs (b) Phase of the STFs

Fig. 6. Amplitude and phase of one default STS vs. two *eP-Mod* variants.

4.1 SIMO Diversity Gain under *eP-Mod*

4.1.1 *eP-Demod with Rx-side Combining*

In a SIMO system, the Tx generates an *eP-Mod* L-STF and VHT-STF pair as described in Section 3.2. Once the Rx synchronizes with the Tx in time and frequency, it starts processing the STFs received on each Rx antenna. The Rx does not process received STFs as a set of $4\,\mu s$-long OFDM symbols like it typically does to payload data [17], but rather processes each $0.8\,\mu s$ STS within the STFs independently. One benefit of such an approach is that the required lower-order FFT and matrix multiplication for equalizing STSs are computationally lighter. As the frequency-domain symbols $\tilde{S}_{l,k}$'s obtained by FFT correspond to the S_k's transmitted on $4k$th subcarriers, only the estimated CSI components on these subcarriers are needed to equalize the STSs.

The Rx first equalizes and then combines the STSs received on different antennas using the estimated CSI (see Section 2.2.3) to achieve diversity gain. This improves the effective SNR. We adapt two diversity schemes, MRC and EGC, to integrate with *eP-Mod*. Assuming a 1×2 SIMO system operating on i.i.d. Rayleigh fading channels, at the lth STSs of two Rx antennas, the transmitted S_k is received as $\tilde{\mathbf{S}}_{l,k} \in \mathbb{C}^{2\times1}$,

$$\tilde{\mathbf{S}}_{l,k} = \mathbf{h}_k S_k + \mathbf{z}_k, \tag{8}$$

where $\mathbf{z}_k \sim \mathcal{CN}(0,1)$ is zero mean circularly symmetric complex Gaussian noise vector, and $\mathbf{h}_k$ is the channel vector of the $4k$th subcarrier (recall that S_k is sparse in frequency domain). To get the symbol for *eP-Demod*, MRC and EGC can be applied to get a weighted sum of the elements in $\tilde{\mathbf{S}}_{l,k}$. The combining weight $\mathbf{w}_k = [w_{1,k}\ w_{2,k}]^T$ for EGC and MRC are: $\mathbf{w}_k^{EGC} = e^{-j\mathbf{\Phi}_k}$ and $\mathbf{w}_k^{MRC} = \mathbf{h}_k^*$, respectively, where $\mathbf{\Phi}_k = \measuredangle\mathbf{h}_k$, and $(\cdot)^*$ is the conjugate operator. Multiplying $\mathbf{w}_k$ on both sides of (8) returns the equalized version of S_k. Applying MRC or EGC for each STS over all the indices

TABLE 4
Effective E_s/N_0 gain for Q-DPSK and M-PSK symbols in 1×2 SIMO *eP-Mod* over a BPSK symbol without Rx-side combining.

	MRC	EGC
Q-Seq	$\frac{16}{3} \times \frac{9}{4} \times \frac{23}{2} \times 2 = 276$	$\frac{16}{3} \times \frac{9}{4} \times \frac{23}{2} \times 1.79 = 247$
M-Seq	$\frac{16}{3} \times \frac{9}{4} \times 24 \times 2 = 576$	$\frac{16}{3} \times \frac{9}{4} \times 24 \times 1.79 = 514$

k, the Rx eventually gets multiple equalized STSs that are ready to be demodulated and decoded as in Section 3.3.

4.1.2 Capacity and Reliability Analysis

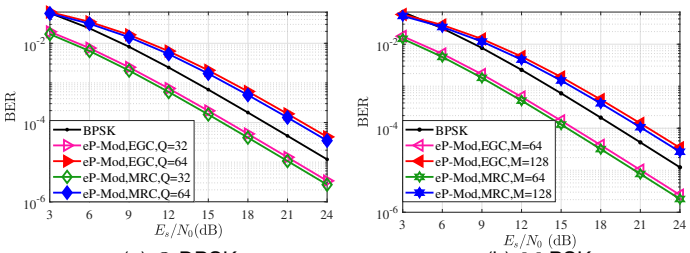
Because the number of embedded bits (capacity) determines Q and M, and in turn, the orders of DPSK and PSK, respectively, increasing it will automatically decrease the demodulation performance (and vice versa). We consider the BER performance of the SIG fields, which always use BPSK, as our reliability benchmark to study the capacity of *eP-Mod*.

First, we note that compared to a BPSK symbol on one of the subcarriers in the SIG or payload, Q-DPSK and M-PSK symbols in *eP-Mod* enjoy a higher energy-per-symbol to noise-power-spectral-density (E_s/N_0) under the same OFDM symbol SNR. This is due to the following factors:

1) The amplitude of each S_k in L-STF is $\sqrt{128/24}$ times higher than a BPSK symbol on a SIG/payload subcarrier—see (1), resulting in a gain of $128/24 = 16/3$ for both Q-DPSK and M-PSK symbols;
2) To estimate ν, *eP-Mod* uses the patterns over 9 STSs and 23 mutually dependent differential phases among 24 subcarriers within each STS to average noise out, where each STS lasts for $1/4$ of the duration of a typical OFDM symbol. That brings about an additional gain of $9/4 \times 23/2$ for Q-DPSK symbols;
3) Similarly, the M-PSK symbol is estimated using 9 STSs and 24 symbols S_k within each STS—see (7), which add up to an additional gain of $9/4 \times 24$.

Both *eP-Mod* and BPSK symbols have an array gain of 2 and $1 + \pi/4$ under 1×2 MRC and EGC, respectively [18, Eq.1.33, 1.35]. The effective E_s/N_0 gain for Q-DPSK and M-PSK symbols is summarized in Table 4. To approximate the resulting BER of Q-DPSK and MPSK under fading channels and search sequentially to find the maximum M and Q that achieve a comparable BER to that of our benchmark, we use the MATLAB function *berfading* because there is no closed-form expression for their BER under such channels. Here, we assume accurate time and frequency synchronization, and perfect CSI estimation to derive the maximum M and Q.

In Fig. 7(a), we contrast the BER of an uncoded Q-DPSK symbol of *eP-Mod* to that of uncoded BPSK, as a function of E_s/N_0. It can be observed that the Q-DPSK symbol with $Q = 32$ (5 bits) has lower BER than BPSK. We also compare the performance when $M = 64$ and $M = 128$ (see Fig. 7(b)). In this case, $M = 64$ (6 bits) guarantees lower M-PSK symbol BER than BPSK, and MRC slightly outperforms EGC. In summary, it is feasible to embed a total of $5+6 = 11$ bits in the L-STF of 1×2 SIMO systems while achieving the same (or better) reliability as BPSK. Additionally, we embed a second sequence of $4 + 5 = 9$ bits in the VHT-STF with half of the effective gain of the L-STF because its duration is



(a) Q-DPSK (b) M-PSK

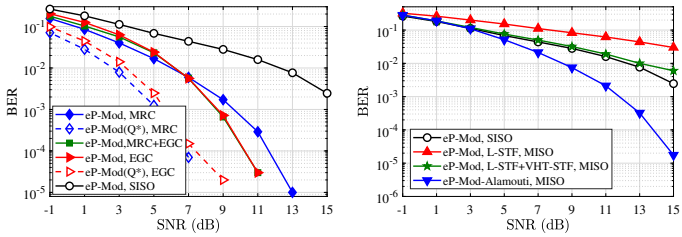
Fig. 7. BER of uncoded BPSK and *eP-Mod* for L-STF with different combining schemes vs. E_s/N_0 of the payload subcarrier .

half of the L-STF. Therefore, we manage to embed a total of 20 bits in the preamble of 802.11ac 1×2 SIMO systems.

However, in practical systems, the Rx does not have accurate CSI but instead uses estimated channel response $\tilde{\mathbf{h}}_k$ for MRC and EGC. Fig. 8(a) compares the BER of *eP-Mod* with MRC and/or EGC against OFDM-symbol SNR in the simulations under Rayleigh multipath channel. Most notably, the benefit of diversity gain that *eP-Mod* obtaines in SIMO is significant when compared with SISO scenario. Due to lack of perfect CSI, only 10 bits can be embedded in the L-STF to beat the BER of BPSK. Interestingly, MRC performs better than EGC at low SNRs, but falls behind EGC at high SNRs. Meanwhile, MRC always outperforms EGC for demodulating *Q-Seq*. This implies that for demodulating *M-Seq*, EGC often performs better than MRC. It is attributed to the fact that MRC achieves higher effective SNR that helps a lot at low SNR regime, but when it comes to high SNR regime, the weaker signal branch is reliable enough to contribute to the phase of combined signal. However, its contribution is crippled by the stronger signal branch that has higher weight in the combined signal. Such impairments impact the absolute phase much more than the phase difference. Therefore, we propose to use (1) MRC for the detection of dependency pattern corresponding to *Q-Seq* and (2) EGC for the phase-shift detection corresponding to *M-Seq*.

4.2 MISO Diversity Gain under *eP-Mod*

In the previous section, we considered a single-antenna Tx configuration that modulates user bits in the L-STF and VHT-STF independently. Next, we consider a 2×1 MISO system and explore how in the absence of any CSI feedback, one can use space-time block codes (STBC) to achieve Tx-side diversity gain under *eP-Mod*.



(a) SIMO vs. SISO. (b) MISO vs. SISO.

Fig. 8. BER of *eP-Mod* with diversity gain, $Q = M = 32$, Rayleigh channel simulation.

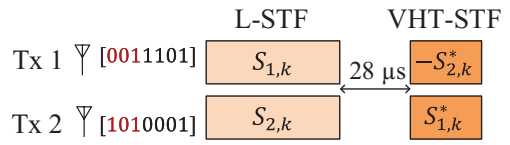


Fig. 9. *eP-Mod*-Alamouti scheme.

4.2.1 Challenges and Proposed Approach

The simplest form of STBCs is Alamouti code. For a two-transmit-antenna system, Alamouti code generates two orthogonal symbol vectors to be transmitted in two time slots. However, L-STF cannot be divided into two slots for two distinct symbols because, as discussed before, an STF must be regarded as a single modulation symbol. Otherwise, its DR, PAPR, and especially periodicity will be impacted.

The 802.11ac preamble format (see Fig. 3) reveals that we can actually take advantage of two different STFs, the L-STF and the VHT-STF, to create two time slots for the Alamouti code. The standard does not have any operational restriction on the two STF fields to be identical. The L-STF is crucial for frame detection, CFO estimation, and coarse AGC, whereas the VHT-STF is only used for fine AGC. Besides, conjugation and negation operations in Alamouti code keep the DR, PAPR and period of the waveform intact.

Another requirement for Alamouti code is that the channel does not change over the two time slots. The L-STF and VHT-STF are separated by $28\,\mu$s, rather than being consecutive as in typical Alamouti code systems. Nonetheless, our derivation in Section 2.2.3 indicates that the coherence time in typical WLANs is almost always larger than $40\,\mu$s. So the channel is time-invariant over the duration from L-STF to VHT-STF. It is safe to implement Alamouti code in these two fields assuming that they experience the same channel.

4.2.2 eP-Mod with Tx-Side Space-Time Coding

Consider two Tx antennas Tx_1 and Tx_2 and two bit-sequences, each of length L, to be collaboratively modulated at the two antennas using the L-STF and VHT-LTF. First, as shown in Fig. 9, using the scheme in Section 3.2, the two bit sequences can be modulated independently by Tx_1 and Tx_2 on their L-STFs to get frequency-domain symbol sequences, $\mathcal{S}_1 = \{S_{1,k}\}$ and $\mathcal{S}_2 = \{S_{2,k}\}, k \in \Lambda$ to be transmitted simultaneously. These two sequences are then exchanged between the two antennas and used to generate VHT-STFs according to Alamouti code. Specifically, when it comes to the VHT-STF period, Tx_1 uses $-S^*_{2,k}$ to generate its VHT-STF waveform, while Tx_2 uses $S^*_{1,k}$ for its VHT-STF to realize *eP-Mod-Alamouti*. It is worth to mention two distinctions between *eP-Mod-Alamouti* and conventional Alamouti: 1) the frequency domain Alamouti is applied to L-STF and VHT-STF, which are separated by $28\,\mu$s; 2) the two fields of the resulted Alamouti code are of different duration.

Similar to the SIMO case, the *eP-Demod* at the Rx side processes each STS rather than each OFDM symbol within the received STFs. Let $Y_{1,k}$ and $Y_{2,k}$ denote the received STS signals at time t and $t+T$, respectively. Considering the channel vector $\mathbf{h}_k = [h_{1,k}\ h_{2,k}]$ of the $4k$th subcarrier, where S_k's reside in, then it is straightforward to get:

$$\begin{bmatrix} Y_{1,k} \\ Y^*_{2,k} \end{bmatrix} = \underbrace{\begin{bmatrix} h_{1,k} & h_{2,k} \\ h^*_{2,k} & -h^*_{1,k} \end{bmatrix}}_{\mathbf{H}_e} \begin{bmatrix} S_{1,k} \\ S_{2,k} \end{bmatrix} + \begin{bmatrix} z_{1,k} \\ z^*_{2,k} \end{bmatrix}, \tag{9}$$

where $z_{1,k}$ and $z_{2,k}$ are the additive noise terms at time t and $t+T$, respectively. Now, $S_{1,k}$ and $S_{2,k}$ could be easily estimated by multiplying both sides of (9) by $\mathbf{H}^H_e$, and repeat the process over all subcarriers and STSs to eventually reconstruct $\mathcal{S}_1$ and $\mathcal{S}_2$ by combining $Y_{1,k}$'s and $Y_{2,k}$'s. To match the length of L-STF, VHT-STF is duplicated once at the Rx to get 10 STSs for Alamouti decoding. Then, we could directly apply the demodulation procedure as in Section 3.3.

The gain of *eP-Mod-Alamouti* is validated by our simulation results in Fig. 8(b). A naive MISO *eP-Mod* scheme where two Tx antennas embed the same sequence of 10 bits ($Q = M = 32$) in their L-STFs performs worse than SISO, because the naive MISO cannot decompose the two interfering signal branches. Embedding the same bit sequence in the VHT-STFs does not help much either for the same reason. In contrast, when *eP-Mod-Alamouti* is employed with two different sequences, not only the BER is improved significantly, but also the number of embedded bits per frame is doubled.

4.3 MIMO Diversity Gain under *eP-Mod*

We now generalize the ***eP-Mod-Alamouti*** scheme above to any $2 \times N_r$ MIMO system. The Tx-side procedure is the same as for MISO, and the Rx has to go through a similar combining as Section 4.1. Yet $Y_{1,k}, Y_{2,k}, h_{1,k}, h_{2,k}$ in (9) are substituted by corresponding vectors of dimension $N_r \times 1$, representing L-STFs, VHT-STFs, channel matrices at time t and $t+T$ over all Rx antennas. The overall diversity gain increases linearly with N_r.

We note that, the standardized STFs already exploit one type of diversity gain named cyclic shift diversity (CSD). CSD adds unequal delays to identical STFs from multiple Tx antennas to mitigate unintentional beamforming or deep fading. Besides CSD for different Tx antennas (space domain), we apply CSD between the L-STF and VHT-STF on the same Tx antenna (time domain). This will reduce the possibility that both STFs are deteriorated equally by the channel.

Accordingly, we embed the same bit sequence $\mathbf{m}$ of length L in all the STFs over all Tx antennas by applying CSD. This *eP-Mod* is called ***eP-Mod-Div*** as it achieves the full diversity gain. On the Rx side, MRC could be utilized when performing equalization. The equalized L-STFs and VHT-STFs over all Tx antennas are then equally combined. Finally, we get a single STF to run preamble demodulation and retrieve the embedded bit sequence $\mathbf{m}$. By exploiting multiple diversity gain techniques in this scheme, the reliability is improved significantly.

5 *eP-Mod* WITH CHANNEL-ADAPTIVE GAINS

Besides achieving diversity gain, MIMO is also considered as a means to increase data rate via multiplexing gain. To this end, we propose spatial multiplexing (space domain) and STF fields multiplexing (time domain) to *eP-Mod* to improve the embedding capacity. Also, we explore variants of *eP-Mod* to capture both types of gains (diversity and multiplexing) adaptive to the channel. Furthermore, we investigate the trade-off between these gains, as well as

Q vs. M, so as to maximize the capacity while maintaining reasonable reliability and finally improve embedding throughput.

5.1 *eP-Mod* Variants with Multiplexing Gain

With the help of multiplexing gain techniques, we introduce three more variants of *eP-Mod*, some of which may involve diversity gains as well. In the following, a general $N_t \times N_r$ MIMO system is considered except explicitly specified otherwise. We denote $N_m = \min\{N_t, N_r\}, N_M = \max\{N_t, N_r\}$. Denote also the bit sequences to be embedded in the L-STF and VHT-STF of ith Tx antenna as $\mathbf{m}_{i1}$ and $\mathbf{m}_{i2}$, respectively. Depending on the channel status (e.g., frequency-selective), the Tx selects one of the three variants where $\mathbf{m}_{ig}$'s, $1 \leqslant i \leqslant N_t$, $g = 1, 2$ may be independent or correlated for achieving channel-adaptive diversity and multiplexing gain. The channel quality/status can be determined based on the estimated CSI, which can be obtained by any common method used in Wi-Fi systems including those that are based on LTFs and pilots. In our implementation, it is estimated by L-LTF and VHT-LTF (see Section 2.2.3).

5.1.1 eP-Mod with Full Multiplexing Gain

When the channel has high SNR and no time-selective fading, we can safely opt for the transmission of one independent bit sequence per STF. Then, there is no need to sacrifice some STFs for diversity gain unless $N_t > N_r$. For multiplexing purposes, we apply *eP-Mod* to the L-STF and VHT-STF of each Tx antenna independently to embed multiple independent bit sequences. We call this *eP-Mod* as ***eP-Mod-Mux*** because it makes full use of multiplexing gain in both space and time domain to linearly improve the embedding capacity. Yet, the capacity is limited not only by N_t, but also by N_r. Because the Rx could at most retrieve N_r streams if $N_t > N_r$. *eP-Mod-Mux* could embed at most $2\,N_m$ distinct bit sequences $\mathbf{m}_{ig}, 1 \leqslant i \leqslant N_m, g = 1, 2$ in STFs of a frame. The remaining Tx/Rx antennas could be used for diversity gain, such as CSD or MRC. The Rx runs *eP-Demod* in parallel for $2\,N_m$ sequences once it retrieves $2\,N_m$ STFs.

5.1.2 Adaptive Variants of eP-Mod

As for the scenarios in which the channel is relatively but not sufficiently good to achieve a full multiplexing gain, we develop two other variants of *eP-Mod* with different combinations of diversity and multiplexing gains applied to space-time domain. They are developed based on *eP-Mod-Div*. Instead of applying CSD in both space and time domains, they only apply CSD in one domain and leave the other domain for multiplexing purposes.

***eP-Mod-vhtDiv*.** In a time-selective fading channel, we only apply CSD between the L-STF and VHT-STF on the same antenna to improve reliability, meaning that $\mathbf{m}_{i1} = \mathbf{m}_{i2}, 1 \leqslant i \leqslant N_m$, but then apply multiplexing gain over different Tx antennas. The constraint of spatial multiplexing gain is similar as *eP-Mod-Mux*, i.e., at most N_m bit sequences could be embedded by this scheme (c.f. $2\,N_m$ in *eP-Mod-Mux*).

***eP-Mod-TxDiv*.** On the contrary, in a frequency-selective fading channel, we only apply CSD between a pair of Tx antennas for higher reliability, and exploits VHT-STFs for time domain multiplexing gain. The procedures are illustrated in Algorithm 1. By this approach, a total of N_t bit sequences can be embedded.

Algorithm 1 eP-Mod with CSD over Tx Antennas

```
procedure eP-Mod-TxDiv
    for i ← 1, N_t do
        if i is odd then
            eP-Mod with 2 bit sequences m_{i1}, m_{i2} to get L-STF and VHT-STF for ith Tx antenna;
        else
            Cyclic shifts the L-STF and VHT-STF of (i − 1)th Tx antenna to get counterparts for ith Tx antenna;
        end if
    end for
end procedure
```

5.1.3 Design Limitations

There are two minor limitations that constrain the capacity of *eP-Mod* variants when applying MIMO gains.

First, the VHT-STF only contains 5 STSs while the L-STF contains 10. When these two STFs are designed for diversity gain, the Rx has to repeat the received VHT-STF once to get 10 STSs and then perform combining with L-STF for demodulating. Alternatively, when these two STFs are designed for multiplexing gain, the preamble modulation/demodulation processing for the VHT-STF is almost the same as the L-STF, except that only the last 4 STSs of VHT-STFs are useful after dropping the first STS. In both cases, the number of bits embedded by *eP-Mod* is restricted by the capacity of VHT-STF, because the actual effective gain in VHT-STF is half of the L-STF.

Second, the transmitted STF waveforms from a pair of antennas should be designed in a way that they do not end up with too close cyclic shifts. Otherwise, it creates unintentional beamforming, which would deteriorate the received signals. More specifically, the time shift applied to the STFs of multiple Tx antennas should keep a minimum difference of $50\,\text{ns}$, inline with the 802.11ac standard. For instance, given two Tx antennas, since the STS period is $T_s = 800\,\text{ns}$, if Tx_1 embeds $\log_2 Q$ bits by a time shift of $t_{s1} \in [0, 800]\,\text{ns}$ in its STFs, Tx_2 should not cyclically shift its STFs by $t_{s2} \in [t_{s1} - 50, t_{s1} + 50]\,\text{ns}$. This results in an effective number of bits of $\log_2(\frac{7}{8}Q)$, reducing the effective total bits for embedding.

5.2 Trade-off Analysis

The main objective behind the design of aforementioned *eP-Mod* variants is to jointly optimize the reliability and capacity. We study two trade-offs that impact this objective.

5.2.1 Diversity–multiplexing Trade-off

Lizhong *et al.* concluded in [19] that the maximum diversity and spatial multiplexing gain for an $N_t \times N_r$ MIMO system is $N_t \times N_r$ and N_m, respectively, when the other type of gain is 0. More generally, if the spatial multiplexing gain is r, then, the optimal diversity gain would be $(N_t - r) \times (N_r - r)$ and vice versa. As a special case, the maximium diversity gain obtained by Alamouti coding in a $2 \times N_r$ MIMO

TABLE 5
Diversity-multiplexing gain of *eP-Mod* variants with example.

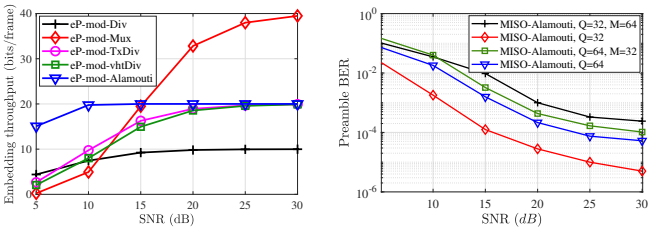
Variant	Diversity		Multiplexing		2×2 MIMO	
	Space	Time	Space	Time	Div.	Mux.
eP-Mod-Div	$2N_r$	2	—	—	8	—
eP-Mod-Alamouti	N_r	2	—	—	4	—
eP-Mod-vhtDiv	—	2	N_m	—	2	2
eP-Mod-TxDiv	2	—	$N_m/2$	2	2	2
eP-Mod-Mux	—	—	N_m	2	—	4
NOTE: MRC and EGC can be applied together with these variants.						

system is $2N_r$ [20]. However, these conclusions may not hold under *eP-Mod* for two reasons. First, CSD and STF-based multiplexing used in our schemes have not been studied in the literature. In addition, *eP-Mod* is not a conventional digital modulation and its performance sensitivity to channel and device impairments is much more complex. Hence, we develop an approximation to analyze the gains that our *eP-Mod* variants achieve.

In fact, when we apply CSD between a pair of Tx antennas, it is equivalent to repetition coding in a $2 \times N_r$ system. So, it attains a maximum space-domain diversity gain of $2\,N_r$ if spatial multiplexing is not applied. The CSD between the L-STF and VHT-STF of the same antenna further attains a maximum diversity gain of 2. With STFs multiplexing, a maximum time-domain multiplexing gain of ≈ 2 can be attained. Based on the above analysis, we summarize the various gains attained by each *eP-Mod* variant in Table 5. We know that the time-domain gain is usually independent of space-domain gain, except in the space-time coding case. So the total diversity/multiplexing gain of each variant is the product of space-domain and time-domain gains. Taking 2×2 MIMO as an example, we also show the exact value of diversity (Div.) and multiplexing (Mux.) gains in Table 5.

As seen in Table 5, with the same number of STF fields and Tx/Rx antennas, the *eP-Mod* variants have to make the diversity-multiplexing trade-off to improve the throughput of *eP-Mod*. *eP-Mod-Div* attains the highest diversity gain while *eP-Mod-Mux* attains the highest multiplexing gain. The higher the diversity gain, the more reliable the *eP-Mod* decoding will be; and the higher the multiplexing gain, the larger the *eP-Mod* capacity we achieve. To maximize the embedding capacity while maintaining a reasonable reliability level, *eP-Mod-vhtDiv* and *eP-Mod-TxDiv* trade off between diversity and multiplexing gain and they obtain comparable gains. Moreover, although *eP-Mod-Alamouti* has no multiplexing gain, its diversity gain is moderate.

The trade-off could also be evaluated by the *embedding throughput*, defined as the product of number of embedded bits per frame and preamble decode ratio (PDR). PDR is the ratio of the number of packets whose all *eP-Mod* bits are successfully decoded (i.e., BER $= 0$) to the total number of transmitted packets, indicating reliability. The *embedding throughput* of five *eP-Mod* variants in a 2×2 MIMO system is compared in Fig. 10(a) under Rician channel. Because of moderate diversity gain, *eP-Mod-Alamouti* achieves the highest *embedding throughput* when SNR < 15 dB; and the *embedding throughput* of *eP-Mod-Mux* increases sharply after 15 dB thanks to high multiplexing gain. In contrast, *eP-Mod-Div* has the lowest throughput due to lack of multiplexing gain.



(a) Diversity vs. multiplexing gain, 2×2 MIMO ($Q = M = 32$).

(b) Q vs. M, 2×1 MISO.

Fig. 10. eP-Mod trade-offs in Rician channel simulation.

5.2.2 *Trade-off between Q and M*

Recall that Q is the order of Q-DPSK symbol for pattern generation and M is the order of M-PSK symbol for phase shift generation in *eP-Mod*. The selection of Q and M directly determines the reliability and capacity of *eP-Mod*. We investigate the trade-off between Q and M that depends on two aspects of channel impairments.

The first impairment comes from channel noise. It is well known that DPSK requires 3 dB more E_s/N_0 than M-PSK of the same order. Apart from that, in *eP-Mod*, Q-DPSK symbol has only half the effective gain of M-PSK symbol as explained in Section 4.1.2. Noise also undermines both the time and frequency synchronization accuracy. The timing error impacts the demodulation of Q-DPSK symbol since it is mapped from the dependency pattern decided by time shift whereas the phase shift estimation for M-PSK symbol demodulation is impacted both by frequency error and wrong pattern detection (possibly caused by timing error). Thus, a slightly smaller or comparable Q than M seems necessary to maximize reliability.

However, it is not always true because of the other channel impairment: channel phasor. Multipaths, Doppler shift and device impairment could cause channel phasor to which high order M-PSK symbols are sensitive. While Q-DPSK symbols are demodulated by phase differences between successive symbols, hence robust to channel phasor. So high order Q-DPSK outperforms M-PSK in the context of channel phasor. Therefore, slightly greater or comparable Q than M is preferred for reliability.

The trade-off between Q and M from simulation results in Fig. 10(b) is consistent with our analysis. The Rician channel model has rich multipaths, and the main impairment at low SNR is noise, while multipath effect dominate at high SNR. Therefore, at low SNR, slightly smaller Q than M performs better, and vice versa at high SNR. We also see

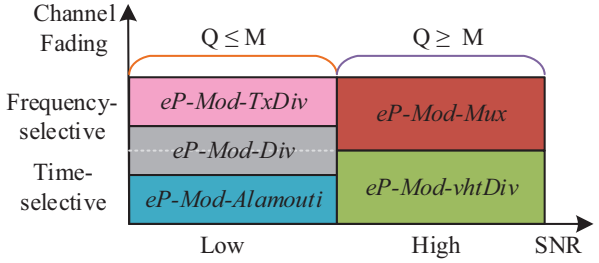


Fig. 11. *eP-Mod* variants corresponding to different SNRs and channel fading conditions.

TABLE 6
RMS delay and coherence bandwidth of TGac channel models [21]

Model	A	B	C	D	E	F
σ_τ (ns)	0	15	30	50	100	150
B_c (MHz)	/	10.6	5.3	3.2	1.6	1.1

that the BER gap between $Q = 32$ and $Q = 64$ is narrower at low SNR than high SNR.

In Fig. 11, we provide a summary of different eP-Mod variants corresponding to different SNRs and channel fading conditions.

5.3 Feasibility Analysis

Now, we study the feasibility of *eP-Mod* under a general MIMO setting where there is no restriction on the number of antennas at the Tx or Rx. When implementing *eP-Mod* under a general MIMO scenario, we must address the impact of STF multiplexing on the preamble's primary functions. Additionally, as MIMO requires a rich multipath environment to perform well, the Rx would need to resolve multipath effects on preamble demodulation. Different standard cyclic shifts on different Tx antennas and different preamble fields result in different incorporated CSI, which is another issue we have to account for when equalizing STF.

5.3.1 STF Functionality with Multiplexing

The most important property we want to ensure under *eP-Mod* is that the superposed STF waveforms at the Rx continue to exhibit a repetitive structure (see Section 2.2). Thanks to the *eP-Mod* design, the $0.8\,\mu s$ period holds for each individual STF, and hence, the superposed signal. Having said that, the multiple STFs received at one Rx antenna may not be aligned at the beginning due to multipath delays. The delay spread σ_τ of TGac channel models [21]–[23] can be as large as $150\,ns$ (see Table 6). So the first cycle of the superposition may not be identical to the rest of them. Nonetheless, the resulted superposed STF is still periodic with a period of $0.8\,\mu s$, though with fewer than 10 identical STSs, and it will not be essentially different than existing systems. Consequently, the auto-correlation-based frame detection and coarse time synchronization could still work with such repetitive structure because the Rx needs to know the period only, and not necessarily the transmitted STSs. To resolve possible timing errors, the Rx usually uses the known LTF to fine tune the frame detection.

Next, we study the DR and PAPR of superposed STFs and contrast them with those of the default 802.11ac STF to make sure the *eP-Mod* STFs are suitable for ADC/DAC and AGC. We keep the standard cyclic shift and subcarrier phase rotation that IEEE 802.11 standards apply to limit the DR and PAPR. In our *eP-Mod* design, additional cyclic shift is applied, which varies by frame and Tx antenna. This will further decrease the chances of aligned peaks or valleys in the superposed STF waveforms. In other words, theoretically, we should not expect much higher DR or PAPR than the default STF. Indeed, we compare the average DR and PAPR for *eP-Mod* STFs and standardized STFs under three typical 4×4 MIMO TGac channel models [21]–[23] in Table 7. From the table, we conclude that *eP-Mod* STFs at the Rx have comparable DR and PAPR to default STFs. Consequently, *eP-Mod* STFs are appropriate for ADC/DAC and amplifiers of existing commercial Wi-Fi devices, hence, their AGC function will not be impaired.

5.3.2 Preamble Demodulation under Impact

Beyond the primary functions of *eP-Mod* STFs, we also need to address preamble demodulation under multipath and standard cyclic shift.

Counter Multipath Effects. Because the delay spread σ_τ of TGac channels is always smaller than the $0.8\,\mu s$ (i.e., one STS duration), only the first received STS would be different from others. So, in the preamble demodulation, we drop the first STS and only consider the other 9 STSs. Moreover, multipath propagation causes interference, and shifts the phase of the signal. Given that *eP-Mod* relies on the phase of frequency domain sequences of STFs, it is necessary to examine such impact. Table 6 lists the coherence bandwidth B_c of typical TGac channel models. B_c in most models is greater than the frequency spacing (i.e., $0.3125 \times 4 = 1.25\,\text{MHz}$) between two successive S_k's. So, two successive S_k's experience the same channel, which means the phase differences between them (i.e., θ_i's) will not be changed by the multipath. In spite of that, as $B_c < 40\,\text{MHz}$, the coherence bandwidth is smaller than the channel bandwidth, the phase shift of subcarriers caused by multipath effect would be different within and out of the B_c range. As a result, the phase shift estimation of S_k's would be impacted. Nevertheless, the frequency domain CSI estimation by VHT-LTFs already includes multipath effect that will be equalized for the whole frame. Hence, the Rx could get a relatively "clean" version of the transmitted STFs. Even if there are errors, they will be averaged out when we sum over all elements of multiple STSs in (6) and (7) for preamble demodulation.

Remove Standard Cyclic Time Shift. Aside from the cyclic time shift we apply via embedding *Q-Seq* in *eP-Mod*, the L-STFs and VHT-STFs also have standard cyclic shifts. It is essential to remove the standard cyclic shifts before decoding *Q-Seq*. Let $H_{ji,k}$ denote the actual frequency domain channel response of kth subcarrier between ith Tx antenna and jth Rx antenna. As mentioned in Section 2.2.3, we could combine the standard cyclic shift and the phase rotation γ_k with $H_{ji,k}$ to get the incorpporated CSI. However, the cyclic shift T^i_{hcs} applied to the VHT portion of the preamble is different from the cyclic shift T^i_{lcs} for the legacy portion. Applying the CSI estimated by VHT-LTF could not remove

TABLE 7
DR and PAPR of received STFs for 802.11ac and *eP-Mod* under three typical 4×4 MIMO TGac channel models [21]

	802.11ac				eP-Mod			
	L-STF		HT-STF		L-STF		HT-STF	
Channel model	DR	PAPR	DR	PAPR	DR	PAPR	DR	PAPR
B	28.04	6.48	28.16	6.49	28.31	6.35	28.36	6.34
D	27.69	6.61	27.58	6.56	28.62	6.50	28.64	6.49
E	27.79	6.39	27.75	6.39	28.84	6.60	28.83	6.60

the standard cyclic shift from L-STFs. Denote the incorporated CSI for the legacy portion and the VHT portion of preamble as $\tilde{H}'_{ji,k}$ and $\tilde{H}_{ji,k}$, respectively. It is easy to derive the relationship between them and $H_{ji,k}$, that is:

$$\begin{cases} \tilde{H}'_{ji,k} = H_{ji,k}\exp(j2\pi k\Delta_F(-T^i_{lcs})) \\ \tilde{H}_{ji,k} = H_{ji,k}\exp(j2\pi k\Delta_F(-T^i_{hcs})) \\ \tilde{H}'_{ji,k} = \tilde{H}_{ji,k}\exp(j2\pi k\Delta_F(T^i_{hcs} - T^i_{lcs})) \end{cases} \quad (10)$$

Based on above analysis, the Rx estimates $\tilde{H}_{ji,k}$ by VHT-LTF and utilizes above equations to get $\tilde{H}'_{ji,k}$. Then, it treats STFs as data, and uses $\tilde{H}_{ji,k}$ to equalize VHT-STF, while uses $\tilde{H}'_{ji,k}$ to equalize L-STF. In this way, the standard cyclic shifts are removed so that the dependency pattern is detected correctly to decode the embedded *Q-Seq*.

6 Extensibility and Complexity

The schemes in the preceding sections are for special MIMO systems operating in 40 MHz channels. We now extend and generalize our design from a 20 MHz bandwith to any higher bandwidths. We also present the computational complexity and its reduction techniques we applied.

6.1 Extension to Higher Bandwidths

In OFDM-based 802.11 protocols, the STFs for higher bandwidths are generated using the sequence $\ddot{S}$ for 20 MHz bandwith [24, Eq.(6)] with replication, subcarrier phase rotation, and CSD. Denote the parent pattern of $\ddot{S}$ as $\ddot{\Theta} = [\theta_1, \theta_2, \ldots, \theta_{11}]$. We extend our proposed modulation scheme to any bandwidth by leveraging the repetitive structure of the STFs, as described in Algorithm 2.

Algorithm 2 Extensible Preamble-Modulation (*eP-Mod*)

1: **procedure** *eP-Mod*(*Q-Seq*||*M-Seq*)
2: Generate the parent pattern $\Theta^{(0)}$ by repeating $\ddot{\Theta}$
3: $\nu = 2\pi q/Q$, where q is the Gray-coded *Q-Seq*'s index
4: Construct child pattern $\Theta^{(\nu)}$ from $\Theta^{(0)}$ and ν
5: Set S_{-K} as the M-PSK symbol of Gray-coded *M-Seq*
6: Derive STF symbols S_k's with S_{-K} and $\Theta^{(\nu)}$
7: Apply the phase rotation for each 20 MHz block
8: Apply IFFT, add CP, and impose CSD
9: **end procedure**

The repetitive structure of the STFs on higher bandwiths provides more diversity that improves the reliability and capacity of *eP-Mod*. The above algorithm is for a single STF, combining this algorithm with aforementioned MIMO gain techniques, the Tx could embed more bits in L-STFs and VHT-STFs over multiple antennas. Considering the performance of BPSK modulation as our benchmark, through similar numeric analysis as in Section 4.1.2, we conclude that *eP-Mod* can embed up to 21 bits per frame in the preamble of a single antenna in an 80 MHz Wi-Fi channel.

6.2 Computational Complexity Reduction

According to the system architecture in Fig. 4, the computational complexity of *eP-Mod* is introduced by the shaded blocks, including *eP-Mod* and *eP-Demod* modules. Suppose we have an $N_t \times N_r$ MIMO system operating on 20β MHz channel width, $\beta = 1, 2, 4, 8$, then the number of samples in one STS as $l = 16\beta$. There are two approaches that we can use to reduce the computational complexity: 1) since S_k's are sparse with 4 subcarrier spacings, we run l-point IFFT/FFT to process each STS, rather than directly process each STF by $4l$-point IFFT/FFT; 2) the zero-forcing equalization is also done for each STS, which applies lower-order matrix multiplication. Taking the most complex *eP-Mod* variant—*eP-Mod-Mux* as an example, the computational complexity for each block is as follows:

1) *eP-Mod* blocks.
   - Pattern Generation: $\mathcal{O}(N_m)$.
   - Generating $\mathcal{S}$: $\mathcal{O}(N_m l)$
   - IFFT of STSs: $\mathcal{O}(N_m l \log l)$
2) FFT and Equalizer blocks.
   - Compensating for the estimated CFO: $\mathcal{O}(N_r l)$
   - FFT for individual STSs: $\mathcal{O}(N_r l \log l)$
   - Compensating $\tilde{H}_{ji,k}$ to get $\tilde{H}'_{ji,k}$: $\mathcal{O}(N_t N_r l)$
   - Channel equalization: $\mathcal{O}(N^3_M l)$
3) *eP-Demod* blocks.
   - Computing MSE for each possible value ν and finding the minimum MSE: $\mathcal{O}(N_m Q l)$
   - Rebuilding $\tilde{S}$: $\mathcal{O}(N_m l)$
   - Calculating the phase shifts from the reference elements, summing up for estimating $\Delta\varphi$: $\mathcal{O}(N_m l)$

Obviously, the signal processing accounts for a significant portion of the computational complexity of the whole scheme. Roughly, the total computational complexity at the Tx is $\mathcal{O}(N_m l \log l)$, lower than the one at the Rx side, which is $\mathcal{O}(N^3_M l)$.

7 Performance Evaluation

To evaluate the reliability and capacity of our proposed scheme, as well as its impact on the the primary preamble functions, we conduct extensive LabVIEW simulations and indoor experiments using NI-USRP RIOs. For both simulations and experiments, we implement the same 802.11ac PHY-layer system with our proposed *eP-Mod* schemes along with the standard 802.11ac PHY-layer system as our benchmark. We further implement an uncoded BPSK or QAM payload of 20 OFDM symbols for both systems.

7.1 Experiments

We prototype proposed variants of *eP-Mod* on two NI-USRP 2942Rs. The experiments are conducted in a realistic indoor environment with rich multipaths reflected/refracted by walls, ceilings, cubicles, furniture, and appliances, as shown

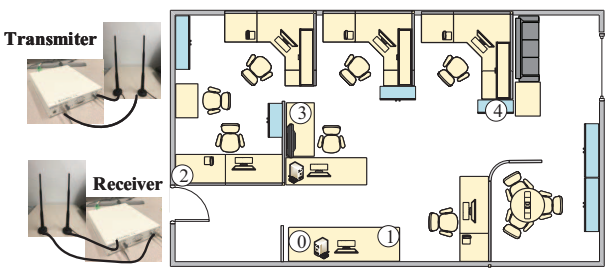


Fig. 12. Experimental setup with USRP 2942R and antennas.

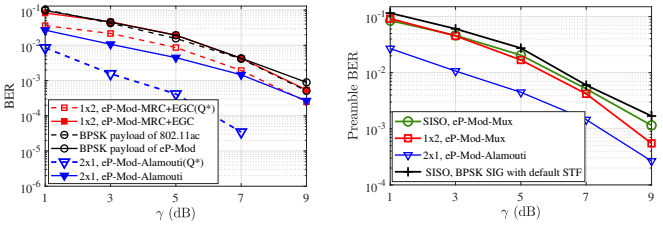


Fig. 13. BER in USRP experiments over 400 kHz bandwidth for *eP-Mod* ($Q = 64, M = 16$) in SISO, SIMO, MISO. (a). *eP-Mod* with diversity gain (10 bits/Tx antenna/frame) vs BPSK payload. (b). *eP-Mod* with diversity/multiplexing gain (20 bits/frame)

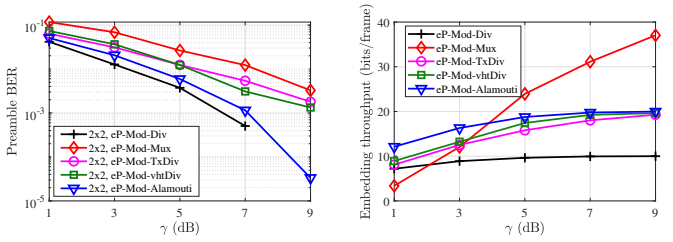


Fig. 14. Performance of eP-Mod variants in 2×2 MIMO experiments over 400 kHz bandwidth ($Q = 64, M = 16$).

in Fig. 12. The Tx is fixed at location 0, and the Rx is placed at four locations with location 1 as the default unless specified otherwise. The Rx location 1 or 4 offers a channel with a dominant line-of-sight path, whereas the Rx location 2 or 3 offers a channel with rich multipaths because of obstacles. Each of the two end devices is equipped with two 8 dBi omnidirectional antennas operating at 2.5 GHz frequency band to avoid interference. Due to the real-time processing limitations of our host CPU and LabVIEW (not the connection), we were not able to experiment over the ideal 40 MHz bandwidth without data overflow at USRP. The highest achievable bandwidths for frames with and without payload in our setup are 400 kHz and 8 MHz, respectively. Having said that, it is equivalent to 40 MHz as we have the same 128 subcarriers as the standard specifies. Since the environmental noise floor at the Rx is too low, a synthetic Gaussian noise is added at the Tx so as to lower the overall SNR. In the following, γ refers to the the synthetic SNR at the Tx (not the exact SNR at the Rx).

Preamble BER and Embedding Throughput. In the SIMO scenario, we only embed bits in the L-STF by *eP-Mod-MRC+EGC*, whereas in the MISO scenario, we apply *eP-Mod-Alamouti* where two Tx antennas jointly embed bits in their L-STFs and VHT-LTFs. Here, for simplicity, we do not impose the restriction with regard to unintentional beamforming of different STFs at the Rx, and let each antenna select its bit sequence independently. Similar to the simulation results shown in Fig. 8, we have been able to reliably communicate 10 bits per Tx antenna in both SIMO and MISO scenarios with comparable or lower BER as BPSK, as shown in Fig. 13(a). However, different from the simulations, we need to set $Q = 64$ and $M = 16$ to get the best performance, which means the demodulation of *Q-Seq* is much more robust in practice than detecting the phase shift for *M-Seq*. Although diversity gain by Alamouti in MISO is comparable as MRC+EGC in SIMO, additional effective gain in VHT-STF guarantees lower BER in MISO scheme than SIMO scheme.

Additionally, we apply *eP-Mod-Mux* to SISO and SIMO systems such that they can also embed 20 bits in their preambles. Fig. 13(b) reveals that the BER of 1×2 *eP-Mod-Mux* retains the same level as the one of 1x2 eP-Mod-MRC-EGC in Fig. 13(a), where the VHT-STF is not used for multiplexing gain. It means the VHT-STF has comparable capacity as the L-STF although the former has lower effective gain. Note that the BER difference between SISO and SIMO is tiny when $\gamma < 7$ dB, because the SIMO receiver with 2 antennas has more heat noise and worse synchronization accuracy than SISO. Anyway, all three *eP-Mod* schemes in this figure have lower preamble BER than BPSK SIG with default STF in SISO.

Next, we consider the 2×2 MIMO system with five variants of *eP-Mod*. When $Q = 64, M = 16$, the preamble BER plotted in Fig. 14(a) decreases as the diversity gain increases over different *eP-Mod* variants. However, among *eP-Mod-Alamouti*, *eP-Mod-TxDiv* and *eP-Mod-vhtDiv* that embed the same number of bits (20 bits in this setup), *eP-Mod-Alamouti* outperforms the other two. It demonstrates the advantage of integrated space-time diversity over independent diversities in two domains. Even though the *eP-Mod-Mux* has the worst preamble BER, it obtains the highest *embedding throughput* when $\gamma > 4$ dB as seen from Fig. 14(b) . In contrast, *eP-Mod-Alamouti* achieves the highest *embedding throughput* at low SNR. Inspired by this, we focus on these two schemes to investigate the capacity of *eP-Mod* for 2×2 MIMO.

Capacity and PDR. For the purpose of capacity test, we set the PDR of 90 percent as the reliability threshold and go over various systems with no more than 2 Tx/Rx antennas. Fig. 15 shows the capacity together with PDR, where SISO, 1×2 SIMO and 2×1 MISO achieve a capacity of 22 (bits/frame) at $\gamma = 9$ dB, and 2×2 MIMO reaches a capacity of 40 (bits/frame). As can be seen, SISO and 1×2 SIMO have exact the same capacity over the tested SNR range, the same for 2×1 *eP-Mod-Alamouti* and 2×2 MIMO *eP-Mod-Alamouti*. Nonetheless, the PDR of SIMO and MIMO benefit from one extra Rx antenna.

To evaluate the performance of *eP-Mod* under more realistic channel bandwidths, we optimized our LabVIEW implementation to be able to transmit preambles without

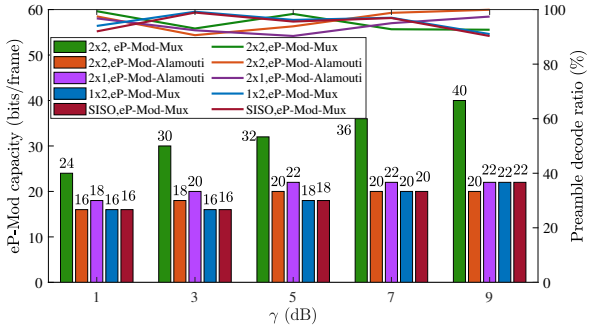


Fig. 15. Capacity and preamble decode ratio (PDR) of *eP-Mod* variants (400 kHz bandwidth).

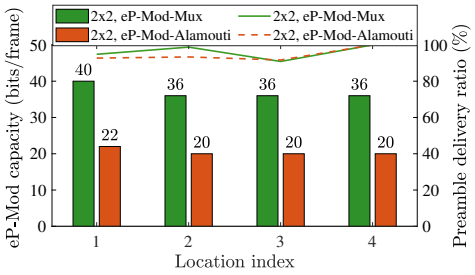


Fig. 16. Capacity and preamble decode ratio (PDR) of 2×2 MIMO *eP-Mod* variants over 8 MHz bandwidth at different locations.

payload over a channel bandwidth of 8 MHz. Considering different channel conditions (multipath and line-of-sight propagation), the Tx location is fixed at location 0 in Fig. 12, and the Rx is placed at locations 1, 2, 3 or 4. Here, we only evaluate the performance of two 2×2 MIMO variants, *eP-Mod-Mux* and *eP-Mod-Alamouti*, as these variants are more sensitive to multipaths than others. We can see from Fig. 16 that, while maintaining a PDR higher than 90 percent, the two variants achieve higher embedding capacity of 40 and 22 bits/frame at location 1 than other locations, where the embedding capacity of the two variants are 36 and 20 bits/frame, respectively. The results are comparable as those in Fig. 15 obtained for 400 kHz bandwidth. So we conjecture that *eP-Mod* would also perform well under a standard 40 MHz channel.

The overall design of *eP-Mod* is strictly compliant with criteria that maintain the preamble functions. To validate the resulting STFs' functionality, we evaluate the impact of *eP-Mod* with respect to CFO estimation accuracy, frame detection accuracy and payload BER.

CFO Estimation Accuracy. Specifically, as shown in Fig. 17(a), the average of estimated CFO under various *eP-Mod* variants are very close to that estimated by the default 802.11ac preamble. As this estimation of overall CFO is done by L-STF followed by L-LTF, it may not prove the functionality of *eP-Mod* STFs. Thus, we focus on the coarse CFO purely estimated by L-STF, and show the average and standard deviation against SNR γ in Fig. 17(b). For clarification, we only compare two typical schemes *eP-Mod-Alamouti* and *eP-Mod-vhtDiv* with default STF. Despite some fluctuation due to unstable USRP devices, either our novel L-STFs or default L-STF get the coarse-estimated CFO of $38\% \sim 41\%$ Δ_F. Meanwhile, the standard deviation of estimated CFO by *eP-Mod* L-STF is lower than the one by the default L-STF. This is because the diverse *eP-Mod* STFs are more robust to channel impairments over time. Also, the standard deviation decreases as SNR γ increases because

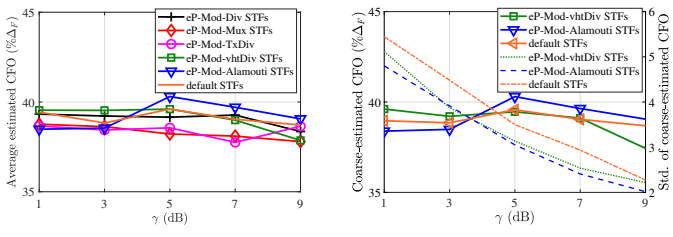
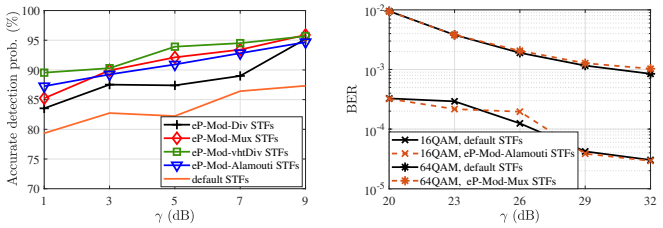


Fig. 17. Estimated CFO in MIMO experiments, *eP-Mod* ($Q = 64, M = 16$). (a) Average of overall estimated CFO by L-STF and L-LTF; (b)Average and standard deviation of coarse estimated CFO by L-STF.



(a) Frame detection accuracy. (b) BER of QAM payloads.

Fig. 18. *eP-Mod* STFs vs. default STFs w.r.t frame detection accuracy and QAM Payload BER in MIMO, ($Q = 64, M = 16$).

lower noise makes more accurate auto-correlation based CFO synchronization. Notably, during the experiments, we found out that the estimated CFO increased as operation time passed by. So the fluctuation is negligible compared to the total amount of estimated CFO.

Frame Detection Accuracy. It turns out that the accurate frame detection probability in SIMO is 91.8 percent, higher than default STFs but worse than 99.5 percent in MISO, because two Rx antennas detecting independently suffer from fading. As for MIMO, our novel STFs by different *eP-Mod* variants achieve relatively accurate detection probability greater than 85 percent, higher than default STFs, illustrated in Fig. 18(a).

Payload BER. We first observed in Fig. 13(a) that the BER of BPSK payload is not impacted by *eP-Mod*. We then checked the BER of high-order QAM payload of *eP-Mod-Alamouti* STFs, neither the BER of 16-QAM nor 64-QAM is degraded compared to when the default STFs are used. Therefore, our scheme ensures the whole system perform as normal.

7.2 High-order MIMO Simulations

Due to limited scalability of our testbed, we evaluate the performance of *eP-Mod* in high-order MIMO systems by LabVIEW simulations. The Rician multipath channels simplified from TGac channel model C are adopted in our simulations. First, we add more Rx antennas to a 2×2 MIMO system to increase the diversity gain. In Fig. 19(a), *eP-Mod-Alamouti* for 2×4 MIMO and *eP-Mod-Mux* for 2×8 MIMO reach BER of 10^{-5} before 12.5 dB SNR when $Q = M = 32$. And they keep almost steady *embedding throughput* around the maximum achievable capacity of 20 and 40 (bits/frame). Similarly, Fig. 19(b) depicts the simulation results of *eP-Mod-Mux* for 4×6 and 4×8 MIMO. They reach BER of 10^{-4} before 20 dB SNR and achieve *embedding throughput* approximate to 80 (bits/frame), which is the maximum achievable capacity for $Q = M = 32$.

Then, we evaluate the capacity of three typical MIMO systems in Fig. 20, where 2×2, 4×4 and 8×8 MIMO reach a capacity of 48, 80 and 144 (bits/frame), respectively at 30 dB SNR. All of them achieve high capacity by *eP-Mod-Mux* variant, and the capacity of 2×2 MIMO for 5 dB and 10 dB is obtained by *eP-Mod-Alamouti*, while the other two systems achieve the lowest capacity by *eP-Mod-Div*. The corresponding PDRs achieved are also shown in the figure.

Overall, these evaluations show that *eP-Mod* can successfully embed user-defined bits in the STF waveform of the

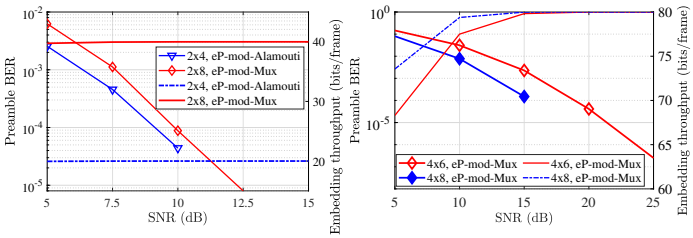
preamble utilizing MIMO gain techniques to improve capacity and reliability. Most importantly, our schemes would not comprise the primary functions of preamble such as frame detection, CFO estimation, and hence, the normal transmission of payload would not be impacted.

8 Related Work

In many standards for wireless networks (e.g., IEEE 802.11, 802.15.4, ...), the preamble is of great importance for decoding a received frame. Accordingly, any (secondary) use of the preamble, e.g., to communicate user-defined bits, should not hinder the preamble's primary functions and must be backward compatibile with legacy systems.

In [5], the authors investigated embedding a sequence of bits in the preamble of the non-OFDM Wi-Fi systems (single carrier, specified by the IEEE 802.11b standard). This preamble has a different structure than the one used for the OFDM-based Wi-Fi systems. The embedding scheme in [5] requires the modified preamble to be known a priori to the Rx so as to be used for CSI estimation. A covert channel for SISO OFDM-based Wi-Fi systems was proposed in [25], where the stegotext was camouflaged in the preamble by shifting its phase. However, the sensitivity of PSK to multipath fading and noise constrains the throughput of this covert channel to 4 bits at a moderate BER. The authors of [26] proposed embedding bits in the preamble of IEEE 802.11a systems by imposing phase and time shifts on the L-STF of the preamble for transmission over a 20 MHz channel. And this work is further extended to IEEE 802.11ac systems where higher bandwidth and multiple antennas are considered [27]. But the schemes in [27] are tailored for 1×2 and 2×1 MIMO systems. The design of *eP-Mod* in general MIMO systems is more challenging because in our case we must: (1) propose algorithms of reasonable computational complexity and applicable to any MIMO systems; (2) trade off between the diversity and multiplexing gain of *eP-Mod* to reliably embed as many bits as possible; and (3) improve the synchronization accuracy under complicated MIMO channels. We address these challenges in a more general context, where we consider OFDM-based 802.11n/ac/ax preambles for embedding bits.

Several methods have been proposed for embedding bits in OFDM-based Wi-Fi frames by utilizing a redundant portion of the frame to constrain the negative impact of embedding. WiPad (Wireless Padding) [28] is one such method, where bits are embedded in the padding of frames (as opposed to conventional zero padding). But WiPad changes



(a) $N_t = 2, N_r = 4, 8.$ (b) $N_t = 4, N_r = 6, 8.$

Fig. 19. Preamble BER and embedding throughput of eP-Mod in high-order MIMO, Rician channel simulation ($Q = M = 32$)

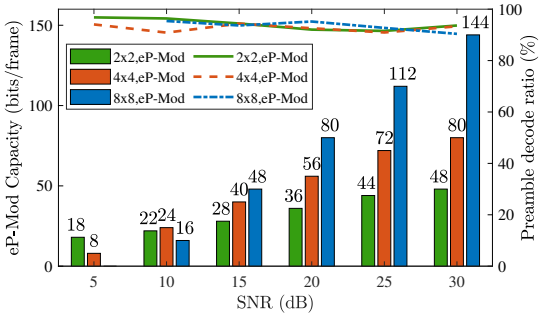


Fig. 20. Capacity and PDR(preamble decode ratio) of *eP-Mod* in higher-order MIMO, Rician channel simulation.

the Service and Tail fields that impact their functions critical for scrambler and encoder initialization. Moreover, padding the data would extend the frame length. The Cyclic Prefix (CP), another redundant field, is a concatenation of the last several samples of an OFDM symbol to its beginning. Szymon *et al.* suggested replacing the conventional CP with information symbols [29]. However, their method is vulnerable to inter-symbol interference (ISI). In fact, it degrades the primary CP function and may also impact data symbols. Frequency offset embedding for authenticating transmitters [30] is another embedding scheme in which the Rx must receive multiple frames to extract the embedded bits. In contrast, our scheme does not even require one full frame to extract these bits and achieves a higher per-frame embedding capacity than the $1 \sim 4$ bits in [30].

More recently, researchers have been able to embed bits in more critical portions of a frame. In [31], covert symbols are mapped by changing the amplitude of primary data symbols. But it is restricted to PSK modulated data, while QAM is widely used in Wi-Fi nowadays. Authors of [32] embedded access point (AP) discovery-related information in the LTF of the preamble to facilitate AP discovery. This would impact the primary functions of the preamble which is not evaluated in the paper. Besides, it is infeasible for MIMO systems since the modified LTFs superposed at the Rx make it difficult to estimate the CSI and retrieve the embedded information.

9 Conclusion and Future Work

We presented the extensible design and implementation of a novel preamble in MIMO-OFDM based 802.11 systems, named *eP-Mod*, enabling the 8×8 MIMO system operating in 40 MHz channel bandwidth to embed up to 144 frame-specific bits for anticipated applications. To do this, we encode time shift and phase shift of STF waveform into the characteristics of STF frequency domain symbols. MIMO diversity and multiplexing gain techniques are customized and applied to *eP-Mod* to improve its capacity and reliability. Its extensibility to higher bandwidth and any 802.11 systems with OFDM PHY protocols could further improve the capacity. Most importantly, our schemes take accounts of standardized preamble properties, thus maintain all the functions of the preamble for backward compatibility and interoperability. Our extensive evaluation of its performance demonstrates that it could be utilized for PHY-layer sig-

naling and PHY-layer security in WLANs. We also outline directions of future work here.

Real World Devices. So far, our proof-of-concept implementation is based on software-defined radio USRPs. To validate the practicality of our method, *eP-Mod* implementation on real world devices is preferred. So hardware complexity and energy efficiency can be evaluated.

Applications. Because of limited time and energy, we have not applied our technique to advertised applications in this paper. As such, one can consider embedding bits in the preamble as BSS color for throughput improvement, or device ID for authentication on PHY layer.

ACKNOWLEDGMENT

This work was supported in part by NSF (grants CNS-1910348, CNS-1563655, CNS-1731164, CNS-1813401, and IIP-1822071) 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.

REFERENCES

[1] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 197–216, 2019.

[2] 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 Sel. Areas in Commun.*, vol. 35, no. 1, pp. 20–29, Jan. 2017.

[3] H. Rahbari and M. Krunz, "Secrecy beyond encryption: obfuscating transmission signatures in wireless communications," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 54–60, Dec. 2015.

[4] M. Vanhoef and F. Piessens, "Advanced Wi-Fi attacks using commodity hardware," in *Proc. 30th Ann. Computer Security Applications Conf. (ACSAC)*, New Orleans, Louisiana, USA, 2014, pp. 256–265.

[5] H. Rahbari and M. Krunz, "Full frame encryption and modulation obfuscation using channel-independent preamble identifier," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 12, pp. 2732–2747, Dec. 2016.

[6] M. Vanhoef and F. Piessens, "Key reinstallation attacks: forcing nonce reuse in WPA2," in *Proc. ACM SIGSAC Conf. Computer and Commun. Security (CCS)*, Dallas, Texas, USA, 2017, pp. 1313–1328.

[7] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 1: Enhancements for High Efficiency WLAN*, IEEE Std. IEEE 802.11ax, 2021.

[8] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, June 2008.

[9] H. Song *et al.*, "Secure cooperative transmission with imperfect channel state information based on BPNN," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10 482–10 491, 2018.

[10] N. O. Tippenhauer, L. Malisa, A. Ranganathan, and S. Capkun, "On limitations of friendly jamming for confidentiality," in *Proc. IEEE Symp. Security and Privacy*, 2013, pp. 160–173.

[11] B. Danev, D. Zanetti, and S. Capkun, "On physical-layer identification of wireless devices," *ACM Computer Survey*, vol. 45, no. 1, pp. 6:1–6:29, Dec. 2012.

[12] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Fingerprints in the ether: Using the physical layer for wireless authentication," in *Proc. IEEE Int. Conf. Commun.*, 2007, pp. 4646–4651.

[13] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz*, IEEE Std. IEEE 802.11ac, 2013.

[14] Y. N. Leonardo Lanante Jr., "Phase rotation for the 80 MHz 802.11ac mixed mode packet," IEEE, Report Doc. IEEE 802.11-10/0791r0, July 2010.

[15] E. Perahia and R. Stacey, *Next generation wireless LANs: 802.11n and 802.11ac*. Cambridge Univ. Press, 2013.

[16] R. Boehnke and T. Doelle, "Alternative proposal for BRAN SYNCH preamble," IEEE, Report Doc. IEEE 802.11-99/048, Mar. 1999.

[17] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*. John Wiley & Sons, 2010.

[18] C. Oestges and B. Clerckx, *MIMO wireless communications: From real-world propagation to space-time code design*. Academic Press, 2010.

[19] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. on Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[20] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[21] V. Erceg, L. Schumacher, P. Kyritsi *et al.*, "TGn Channel Models," IEEE, Report Doc. IEEE 802.11-03/940r4, may 2004.

[22] G. Breit, H. Sampath, S. Vermani *et al.*, "TGac Channel Model Addendum," IEEE, Report Doc. IEEE 802.11-09/0308r12, mar 2010.

[23] G. Breit, H. Sampath, S. Verman *et al.*, "TGac Channel Model Addendum Supporting Material," IEEE, Report Doc. IEEE802.11-09/0569r0, may 2009.

[24] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std. IEEE 802.11a, 1999.

[25] J. Classen, M. Schulz, and M. Hollick, "Practical covert channels for WiFi systems," in *Proc. IEEE Conf. on Commun. and Network Security (CNS)*, 2015, pp. 209–217.

[26] H. Rahbari and M. Krunz, "Exploiting frame preamble waveforms to support new physical-layer functions in OFDM-based 802.11 systems," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 6, pp. 3775–3786, June 2017.

[27] Z. Zhang, H. Rahbari, and M. Krunz, "Expanding the role of preambles to support user-defined functionality in MIMO-based WLANs," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, 2020, pp. 1191–1200.

[28] K. Szczypiorski and W. Mazurczyk, "Steganography in ieee 802.11 ofdm symbols," *Security and Communication Networks*, vol. 9, no. 2, pp. 118–129, 2016.

[29] S. Grabski and K. Szczypiorski, "Steganography in OFDM symbols of fast IEEE 802.11n networks," in *Proc. 2013 IEEE Security and Privacy*, May. 2013, pp. 158–164.

[30] V. Kumar, J.-M. Park, and K. Bian, "Blind transmitter authentication for spectrum security and enforcement," in *Proc. ACM Conf. Computer and Commun. Security (CCS)*, Scottsdale, Arizona, USA, 2014, pp. 787–798.

[31] S. D'Oro, F. Restuccia, and T. Melodia, "Hiding data in plain sight: undetectable wireless communications through pseudo-noise asymmetric shift keying," in *Proc. IEEE Int. Conf. on Computer Commun (INFOCOM).*, 2019, pp. 1585–1593.

[32] D. J. K. Sankhe and K. Chowdhury, "Csiscan: Learning CSI for efficient access point discovery in dense WiFi networks," in *IEEE Int. Conf. on Network Protocols (ICNP)*, 2020, pp. 1–11.

Zhengguang Zhang received the B.S. and M.S. degrees in communication and information engineering from Univ. of Elec. Sci. and Tech. of China, in 2014 and in 2017, respectively. She is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, the University of Arizona. Her research interests include wireless and spectrum sharing security, wireless communications and networking with emphasis on designing cross-layer and cross-technology protocols.

Hanif Rahbari is an assistant professor of the Department of Computing Security and a member of Global Cyber Security Institute, Rochester Institute of Technology (RIT). He received the Ph.D. degree in electrical and computer engineering from the University of Arizona (UA) in 2016. He joined RIT in Spring 2018 after a short-term affiliation with UA as a Senior Research Specialist and a brief experience as a post-doctoral associate at Virginia Tech. His broad research area is wireless security and wireless communications, with emphasis on jamming and privacy (obfuscation) at the physical layer, connected vehicles security, Internet of Things (IoT), and spectrum sharing, and is a co-inventor on three US patents.

Marwan Krunz is the Regents Professor and Kenneth VonBehren Endowed Professor in the ECE Department at the University of Arizona. He directs the Broadband Wireless Access and Applications Center, a multi-university industry-focused NSF center that includes affiliates from industry and government labs. Dr. Krunz's research interests are in wireless communications and networking, with emphasis on resource management, adaptive protocols, and security issues. He has published more than 310 journal articles and peer-reviewed conference papers, and is a co-inventor on several US patents. He is an IEEE Fellow, an Arizona Engineering Faculty Fellow $(2011-2014)$, and an IEEE Communications Society Distinguished Lecturer $(2013$ and $2014)$. He received the NSF CAREER award in 1998. He served as the Editor-in-Chief for the IEEE Transactions on Mobile Computing, and was on the editorial boards for numerous journals.