Extrinsic Channel-Like Fingerprinting Overlays Using Subspace Embedding

Nate S. Goergen, *Student Member, IEEE*, W. Sabrina Lin, *Member, IEEE*, K. J. Ray Liu, *Fellow, IEEE*, and T. Charles Clancy, *Senior Member, IEEE*

***Abstract*—We present a physical-layer fingerprint-embedding scheme for orthogonal frequency-division-multiplexing (OFDM) transmissions, where the fingerprint signal conveys a low capacity communication suitable for authenticating the transmission and further facilitating secure communications. Our system strives to embed the fingerprint message into the noise subspace of the channel estimates obtained by the receiver, using a number of signal spreading techniques. When side information of the channel state is known and leveraged by the transmitter, the performance of the fingerprint embedding can be improved. When channel state information is not known, blind spreading techniques are applied. The fingerprint message is only visible to aware receivers who explicitly preform detection of the signal, but is invisible to receivers employing typical channel equalization. A taxonomy of overlay designs is discussed and these designs are explored through experiment using time-varying channel-state information (CSI) recorded from IEEE802.16e Mobile WiMax base stations. The performance of the fingerprint signal as received by a WiMax subscriber is demonstrated using CSI measurements derived from the downlink signal. Detection performance for the digital fingerprint message in time-varying channel conditions is also presented via simulation.**

***Index Terms*—Cognitive radio, communication system security, orthogonal frequency-division-multiplexing (OFDM) modulation, watermarking, WiMax.**

I. Introduction

WITH THE widespread adoption of wireless communication, the security of wireless systems has become an extensively researched topic. While cryptographic methods at higher layers have been widely used to authenticate wireless users and prevent interception of original pre-encryption transmissions by malicious or unintended users, the ability to authenticate and classify wireless transmissions at the physical (PHY) layer has a number of advantages over higher-layer approaches. Authentication at the PHY-layer, before demodulating and decoding the signal, can prevent wasteful processing of unintended transmissions and allows nodes to quickly authenticate legitimate users and implicate charlatans.

In general, robust authentication devices are crucial to securing wireless systems against message forgery and the malicious actions of impostors, thereby preventing a number of identity attacks to next-generation wireless systems [1]–[4]. Additionally, PHY-layer approaches provide a completely independent authentication mechanism decoupled from higher-layer authentication devices or protocols, allowing the authentication mechanism to be invariant of higher-level protocol changes or revisions.

Message fingerprinting is the practice of appending or embedding a secondary message conveying the credentials of a data source into a signal. Numerous fingerprinting methods have been successfully applied to multimedia systems allowing for secure transmission of multimedia content [5]. When applied to wireless transmissions, robust physical-layer fingerprints can enable signal authentication even when the signal itself is unrecoverable due to low signal-to-noise ratio (SNR) or fading conditions.

In this paper, we present a physical-layer (PHY) fingerprinting method for orthogonal frequency-division-multiplexed (OFDM) transmissions, where side information about anticipated channel conditions is incorporated into the fingerprint design. A number of PHY-layer fingerprinting approaches for wireless communications have been investigated using basic signal superposition methods [6]–[9]. The main disadvantage of blind superposition is that the fingerprint signal appears as interference to the original signal and is fully present when the signal is decoded, resulting in decreased SNR for the original signal. Instead we investigate fingerprint designs that consider how the signal will be perceived by the receiver, and side information of the channel distortions that the signal will experience. By leveraging channel side information and considering the receiver's perception of the signal, improved fingerprint designs [10] are possible as the undesirable effects of the fingerprint signal associated with blind superposition approaches [7] and [8] are partially removed by the receiver when preprocessing the signal.

In [11], it was demonstrated that robust PHY-layer fingerprints can be obtained from intrinsic features characteristic of wireless channels, such as unique scattering environments and spatial variability. However, when channel conditions are not conducive to intrinsic fingerprint recognition, due to either highly correlated multipath profiles between transmitters or rapidly varying channel conditions, a more robust fingerprint is required to authenticate wireless nodes. We consider augmenting current intrinsic channel-based authentication mechanisms with an extrinsic synthetically generated channel-like signal that is applied by the transmitter. This

Manuscript received September 16, 2010; revised August 27, 2011; accepted September 08, 2011. Date of publication October 13, 2011; date of current version November 18, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Wade Trappe.

N. S. Goergen, W. S. Lin, and K. J. R. Liu are with the Electrical and Computer Engineering Department, University of Maryland, College Park, MD 20742 USA.

T. C. Clancy is with the Ted and Karyn Hume Center for National Security and Technology, Virginia Tech, Arlington, VA 22203 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2011.2172208

signal conveys a cryptographically secure digital signature and authentication message digest along with the original transmission. Channel-like fingerprint signals are designed in ways that the fingerprints can be modeled as time-varying channel distortions. The primary advantage of channel-like fingerprinting approaches is that distortions incurred by the fingerprints can be subsequently corrected by the receiver through traditional channel equalization and synchronization methods.

In [12], an extrinsic channel-like fingerprint for narrowband single-input single-output (SISO) digital television signals was considered, where the fingerprint message is applied at the transmitter by emulating nominal multipath channel responses. In [13], an extrinsic channel-like fingerprint signal is considered for multiple-input multiple-output (MIMO) systems using space–time block codes (STBC). In this work, we extend these previous approaches to OFDM signals, and we incorporate previous channel state knowledge as side information into the design of the fingerprint signal.

We propose a number of techniques to spread the fingerprint message in both frequency and time domain using subspace decomposition. When the full channel-state information (CSI) is known and leveraged by the transmitter, the fingerprint can be embedded into the noise subspace of the receiver's channel estimates to ensure that the fingerprint signal incurs minimal distortion during transmission. When CSI is not known by the transmitter, blind spreading designs using orthogonal codes, such as Walsh codes, can be applied. We present a number fingerprint spreading designs that incorporate various amounts of previous CSI into the design of the spread fingerprint signaling bases. Our designs demonstrate that additional CSI knowledge can be leveraged by the transmitter to improve the performance of the fingerprint embedding.

This paper is organized as follows: Section II describes the OFDM system and presents a framework for introducing the channel-like fingerprint. In Section III, an analysis of the extrinsic fingerprint overlay design is given, and the embedding and recovery of the fingerprint message is demonstrated. A taxonomy of overlay designs is also presented. In Section IV, we present experimental results, where one fingerprint overlay design is evaluated using CSI extracted from an IEEE802.16e WiMax transmission. In Section V, we present simulation results for two of the fingerprint overlay designs. In Section VI, we present our conclusions. The following is a list of the most frequently used notation in Sections II, III, and VI.

1) $\boldsymbol{\xi}[k]$: Extrinsic fingerprint signal vector;
2) $\mathbf{f}[k]$: Extrinsic fingerprint signaling vector to be applied the TFB transmission at time k;
3) $\hat{\mathbb{H}}[k]$: Block of frequency domain intrinsic channel estimate vectors;
4) $\mathbb{F}[k]$: Block of frequency domain extrinsic fingerprint vectors;
5) $\hat{\mathbb{Q}}[k]$: Block of frequency domain aggregate channel estimate vectors;
6) $\hat{\mathbb{L}}[k]$: Block of frequency domain intrinsic channel estimate vectors;
7) $\mathbb{U}_2[k]$: Time-spreading fingerprint signaling basis;
8) $\mathbb{V}_2[k]$: Frequency-spreading fingerprint signaling basis;
9) $\mathbb{K}[k]$: Extrinsic fingerprint overlay block;
10) $\mathbb{E}[k]$: Intrinsic channel model mismatch error block.

II. System Model

We consider an OFDM system where the transmission is subjected to a linear time-domain channel response $g(t)$, given as

$$g(t) = \sum_c A_c(t)\delta(t - \Delta\tau_c) \tag{1}$$

where Δ is the sampling interval, τ_c are the delays for each channel component, and $A_c(t)$ are the complex valued delay-spread gains at time t for multipath component c. The OFDM system is modulated using an N-point discrete-time inverse Fourier transform (IDFT_N), and then subsequently demodulated using an N-point discrete-time Fourier transform (DFT_N). The matrix representation of the OFDM system is given as

$$\mathbf{y} = \mathbf{gWX} + \mathbf{n} \tag{2}$$

where $\mathbf{y} = [y_0\ y_1 \cdots y_{N-1}]$ is the received band-limited signal vector after DFT_N, where elements $y_n \in \mathcal{C}$, the matrix $\mathbf{X} = \mathrm{diag}([x_0 x_1 \cdots x_{N-1}])$ is of size $\mathcal{C}^{(N\times N)}, \mathbf{x} = [x_0\ x_1 \cdots x_{N-1}]$ are the data symbols to be fingerprinted and transmitted, $\mathbf{W}$ is a $\mathcal{C}^{(N\times N)}$ DFT-matrix with elements $W_N^{nk} = (1/\sqrt{N})e^{-j2\pi(nk)/(N)}$, using row index $n = 0,\ldots,N-1$ and column index $k = 0,\ldots,N-1$, the vector $\mathbf{g} = [g_0\ g_1 \cdots g_{N-1}]$ is the sampled channel impulse response, where each element $g_i \in \mathbf{g}$ is defined as $g_i = \sum_c A_c e^{j(\pi/N)(i+(N-1)\tau_c)}(\sin(\pi\tau_c))/(\sin((\pi/N)(\tau_c - i)))$, and $\mathbf{n} = [n_0\ n_1 \cdots n_{N-1}]$ is the frequency-domain representation of complex Gaussian noise.

To recover the data transmission, the receiver must estimate the channel response $\mathbf{g}$ or its frequency-domain equivalent $\mathbf{h} = \mathbf{gW}$. A number of channel estimation techniques have been considered for OFDM systems, including the minimum mean-squared error (MMSE) estimator, and the least-squares (LS) estimator. These estimators, with some improvements, are discussed in [14]. A discussion of particular channel estimation techniques is beyond the scope of this paper, therefore without loss of generality, we use the least-square (LS) channel estimator [14] in Sections IV and V due its widespread adoption in OFDM systems and low computational complexity.

We now augment the OFDM transmission system with our extrinsic fingerprint function $\mathbf{f} = [f_0\ f_1 \cdots f_{N-1}]^T$, and its matrix equivalent $\mathbf{F}^{(N\times N)} = \mathrm{diag}([f_0\ f_1 \cdots f_{N-1}])$. The transmitted OFDM symbol of (2) with fingerprinting function applied after the modulating IDFT matrix becomes

$$\mathbf{y} = \mathbf{gWFX} + \mathbf{n} = \mathbf{u}_{\mathbf{x},\mathbf{f},\mathbf{g}} + \mathbf{n} \tag{3}$$

where $\mathbf{u}_{\mathbf{x},\mathbf{f},\mathbf{g}}$ is the received noiseless OFDM transmission.

To facilitate channel estimation, we use pilot-aided channel estimation, where pilot signals are periodically embedded into the transmitted signal. To represent periodic preamble and pilot signals, we expand the OFDM symbol transmission system above (3) to a block-based system consisting of M consecutive OFDM symbol vectors in time. The resulting time-frequency block (TFB) received by the receiver at time index $t = kM$ is represented by the matrix $\mathbf{U}[k] \in \mathcal{C}^{(N\times M)}$, where each column of $\mathbf{U}[k]$ is an OFDM symbol vector $\mathbf{u}^m_{\mathbf{x},\mathbf{f},\mathbf{g}}$ received at time $(k-1)M + m$, $m = 0, 1, \ldots, M-1$. The TFB received at time index k becomes $\mathbf{Y}[k]$, i.e.,

$$\mathbf{Y}[k] = \mathbf{U}[k] + \mathbf{N}[k] \tag{4}$$

with

$$\mathbf{U}[k] = \left[\left\{\mathbf{u}^0_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T \; \left\{\mathbf{u}^1_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T \; \ldots \left\{\mathbf{u}^{M-1}_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T\right] \tag{5}$$

and $\mathbf{N}[k] = [\mathbf{n}_0^T \; \mathbf{n}_1^T \; \ldots \; \mathbf{n}_{M-1}^T]_k$ are the noise vectors from (2).

If we assume the elements of $\mathbf{g}$ to be independent Rayleigh block-stationary and quasi-static, then $\mathbf{g}$ in (3) remains constant over the entire TFB for a total of M symbols. Similarly, if the fingerprinting function is also designed to be block-stationary, then (4) can be written as

$$\mathbf{Y}[k] = \mathbf{H}[k]\mathbf{F}[k]\mathbf{X}[k] + \mathbf{N}[k] \tag{6}$$

where $\mathbf{H}[k] = \mathrm{diag}(\mathbf{h}[k])$, $\mathbf{F}[k] = \mathrm{diag}(\mathbf{f}[k])$, $\mathbf{f}[k]$ is the fingerprinting function applied to the entire TFB, and $\mathbf{g}[k]$ is the block-stationary channel response experienced by the received TFB, $\mathbf{Y}[k]$.

We construct $\mathbf{X}[k]$ as a composite signal composed of two components: the user-data signal and the embedded preamble and pilot signals used for channel estimation and equalization. Such a scheme enables the channel-like fingerprints to be detected using the known pilot signals. The frame preamble occupying $M - L$ time slots is followed by a section containing user-data symbols mixed with pilot signals occupying the remaining L time slots [15]. The TFB to be transmitted, augmented with embedded pilots signals becomes

$$\mathbf{X}[k] = \mathbf{D}[k]\mathbf{A} + \mathbf{P} \tag{7}$$

where $\mathbf{D}[k] \in \mathcal{C}^{(N\times L)}$ is the TFB data matrix, $\mathbf{A} \in \mathcal{R}^{(L\times M)}$ is the data-projection matrix, and $\mathbf{P} \in \mathcal{R}^{(N\times M)}$ is the pilot signal matrix. The data-projection and pilot matrix satisfy the following properties:

$$\begin{aligned} &\mathbf{A}\mathbf{P}^T = \mathbf{0} \in \mathcal{R}^{(L\times N)}, \\ &\mathbf{A}\mathbf{A}^T = \mathbf{I} \in \mathcal{R}^{(L\times L)}, \quad \mathbf{P}\mathbf{P}^T = \mathbf{I} \in \mathcal{R}^{(N\times N)}. \end{aligned} \tag{8}$$

The properties (8) of the data-projection matrix $\mathbf{A}$ and pilot matrix $\mathbf{P}$ essentially allow $\mathbf{A}$ to project the data component $\mathbf{D}[k]$ onto the orthogonal subspace of the pilot matrix $\mathbf{P}$, allowing for signal demodulation by means of a maximum likelihood (ML) receiver. In the simulation in Section V, we will use a time-multiplexed (TM) single symbol preamble structure [15], which is given as

$$\mathbf{A} = \left[\mathbf{0}^{(L\times 1)}; \mathbf{I}^{(L\times M-1)}\right] \; \mathbf{P} = \left[\mathbf{I}^{(N\times 1)}; \mathbf{0}^{(N\times M-1)}\right] \tag{9}$$

where $\mathbf{I}^{(\,\cdot\,)}$ and $\mathbf{0}^{(\,\cdot\,)}$ are the identity matrix and the zero matrix, respectively, with dimensionality denoted in the superscript $(\,\cdot\,)$.

The expanded form of the TFB signal at the receiver (4), using (3), (6), and (7), becomes

$$\mathbf{Y}[k] = \mathbf{Q}[k](\mathbf{D}[k]\mathbf{A} + \mathbf{P}) + \mathbf{N}[k] \tag{10}$$

where $\mathbf{Q}[k] = \mathrm{diag}(\mathbf{H}[k]\mathbf{F}[k]) = \mathrm{diag}(\mathbf{q}[k])$ is the aggregate channel-like distortion.

In [13], delineation of the intrinsic channel response and the extrinsic fingerprint signal was accomplished through an even–odd time-division delivery (TDD) scheme that implemented a differential modulation scheme for transmission of the fingerprint signal. In the even–odd transmission scheme, the fingerprint transmission $\mathbf{F}[k]$ is omitted during even block transmissions, i.e., $k = [0, 2, 4, \ldots]$, and is present during odd block transmissions, i.e., $k = [1, 3, 5, \ldots]$. To sound the physical channel, the fingerprint function is omitted by simply replacing $\mathbf{F}[k]$ with the identity matrix. This scheme allows for periodic sampling of intrinsic channel distortions, denoted $\mathbf{g}[k]$ in this work, when a generalized pilot embedding scheme, (7), (8), and (10), is employed. In Section III, we will consider a similar fingerprint transmission scheme that enables periodic channel-sounding.

In this paper, we aim to design channel-like fingerprint signaling schemes that result in minimal degradation to the primary data signal $\mathbf{D}[k]$ and the detection thereof, when the fingerprint is present.

III. Subspace Extrinsic Channel-Like Fingerprinting

In this section, we consider the subspace decomposition of CSI and describe how a fingerprint message can be embedded in the noise subspace of these channel estimates. While a perfectly embedded fingerprint will occupy only the noise subspace of the received fingerprint message, we will use the framework presented in this section to create a number of suboptimal fingerprint designs that leverage various amounts of CSI.

In essence, an optimal extrinsic fingerprint signaling scheme, embodied by $\mathbf{f}[k]$, would adapt to the intrinsic channel response $\mathbf{g}[k]$, and by means of spectral water-filling these two processes, work in consort to produce the received band-limited signal, $\mathbf{Y}[k]$, given in (10). Traditional transmission precoding by way of water-filling methods typically strive to increase the capacity of the primary-signal, represented by $\mathbf{D}[k]$, using partial CSI at the transmitter. Instead, in this work, we consider the case where any additional capacity obtained via water-filling is provisioned to the fingerprint signal which operates independently of the primary-signal.

A. Subspace Decomposition of Channel Information

We consider a sequence of P previous block-stationary channel estimate vectors $\hat{\mathbf{h}}[k]$, obtained via channel estimation and arranged as column vectors in a matrix $\hat{\mathbb{H}}[k] \in \mathcal{C}^{(N\times P)}, N \geq P$, i.e.,

$$\hat{\mathbb{H}}[k] = [\hat{\mathbf{h}}^T[k-P+1]\hat{\mathbf{h}}^T[k-P+2]\ldots\hat{\mathbf{h}}^T[k-1]]. \tag{11}$$

We note from this point forward that $\hat{\mathbb{H}}[k]$ and any derivations are based on the estimate of the true channel gain $\mathbb{H}[k] = [\mathbf{h}^T[k-P+1]\ldots\mathbf{h}^T[k-1]]$ since both the receiver and transmitter only have the information of the estimated channel gain. We consider the case where the fingerprinting function is designed using a block-based scheme such that $\mathbf{f}[k]$ in (10) is applied by the transmitter to a sequence of TFBs as a matrix denoted $\mathbb{F}[k]$. We select the length of the fingerprint block to be P, so that manipulations of one fingerprint block over P consecutive TFBs will coincide with the evolution of CSI at the receiver, i.e., $\hat{\mathbb{H}}[k]$ which is also of length P TFBs. Let the block-based fingerprinting function, $\mathbb{F}[k] \in \mathcal{C}^{(N \times P)}$, be the matrix-representation of P fingerprinting functions $\mathbf{f}^T[k]$ applied by the transmitter, such that

$$\mathbb{F}[k] = [\mathbf{f}^T[k-P+1]\mathbf{f}^T[k-P+2] \dots \mathbf{f}^T[k-1]]. \tag{12}$$

For the fingerprinting transmission scheme described by (10), we arrange the aggregate block-stationary channel estimate vectors $\hat{\mathbf{q}}[k]$ that are observed by the receiver as columns in $\hat{\mathbb{Q}}[k]$, which are related to $\hat{\mathbb{H}}[k]$ and $\mathbb{F}[k]$ via the Hadamard product, i.e.,

$$\hat{\mathbb{Q}}[k] = \hat{\mathbb{H}}[k] \circ \mathbb{F}[k] \tag{13}$$

where $(\circ)$ represents the Hadamard product. The intrinsic time-varying channel measurement $\hat{\mathbb{H}}[k]$ is easily obtained by omitting the fingerprint component and replacing the fingerprinting function $\mathbf{f}[k]$ with the ones vector $\mathbf{1}^{(1 \times N)}$. Thus, $\mathbb{F}[k]$ in (13) becomes the identity matrix for the Hadamard product, which is the ones matrix $\mathbf{1}^{(N \times P)}$. We denote this process as the *channel-sounding phase* with accompanying channel-sounding fingerprint function $\mathbb{F}_{snd}[k]$. The aggregate distortion perceived by the receiver during the channel-sounding phase, denoted $\hat{\mathbb{Q}}_{snd}[k]$, is simply

$$\hat{\mathbb{Q}}_{snd}[k] = \hat{\mathbb{H}}[k] \circ \mathbb{F}_{snd}[k] = \hat{\mathbb{H}}[k] \circ \mathbf{1}^{(N \times P)} = \hat{\mathbb{H}}[k]. \tag{14}$$

CSI used by the transmitter must first be estimated by the receiver and then conveyed to the transmitter as feedback, resulting in a delay. Thus, we denote CSI obtained by the transmitter during the channel-sounding phase as $\hat{\mathbb{Q}}_{snd}[l-\epsilon] = \hat{\mathbb{H}}[l-\epsilon]$, where ϵ is the number of OFDM symbols of delay experienced by the channel-sounding CSI, as received by the transmitter, and the current fingerprint transmission at time $l = \lfloor k/P \rfloor$. The diagonalization of $\hat{\mathbb{Q}}_{snd}[l-\epsilon]$ in (14) via singular-value decomposition (SVD) for the case $N \geq P$ yields

$$\hat{\mathbb{H}}^T[l-\epsilon] = \hat{\mathbb{Q}}^T_{snd}[l-\epsilon] = \mathbb{U}[l-\epsilon]\mathbb{S}[l-\epsilon]\mathbb{V}^H[l-\epsilon] \tag{15}$$

where $\mathbb{U}[l-\epsilon] \in \mathcal{C}^{(P \times P)}$ is the *left* unitary matrix of the decomposition with orthonormal columns representing the *left* singular-vectors of $\hat{\mathbb{H}}[l-\epsilon]$, and $\mathbb{V}[l-\epsilon] \in \mathcal{C}^{(N \times N)}$ contains the *right* orthonormal columns of the singular-vectors of $\hat{\mathbb{H}}[l-\epsilon]$, and $\mathbb{S}[l-\epsilon] \in \mathcal{R}^{(P \times N)}$ contains the diagonalized eigenvalues of the decomposition. The column vectors of $\mathbb{U}[l-\epsilon]$ are the projections of the eigenvectors of $\hat{\mathbb{H}}[l-\epsilon]$ in the time dimension for the previous P TFBs, while $\mathbb{V}[l-\epsilon]$ are the projections of the eigenvectors of $\hat{\mathbb{H}}[l-\epsilon]$ in the frequency dimension.

According to our time-varying channel model, we surmise that linearly correlated time-varying channel components can be separated from the uncorrelated noise components via subspace decomposition, if P is large enough [16]. Hence, $\mathbb{U}[l-\epsilon]$ and $\mathbb{V}[l-\epsilon]$ form a basis describing the linear dependencies in $\hat{\mathbb{H}}[l-\epsilon]$ in both time and frequency. We consider the case where P is selected to be large enough so that the decomposition of $\hat{\mathbb{H}}[l-\epsilon]$ is over-determined. Via subspace decomposition we bifurcate each matrix of (15) into two subspaces:

1) A correlated time-varying *signal subspace* characterizing the linear dependencies in $\hat{\mathbb{H}}[l-\epsilon]$ spanning both time and frequency, thus forming a temporal-spectral model for the linearly correlated components present in previous CSI estimates.
2) An uncorrelated *noise subspace*, also spanning $\hat{\mathbb{H}}[l-\epsilon]$ in both time and frequency, and thus characterizing the noise component present in previous CSI.

The correlated signal subspace will also be referred to as the *intrinsic* distortion subspace, as this subspace characterizes the principle linear relationships between consecutive channel estimates intrinsic to the linear time-varying fading channel. The bifurcation of $\mathbb{U}[l-\epsilon]$, $\mathbb{S}[l-\epsilon]$, and $\mathbb{V}[l-\epsilon]$ into principal components and noise, according to a model parameter p, becomes

$$\begin{aligned}\hat{\mathbb{H}}^T[l-\epsilon] &= \{\mathbb{U}\mathbb{S}\mathbb{V}^H\}_{l-\epsilon} \\ &= \left[\mathbb{U}_1^{(P \times p)} \quad \mathbb{U}_2^{(P \times (P-p))}\right]_{l-\epsilon} \\ &\quad \times \left[\mathbf{\Lambda}^{(P \times P)} \quad \mathbf{0}^{(P \times (N-P))}\right]_{l-\epsilon} \\ &\quad \times \left[\mathbb{V}_1^{(N \times p)} \quad \mathbb{V}_2^{(N \times (N-p))}\right]^H_{l-\epsilon}\end{aligned} \tag{16}$$

where the columns of $\mathbb{U}_1$ and $\mathbb{V}_1$ represent the singular-vectors of the left and right unitary matrices $\mathbb{U}$ and $\mathbb{V}$, respectively, spanning the correlated time-varying signal subspace, $\mathbb{U}_2$ and $\mathbb{V}_2$ represent the singular-vectors of $\mathbb{U}$ and $\mathbb{V}$, respectively, spanning the noise subspace, and $\mathbf{\Lambda} \in \mathcal{R}^{(P \times P)} = \mathrm{diag}(\lambda_0, \lambda_1, \dots, \lambda_{P-1})$, are the eigenvalues $\lambda_i, i = 0, 1, \dots, P-1$ of $\mathbb{S}$. The dimensionality of these subspaces, determined by p, should be chosen to be equal to the effective rank of $\hat{\mathbb{H}}[l-\epsilon]$, for optimal bifurcation of the subspaces. That is, if we sort the eigenvectors of $\mathbb{S}[l-\epsilon]$ according to magnitude in descending order, i.e., $\lambda_0 \geq \lambda_i \geq \lambda_{P-1}, i = 1, \dots, P-2$, the effective rank of $\mathbb{H}[l-\epsilon]$, and thus the optimal value for p, is equal to the number of eigenvalues of $\mathrm{diag}(\mathbf{\Lambda})$ that are not equal to σ_H^2, i.e.,

$$\begin{gathered}\lambda_0 \geq \lambda_i \geq \lambda_j = \sigma_H^2, \\ i = 1, \dots, p-1, \quad j = p, \dots, P-1\end{gathered} \tag{17}$$

where σ_H^2 is the noise power of $\mathbf{N}[l-\epsilon]$ projected onto the pilot matrix $\mathbf{P}$. In practice, the eigenvalues of the noise subspace may not all be equal, making the estimation of the effective rank of $\hat{\mathbb{H}}[l-\epsilon]$ difficult. However, these values will be very close to σ_H^2 [17]. Additionally, the spectrum of the signal spanned by the noise subspace is orthogonal to the intrinsic channel disturbance spectrum. We will exploit this orthogonality property later in our fingerprinting designs to be discussed in Section III-B. Various criteria have been proposed to estimate p in these cases [17]; however, this discussion is beyond the scope of this paper. For the sake of exposition, we will assume that p is perfectly selected to be the effective rank of $\mathbb{H}[l-\epsilon]$, $\lambda_j = \sigma_H^2, j = p, \dots, P-1$, and note that improper estimation of p will result in degraded performance as the orthogonality between the two subspaces is degraded in this case.

For sufficiently stationary and nontrivial time-varying channels, the channel estimates $\hat{\mathbf{h}}[k-i], i = 1, \dots, P+1$ obtained using embedded pilot signals are correlated in both frequency and in time. Thus, when P is properly selected, the resulting decomposition of $\mathbb{H}[l-\epsilon]$ in (15) will be over-determined and, therefore, $p < P$ and the size of both subspaces will be nonzero. For the case $p \geq P$, the system is under-determined and an accurate delineation of both subspaces in (16) is not possible. Therefore, for the sake of exposition, we consider here the case

$$1 \leq p < P \quad \text{where} \quad P \geq 2 \tag{18}$$

which is the most interesting case for our extrinsic fingerprinting method, as both the intrinsic time-varying channel distortion subspace and the noise subspace are of nonzero size.

B. Fingerprint Design Using Subspace Modeling

We now describe how the fingerprinting function $\mathbf{F}[k]$ is embedded into the transmission to produce a time-frequency fingerprint that leverages knowledge of the channel, $\mathbf{H}[k]$.

According to a spectral water-filling observation, the capacity of a sequence of channel estimate data, i.e., the amount of information conveyed by $\hat{\mathbb{Q}}[k]$ to the receiver using the embedded pilot signals $\mathbf{P}$ of (7) to drive $\hat{\mathbb{Q}}[k]$, can be maximized by introducing a Gaussian process with energy where the spectrum of $\mathbb{H}[k]$ is lowest. The transmitter can first estimate the noise subspace of $\mathbb{H}[k]$ by decomposing $\hat{\mathbb{H}}[k]$ as in (17), in which $\mathrm{diag}(\mathbf{\Lambda})$ denotes spectral magnitudes and the smallest eigenvalues represent the power spectrum of the noise subspace of $\hat{\mathbb{H}}[k]$. We introduce our fingerprinting signal using a *fingerprint overlay* onto these parts of the spectrum.

Hence, the fingerprint function should be in the noise subspace, and the transmitter has to ensure that the receiver can detect the fingerprint from the aggregate channel estimate $\hat{\mathbb{Q}}[k]$. To aid in analysis, we split the subspace decomposition of the CSI in possession by the transmitter (16) into a summation of significant channel gains and noise, i.e.,

$$\hat{\mathbb{H}}^T[l-\epsilon] = \hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon] \tag{19}$$

where $\hat{\mathbb{L}}[l-\epsilon]$ is an estimate of the significant intrinsic channel gains $\mathbb{L}[l-\epsilon]$ on the signal subspace, and $\hat{\mathbb{N}}[l-\epsilon]$ is an estimate of noise on the noise subspace $\mathbb{N}[l-\epsilon]$, of $\mathbb{H}[l-\epsilon]$. From this definition, the principle linear components of the intrinsic time-varying channel fading patterns are parameterized by the basis $\mathbb{L}[l]$. The definition of $\mathbb{L}[l-\epsilon]$ and $\mathbb{N}[l-\epsilon]$ according to (16) and (17) is

$$\begin{aligned}\hat{\mathbb{L}}[l-\epsilon] = {} & [\,\mathbb{U}_1 \quad \mathbf{0}^{(P\times(P-p))}\,]_{l-\epsilon} \\ & \times \left[\mathrm{diag}\left(\boldsymbol{\lambda}^{(p)}, \mathbf{0}^{(P-p)}\right) \quad \mathbf{0}^{((N-p)\times P)}\right]_{l-\epsilon} \\ & \times \begin{bmatrix} \mathbb{V}_1^H \\ \mathbf{0}^{((N-p)\times N)} \end{bmatrix}_{l-\epsilon}\end{aligned} \tag{20}$$

where $\boldsymbol{\lambda}[l-\epsilon]$ is defined as

$$\boldsymbol{\lambda}[l-\epsilon] = [\lambda_0\lambda_1\ldots\lambda_{p-1}]_{l-\epsilon} \tag{21}$$

and

$$\begin{aligned}\hat{\mathbb{N}}[l-\epsilon] = {} & [\,\mathbf{0}^{(P\times p)} \quad \mathbb{U}_2,\,]_{l-\epsilon} \\ & \times [\,\mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\sigma}^{(P-p)}\right) \quad \mathbf{0}^{(P\times(N-P))}\,]_{l-\epsilon} \\ & \times \begin{bmatrix} \mathbf{0}^{(p\times N)} \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon}\end{aligned} \tag{22}$$

where $\boldsymbol{\sigma}[l-\epsilon]$ is defined as

$$\begin{aligned}\boldsymbol{\sigma}[l-\epsilon] &= [\lambda_p\lambda_{p+1}\ldots\lambda_{P-1}]_{l-\epsilon}, \\ \lambda_p &= \lambda_{p+1} = \cdots = \lambda_{P-1} = \sigma_H^2.\end{aligned} \tag{23}$$

This low-rank modeling of intrinsic channel conditions by $\mathbb{L}[l]$ will help reduce feedback overhead when conveying CSI to the transmitter, as this feedback decreases system efficiency.

1) Subspace Fingerprint Design: The transmitter designs $\mathbb{F}[l]$, which will be recovered by the receiver using the estimate of the aggregate channel, $\hat{\mathbb{Q}}[l]$. Since $\hat{\mathbb{Q}}[l] = \mathbb{F}[l] \circ \hat{\mathbb{H}}[l]$, for the transmitter, designing $\mathbb{F}[l]$ is the same as designing $\hat{\mathbb{Q}}[l]$ if the transmitter has the current channel estimate, $\hat{\mathbb{H}}[l]$. However, the transmitter possesses only a delayed version of the channel estimate data, $\hat{\mathbb{H}}[l-\epsilon]$. If prediction of future channel state is not employed, the transmitter must approximate $\hat{\mathbb{H}}[l]$ using $\hat{\mathbb{H}}[l-\epsilon]$. Therefore, the transmitter will approximate the aggregate channel estimate $\hat{\mathbb{Q}}[l]$ at the receiver's side by $\hat{\mathbb{Q}}_{\mathrm{des}}[l]$, where $\hat{\mathbb{Q}}_{\mathrm{des}}[l] = \mathbb{F}[l] \circ \hat{\mathbb{H}}[l-\epsilon]$.

Similarly, we can decompose $\hat{\mathbb{Q}}_{\mathrm{des}}[l]$ using a summation model

$$\hat{\mathbb{Q}}_{\mathrm{des}}[l] = \mathbb{P}[l] + \mathbb{K}[l] \tag{24}$$

where $\mathbb{P}[l]$ is the projection of $\hat{\mathbb{Q}}_{\mathrm{des}}[l]$ onto the intrinsic channel subspace corresponding to $\hat{\mathbb{L}}[l]$ in (19), and $\mathbb{K}[l]$ is the extrinsic fingerprinting overlay matrix that we will design to overlay the noise component $\hat{\mathbb{N}}[l]$ in (19).

Consequently, the transmitter will design $\mathbb{F}[l]$ according to $\hat{\mathbb{Q}}_{\mathrm{des}}[l]$, such that

$$\mathbb{F}[l] = (\mathbb{P}[l] + \mathbb{K}[l]) \circ (\hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon])^{(-1)} \tag{25}$$

where $(\,\cdot\,)^{(-1)}$ is the Hadamard inverse operation.

2) Fingerprint Extraction: The aggregate channel estimate $\hat{\mathbb{Q}}[l]$ that the receiver will obtain can be formulated by substituting (25) into (13), i.e., $\hat{\mathbb{Q}}[l]$, becomes

$$\begin{aligned}\hat{\mathbb{Q}}[l] &= \hat{\mathbb{H}}[l] \circ (\mathbb{P}[l] + \mathbb{K}[l]) \circ (\hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon])^{(-1)} \\ &= (\hat{\mathbb{L}}[l] \circ \mathbb{P}[l] + \hat{\mathbb{L}}[l] \circ \mathbb{K}[l] + \hat{\mathbb{N}}[l] \circ \mathbb{P}[l] + \hat{\mathbb{N}}[l] \circ \mathbb{K}[l]) \\ &\qquad \circ (\hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon])^{(-1)}.\end{aligned} \tag{26}$$

We will now show that $\hat{\mathbb{Q}}[l]$ is an unbiased estimate of $\hat{\mathbb{Q}}_{\mathrm{des}}[l] = \hat{\mathbb{P}}[l] + \hat{\mathbb{K}}[l]$ enabling recovery of the fingerprint $\hat{\mathbb{F}}[l]$ by the receiver, without bias. Equivalently, we want to show that $E[\hat{\mathbb{Q}}[l]] = \mathbb{P}[l] + \mathbb{K}[l]$.

Assuming that $\hat{\mathbb{L}}[l]$ and $\hat{\mathbb{L}}[l-\epsilon]$ are unbiased estimates of their respective channel gain components, i.e.,

$$E[\hat{\mathbb{L}}[l]] = \mathbb{L}[l] \quad \text{and } E[\hat{\mathbb{L}}[l-\epsilon]] = \mathbb{L}[l-\epsilon] \tag{27}$$

then the expectation of (26) yields

$$\begin{aligned}E[\hat{\mathbb{Q}}[l]] = {} & (\mathbb{L}[l] \circ \mathbb{P}[l] + \mathbb{L}[l] \circ \mathbb{K}[l] \\ & + E[\hat{\mathbb{N}}[l]] \circ \mathbb{P}[l] + E[\hat{\mathbb{N}}[l]] \circ \mathbb{K}[l]) \\ & \circ (\hat{\mathbb{L}}[l-\epsilon] + E[\hat{\mathbb{N}}[l-\epsilon]])^{(-1)}.\end{aligned} \tag{28}$$

In the derivation of (28), we recall that $(\circ)$ is the Hadamard product, therefore regular matrix multiplication and inversion is not used in this result.

We recall that $\hat{\mathbb{N}}[l]$ is the projection of $\mathbf{N}[k]$ on the noise subspace of channel estimates $\hat{\mathbb{H}}[l]$. As $\mathbf{N}[k]$ is a matrix of zero-mean Gaussian random variables and the basis of the noise subspace is formed from Gaussian random variables $\mathbf{N}[k]$ projecting on the pilot signals $\mathbf{P}$, then $\hat{\mathbb{N}}[l]$ is also Gaussian with each element having zero mean, i.e., $E[\hat{\mathbb{N}}[l]] = \mathbf{0}$. The elements of $\mathbf{N}[k]$ are also i.i.d. Gaussian, thus $E[\hat{\mathbb{N}}[l-\epsilon]]$ is also zero-mean Gaussian and $\hat{\mathbb{N}}[l-\epsilon]$ is uncorrelated with $\hat{\mathbb{N}}[l]$, therefore, by a similar argument $E[\hat{\mathbb{N}}[l-\epsilon]] = \mathbf{0}$, and (28) becomes

$$E[\hat{\mathbb{Q}}[l]] = (\mathbb{P}[l] + \mathbb{K}[l]) \circ \mathbb{L}[l] \circ \mathbb{L}[l-\epsilon]^{(-1)}. \tag{29}$$

The above equation demonstrates that if the intrinsic channel is stationary over ϵ blocks, i.e., $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, then the aggregate channel estimate at the receiver's side $\hat{\mathbb{Q}}[l]$ is an unbiased estimate of the information that the transmitter conveyed, $\hat{\mathbb{Q}}_{\mathrm{des}}[l]$.

Moreover, to obtain an estimate for only the extrinsic fingerprinting overlay $\mathbb{K}[l]$ from $\hat{\mathbb{Q}}[l]$, in (29) we immediately see that our estimate $\hat{\mathbb{Q}}[l]$ will be biased by $\mathbb{P}[l]$. By a water-filling argument, the extrinsic fingerprinting signal should contribute energy to the noise subspace of $\mathbb{H}[l]$ to maximize the information conveyed by $\mathbb{H}[l]$ to the receiver, as this basis represents the spectral elements of $\mathbb{H}[l]$ with the lowest energy. Thus, we design $\mathbb{F}[l]$ to contribute only to the noise subspace while leaving the intrinsic channel subspace unperturbed by setting $\mathbb{P}[l]$ to be the identity matrix for the Hadamard product, i.e.,

$$\mathbb{P}[l] = \mathbf{1}. \tag{30}$$

With (30), (29) becomes

$$\begin{aligned} E[\hat{\mathbb{Q}}[l]] &= \mathbb{L}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon] + \mathbb{L}[l] \circ \mathbb{K}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon] \\ &= (\mathbf{1} + \mathbb{K}[l]) \circ \mathbb{M}[l] \end{aligned} \tag{31}$$

where we introduce the definition

$$\mathbb{M}[l] = \mathbb{L}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon]. \tag{32}$$

In (32), $\mathbb{M}[l]$ is the intrinsic-channel model mismatch error in estimating $\mathbb{K}[l]$, and is the Hadamard product between $\mathbb{L}[l]$ and the Hadamard inverse of the previous intrinsic channel estimate, $\mathbb{L}^{(-1)}[l-\epsilon]$. From this definition, the model error mismatch matrix, $\mathbb{E}[l]$, is simply $\mathbb{E}[l] = \mathbf{1}^{(N \times P)} - \mathbb{M}[l]$.

From (32), we readily see that when the low-rank subspace approximation of $\mathbb{H}[l]$, $\mathbb{L}[l-\epsilon]$, is a perfect match of the low-rank approximation of the current channel conditions $\mathbb{L}[l]$, then $\mathbb{M}[l] = \mathbf{1}$, and by removing the bias introduced by (30), (31) becomes

$$E[\hat{\mathbb{Q}}[l]] - \mathbf{1} = \mathbb{K}[l] \tag{33}$$

thus an unbiased estimate for $\mathbb{K}[l]$ can be obtained from $\hat{\mathbb{Q}}[l]$. When the error matrix $\mathbb{E}[l]$ has nonzero elements, additional model mismatch error will result in degraded performance when detecting the fingerprint signal.

Using (30), the design of the fingerprinting function $\mathbb{F}[l]$ from (25) becomes simply

$$\mathbb{F}[l] = \mathbb{K}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon]. \tag{34}$$

We note from (34) that either the transmitter or the receiver may apply $\mathbb{L}^{(-1)}[l-\epsilon]$ before recovering $\mathbb{K}[l]$ as the Hadamard product is commutative. The case where $\mathbb{L}^{(-1)}[l-\epsilon]$ is applied by the transmitter is analogous to linear OFDM block precoding [18], for the purpose of channel fade mitigation. If we assume that the receiver has memory and can store $\mathbb{L}^{(-1)}[l-\epsilon]$ for future computation, this would allow the receiver to preform this computation, eliminating the need to transmit $\mathbb{L}^{(-1)}[l-\epsilon]$ to the transmitter, thus decreasing the amount of CSI feedback required and reducing overhead.

C. Subspace Fingerprinting Overlays

We now discuss a methodology for designing the extrinsic fingerprint overlay $\mathbb{K}[l]$ that will allow the authentication signal to overlay the noise subspace $\mathbb{N}[l]$ of $\mathbb{H}[l]$. Additionally, we demonstrate how $\mathbb{K}[l]$ can be used to modulate the extrinsic fingerprinting signal.

Similar to the definition of $\hat{\mathbb{L}}[l]$ in (20) and $\hat{\mathbb{N}}[l]$ in (22), according to (16) and (17) we define the extrinsic fingerprinting overlay matrix $\mathbb{K}[l]$ as

$$\begin{aligned} \mathbb{K}^T[l] &= \mathbb{U}_{l-\epsilon} \mathbb{S}_l \mathbb{V}_{l-\epsilon}^H \\ &= \begin{bmatrix} \mathbf{0}^{(P \times p)} & \mathbb{U}_2 \end{bmatrix}_{l-\epsilon} \\ &\quad \times \begin{bmatrix} \mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}^{(P \times (N-P))} \end{bmatrix}_l \\ &\quad \times \begin{bmatrix} \mathbf{0}^{(N \times p)} \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon} \end{aligned} \tag{35}$$

where $\boldsymbol{\xi}[l] \in \mathcal{R}^{((P-p) \times (P-p))}$ is defined as

$$\boldsymbol{\xi}[l] = [\xi_p \xi_{p+1} \ldots \xi_{P-1}]_l, \quad \xi_p, \xi_{p+1}, \ldots, \xi_{P-1} > 0 \tag{36}$$

and $\mathbb{U}_2$ and $\mathbb{V}_2$ are the left and right singular bases, respectively, that are constructed using CSI obtained during the channel-sounding phase at time $l-\epsilon$, and will be used as an orthonormal basis to signal our extrinsic fingerprinting function. Using this signaling basis, the vector $[\xi_p \xi_{p+1} \ldots \xi_{P-1}]_l$ will convey the extrinsic fingerprint message to the receiver.

In general, the channel stationary conditions will not hold, thus some model mismatch error between $\mathbb{L}[l-\epsilon]$ and $\mathbb{L}[l]$ will occur. This model mismatch error manifests itself as $\mathbb{M}[l]$, defined in (32). We will first consider system design and performance using the assumption of quasi-stationary behavior between $\mathbb{L}[l-\epsilon]$ and $\mathbb{L}[l]$ without attempting to predict $\mathbb{L}[l]$.

D. Fingerprint Recovery and Modulation

We now describe how the fingerprint signal vector $[\xi_p \ \xi_{p+1} \ldots \xi_{P-1}]_l$ may be recovered from $\mathbb{K}[l]$, of which $\hat{\mathbb{Q}}[l] - \mathbf{1}$ is an unbiased estimate. Substituting (35) into (31) we obtain

$$\begin{aligned} E[\mathbb{Q}^T[l]] - \mathbf{1} &= (\mathbf{1} + \mathbb{K}^T[l]) \circ (\mathbf{1} - \mathbb{E}^T[l]) - \mathbf{1} \\ &= \begin{bmatrix} \mathbf{0}^{(P \times p)} & \mathbb{U}_2 \end{bmatrix}_{l-\epsilon} \\ &\quad \times \begin{bmatrix} \mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}^{(P \times (N-P))} \end{bmatrix}_l \\ &\quad \times \begin{bmatrix} \mathbf{0}^{(p \times N)} \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon} - \mathbb{E}^T[l] - \mathbb{K}^T[l] \circ \mathbb{E}^T[l] \end{aligned} \tag{37}$$

where the terms $\mathbb{E}[l]$ and $(\mathbb{K}[l] \circ \mathbb{E}[l])$ represent model mismatch error in (37). These model mismatch error terms degrade the performance of the fingerprint detector. To recover the fingerprint signal $[\xi_p \xi_{p+1} \ldots \xi_{P-1}]_l$, we must first estimate $\mathbb{S}[l]$ from (35). To produce the estimate $\hat{\mathbb{S}}[l]$, the receiver premultiplies (37) by $[\,\mathbf{0} \quad \mathbb{U}_2\,]_{l-\epsilon}^H$ and postmultiplies by $[\,\mathbf{0} \quad \mathbb{V}_2\,]_{l-\epsilon}$, and (37) becomes

$$\begin{aligned}\hat{\mathbb{S}}[l] &= \begin{bmatrix}\mathbf{0}\\ \mathbb{U}_2^H\end{bmatrix}_{l-\epsilon} (E[\mathbb{Q}^T[l]] - \mathbf{1})[\,\mathbf{0} \quad \mathbb{V}_2\,]_{l-\epsilon}\\ &= \mathbb{R}_u[l-\epsilon]\left[\operatorname{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}\right]_l \mathbb{R}_v[l-\epsilon] - \mathbb{B}[l]\end{aligned} \tag{38}$$

where we define the fingerprint model mismatch error component $\mathbb{B}[l]$

$$\mathbb{B}[l] = [\,\mathbf{0} \quad \mathbb{U}_2\,]_{l-\epsilon}^H (\mathbb{K}[l] \circ \mathbb{E}[l] - \mathbb{E}[l])^T [\,\mathbf{0} \quad \mathbb{V}_2\,]_{l-\epsilon} \tag{39}$$

and

$$\mathbb{R}_u[l-\epsilon] = \begin{bmatrix}\mathbf{0}\\ \mathbb{U}_2^H\end{bmatrix}_{l-\epsilon} [\,\mathbf{0} \quad \mathbb{U}_2\,]_{l-\epsilon} \tag{40}$$

and

$$\mathbb{R}_v[l-\epsilon] = \begin{bmatrix}\mathbf{0}\\ \mathbb{V}_2^H\end{bmatrix}_{l-\epsilon} [\,\mathbf{0} \quad \mathbb{V}_2\,]_{l-\epsilon}. \tag{41}$$

The left and right correlation matrices, $\mathbb{R}_u[l-\epsilon]$ and $\mathbb{R}_v[l-\epsilon]$, respectively, can be used as a measure of the *closeness* of the left and right extrinsic signaling basis $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, to a true unitary basis for their respective subspaces. An optimal selection of extrinsic fingerprinting bases, yielding the correlation matrices $\mathbb{R}_u^*[l-\epsilon]$ and $\mathbb{R}_v^*[l-\epsilon]$, respectively, would preserve the orthogonality of the intrinsic and extrinsic subspaces. Thus, an optimal selection of bases would yield $\mathbb{R}_u^*[l-\epsilon] = [\,\mathbf{0}^{(P\times p)} \quad \mathbf{I}^{(P\times P-p)}\,]$ and $\mathbb{R}_v^*[l-\epsilon] = [\,\mathbf{0}^{(N\times p)} \quad \mathbf{I}^{(N\times N-p)}\,]$, respectively. We define deviation from a true orthonormal signaling basis for the left and right bases, e_{R_u} and e_{R_v}, respectively, as the Frobenius norm of the difference between $\mathbb{R}_v[l-\epsilon]$ and $[\,\mathbf{0}^{(P\times p)} \quad \mathbf{I}^{(P\times P-p)}\,]$, and $\mathbb{R}_u[l-\epsilon]$ and $[\,\mathbf{0}^{(N\times p)} \quad \mathbf{I}^{(N\times N-p)}\,]$, respectively, i.e.,

$$e_{R_u} = \|\mathbb{R}_u - [\,\mathbf{0}^{(P\times p)} \quad \mathbf{I}^{(P\times P-p)}\,]\|_F \tag{42}$$

and

$$e_{R_v} = \|\mathbb{R}_v - [\,\mathbf{0}^{(N\times p)} \quad \mathbf{I}^{(N\times N-p)}\,]\|_F \tag{43}$$

which are both nonnegative values.

When both $\mathbb{R}_v[l-\epsilon]$ and $\mathbb{R}_u[l-\epsilon]$ are perfectly unitary, a condition which we will denote with the subscripts (ru) and (rv), respectively, (38) becomes simply

$$\begin{aligned}\hat{\mathbb{S}}_{ru,rv}[l] &= \begin{bmatrix}\mathbf{0}\\ \mathbb{U}_2^H\end{bmatrix}_{l-\epsilon} [\,\mathbf{0} \quad \mathbb{U}_2\,]_{l-\epsilon} \left[\operatorname{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}\right]_l\\ &\quad \times \begin{bmatrix}\mathbf{0}\\ \mathbb{V}_2\end{bmatrix}_{l-\epsilon} [\,\mathbf{0} \quad \mathbb{V}_2^*\,]_{l-\epsilon} - \mathbb{B}[l]\\ &= \left[\operatorname{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}\right]_l - \mathbb{B}[l].\end{aligned} \tag{44}$$

From (44), we note that the channel model mismatch error term $\mathbb{B}[l]$ is still present. However, under perfect channel estimation conditions when $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, a condition which we will denote with the subscript (m), the channel model-mismatch term becomes the zero matrix and no model mismatch error is present. Thus (44) is simply

$$\hat{\mathbb{S}}_{ru,rv,m}[l] = \mathbb{S}[l] = \left[\operatorname{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_l \tag{45}$$

and the extrinsic fingerprint signal of a fingerprinted block transmitted at time l may be recovered from $\hat{\mathbb{S}}[l]$ by simply extracting the elements $[\xi_p \xi_{p+1} \ldots \xi_{P-1}]$ using (45) and (36). From (44), we observe that any model mismatch will degrade the fingerprint statistics $[\xi_p \xi_{p+1} \ldots \xi_{P-1}]$ as interference.

E. Extrinsic Fingerprint Overlay Design

In this section, we evaluate various methodologies for incorporating previous CSI into the design of the left and right signaling bases $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, and discuss the performance trade-offs of these designs. The design taxonomy we present will be ordered according to the amount of CSI required, in descending order. Therefore, we will lead our discussion with designs that require the greatest amount of CSI, and end our discussion with designs that do not require CSI at the transmitter at all.

1) Direct Fingerprint Overlay Using Full CSI: We first consider the possibility of directly using the left and right singular vectors $\mathbb{U}_2$ and $\mathbb{V}_2$ from (35) to implement $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, and denote this design $\mathbb{K}_{\text{direct}}$, i.e.,

$$\begin{aligned}&\mathbb{K}_{\text{direct}}[l] = [\,\mathbf{0}^{(P\times p)} \quad \mathbb{U}_2^H\,]_{l-\epsilon}\\ &\times \left[\operatorname{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_l \begin{bmatrix}\mathbf{0}^{(p\times N)}\\ \mathbb{V}_2^H\end{bmatrix}_{l-\epsilon}.\end{aligned} \tag{46}$$

While $\mathbb{K}_{\text{direct}}$ demonstrates that $\mathbb{U}_2$ and $\mathbb{V}_2$ can be used directly to implement an orthonormal basis for signaling $[\xi_p \;\, \xi_{p+1} \ldots \xi_{P-1}]$, there are a number of downfalls to this approach. To recover $\mathbb{K}_{\text{direct}}$ via (38), both $\mathbb{U}_2$ and $\mathbb{V}_2$ must be communicated to the transmitter from the receiver, requiring a total of $P(P-p) + N(N-p)$ units of feedback information. Also, if $\mathbb{V}_2[l-\epsilon]$ accurately models the noise subspace in the frequency dimension under particular channel conditions while the singular vectors of $\mathbb{U}_2[l-\epsilon]$ inaccurately model the noise subspace in the time dimension, $\mathbb{U}_2[l-\epsilon]$ will predominately contribute to model mismatch error component, $\mathbb{E}[l]$. This would be the case when $\mathbb{U}_2[l-\epsilon]$ captures transient fading or other irrelevant temporal information. Vice-versa, a similar argument may be made for $\mathbb{U}_2[l-\epsilon]$ under some channel conditions, where in this case $\mathbb{U}_2[l-\epsilon]$ is an accurate model of $\mathbb{U}_2[l]$ but $\mathbb{V}_2[l-\epsilon]$ has captured some inaccurate information and thus predominately contributes to model-mismatch error. To ameliorate these effects we will now consider the possibility of using a standard, uniform orthonormal basis for $\mathbb{U}_2[l-\epsilon]$ and/or $\mathbb{V}_2[l-\epsilon]$ when designing the overlay.

2) Uniform Fingerprint Overlays Using Partial CSI: In general, the columns of $\mathbb{U}_2$ forming the left signaling basis for our fingerprint message in the time dimension can be selected from any unitary matrix that is at least size $\mathcal{C}^{(P\times(P-p))}$. Similarly, the columns of $\mathbb{V}_2$ forming the right signaling basis for our fingerprint message in the frequency dimension can be selected from any unitary matrix that is at least size $\mathcal{C}^{(N\times N-p)}$, and further, this basis may be selected independently from $\mathbb{U}_2$. We note that deviation from $\mathbb{U}_2$ and/or $\mathbb{V}_2$ as used in $\mathbb{K}_{\mathrm{direct}}$ necessarily degrades the orthogonality between the intrinsic bases $\mathbb{U}_1[l-\epsilon]$ and $\mathbb{V}_1[l-\epsilon]$ and the extrinsic fingerprinting overlay formed by $\mathbb{U}_2$ and $\mathbb{V}_2$. When the orthogonality between these subspaces is degraded $\mathbb{U}_1$ will partially project on $\mathbb{U}_2$ as interference, $\mathbb{V}_1$ will partially project onto $\mathbb{V}_2$ as interference, and vice-versa.

A number of matrices with the unitary property exist in the literature that would suffice for selecting $\mathbb{V}_2$ and/or $\mathbb{U}_2$; however, three typical unitary matrices will be considered here: the identity matrix, the discrete Fourier transform matrix (DFT) matrix, and the Walsh–Hadamard matrix. Since $\mathbb{Q}[k]$ in (13) is the Hadamard product between intrinsic channel estimate matrix $\mathbb{H}[k]$ and the fingerprinting matrix $\mathbb{F}[k]$, another desirable property of $\mathbb{F}[k]$ is that it does not unduly bias particular elements of $\mathbb{Q}[k]$ in either the time or frequency dimensions when conveying $[\xi_p\ \xi_{p+1}\ldots\xi_{P-1}]$ to the receiver. Both the Walsh–Hadamard matrix, denoted $\mathcal{H}$ and the DFT matrix, denoted $\mathcal{W}$, are felicitous choices for selecting $\mathbb{U}_2$ and/or $\mathbb{V}_2$, since the majority of elements in these matrices are nonzero. This property effectively allows these bases to *spread* the extrinsic fingerprint signal in the respective dimension, i.e., *frequency spreading* for $\mathbb{V}_2$ and *time spreading* for $\mathbb{U}_2$. We define the following design criteria when selecting $\mathbb{U}_2[k]$ and $\mathbb{V}_2[k]$:

1) If $\mathbb{U}_2[k]$ is to be designed from a uniform signaling basis, select the $P-p$ columns of $\mathbb{U}_2[k]$ from a column subset of a unitary matrix of size $\mathcal{C}^{(P\times P)}$.
2) Similarly, if $\mathbb{U}_2[k]$ is to be designed from a uniform signaling basis, select the $N-p$ columns of $\mathbb{V}_2[k]$ from a column subset of a unitary matrix of size $\in\mathcal{C}^{(N\times N)}$.

Using the criteria above, $\mathbb{U}_2[k]$ can be designed using either a Walsh–Hadamard matrix or a DFT matrix as a basis, and $\mathbb{V}_2[k]$ can use either a Walsh–Hadamard matrix or a DFT matrix as a basis, and the basis selections for $\mathbb{U}_2[k]$ and/or $\mathbb{V}_2[k]$ may be made independently.

With this design criteria, we define an extrinsic fingerprint overlay design for $\mathbb{K}[l]$, where $\mathbb{U}_2$ is drawn from a Walsh–Hadamard basis such that

$$\mathbb{K}_{hu}^{T}[l] = \left[\mathbf{0}^{(P\times p)}\quad \mathcal{H}^{(P\times P-p)}\right]_{l-\epsilon} \times\left[\mathrm{diag}\left(\mathbf{0}^{(p)},\boldsymbol{\xi}^{(P-p)}\right)\quad \mathbf{0}^{(P\times N-P)}\right]_l \begin{bmatrix}\mathbf{0}^{(N\times p)}\\ \mathbb{V}_2\end{bmatrix}_{l-\epsilon}^{H} \tag{47}$$

where the subscript (hu) on $\mathbb{K}_{hu}$ denotes that the columns of $\mathbb{U}_2$ are selected from a subset of columns of a Walsh–Hadamard matrix of size $\mathcal{R}^{(P\times P)}$, while $\mathbb{V}_2$ are the original columns of the noise subspace projected in the frequency dimension derived from channel-sounding information obtained from $\mathbb{H}[l-\epsilon]$. While $\mathbb{K}_{hu}$ can be an improvement over $\mathbb{K}_{\mathrm{direct}}$ for some channel conditions, it still requires transmission of $\mathbb{V}_2$ to the transmitter using $N(N-p)$ resources of feedback.

To improve on the feedback requirement of (47), we also consider the case where $\mathbb{V}_2$ is selected from a standard, uniform basis, i.e.,

$$\mathbb{K}_{hv}[l] = \left[\mathbf{0}^{(P\times p)}\quad \mathbb{U}_2\right]_{l-\epsilon} \times\left[\mathrm{diag}\left(\mathbf{0}^{(p)},\boldsymbol{\xi}^{(P-p)}\right)\quad \mathbf{0}^{(P\times N-P)}\right]_l \times\begin{bmatrix}\mathbf{0}^{(N\times p)}\\ \mathcal{H}^{(N\times N-p)}\end{bmatrix}_{l-\epsilon}^{H} \tag{48}$$

where the subscript hv on $\mathbb{K}_{hv}$ denotes that $\mathbb{V}_2[k]$ is selected from a subset of columns of a Walsh–Hadamard matrix of size $\mathcal{R}^{(N\times N)}$, while $\mathbb{U}_2$ are the original columns of the noise subspace projected in the time dimension and derived from channel-sounding information obtained from $\mathbb{H}[l-\epsilon]$. While $\mathbb{K}_{hv}$ improves on the feedback required by $\mathbb{K}_{wu}$, requiring transmission of only $\mathbb{U}_2$ to the transmitter using $P(P-p)$ resources for feedback, $\mathbb{K}_{hv}$ leverages much less CSI in the design of $\mathbb{K}[l]$. Additionally, if the information captured by $\mathbb{U}_2$ represents transient temporal information while $\mathbb{V}_2$ captures accurate frequency-selective fading behavior, $\mathbb{K}_{hu}$ may yield greater model mismatch error than $\mathbb{K}_{hv}$ because the CSI used in the design of $\mathbb{K}_{hv}$ may be an inaccurate representation of channel state during $\mathbb{H}[l]$.

3) Fingerprint Overlays Requiring Zero CSI: For comparison, we also consider the *blind* orthonormal signaling basis overlay, where both $\mathbb{U}_2$ and $\mathbb{V}_2$ are selected from standard, uniform signaling bases and previous CSI is not needed or used by the transmitter. When both $\mathbb{U}_2$ and $\mathbb{V}_2$ are replaced with columns of the Walsh–Hadamard matrix, we denote the result $\mathbb{K}_{hu,hv}$

$$\mathbb{K}_{hu,hv}[l] = \left[\mathbf{0}^{(P\times p)}\quad \mathcal{H}^{(P\times P-p)}\right]_{l-\epsilon} \times\left[\mathrm{diag}\left(\mathbf{0}^{(p)},\boldsymbol{\xi}^{(P-p)}\right)\quad \mathbf{0}^{(P\times N-P)}\right]_l \begin{bmatrix}\mathbf{0}^{(N\times p)}\\ \mathcal{H}^{(N\times N-p)}\end{bmatrix}_{l-\epsilon}^{H}. \tag{49}$$

One advantage of using $\mathbb{K}_{hu,hv}[l]$ is that no previous CSI is required in the design of the extrinsic fingerprint overlay, as receiver feedback is not required. The primary disadvantage of $\mathbb{K}_{hu,hv}[l]$ is that orthogonality between the intrinsic channel distortions and the extrinsic fingerprinting subspace is not present, thus distortions indicative of the intrinsic time-varying channel will act as interference in the detection of the extrinsic fingerprint signal. Since the $\mathbb{K}_{hu,hv}[l]$ design does not need or use CSI, in (49) p is not the effective rank but merely determines the dimensionality of the fingerprint. We will consider this design for comparison in our experiments in Section IV and simulations in Section V.

Similarly, (47), (48), and (49) may select columns from the DFT matrix for $\mathbb{U}_2$ and/or $\mathbb{V}_2$, yielding $\mathbb{K}_{wu}[l]$, $\mathbb{K}_{wv}[l]$, and $\mathbb{K}_{wu,wv}[l]$, respectively; however, the delineation of these designs will be omitted as they are similarly defined. The bases for $\mathbb{U}_2$ and $\mathbb{V}_2$ may be selected independently, yielding the following possible designs:

$$\mathbb{K}_{x1,x2}[l],\ x1\in\{-,hu,wu\}, x2\in\{-,hv,wv\}. \tag{50}$$

F. Discussion of the Authentication Message

We conclude this section by providing an example multibit digital fingerprint message that can be conveyed via the fingerprint symbols, $[\xi_p\ \xi_{p+1} \ldots \xi_{P-1}]_l$, which are embedded using the fingerprint function, $\mathbf{f}[k]$. We include this example to illustrate our fingerprinting scheme in the context of a complete authentication system.

To modulate a multibit digital authentication message, the elements $[\xi_p\ \xi_{p+1} \ldots \xi_{P-1}]_l$ can be selected by the transmitter as symbols from a typical pulse amplitude modulation (PAM) signal constellation, and using an appropriate bit-to-symbol mapping the receiver can recover the digital authentication message from $[\xi_p \xi_{p+1} \ldots \xi_{P-1}]_l$ after extracting these symbols. While the vector $[\xi_p\ \xi_{p+1} \ldots \xi_{P-1}]_l$ is only $P-p$ symbols long, the concatenation a number of consecutive vectors over ω fingerprinted blocks will yield a digital authentication message of length $\rho = \omega(P-p)$ symbols long. For example, for the values $P = 128, p = 8\omega = 10$ a digital authentication message of $\rho = 1200$ bits can be embedded.

To address the needs of dynamic spectrum applications (DSA), for example, the digital authentication message embedded in each node's transmission should contain bit fields for the basic self-verifying information of the signal such as the frequency, location, and time the signal is authorized for transmission. We will denote these fields as $F, L,$ and T, respectively. A message hash of these parameters is then digitally signed using a secret key owned by the transmitter and included in the message, while a time stamp denoted TS is also included with the authentication message to prevent future replay of the message by malicious users. The time stamp allows for enforcement of an expiration deadline on the content of the message, and in the event that an authentication message is received with a time stamp that has passed the expiration deadline, it will be discarded by the receiver. The authentication message for an authorized user U_j, denoted $msg_{U_j,A}$, is given as

$$msg_{U_j,A} = \left\{TS, F, L, T, K_A^+, [\mathrm{Hash}_m[TS, F, L, T]]_{K_A}\right\} \tag{51}$$

where $[\cdot]_{K_A}$ is a digital signature of the content within $[\cdot]$ using the private key owned by the authorized users group, the subscript A is used to denote that user U_j is a member of the authorized users group A, K_A^+ is the public key of the authorized users group, and $\mathrm{Hash}_m[\cdot]$ is message digest of length m for the content within $[\cdot]$. The hash algorithm $\mathrm{Hash}_m[\cdot]$ can be any of a number of widely used collision-resistant hash algorithms, such as MD5 or SHA-1 [19], which provide reasonable security against the malicious fabrication of messages.

Since the authentication message $msg_{U_j,A}$ is transmitted as a multibit digital signal, the probability of a fingerprint detection miss is the same as the probability of receiving the entire authentication message with one or more bit errors. Because a single bit error in either the authentication message or the signature will cause the authentication to fail, the probability of missing the authentication message is the same as the probability of at least one bit error in the message. Therefore, for an uncoded binary transmission, the probability that the received authentication message is in error is simply

$$P[m\hat{s}g_A \neq msg_A] = 1 - (1 - P_e)^{B+C} \tag{52}$$

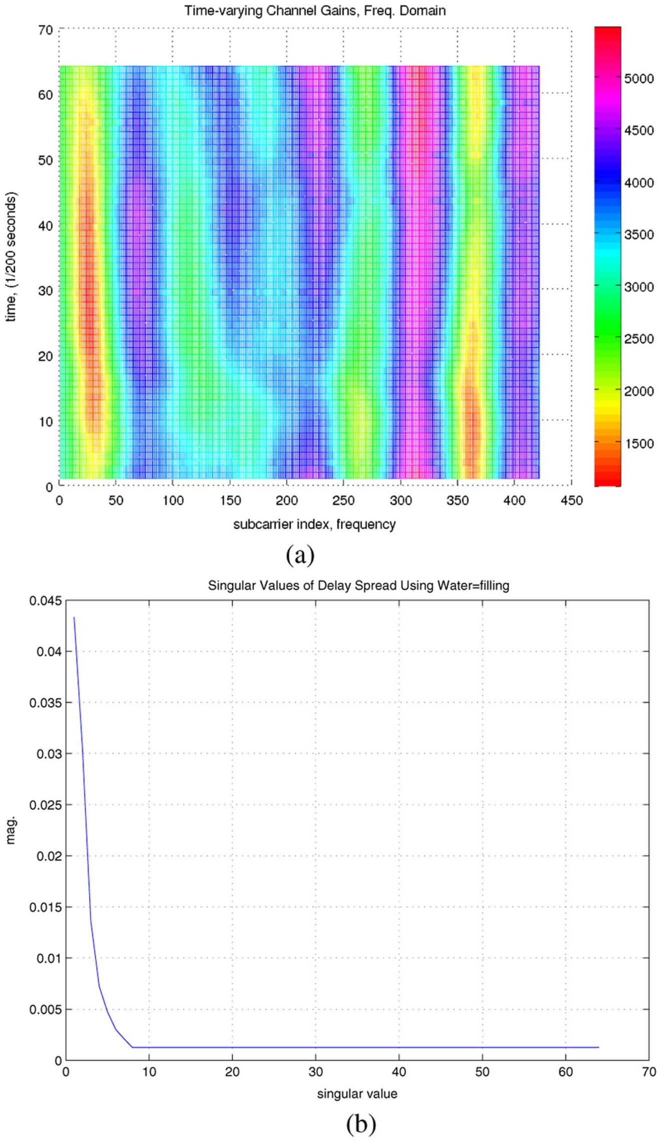


Fig. 1. Time-varying channel gains $\mathbb{H}[l-\epsilon]$. (a) Top view. (b) Sorted singular values.

where P_e is the probability of a bit error in the authentication signal, $B = \mathrm{length}\{TS, F, L, T, K_A\}$ and $C = \mathrm{length}\{[\mathrm{Hash}_m[TS, F, L, T]]_{K_A}\}$. The use of forward-error correction (FEC) on the authentication signal, combined with a continuously repeated message (i.e., repetition encoding), can further decrease the probability of an authentication miss.

The authentication message in (51) also includes the frequency F that the transmitter is authorized to transmit on, which would presumably be associated with the transmitter's key and recorded by a CA like the FCC. Therefore, even if we assume that an adversary can compromise an authorized user's key and forge $\mathbf{F}[k]$ at the PHY-layer, the attacker will be constrained to the frequency or frequencies prescribed by the compromised key. Using a forged $\mathbf{F}[k]$ on a frequency other than the original frequency prescribed by the key will implicate the transmission as a forgery when validating the credentials of the key against the CA's records.

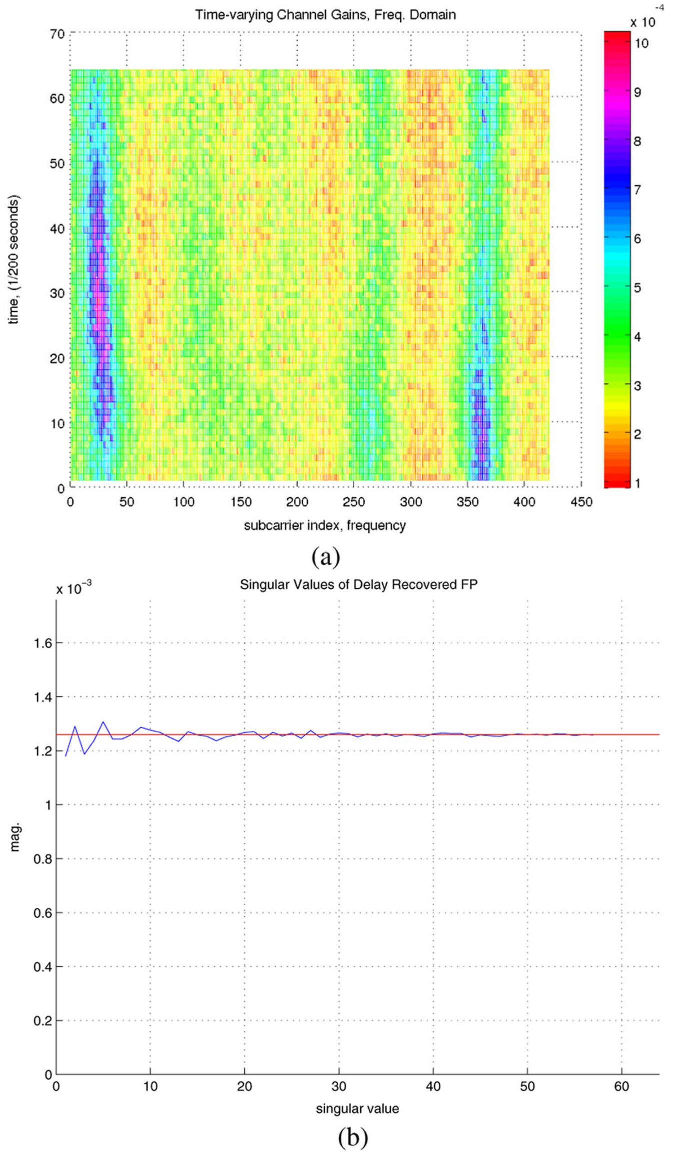


Fig. 2. Results for $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$ using the $\mathbb{K}_{hu}$ overlay. (a) Fingerprinted time-varying CSI. (b) Singular values.

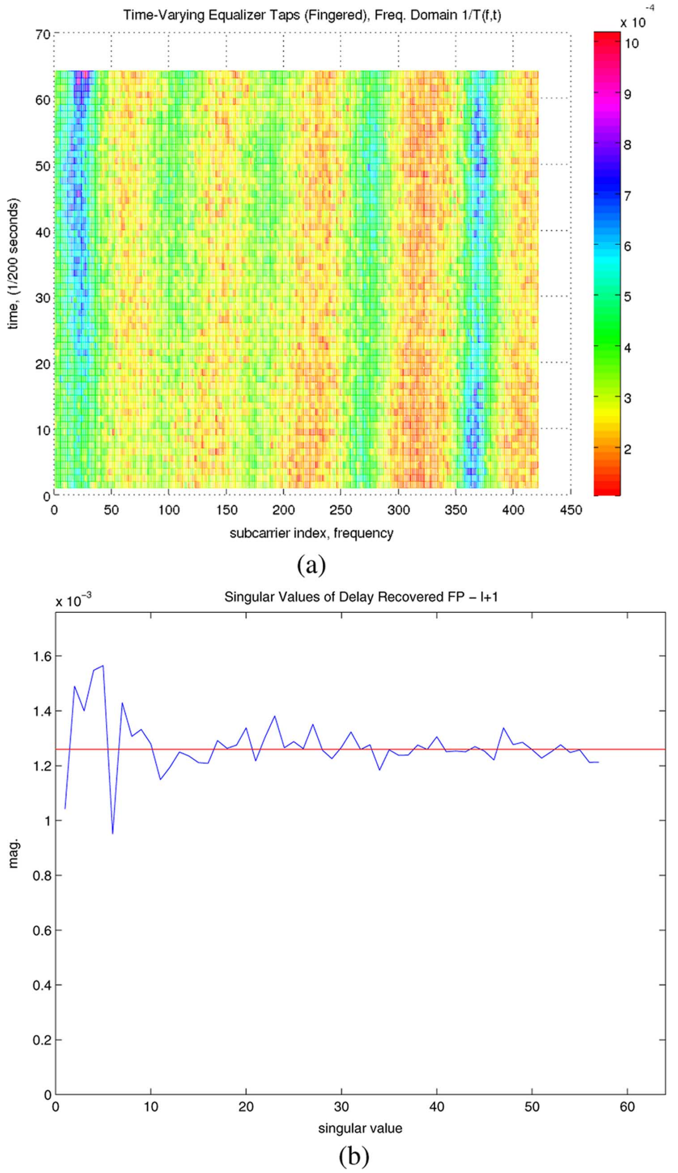


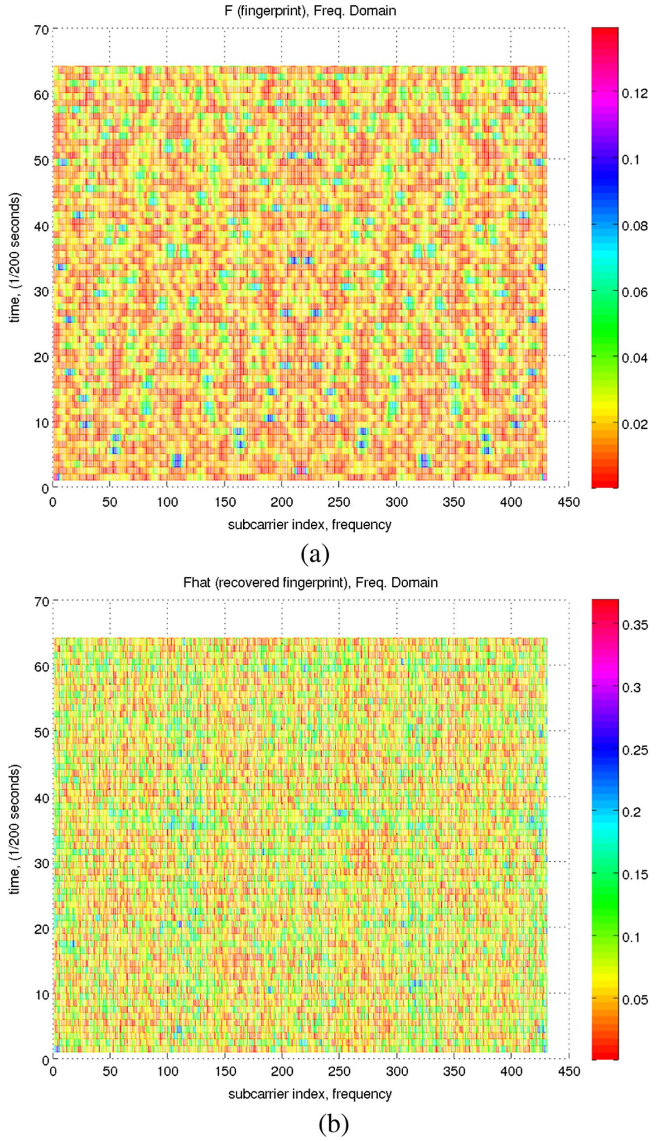
Fig. 3. $\mathbb{K}_{hu}$ overlay with model mismatch error. (a) Fingerprinted CSI $\mathbb{Q}$. (b) Singular values of the fingerprinted symbol $\hat{\mathbb{S}}$.

IV. Experimental Results

We now present experimental results for the fingerprinting method described, using time-varying channel estimate data collected from the IEEE 802.16e WiMax waveform. Raw signals were collected from a 5-MHz WiMax base-station configured to use a $N = 512$ subcarrier FFT, and a sequence of channel estimates were obtained using the training data present in each 5-ms frame preamble. This experiment data represents the channel conditions of a handheld mobile unit, where other than subtle hand movement, the mobile user is stationary.

The WiMax preamble uses a known data sequence that is duplicated three consecutive times in the time domain, therefore, a $3\times$ sinc interpolating filter was used to complete the channel estimate. A sequence of $P = 64$ frames was selected to form $\mathbb{H}[l-\epsilon]$, and the magnitudes of the equalizer-tap gains associated with each of the N subcarriers are presented in the top view in Fig. 1(a). From Fig. 1(a), we readily observe the slow frequency-selective fading behavior of this channel where areas of deepest fade are lightest, while the areas with the least fading are darker. We observe that the locations of frequency-selective fades are highly correlated in time, and that the environment is slightly changing since the locations of frequency-selective fades drift slightly.

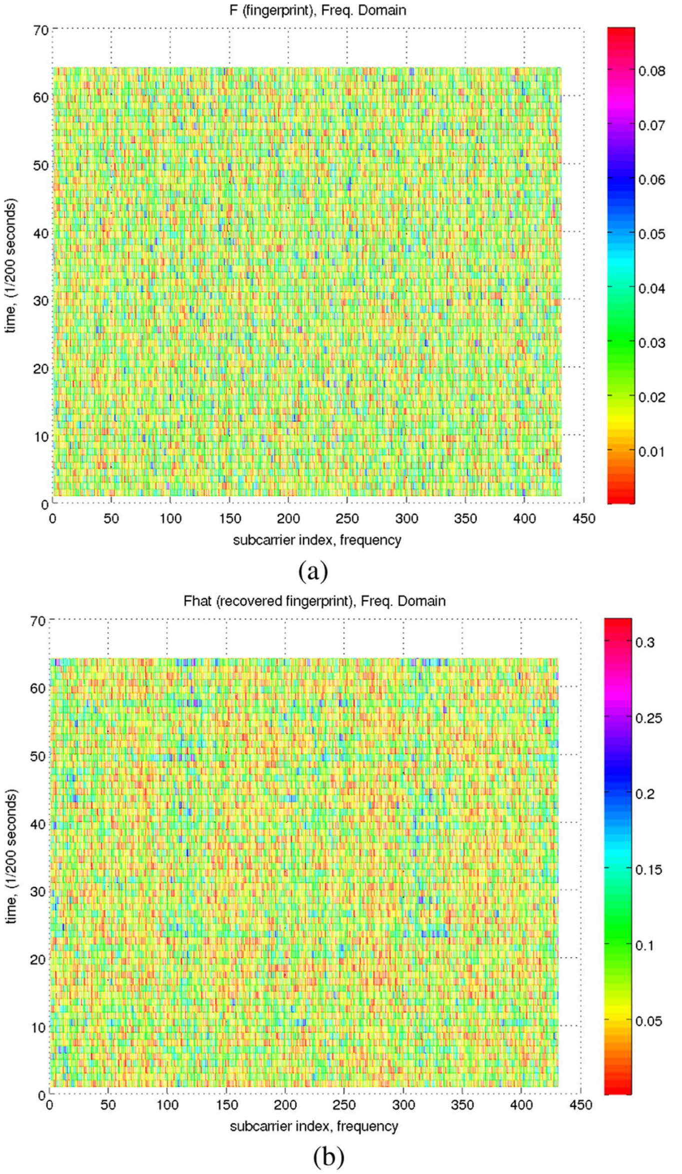
After diagonalizing the time-varying CSI of $\mathbb{H}[l-\epsilon]$ presented in Fig. 1(a) via SVD, a plot of the sorted singular values of $\mathrm{diag}(\mathbf{\Lambda}[l-\epsilon])$ is given in Fig. 1(b). From Fig. 1(b), we estimate the effective rank of $\mathbb{H}[l-\epsilon]$ to be approximately 8, and thus we select $p = 8$ accordingly, yielding a fingerprinting subspace of size $|\boldsymbol{\xi}[l]| = P - p = 56$, i.e., 56 eigenvectors and accompanying eigenvalues may be used for embedding the fingerprint message. The water-level used in this experiment was selected to be equal to $\xi_p = 0.0013$. In a full fingerprinting system, the elements $\xi_i, i = p+1, \ldots, P$ will be selected from a uniform PAM-like constellation to signal the digital fingerprint message.



(a)

(b)

Fig. 4. Simulated fingerprint embedding using the $\mathbb{K}_{hu,wu}$ overlay design with $p = 0$. (a) Original fingerprint. (b) Recovered fingerprint.



(a)

(b)

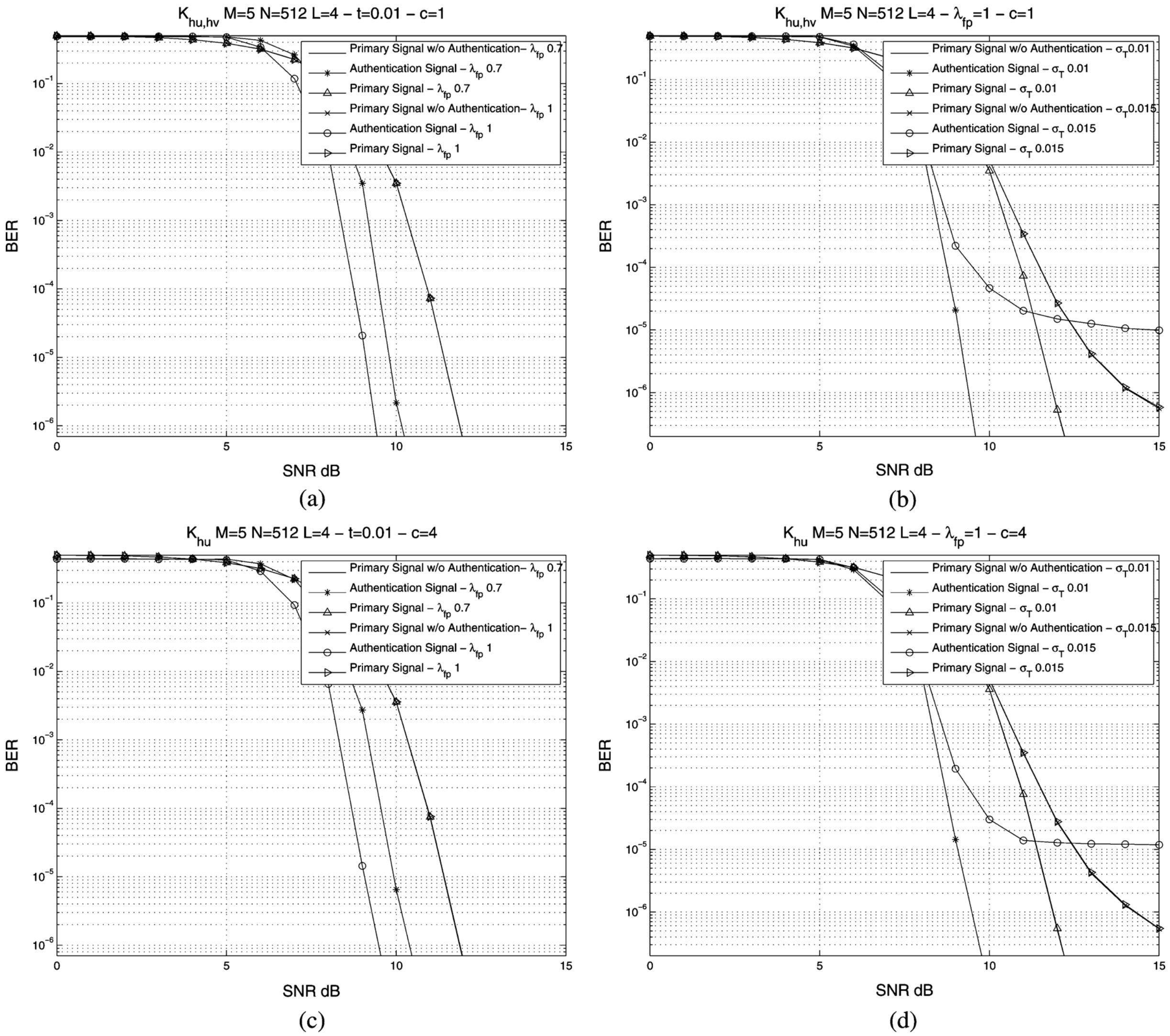
Fig. 5. Simulated fingerprint embedding using the $\mathbb{K}_{hu}$ overlay design with $p = 8$. (a) Original fingerprint. (b) Recovered fingerprint.

In this experiment, the $\mathbb{K}_{hu}[l]$ fingerprint overlay design was used, where the left singular-vectors of $\mathbb{U}_2[l-\epsilon]$ spanning $\mathbb{K}[l]$ in time are replaced with columns from a Hadamard matrix of size $\mathcal{H} \in \mathcal{R}^{(P\times P)}$ yielding the augmented version of this basis denoted $\mathbb{U}_2[l-\epsilon]$. According to the $\mathbb{K}_{hu}[l]$ design, the right singular-vectors are used directly, i.e., $\mathbb{V}_2[l-\epsilon] = \mathbb{V}_2[l-\epsilon]$. This design effectively *spreads* the fingerprinting signal in the time dimension using a time-uniform basis consisting of Walsh codes, in a way similar to frequency spreading via Walsh codes in CDMA systems. When a nonunitary basis is used in the construction of $\mathbb{U}[l-\epsilon]$, the innate orthogonality of the columns of the original matrix is degraded.

We first consider the case of no model-mismatch error, i.e., when $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, by applying the fingerprint to the same block of CSI used in the construction of $\mathbb{K}_{hu}[l]$. The magnitude of the fingerprinted time-varying CSI, $\mathbb{Q}[l]$, using the same intrinsic channel distortions of Fig. 1(a) and precoding using the $\mathbb{K}_{hu}[l]$ fingerprinting design, is depicted in Fig. 2(a). We note that the fingerprinted CSI of Fig. 2(a) and the original CSI of Fig. 1(a) are very similar, and that the fingerprinted version is visually a *noisier* version of the original intrinsic time-varying channel distortions. By preforming the fingerprint recovery steps of (38), $\hat{\mathbb{S}}[l-\epsilon]$ can be recovered and the fingerprint signal elements $[\xi_{p+1} \ldots \xi_{P-1}]$ can then be recovered from the diagonal elements of $\hat{\mathbb{S}}[l-\epsilon]$, as depicted in Fig. 2(b), while the elements $[\xi_i\xi_0 \ldots \xi_P]$ are omitted.

From Fig. 2(b), we readily observe the effect of receiver noise on the received fingerprint signal, as the values $[\xi_{p+1} \ldots \xi_{P-1}]$ should all be identically equal to $\xi_i = \xi_p = 0.0013, \forall i = p+1, \ldots, P-1$. We note that even in the absence of model-mismatch error these values are distorted by noise. It is clear, however, that in this example an ML receiver can recover the digital fingerprint, all 1's in this case, without bit errors using two-level PAM signaling and a symbol decision region that is half way between 0 and 0.0013.

While Figs. 2(a) and (b) demonstrate fingerprint application in the absence of model mismatch error, we now consider the



(a) (b) (c) (d)

Fig. 6. BER of the primary transmission and the fingerprint signal with different schemes. (a) $\mathbb{K}_{hu,wu}$ design versus λ_{fp}; (b) $\mathbb{K}_{hu,wu}$ design versus σ_T^2; (c) $\mathbb{K}_{hu}$ design versus λ_{fp}; (d) $\mathbb{K}_{hu}$ design versus σ_T^2.

performance of a fingerprint overlay when $\mathbb{K}[k]$ is applied to a future block of data transmissions. For this result, we select the next $P = 64$ channel estimates from the same signal used to create Fig. 1(a). The resulting fingerprinted time-varying CSI, $\mathbb{Q}[l]$, for the more general case when $\mathbb{L}[l] \neq \mathbb{L}[l-\epsilon]$ is presented in Fig. 3(a), using the same $\mathbb{K}_{hu}[k]$ design delineated in Fig. 2(a). We note that the fading behavior depicted in Fig. 3(a) is highly correlated with the fading behavior shown in Fig. 2(a), as this channel estimate data was obtained from the same WiMax signal and these blocks of CSI are exactly $P = 64$ OFDM frames, or 320 ms, apart. This adjacent block of CSI also demonstrates frequency-selective fade locations that are highly correlated in time. We also present the extracted fingerprinting signal elements $[\xi_{p+1} \ldots \xi_{P-1}]$ in Fig. 3(b).

Comparing the results of Fig. 3(b) to 2(b), the recovered fingerprinting signal elements $[\xi_{p+1} \ldots \xi_{P-1}]$ are even more distorted due to the additional model miss-match error.

V. Simulation Results

In this section, we present simulation results for one of the fingerprint overlay designs in (50), using the intrinsic time-varying channel model described in [11]. In [11], each channel gain of the delay profile is modeled using an Auto Regressive process of order 1 (AR-1), and the driving process for each AR-1 model is a Gaussian process with variance σ_T^2. As σ_T increases, the magnitude of the fluctuation of each channel gain in the model increases, and channel conditions change more violently.

Through simulation, we can quantitatively compare the performance of overlay designs and measure any degradation experienced by the primary signal from the embedding of the fingerprint. To measure primary signal degradation, we compare the BERs of the primary signal with and without the fingerprint present. Additionally, we use BER to compare the efficacy of each overlay design.

To simulate the embedding of various fingerprint designs, a full OFDM system and accompanying channel simulator were created in Matlab. The OFDM signal generated uses a 512-point FFT with 430 occupied subcarriers and 41 left and right guard subcarriers. A BPSK-modulated preamble occupies the first symbol of each frame using a 2× time-domain repetition. The preamble is represented in (9) by $\mathbf{P}$, while the primary signal payload is represented by $\mathbf{D}[k]$. In the simulated signal, every odd frame is fingerprinted while every even frame is used for channel sounding.

The OFDM frame is then subjected to a simulated time-varying channel by applying $g(t)$ in the time domain using a transversal filter, according to (1). Timing jitter was also added to test the fingerprint's affect on typical frame synchronization algorithms. For the channel delay-spread gains, the simplified *Typical Urban* delay spread profile of Section B.1 in the 3GPP specification [20] was used, while the values for the AR-1 model coefficients were chosen empirically.

To decode the signal, the Schmidl and Cox algorithm [21] was first applied for course timing synchronization. The LS channel estimator [14] was applied to estimate the aggregate channel distortion using the frame preamble as training data, and the resulting estimate was then up-sampled using 2× sinc interpolation. The channel sounding symbol and fingerprinted symbol were then equalized independently, and $\hat{\mathbf{q}}[l-\epsilon]$ and $\hat{\mathbf{h}}[l]$ are recorded for each frame for later use during the fingerprint recovery phase. Both the BPSK preamble data and QPSK payload data are demodulated into bits, and bit errors for the fingerprinted frames and nonfingerprinted frames are recorded.

In Figs. 4(a) through (b), the $\mathbb{K}_{hu,wu}$ overlay design was used using $p=0$ and the *Typical Urban* channel delay spread profile. By selecting $\mathbb{K}_{hu,wu}$ and $p=0$, CSI is not used in the design of $\mathbb{K}_{hu,wu}$. We will use this design for comparison when considering designs that do leverage knowledge of previous CSI.

The recovered fingerprint signal, $\hat{\mathbb{F}}[l]$, is depicted in Fig. 4(b), where we see the effects of the intrinsic time-varying channel on the blind superposition design in the additional distortions in Fig. 4(b) which are not present in Fig. 4(a). This is caused by the lack of orthogonality between the intrinsic and extrinsic subspaces discussed in Section III, thus demonstrating the susceptibility of the blind fingerprint approach to intrinsic channel distortions. For comparison, the same plots are presented for the $\mathbb{K}_{hu}$ overlay design using $p=8$ in Fig. 5(a) and (b). By comparing Figs. 4(a) to 5(a), we see that the fingerprint in Fig. 5(a) is more noise-like since its basis incorporates CSI derived from the noise subspace.

The BER results for the $\mathbb{K}_{hu,wu}, p=0$ design are given in Fig. 6(a) with $\sigma_T^2=0.01$ and the values $\lambda_{fp}=0.7$ and $\lambda_{fp}=1.0$, where λ_{fp} is the signal magnitude of ξ_i representing transmission of a 1 while transmission of a 0 is represented by zero, when using two-level signaling. Fig. 6(b) shows results for the $\mathbb{K}_{hu,wu}, p=0$ design, with $\lambda_{fp}=1.0$ and the values $\sigma_T^2=0.01$ and $\sigma_T^2=0.015$. We observe that the $\mathbb{K}_{hu,wu}$ fingerprint design does yield a BER improvement of 10 to 20 dB over the primary signal for both values of λ_{fp}, for SNR greater than 7 dB. This design also operates with a BER advantage for the low SNR regions of Fig. 6(b); however, for $\sigma_T^2=0.015$, the authentication signal BER flattens out around 10^{-5} as σ_T^2 becomes the predominate noise term and the time-varying distortions further degrade the fingerprint signal.

For comparison, BER results for the $\mathbb{K}_{hu}, p=8$ overlay design are given in Figs. 6(c) and (d). We observe that the $\mathbb{K}_{hu}$ fingerprint design also achieves a BER improvement over the primary signal for SNR greater than 7 dB. The "flattening out" phenomenon of the authentication signal BER for $\sigma_T=.015$ is also apparent, as the fingerprint yields BER slightly lower than 10^{-5} in higher SNR. In Figs. 6(a), (b), (c), and (d), we see zero impact to primary signal BER do to the presence of the fingerprint message, since the primary signal series with and without the fingerprint present completely overlap for a given value of σ_T.

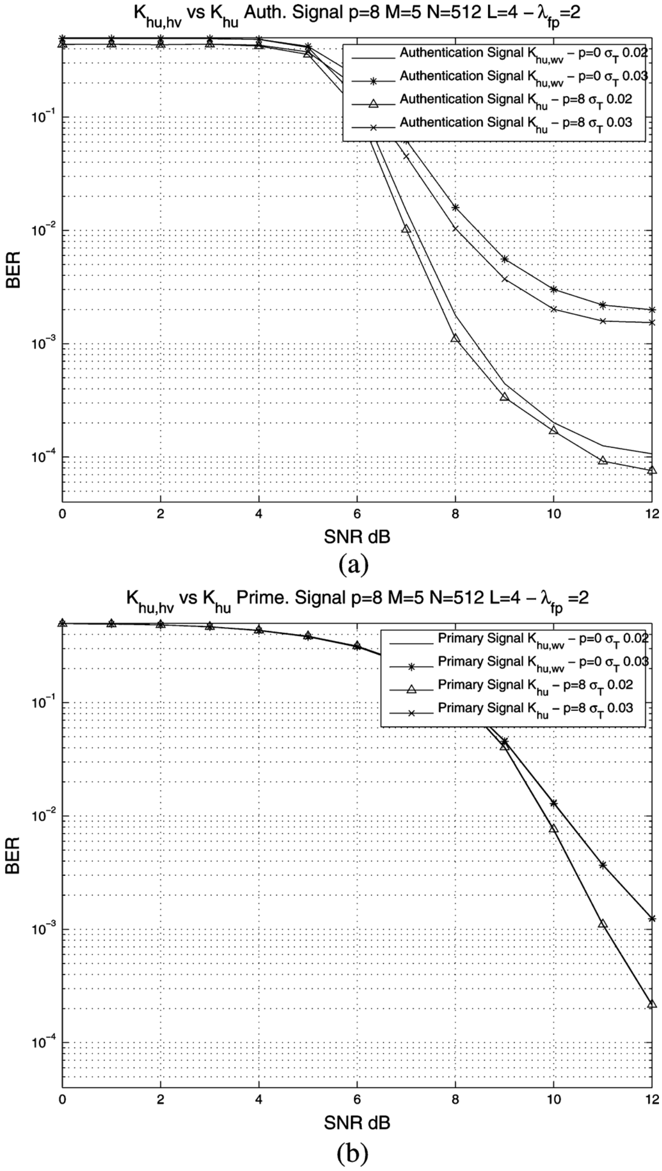


Fig. 7. Comparison between $\mathbb{K}_{hu,wu}$ design and $\mathbb{K}_{hu}$. (a) Fingerprint BER. (b) Primary BER.

To observe the benefits of incorporating previous CSI into the design of the fingerprinting overlay, we display the authentication signal BERs for the $\mathbb{K}_{hu,wu}, p=0$ and $\mathbb{K}_{hu}, p=8$ designs together, for the values $\sigma_T=[0.02, 0.03]$ in Fig. 7(a). The primary signal BER for these simulations is depicted in Fig. 7(b), and from this figure we observe nearly zero impact to primary signal BER for both values of σ_T^2. From 7(a), we observe that the fingerprint overlay design incorporating CSI, i.e., $\mathbb{K}_{hu}$, outperforms the design that does not incorporate previous CSI, i.e., $\mathbb{K}_{hu,wu}$. This advantage is demonstrated by the lower BER of the $\mathbb{K}_{hu}$ design, for all values of σ_T. From 7(a), we also observe that the BER advantage of the $\mathbb{K}_{hu}$ design increases as σ_T increases. This is because an increase in σ_T corresponds to an increase in model mismatch error, which manifests itself as $\mathbb{B}[l]$ in (38). The incorporation of CSI into the $\mathbb{K}_{hu}$ design helps mitigate the distortions caused by model-mismatch error.

VI. Conclusion

In this paper, we presented a new OFDM physical-layer fingerprint embedding scheme that incorporates previous CSI into the design of a overlay signaling basis. The transmitter embeds the fingerprint only onto the noise subspace of the wireless channel in a water-filling manner that maximizes the fingerprint capacity. We have demonstrated the embedding scheme through experimentation using real channel data collected from WiMax base stations, and the presented simulation results demonstrate that the BER of the primary signal is not influenced by the presence of the fingerprint. Also, the BER of the fingerprint signal outperforms the primary transmission by 20 dB, or more, in the channel conditions tested. Additionally, the proposed embedding scheme has demonstrated robustness to time-varying block-stationary fading.

References

[1] T. Clancy and N. Goergen, "Security in cognitive radio networks: Threats and mitigation," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications (Crowncom'08)*, Singapore, May 2008.

[2] R. Shaukat, S. Khan, and A. Ahmed, "Threats identification and their solution in inter-basestation dynamic resource sharing IEEE-802.22," in *Proc. Int. Conf. Convergence and Hybrid Information Technology*, Aug. 2008, pp. 609–614.

[3] R. Chen, J. Park, and J. Reed, "Defense against primary user emulation attacks in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 25–37, Jan. 2008.

[4] J. Burbank, "Security in cognitive radio networks: The required evolution in approaches to wireless network security," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications (Crowncom'08)*, Singapore, May 2008.

[5] C.-S. Lu*, Multimedia Security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property*. Hershey, PA: IGI Publishing, 2004.

[6] X. Wang, Y. Wu, and B. Caron, "Transmitter identification using embedded pseudo random sequences," *IEEE Trans. Broadcast.*, vol. 50, pp. 244–252, Sep. 2004.

[7] M. Morimoto, M. Okanda, and S. Komaki, "A hierarchical image transmission system in fading channel," in *IEEE Proc. 4th IEEE Int. Conf. Universal Personal Communications*, Nov. 1995, pp. 769–772.

[8] L. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1439–1449, Oct. 1993.

[9] P. Yu, J. Baras, and B. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, Mar. 2008.

[10] I. Cox, M. Miller, and A. McKellips, "Watermarking as communications with side information," *Proc. IEEE*, vol. 87, no. 7, pp. 1127–1141, Jul. 1999.

[11] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2571–2579, Jul. 2008.

[12] N. Goergen, T. C. Clancy, and T. R. Newman, "Physical layer authentication watermarks through synthetic channel emulation," in *Proc. New Frontiers in Dynamic Spectrum Access Networks (DySPAN'10)*, Singapore, Apr. 2010.

[13] N. Goergen, W. S. Lin, K. J. R. Liu, and T. C. Clancy, "Authenticating MIMO transmissions using channel-like fingerprinting," in *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 2010.

[14] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesso, "On channel estimation in OFDM systems," in *Proc. 45th IEEE Vehicular Technology Conf. (VTC)*, Chicago, IL, Jul. 1999, pp. 815–819.

[15] C. Pirak, Z. J. Wang, K. J. R. Liu, and S. Jitapunkul, "A data-bearing approach for pilot-embedding frameworks in space-time coded MIMO systems," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3966–3979, Oct. 2006.

[16] R. Everson and S. Roberts, "Inferring the eigenvalues of covariance matrices from limited, noisy data.," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 2083–2091, Jul. 2000.

[17] K. Konstantinides and K. Yao, "Statistical analysis of effective singular values in matrix rank determination," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 5, pp. 757–763, May 1988.

[18] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Taiwan, Mar. 2001, pp. 267–270.

[19] *Secure Hash Standard*, FIPS 180-1, 1995, 180th ed., Federal Information Processing Standards.

[20] *3rd Generation Partnership Project; Technical Specification Group Radio Access Networks; Deployment aspects (Release 7)*, 3GPP TR 25.943, v7.0.0 ed.

[21] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

Nate S. Goergen (S'03) received the B.S. degree in electrical engineering from Rose-Hulman Institute of Technology in 2004, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 2010 and 2011, respectively.

His research interests include cognitive radio, signal processing, and physical-layer security of wireless signals. His current research is in watermarking approaches for wireless communications.

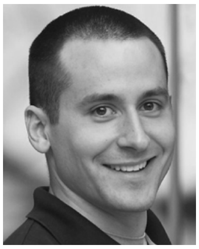
Dr. Goergen was awarded the DoD S.M.A.R.T. Scholarship in 2007, the 2010 University of Maryland Invention of the Year Award, and the Jimmy Lin Award for Invention in 2011.

W. Sabrina Lin (M'06) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, in 2002 and 2004, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park.

She is a Research Associate in the Electrical and Computer Engineering Department, University of Maryland. Her research interests are in the area of information security and forensics, multimedia signal processing, and multimedia social network analysis. She received the University of Maryland Future Faculty Fellowship in 2007.

K. J. Ray Liu (F'03) is named a Distinguished Scholar-Teacher of University of Maryland, College Park, in 2007, where he is the Christine Kim Eminent Professor of Information Technology. He serves as Associate Chair of Graduate Studies and Research of the Electrical and Computer Engineering Department and leads the Maryland Signals and Information Group conducting research encompassing broad aspects of wireless communications and networking, information forensics and security, multimedia signal processing, and biomedical engineering.

Dr. Liu is the recipient of numerous honors and awards including IEEE Signal Processing Society Technical Achievement Award and Distinguished Lecturer. He also received various teaching and research recognitions from University of Maryland including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award, both from A. James Clark School of Engineering. An ISI Highly Cited Author in Computer Science, he is a Fellow of IEEE and AAAS. He is President-Elect and was Vice President–Publications of IEEE Signal Processing Society. He was the Editor-in-Chief of *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of *EURASIP Journal on Advances in Signal Processing*. His recent books include *Cognitive Radio Networking and Security: A Game Theoretical View* (Cambridge Univ. Press, 2010); *Behavior Dynamics in Media-Sharing Social Networks* [Cambridge Univ. Press (to appear)]; *Handbook on Array Processing and Sensor Networks* (IEEE-Wiley, 2009); *Cooperative Communications and Networking* (Cambridge Univ. Press, 2008); *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications* (Cambridge Univ. Press, 2008); *Ultra-Wideband Communication Systems: The Multiband OFDM Approach* (IEEE-Wiley, 2007); *Network-Aware Security for Group Communications* (Springer, 2007); *Multimedia Fingerprinting Forensics for Traitor Tracing* (Hindawi, 2005).



T. Charles Clancy (S'02–M'06–SM'10) received the B.S. degree in computer engineering from the Rose-Hulman Institute of Technology, Terre Haute, IN, the M.S. degree in electrical engineering from the University of Illinois, Urbana-Champaign, and the Ph.D. degree in computer science from the University of Maryland, College Park.

He is the Associate Director of the Ted and Karyn Hume Center for National Security and Technology at Virginia Tech, where he leads the university's educational and research efforts in national security. Prior to joining Virginia Tech, he led a number of wireless research programs at the Laboratory for Telecommunications Science, at the University of Maryland, emphasizing development in commodity use of software-defined radio. His research interests are in the security of wireless communications, particularly spectrum access and waveform robustness.