On Tracking the Physicality of Wi-Fi: A Subspace Approach

MOHAMMED ALLOULAH[ID], ANTON ISOPOUSSU, CHULHONG MIN[ID], AND FAHIM KAWSAR

Nokia Bell Labs, Cambridge CB3 0FA, U.K

Corresponding author: Mohammed Alloulah (alloulah@outlook.com)

ABSTRACT Wi-Fi channel state information (CSI) has emerged as a plausible modality for sensing different human activities as a function of modulations in the wireless signal that travels between wireless devices. Until now, most research has taken a statistical approach and/or purpose-built inference pipeline. Although interesting, these approaches struggle to sustain sensing performances beyond experimental conditions. As such, the full potential of CSI as a general-purpose sensing modality is yet to be realized. We argue that a universal approach with the well-grounded formalization is necessary to characterize the relationship between the wireless channel modulations (spatial and temporal) and human movement. To this end, we present a formalism for quantifying the changing part of the wireless signal modulated by human motion. Grounded in this formalization, we then present a new subspace tracking technique to describe the channel statistics in an interpretable way, which succinctly contains the human modulated part of the channel. We characterize the signal and noise subspaces for the case of uncontrolled human movement and show that these subspaces are dynamic. Our results demonstrate that the proposed channel statistics alone can robustly reproduce the state-of-the-art application-specific feature engineering baseline, however, across multiple usage scenarios. We expect that our universal channel statistics will yield an effective general-purpose featurization of wireless channel measurements and will uncover opportunities for applying CSI for a variety of human sensing applications in a robust way.

INDEX TERMS Channel sensing, interpretable dimensionality reduction, machine learning, multiple-input multiple-output (MIMO).

I. INTRODUCTION

Due to the ubiquity and penetration of Wi-Fi in our homes, workplaces and cities, Wi-Fi traffic can be repurposed as a sensing modality for many potential applications beyond the original intended data-carrier functionality. Indeed, recent compelling research has reimagined a commodity Wi-Fi device as a multi-purpose sensor capable of turning *Wi-Fi traffic*—that is, packets transmitted over a wireless communication channel for either data transfer and/or the judicious probing of the channel—into a rich source of computational information explaining space dynamics, assessing the social environment and even tracking people's posture, and gestures [1]–[5].

However, human-perturbed Wi-Fi channels remain ill-understood. Despite prior art showcasing compelling use cases, ad hoc inference pipeline and careful parameter tuning

are commonplace for arriving at sensing recipes that yield good performance. Essentially, conventional approaches seek to associate patterns in Wi-Fi channel state information (CSI) with human activity through training classifiers on top of often bespoke featurization e.g. statistical distributions in [3] and Doppler variations in [5]. Although these sensing approaches demonstrated the potential of CSI sensing in a brand new class of applications, often, they are sensitive to environmental conditions and thus require controlled setup and development of pre-processing and inference pipelines which do not generalize across tasks (i.e. applications) and deployment environments. As such, CSI as a general-purpose sensing modality has not been adopted widely.

We argue for unleashing the true potential of CSI as a general-purpose human sensing modality; we need to turn our attention to developing sound theories explaining the relationship between spatiotemporal wireless channel modulations and human movement. Such characterization will assist in designing the future Wi-Fi network with stack

---

The associate editor coordinating the review of this manuscript and approving it for publication was Chunlong He.

layers augmented with annotations derived from the wireless propagation medium. These annotations would describe the physicality induced by the dynamic human movement which accompanied the delivery of data, thereby providing added *context*.

To this end, in this paper we present a first formalization for quantification of the changing part of the wireless signal modulated by human motion. Based on established channel models we devise new channel statistics that succinctly characterize the signal modulated by dynamic human movement. We then demonstrate that these channel statistics carry enough information to describe spatiotemporal human movement when observed continuously. This leads us to develop a novel subspace tracking algorithm that continuously analyses signal subspace as a function of dynamic human movement. The application of such metric enables us to precisely describe a set of human movement primitives including presence, motion activities, etc. As a step towards realizing CSI as a general-purpose sensing modality, we showcase how features extracted from the evolution of these subspaces can robustly reproduce state-of-the-art application-specific feature engineering baseline, however, across multiple usage scenarios and environmental conditions. Our research contributions are three-fold:

1) Statistical analysis of the signals to formally devise new statistics characterizing human-perturbed Wi-Fi channels.
2) Formalization of CSI sensing as a subspace tracking problem, demonstrating that the analysis of the dynamics of a signal subspace is the equivalent of sensing human movements.
3) Quantification of the benefits of using features derived from the proposed statistics and corresponding tracking technique concerning bleeding-edge CSI sensing applications.

We start by reviewing the required mathematical background of channel modeling in Section II-C. How the channel model can be used for sensing is explained in Section II-D. We use subspace based statistics to analyze human modulation of wireless channels in Section III. We show that the analysis of the dynamics of a signal subspace is equivalent to sensing human movements. We show by way of example how features extracted from subspace evolution can be used to solve sensing tasks in Section IV. We evaluate our subspace tracking featurization for two applications in Section V, provide a discussion in Section VI, and conclude with Section VII.

II. MEASUREMENT MODEL

A. NOTATION

Vectors and matrices are denoted in bold lowercase **a** and bold uppercase **A**, respectively. We use $||\mathbf{a}||$ to denote the Euclidean norm of a vector and $\angle(\mathbf{a}, \mathbf{b})$ to denote the angle between two vectors. The operator $\mathbf{E}\{\cdot\}$ represents the expectation. The operator $\text{Tr}\{\cdot\}$ represents the trace. The superscript ^{H} denotes the Hermitian transpose. The *i*th row and

*j*th column entry of a matrix **A** is a_{ij}. The all-ones matrix is denoted by **1**. Higher-order tensors are denoted by uppercase calligraphic letters $\mathcal{A}$. The symbol $\sim$ means statistically distributed as. The complex normal distribution is referred to as $\mathcal{CN}$.

B. PROBLEM STATEMENT

Our goal is to take steps towards a systematic study of the human-modulated subspace of CSI measurements. To this end, suppose we have a collection of CSI measurements $\mathcal{H}$. We postulate the existence of a universal decomposition

$$\mathcal{H} = \mathcal{S} + \mathcal{N}, \qquad (1)$$

where $\mathcal{S}$ contains all information of human modulation. Moreover, we assume that at each $k = 1, 2, 3, \ldots$, the signal subspace $\mathcal{S}[k] \subset \mathcal{H}[k]$ at time step k can at least in principle be computed from measurements $\mathcal{D}[l]$, with l "not too far" from k. What this means in practice, is that it is possible to *filter* out the noise subspace $\mathcal{N}$ and to *track* the human modulated subspace $\mathcal{S}[k]$ as time, represented by k, evolves. We make two further hypotheses about the decomposition (1):

Sufficiency of covariance statistics.
> *It is sufficient to consider the covariance statistics of $\mathcal{H}$ along different measurement dimensions independently.*

Dominance of the signal subspace.
> *The human modulation is characterised by magnitudes of variation of the covariance statistics at the appropriate time scales.*

Considering the measurement axes independently leads to an interpretable and effective dimensionality reduction on $\mathcal{H}$. Our approach is to use the eigendecomposition of the covariance matrices derived from the tensor $\mathcal{H}$.

We introduce the *structured channel model* and the *observation model* used in the rest of the paper in Sections II-C and II-D, respectively. The channel model provides a mathematical description of the measurement data. The observation model will be used to derive pre-processing techniques that have a sound physical justification for sensing tasks in Sections III and IV.

C. WIDEBAND MIMO CHANNEL MODEL

The Structured channel model we use belongs to a class of correlative wideband MIMO channel models [6]. Our starting point is the eigendecompositions of the channel model. This approach was first developed by Weichselberger [7], although it had also been developed independently by other works e.g., [8].

In the general case, we assume that CSI data forms a four dimensional dataset, with the four axes being the choice of receive antenna, transmit antenna, delayspread tap during one transmission step in time. We denote these *measurement dimensions* by the subscripts Rx, Tx, and Dy, respectively. We arrange CSI measurements into a tensor $\mathcal{H}$, and for each

time step k, we write $\mathcal{H}[k]$ for the three dimensional tensor of measurements at that timestep. The mth unfolding $\mathcal{H}_{(m)}$ is defined to be a matrix, whose columns are formed of the mth index of $\mathcal{H}$, and rows are formed by ordering the rest of the dimensions lexicographically. The three tensor unfoldings $\mathcal{H}_{(1)}[k]$, $\mathcal{H}_{(2)}[k]$, and $\mathcal{H}_{(3)}[k]$ describe the channel behavior in single-input multiple-output (SIMO), multiple-input single-output (MISO), and delayspread, respectively. Basically, the dimensions in the chosen axis are considered to be variables, and the other dimensions are collapsed in a lexicographical order.

We treat the measurement of the tensor $\mathcal{H}[k]$ as an unknown, random, and *non-stationary* complex Gaussian process. Concretely, in the absence of prior formalism on human-modulated wireless channels, we assume a generally evolving random process at two different time instances of the form [9]

$$\mathcal{H}_{(m)}(t + \tau) = \mathbf{W}\mathcal{H}_{(m)}(t) + \bar{\mathbf{W}}\mathbf{\Xi} \tag{2}$$

where $m \in [1, 2, 3]$, $\mathbf{W}$ (and its complement $\bar{\mathbf{W}} = \sqrt{1 - \mathbf{W}}$) is a correlation matrix whose elements $w_{i,j}$ are the correlation coefficients between the ith and jth RV-modeled entries of $\mathcal{H}_{(m)}$, and the auxiliary matrix $\mathbf{\Xi}$ is independent of $\mathcal{H}_{(m)}$ with i.i.d. entries $\sim \mathcal{CN}(0, \sigma_\xi^2)$.

We now suppress the time step k for simplicity, and assume that all calculations are done at a fixed time. The one-sided correlation matrices at the receive-, transmit-, and delayspread-side are then computed as

$$\mathbf{R}_{\mathrm{Rx}} = \mathbf{E}\left\{\mathcal{H}_{(1)}\mathcal{H}_{(1)}^H\right\},$$
$$\mathbf{R}_{\mathrm{Tx}} = \mathbf{E}\left\{\mathcal{H}_{(2)}\mathcal{H}_{(2)}^H\right\},$$
$$\mathbf{R}_{\mathrm{Dy}} = \mathbf{E}\left\{\mathcal{H}_{(3)}\mathcal{H}_{(3)}^H\right\}, \tag{3}$$

where the expectation is in practice computed by averaging samples across a short time interval, during which human modulation is assumed to be static i.e. wide-sense stationary (WSS).

Eigendecomposition is then applied to Equation (3) in order to extract the channel eigenbases in space (receive and transmit dimensions) and in delay spread according to

$$\mathbf{R}_{\mathrm{Rx}} = \mathbf{U}_{\mathrm{Rx}}\mathbf{\Delta}_{\mathrm{Rx}}\mathbf{U}_{\mathrm{Rx}}^H,$$
$$\mathbf{R}_{\mathrm{Tx}} = \mathbf{U}_{\mathrm{Tx}}\mathbf{\Delta}_{\mathrm{Tx}}\mathbf{U}_{\mathrm{Tx}}^H,$$
$$\mathbf{R}_{\mathrm{Dy}} = \mathbf{U}_{\mathrm{Dy}}\mathbf{\Delta}_{\mathrm{Dy}}\mathbf{U}_{\mathrm{Dy}}^H. \tag{4}$$

D. OBSERVATION MODEL

We observe a sequence of channel tensors $\mathcal{H}[k] \in \mathbb{C}^N$, $k = 1, 2, 3, \ldots, T$. We define a decomposition

$$\mathcal{H}[k] = \mathcal{S}[k] + \mathcal{N}[k] \tag{5}$$

where $\mathcal{S}[k]$ and $\mathcal{N}[k]$ are the *latent* human-modulated channel component, and additive noise tensor uncorrelated with human activity, respectively. The noise tensor need not be additive Gaussian and could account for many effects ranging

from wireless SNR variations,[1] suboptimal channel estimation, and/or quantization.

We write the observation model in terms of unfolding matrices as

$$\mathcal{H}_{(i)}[k] = \mathcal{S}_{(i)}[k] + \mathcal{N}_{(i)}[k] \tag{6}$$

for $i = 1, 2, 3$.

In what follows, we look at the third unfolding, which in our setup corresponds to the Dy dimension. Similar treatment applies to the Rx and Tx dimensions. Since the human-induced modulation and noise are uncorrelated, we can rewrite the one-sided correlations of equation (3) as

$$\mathbf{R}_{\mathrm{Dy}}[k] := \mathbf{E}\left\{\mathcal{H}_{(3)}[k]\mathcal{H}_{(3)}^H[k]\right\} \tag{7}$$
$$= \mathbf{C}_{\mathrm{Dy}}[k] + \sigma_{\mathrm{Dy}}^2[k]\mathbf{I}_{M_h} \tag{8}$$

where $\mathbf{C}_{\mathrm{Dy}}$ is the rank deficient covariance arising from the human-modulating effect, σ_{Dy}^2 is AWGN noise power, and $\mathbf{I}_{M_h}$ is an identity matrix [10].

Dropping k for brevity, the eigendecomposition in equation (4) can now be rewritten for the *observed covariance* $\mathbf{R}_{\mathrm{Dy}}[k]$ as

$$\mathbf{R}_{\mathrm{Dy}} = \mathbf{U}_{\mathrm{Dy}}\mathbf{\Delta}_{\mathrm{Dy}}\mathbf{U}_{\mathrm{Dy}}^H$$
$$\mathbf{R}_{\mathrm{Dy}} = \begin{bmatrix} \mathbf{U}_{\mathrm{Dy}}^s & \mathbf{U}_{\mathrm{Dy}}^n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Delta}}_{\mathrm{Dy}}^s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}_{\mathrm{Dy}}^n \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\mathrm{Dy}}^s{}^H \\ \mathbf{U}_{\mathrm{Dy}}^n{}^H \end{bmatrix} \tag{9}$$

where $\mathbf{U}_{\mathrm{Dy}}^s \in \mathbb{C}^{M_h \times M_s}$ is an orthonormal *signal subspace* basis, $\mathbf{U}_{\mathrm{Dy}}^n \in \mathbb{C}^{M_h \times M_n}$ is an orthonormal *noise subspace* basis, M_s and M_n are respectively the signal and noise sub-dimensions of the M_h-dimensional channel (i.e. $M_n = M_h - M_s$), $\hat{\mathbf{\Delta}}_{\mathrm{Dy}}^s = \mathbf{\Delta}_{\mathrm{Dy}}^s + \sigma_{s,\mathrm{Dy}}^2\mathbf{I}_{M_s} \in \mathbb{R}^{M_s \times M_s}$ is a noisy estimate diagonal eigenvalue matrix for $\mathbf{C}_{\mathrm{Dy}}$ the covariance matrix arising from the *true* human-modulating effect, and $\mathbf{\Delta}_{\mathrm{Dy}}^n = \sigma_{n,\mathrm{Dy}}^2\mathbf{I}_{M_n}$ is a diagonal noise eigenvalue matrix.

Each of the eigendecompositions in Equation (4) define a natural *filtration*, that is, a succession of growing subspaces $V_0 \subset V_1 \subset \cdots \subset V = \mathbb{C}^N$ spanned by the first i eigenvectors $\mathbf{u}_{\mathrm{Dy},j}[k]$, where $j \leq i$ and $i = 1, \ldots, N$. Here N is the dimensionality of the chosen measurement dimension, i.e. the number of delayspread taps (or by duality, frequency bins). By our assumptions in II-B, we may use the subspace V_i as a sufficient statistic for the signal subspace of $\mathcal{H}$ for some $i < N$. For each measurement dimension, we call the subspace defined here the *Tx/Rx/Dy-projected instantaneous signal subspace*, and we denote it by V_{Tx}, V_{Rx}, and V_{Dy}, respectively.

III. SUBSPACE CHARACTERISATION

In this section we hope to justify the claim that the projected signal subspaces introduced in the previous section are useful statistics which preserve human channel-modulating effects, while simultaneously being minimally diluted by noise. This claim is clearly non-trivial: human movements in the signal

---

[1] A "watery" human body in motion gives rise to complex and unconventional wireless propagation properties.

locale exert unconventional effects on the wireless channel which have not seen similar formal treatment in literature compared to more established channel models adopted widely by industry, say typical urban cellular fading channels [11]. The closest kin to human-modulated Wi-Fi channels in prior literature are perhaps body area network (BAN) channel models; consult [12]–[14] and literature therein for further detail. Specific characteristics of the wireless standard 802.11g/n/ac such as bandwidth, carrier frequencies, and air interface, impart modulating effects well beyond those studied for BANs.
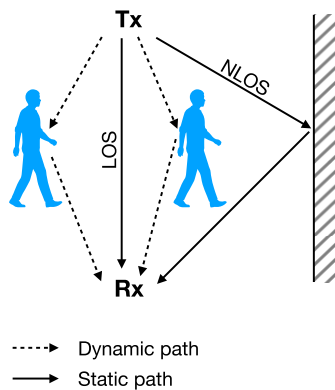
FIGURE 1. Human channel modulation.

A. CSI SENSING MODEL

As illustrated in figure 1, the Wi-Fi-based sensing model consists of placing a pair of transmitter and receiver devices in the environment. There are many paths by which electromagnetic energy travels between the transmitter and receiver. When people move, they disturb the multipath profile in the environment. The multipath profile is the linear superposition of a number of paths. For instance, figure 2 shows two static paths: a direct line-of-sight (LOS) and a reflected

non-line-of-sight (NLOS) paths. When a human subject walks from left to right in the figure, a dynamic path is *modulated* by this movement. By analyzing the temporal pattern of these dynamic paths at the receiver, we are able to build sensing applications.

For each transmitter-receiver pair, the superposition of multipaths in the time domain is described by a N_{sc}-dimensional frequency-domain CSI H corresponding to a sampling of OFDM subcarriers across the bandwidth.[2]

As such, the transmitted signal X can be related to the received signal Y through this input-output channel response relationship according to $Y = HX$. A MIMO system generalizes this input-output relationship for N_{tx} transmitters and N_{rx} receivers. For instance, if we have 3 transmitters and 3 receivers, the channel is described as a $3 \times 3 \times 30$ tensor.

We ask some basic questions:
- How can we characterize the human modulated subspace of the channel?
- How do the dimensionality and direction of the subspace vary in time as a result of human movement?

We take a first step towards providing a formal treatment of these key questions, and present a semi-analytical analysis of the projected signal subspace.

We discuss the theoretical underpinnings of our approach in Section III-B, particularly with a view towards contrasting to seminal prior work in Wi-Fi sensing. We then study data on uncontrolled human movement in Section III-D

B. BACKGROUND ON SUBSPACE TRACKING FOR WIRELESS SIGNALS

In classic signal processing, estimating the relevant subspace of variation in data is a basic building block of a data processing pipeline [15], [16]. In the context of an indoor wireless channel, the human modulated portion of the correlation data (cf. Equation (3)) is unknown with complex temporal dynamics.

Wang *et al.* [5] obtain good sensing results using an ad hoc pipeline starting with the full wideband covariance matrix (see [6]). We believe that this choice necessitates the use of excessive time-averaging of the CSI data. Furthermore, the resulting signal subspace is not easy to interpret. In contrast, the Rx, Tx and Dy correlations defined in Equation (3) are interpretable low dimensional representations. Despite the pioneering sensing approach, two drawbacks come to mind:
- the spatial and temporal behavior of the channel are not easily exposed, and
- the temporally highly averaged subspaces are less reactive to human activities.

The good sensing results aside, the approach of [5] does not conform to wireless theory, according to which human modulation should be quantifiable using subspace tracking.
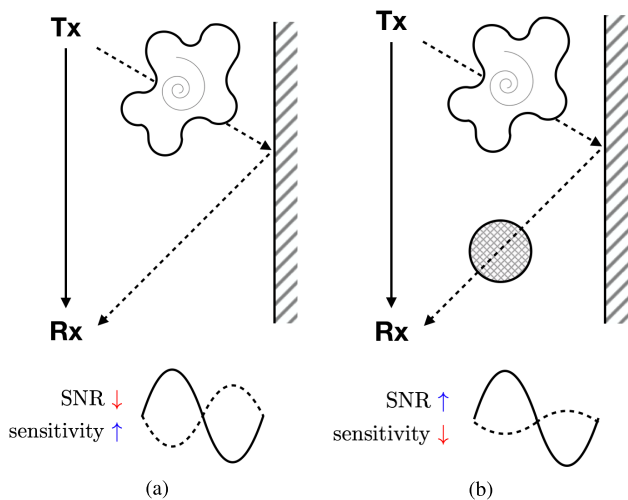

FIGURE 2. Good wireless SNR does not necessarily translate into good sensing sensitivity. for sensing, there is more to designating signal and noise subspaces than meets the eye. (a) Strong reflection. (b) Weak reflection.

---

[2]e.g. $N_{sc} = 30$ for the widely used Intel 5300 chip https://dhalperi.github.io/linux-80211n-csitool/.

Correlative MIMO subspace-based channel models have been shown to estimate capacity [6]–[8], and therefore the physicality of the medium. Intuitively, a model able to conform with a universal information-theoretic measure such as capacity is bound to convey fundamental information about the state of the channel irrespective of what modulates the channel. Further, recent theoretical results suggest that the rate of change of a MIMO OFDM channel can be inferred from the statistical analysis of its first and last eigenvectors [9], which can be viewed as canonical representatives of the signal and noise subspaces, respectively.

To elaborate on the dynamic nature of the signal subspace, consider a multipath component whose phase adds destructively to a main cluster of multipath, as depicted in Figure 2a. If the single multipath were to be shadowed as a result of a transient movement as in Figure 2b, it is clear that SNR would increase momentarily commensurate with the gain in total multipath arrivals energy. However, the sensing scene could have further nuances that are not captured by this simple SNR enhancement. As a further thought experiment, let the single multipath component be probing of a spatial sector in the environment in which a physical activity is unfolding—denoted by a spiral in Figure 2. That is, the single multipath component disproportionately delivers added movement sensitivity over that delivered by the main cluster of multipath. Despite the transient shadowing effect resulting in a *boost* in SNR, the instantaneous combined channel response is rendered *less sensitive* to activities occurring in the aforementioned spatial sector. The reduced motion modulation is manifested in reduced correlation structure in the regions of covariance matrix. Consequently—and perhaps counter-intuitively given the SNR gain—the signal subspace would necessarily "shrink" and noise subspace would "expand" momentarily. Therefore, robust sensing requires that the signal and noise subspaces be tracked explicitly in order to account for nuanced instantaneous channel effects.

The above contrived discussion suggests that a sensing system is required to adapt to dynamic channel effects in order to sustain optimal performance. Until provision for such adaptation is made in CSI-based sensing systems, we argue that models will fall short at being generalizable with guaranteed performance bounds irrespective of the nuances encountered in real-world deployment environments.

STATIONARITY PERIOD

The evolution of the signal subspace can be monitored at different granularities depending on the end-user application. An example of this scenario may be seen in activity recognition applications. Activity recognition requires deriving channel signatures of sufficient discriminatory power as to allow for the unambiguous separation of activities potentially similar in their broad nature e.g. walking versus running. The stationarity period is affected by, besides the application, the sensor configurations such as sampling rate.

For example, while 25ms may be necessary for responsive activity recognition applications, a 100ms or more may suffice for the much coarser presence detection. Note that sensing models may also be possible to realize even with "aliased" channel statistics akin to *compressive* sensing. However, we will not discuss this further here.

C. SENSING COMPLEXITY

The trade-off between sensing sensitivity and generality is a key question when it comes to designing any data processing pipeline. Generality implies flexibility for applying techniques from one sensing application to another. Sensitivity refers to optimality for a fixed sensing task. These are affected mainly by

- sensing pipeline configuration alongside its parameters, and
- the dimensionality of the signal subspace of the data, as it travels through the pipeline.

The latter is of particular importance because the size of the signal subspace allows for a controlled grading of sensing sophistication from the simplest (i.e. a one-dimensional subspace) to the most general (i.e. the entire signal subspace). The simplest extreme is particularly useful when out-of-the-box flexibility and ease of realization are desirable. When optimal performance and sensitivity are required, more elaborate and intricate sensing models can be used on a larger portion of the signal subspace.

We next shed light on the complexity of the human-modulated Wi-Fi signal subspace by way of an empirical study. The aim is to establish that there is more to designating signal and noise subspaces than meets the eye. Future research ought to take this complexity into consideration if Wi-Fi sensing were to be transitioned from controlled setups and into the wild.

D. EMPIRICAL STUDY OF UNCONTROLLED HUMAN MOVEMENT

We proceed to study the statistical effects of human activities on the channel covariances. Specifically, we study the projected signal subspaces, our putative proxies for the signal subspace for human modulation. To this end, we first quantify the information about the physical environment contained in the covariance matrix. This information is dynamic in nature and needs to be quantified instantaneously. One approach to gauging the information content in a series of covariances is to monitor the distortion contributed by the constituent eigenvectors. That is, by successively nulling the respective eigenvectors and measuring the fidelity of the covariance matrix reconstruction, we can quantify in the mean squared error-sense (MSE) the signal and noise boundaries at a given target distortion level (e.g. -12dB).

Concretely, let $\mathbf{R} = \mathbf{U}\mathbf{\Delta}\mathbf{U}^H$ be the eigendecomposition of one of the channel correlation matrices in equation (4). Define $\mathbf{R}'_i = \mathbf{U}\mathbf{\Delta}'_i\mathbf{U}^H$ as a reconstructed channel correlation matrix whose modified diagonal eigenvalue matrix $\mathbf{\Delta}'_i$ nullifies all
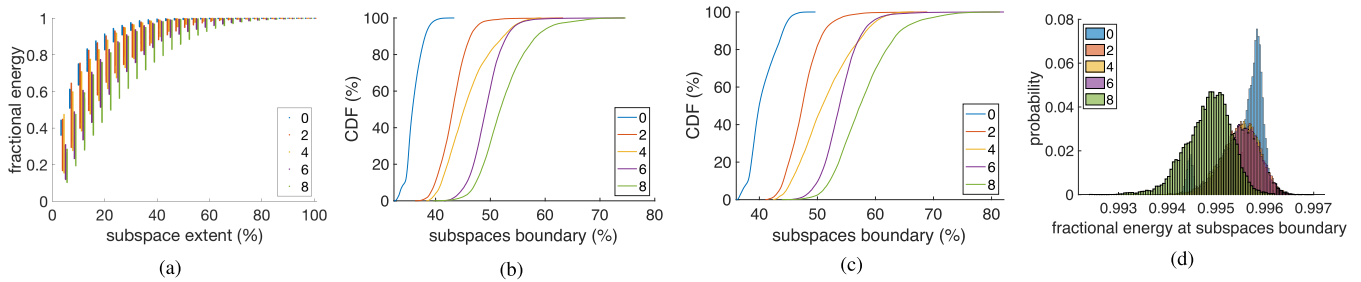
FIGURE 3. Characterizing signal and noise subspaces through an MSE search procedure for 5 uncontrolled human movement scenarios. The numbers in legend 0, 2, 4, 6, and 8 denote how many moving people are present. (a) Fractional energy E_s evaluated across the signal dimensionality. (b) MSE-based subspace boundary at -6 dB reconstruction objective. (c) MSE-based subspace boundary at -12 dB reconstruction objective. (d) Fractional energy at subspace boundary.

diagonal entries beyond index i i.e.

$$\boldsymbol{\Delta}_i' = \begin{bmatrix} \ddots & 0 & 0 & \dots & \dots & 0 \\ 0 & \delta_{i-1} & 0 & 0 & \dots & 0 \\ 0 & 0 & \delta_i & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 0 \end{bmatrix} \quad (10)$$

The reconstruction error matrix is $\mathbf{R}^\epsilon = \mathbf{R} - \mathbf{R}_i' = [r_{kl}^\epsilon]$. The reconstruction MSE error can then be described as

$$\text{MSE} = \frac{1}{N^2} \sum_{k,l} (r_{kl}^\epsilon)^2 \quad (11)$$

The above MSE search allows us to build a time-series picture of the dynamic partitioning of the covariance into signal and noise subspaces. This evolution of signal and noise subspaces is indicative of the evolution in the corresponding propagation conditions and also necessarily human movement. Intuitively, the harsher the dynamics of wireless propagation conditions, the more fluctuating the boundary between signal and noise subspaces is.

Having arrived at a statistical picture of subspaces boundary, we can utilize this knowledge to examine how the fractional signal subspace energy changes throughout human movement. We define the *fractional signal subspace energy* as the ratio between energy in the signal subspace to total energy contained in the channel. Thus, the fractional energy can be written as $E_s = \text{Tr}(\boldsymbol{\Delta}_x^s)/\text{Tr}(\boldsymbol{\Delta}_x)$, where Tr is the trace operator, and $\boldsymbol{\Delta}$ is the unitary eigenvalue matrix (cf. Equation (9)), and x $\in$ [Rx, Tx, Dy]. As such, E_s conveys information about optimum *sensing* SNR dynamics. A *parsimonious* suboptimal sensing system that utilizes instantaneously less of the available $E_s[k]$ at the kth time is effectively throwing away information.

The following discourse considers uncontrolled indoor human movement. This is perhaps the most generic form of activities likely to occur indoors. Naturally, uncoordinated motion components superimpose to modulate the signal

subspace in random ways. Stronger motion components could also mask much weaker ones.

We begin by examining what effect increased human movements has on the signal and noise subspaces. We conduct an experiment in which participants were asked to walk randomly in a room. The number of moving people present was varied from 0 (i.e. empty) to 8. The duration of movement per session was 5 to 10 minutes. An 802.11n $3{\times}3$ MIMO transmitter node was placed outside the room and a receiver node was placed inside. The CSI was sampled at a nominal sampling rate of 500Hz using a 5GHz carrier and 40MHz channel bandwidth. The reported CSI is 30 dimensional for each transmitter-receiver pair sampling the available 40MHz bandwidth coarsely but equidistantly. That is, 1-in-4 OFDM subcarriers are reported, resulting in a measured MIMO CSI $3 \times 3 \times 30$ tensor.

We investigate the effect of increased human movement on signal and noise subspaces by way of searching for the subspaces boundary yielding an objective MSE distortion as outlined earlier. Eigenvectors contributing less to the fidelity of covariance reconstruction will fall within the noise subspace. Conversely, eigenvectors impacting the fidelity of reconstruction more pronouncedly belongs to the signal subspace. The MSE-guided search finds the subspaces boundary that satisfies a desired distortion level in the MSE sense. Owing to the finite subspace resolution of a practical system, we interpolate between two MSE distortion levels produced at adjacent eigenvectors in order to simulate the effect of a smoothly varying MSE distortion and its respective "fractional" subspace index.

Figure 3a illustrates the variability in the fractional subspace energy E_s within the signal subspace extent and across movement scenarios—as denoted by the vertical scatter points. It is evident that the variability increases towards the lower-end of the subspace extent, reflecting the poor SNR contributed. Further, the variability increases markedly with the number of moving people i.e. fractional energy is more diffused in higher occupancy classes. Figure 3c shows the result of the MSE search procedure on the demarcation of the boundary between the signal and noise subspaces. Note, however, the statistical variability corroborating the earlier

hypothesis; namely, that dynamic stresses on the wireless channel would result in equivalent shrinkage or expansion of the signal subspace as needed to satisfy the target reconstruction distortion level. Similar subspace dynamic behavior can be seen when doubling the objective MSE distortion in figure 3b the subspaces boundary demarcation is insensitive to the chosen MSE level. It is further interesting to observe the accompanied effects in figure 3d on the fractional energy at the very same instantaneous demarcations of the signal and noise boundary established by the MSE search. The fractional energy at the *true*[3] instantaneous subspaces boundary is unable to provide a faithful statistical account on the expansion/shrinkage of the signal subspace at least for scenarios 2, 4, and 6 as evident by their density overlap. That is, the fractional energy cannot be called upon to optimally partition the covariance matrix.
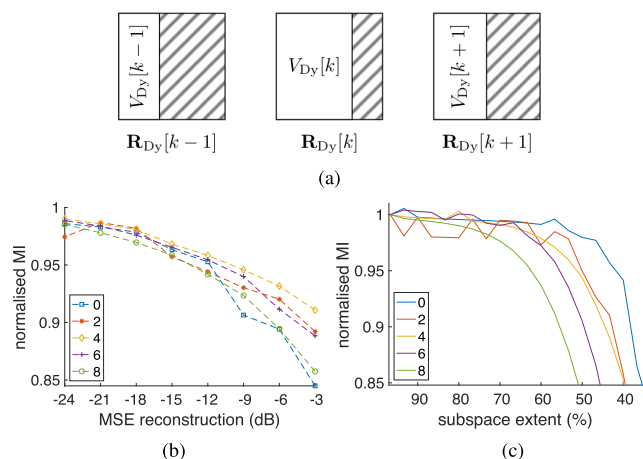
FIGURE 4. Normalized mutual information between covariances and their imperfect reconstructions over time and across 5 uncontrolled human presence scenarios, highlighting that the signal subspace is dynamic in nature. Legend denotes how many moving people are present. (a) Instantaneous subspace boundary. (b) Normalized MI versus MSE reconstruction objective. (c) Normalized MI versus subspace extent.

We conclude this section by qualifying our MSE-search methodology using mutual information (MI). The *instantaneous* subspaces boundary is used to agglomerate series of reconstructions of the covariance matrix as to compare against the groundtruth covariance distribution. We sweep the objective MSE distortion between -24 dB and -3 dB in 3 dB increments. We then measure the normalized mutual information between $V_{\text{Dy}}[k]$ and $R_{\text{Dy}}[k]$ for different occupancy cases as illustrated in figure 4a. Figure 4b shows that, for all human presence scenarios, the normalized MI at the instantaneous subspaces boundary steadily approaches unity as MSE reconstruction fidelity increases towards -24 dB. We observe that in terms of mutual information, our MSE-reconstruction based methodology is consistent under different channel conditions. In order to corroborate this observation, we compute the same normalized mutual information metric for the static (i.e. truncated) subspace extent

across occupancy cases. Figure 4c depicts such MI between $R_{\text{Dy}}[k]$ on the one hand, and $V_0 \subset V_1 \subset \cdots \subset V = \mathbb{C}^N$ on the other hand.[4] A "waterfall" effect can be seen whereby more truncated static subspace is needed at higher occupancy classes in order for MI to approach unity. Such MI waterfall effect is equivalent to the MSE-based subspace boundary shown earlier in figures 3b & 3c, reaffirming the notion of instantaneous subspace expansion and shrinkage as a function of the intensity of human movements.

IV. SUBSPACE TRACKING
In Section III, we established and characterised the notions of signal and noise subspaces within the context of human-induced channel perturbations. We now turn to examples of how to derive features for sensing tasks. Our approach is to track the evolution of the projected signal subspaces (cf. Section II-D).

The subspace-based human sensing we advocate for is in line with foundational work in wireless channels [6]–[9], which is in contrast to prior work on wireless Wi-Fi sensing (see [5]). We show that with a good enough instantaneous estimate of the covariances described in Section II this tracking can be used to capture the effects of human modulation. We present our analysis of the Dy-projected signal subspace, but the same can be easily repeated for the Rx dimension.

A. A GEOMETRIC VIEW OF SUBSPACE EVOLUTION
As an example of subspace tracking, we present the trajectories of the eigenvectors of the covariance matrices (cf. Equation (3)).
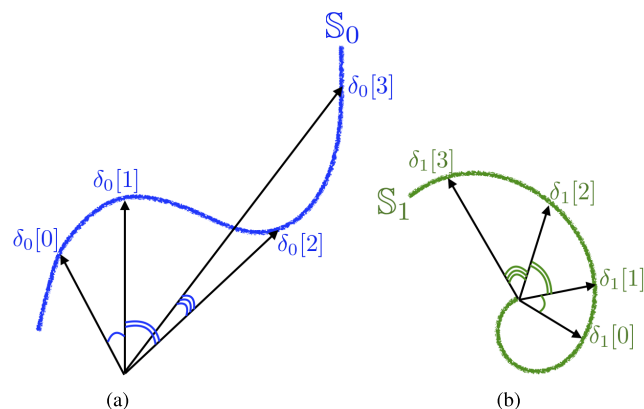
FIGURE 5. Geometric interpretation of subspace evolution. (a) Subspace component 0. (b) Subspace component 1.

Consider the time evolution of subspaces spanned by the first two (unnormalized) eigenvectors $\delta_0[k]$ and $\delta_1[k]$ of $R_{\text{Dy}}[k]$. Let $\mathbb{S}_0$ and $\mathbb{S}_1$ be subspaces spanned by $\delta_0[k]$ and $\delta_1[k]$ for $k = 1, 2, 3, \ldots$—depicted in local coordinate systems—respectively. See figure 5 for a geometric interpretation.

---

[3] we will return to address this claim in due course

[4] see Section II-D for a reminder on the definition of the succession of growing subspaces
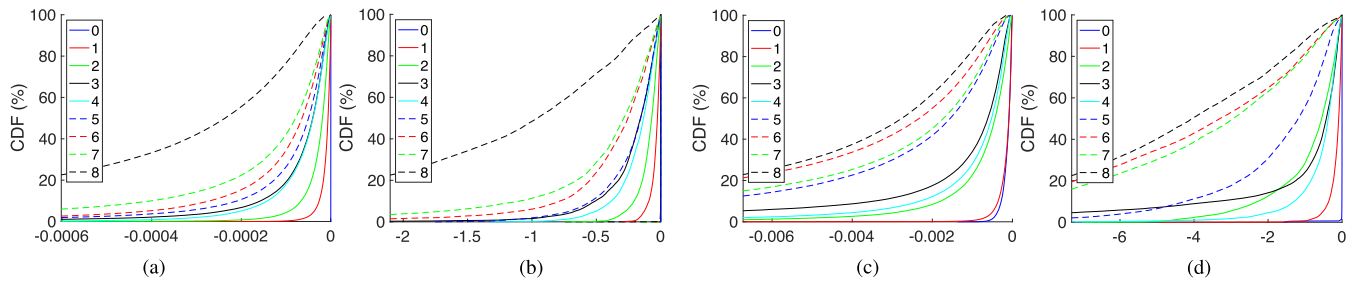
FIGURE 6. Differential unitarity cumulative distribution functions across 9 occupancy scenarios for Rx and Dy. legend denotes how many moving people are present. (a) and (c) correspond to the successive pairwise correlations, while (b) and (d) measure the rate of range. The rate of change is referred to by the ′ operator.

That is, recalling equation (9), each of these subspace components at the kth discrete time would correspond to (1) an ith eigenvector $\mathbf{u}_{\text{Dy},i}[k] \in \mathbf{U}_{\text{Dy}}^{s}[k]$ (i.e. belonging to the signal subspace), and (2) a scaling eigenvalue $\delta_i[k] \in \hat{\boldsymbol{\Delta}}_{\text{Dy}}^{s}[k]$. The empirical signal and noise characterization study reported in Section III has concluded that the fractional energy E_s evaluated at the subspaces boundary is less able to reveal increased multi-user channel variations. That is, when considering the movement of the signal subspace as a result of human-induced channel stresses, less stock should be put in the eigenvalues δ's. This is also intuitive to communications practitioners because phase-modulation, when combined with amplitude-modulation, is what really allows for packing more information efficiently within a finite stretch of bandwidth. The equivalence to the unreliability of power (i.e. eigenvalues) has also been echoed in prior art; namely, that "wireless internal state transitions result in high amplitude impulse and burst noises in CSI streams" [5]. As an example of this noisy state transition, note the bimodal nature of the 0 occupancy density of the fractional energy E_s in figure 3d—as indicated by the transparent underlaying behind the 8 occupancy case density. Another example is recent work on *channel charting* in the context of urban CSI measurements from basestations wherein Studer *et al.* [17] propose CSI scaling part of their feature mapping procedure.

Therefore, referring to figure 5 again, a critical insight emerges: *human effects on the wireless channel can be "demodulated" by observing the corresponding angular movements of the signal subspace.*

B. DIFFERENTIAL SUBSPACE EVOLUTION

The time dependency of the angular movements of the subspace is visualized in figure 5. The (complex) angles, which can be computed as the real part of Hermitian inner product, $\psi[1] = \angle(\mathbf{u}_{\text{Dy},i}[0], \mathbf{u}_{\text{Dy},i}[1]), \ldots,$ and $\psi[3] = \angle(\mathbf{u}_{\text{Dy},i}[2], \mathbf{u}_{\text{Dy},i}[3])$ are depicted for both subspace component 0 and 1.

These angles signify the *differential* movement of a certain signal subspace component between the $k-1$ and k discrete times. Incidentally, these angles have also another interpretation. Note that the diagonalization of the covariance matrix of equation (9) will produce eigenvectors which are by

construction unitary i.e. $\mathbf{u}_{\text{Dy},i}^{H} \mathbf{u}_{\text{Dy},i} = 1 = \cos(0)$. However, a human movement will cause the channel's signal subspace to *evolve* out of its "rest" condition. The resultant deviation in the subspace will be manifested in equivalent deviation in the *unitarity* of its constituent, evolved eigenvectors w.r.t. their original "rest" conditions. Thus, the successive change in *unitarity* for the ith subspace component between time $k-1$ and k is *quantified* by $\mathbf{u}_{\text{Dy},i}^{H}[k] \mathbf{u}_{\text{Dy},i}[k-1] = \cos(\psi[k])$ which coincides with the angular movement of the subspace. Hence we term this angular metric the *differential unitarity*.

In general, our proposed differential unitarity feature for tracking human-modulated signal subspaces is applicable to any channel eigendecomposition formulation commonly encountered in literature. Denote by $\mathbf{u}_{\text{Dy},i}[k]$ the ith delayspread eigenvector at time k. Then the differential unitarity $\hat{u}_{\text{Dy},i}[k] = \cos(\psi_{\text{Dy},i}[k])$ between time k and $k-1$ is formulated as

$$\hat{u}_{\text{Dy},i}[k] = \mathbf{u}_{\text{Dy},i}^{H}[k] \mathbf{u}_{\text{Dy},i}[k-1] \qquad (12)$$

Similarly, for the receive-side eigenbasis

$$\hat{u}_{\text{Rx},i}[k] = \mathbf{u}_{\text{Rx},i}^{H}[k] \mathbf{u}_{\text{Rx},i}[k-1] \qquad (13)$$

Equations (12) & (13) represent two degrees of freedom through which we can measure the volatility in the wireless channel as a result of human stressors: (1) spatial from multiple antennae and (2) temporal across the delayspread (or equivalently bandwidth). We next build intuition for these complementary differential unitarity metrics by presenting a series of concrete numerical examples.

We return to the uncontrolled movement dataset reported in Section III. Further, let us examine the behavior of the differential unitarity for the 1st subspace component of both the receive-side and delayspread subspaces i.e. $\hat{u}_{\text{Rx},1}$ and $\hat{u}_{\text{Dy},1}$, respectively. Figure 6 plots the cumulative distribution functions (CDFs) for $\hat{u}_{\text{Rx},1}$ and $\hat{u}_{\text{Dy},1}$ for all 9 occupancy cases. Specifically, note the dispersive nature of the metric in figures 6a & 6c as a function of increased human-induced channel perturbations. It is clear that the dispersion in the statistics of the magnitude of differential unitarity—corresponding to the 1st subspace components—monotonically increases, generally, with increased human movement.

(a)

(b)
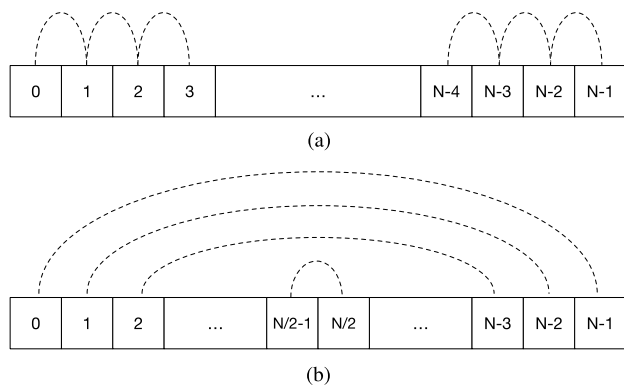
FIGURE 7. Interrogating rate of change of differential unitarity by correlating farther apart eigenvectors with successively decreasing distance. (a) Pairwise. (b) Slope.

The diagram in figure 7b shows that the subspace bases relating to N time periods are buffered so that comparison can be made across a wider time window. Thus, for example, the eigenbases at $t = 0$ can be compared with those at $t = N - 1$, the eigenbases at $t = 1$ can be compared with the eigenbases at $t = N - 2$ and the eigenbases at $t = 2$ can be compared with the eigenbases at $t = N - 3$. Such an arrangement may enable the changes in channel statistics to be viewed across a wider time period and may enable the rate of change of eigenvector unitarity to be determined.

Deriving a measure of the rate of change of differential unitarity has the advantage of increasing the separation of the CDFs of figures 6a & 6c. To this end, we apply the scheme depicted in the lower diagram of figure 7 (termed "slope") to the same experiment and obtain the CDFs shown in figures 6b & 6d. It is readily evident that the CDFs corresponding to the rate of change in the differential unitarity extracted over a window of time experience increased dispersion as a result of human occupancy. This may allow for learning and/or calibrating better discrimination boundaries in the inference logic.

1) SUBSPACE SAMPLING

Recalling equation (8), we note that the expectation operator implies an averaging effect. Earlier we have elaborated on the notion of stationarity period and its connection to CSI sampling and application granularity requirements. Yet another pertinent aspect for consideration lies in how to realize the expectation. Broadly, there are two methods often employed in classic signal processing literature for updating the signal subspace: (i) stochastic approximation, and (ii) batch averaging. These two variants have implications on signal subspace tracking, which we discuss next.

An unbiased stochastic expectation estimator is given by

$$\mathbf{R}_{\mathrm{x}}[k] = (1 - \lambda) \sum_{n=0}^{k} \lambda^{k-n} \mathcal{H}_{(m)}[n] \mathcal{H}_{(m)}^{H}[n] \quad (14)$$

where $\mathrm{x} \in [\mathrm{Rx}, \mathrm{Tx}, \mathrm{Dy}]$ and $m \in [1, 2, 3]$, respectively. This estimator reduces to the recursive expression

$$\mathbf{R}_{\mathrm{x}}[k] = \lambda \mathbf{R}_{\mathrm{x}}[k - 1] + (1 - \lambda) \mathcal{H}_{(m)}[k] \mathcal{H}_{(m)}^{H}[k] \quad (15)$$

where $\lambda \in [0, 1)$ is a forgetting factor often chosen close to 1. This stochastic estimator accounts for a long channel history, albeit while de-emphasizing far away events. Such subspace update tends to "dampen" the effect of abrupt channel changes on the signal subspace. Alternatively, these abrupt changes can also be preserved and admitted into the subspace using the sliding window (a.k.a batch) approach given by

$$\mathbf{R}_{\mathrm{x}}[k] = \frac{1}{L} \sum_{n=k-L+1}^{k} \mathcal{H}_{(m)}[n] \mathcal{H}_{(m)}^{H}[n] \quad (16)$$

where L is the window size determined by the assumed stationarity period.

We now compare and contrast between these two subspace update variants. An activity recognition dataset available publicly is used [1]. The dataset is comprised of 6 single-user activities; namely, standing up, sitting down, lying down, falling, walking, and running. SIMO CSI data from three receivers is sampled at 1 ksps rate. For added tracking responsiveness and resolution, we choose a stationarity period of 25ms and proceed to update the covariance matrix with 95% CSI overlap from previous stationarity period. This results in around 800 Hz subspace update rate.

In Figure 8, we perform time-frequency localization on the pairwise differential unitarity subspace tracking metric. The localization uses a window of 1.28 seconds with 95% content overlap between two windows for finer time-frequency resolution. In the interest of space, only four single-user activities are shown corresponding to falling, lying down, walking, and running. The spectrograms of the upper row of Figure 8 were generated using the batch subspace update variant; while those in the lower row utilized the stochastic variant with a forgetting factor $\lambda = 0.99$. The color coding of the spectrograms in each row was group-harmonized in order to convey correct information about the differential intensity of the time-frequency bins across activities. We therefore safely omit the color maps from the spectrograms. As touched upon previously, the batch subspace update is more responsive to background disturbances in the channel and would admit these into signal subspace. We can readily observe more background variations across all activities in the upper series of spectrograms. Despite this, we can still see distinctly individual behavior across these activities—falling being the most concentrated in time-frequency and running being the most dispersed. However, it is interesting to see how the stochastic update was able to *filter* out much of the background channel disturbances while preserving the discriminative features of the four activities; namely, the increased time-frequency dispersion from falling, lying down, through to walking and running—again the latter being the most dispersed.
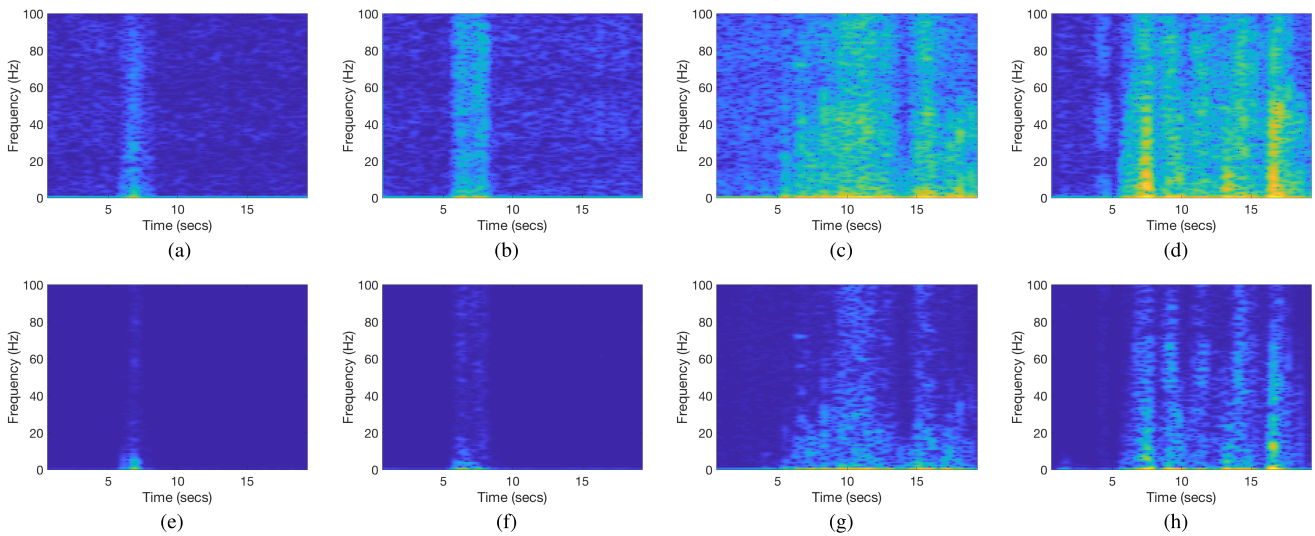
FIGURE 8. Spectrograms for 4 activities (falling, lying down, walking, and running) using two subspace update variants: batch and stochastic. It is readily that stochastic has a filtering effect on the time-frequency localization of pairwise differential unitarity subspace tracking metric. (a) Falling – batch. (b) Lying down – batch. (c) Walking – batch. (d) Running – batch. (e) Falling – stochastic. (f) Lying down – stochastic. (g) Walking – stochastic. (h) Running – stochastic.
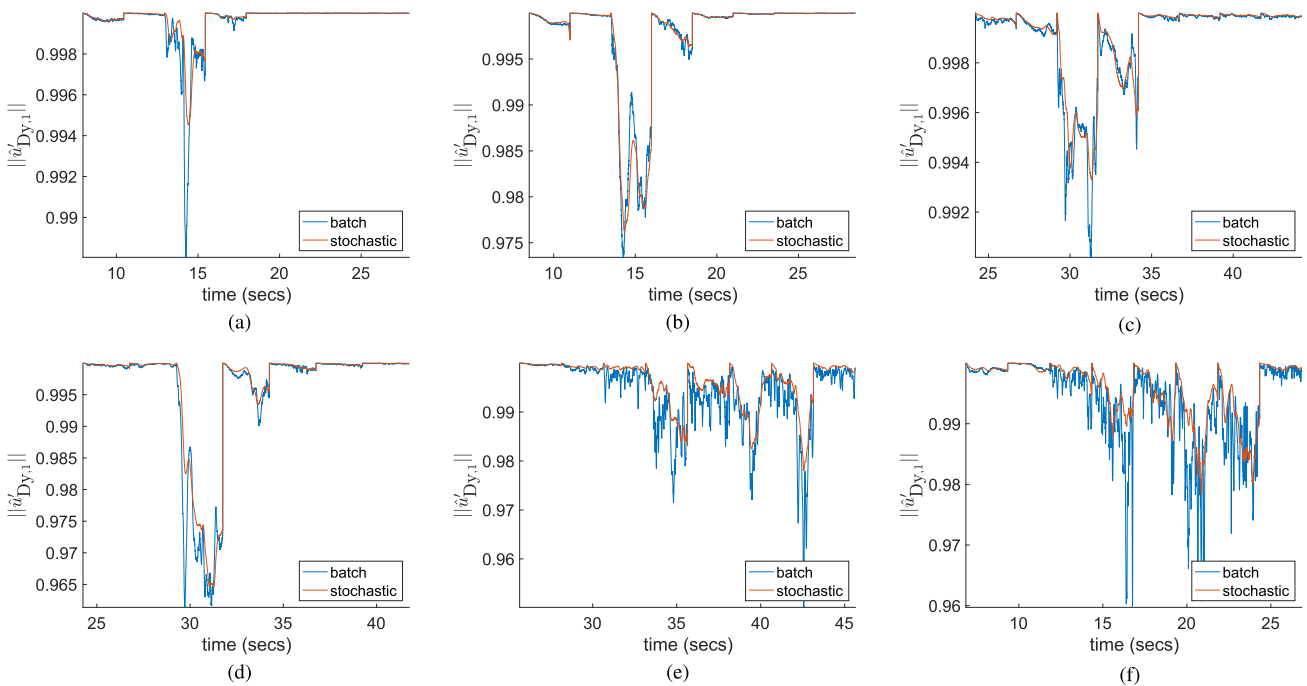
FIGURE 9. Waveforms for 6 activities (standing up, sitting down, lying down, falling, walking, and running) using two subspace update variants: batch and stochastic. it is readily noticeable that stochastic has a filtering effect on the slope differential unitarity subspace tracking metric. (a) Lying down. (b) Sitting down. (c) Standing up. (d) Walking.

We have opted to conduct time-frequency localization on the pairwise subspace tracker owing to its more intuitive association with speed i.e. 1st-order derivative of subspace evolution. A justification for the correspondence between the rate of change in CSI and speed can be found in [5]. Our 1st-order differentiation of the subspace can be viewed as a generalized *fusion* method for extracting information embedded in all subcarriers simultaneously. This fusion is a data-level fusion, rather than feature-level approaches involving ad hoc subcarrier selection strategies [4]. Some reported Wi-Fi sensing systems resort to selecting subcarriers of better SNR since frequency selectivity of wideband Wi-Fi channels causes some subcarriers to fall within the channel nulls—with obvious consequences for their reliability. Our subspace
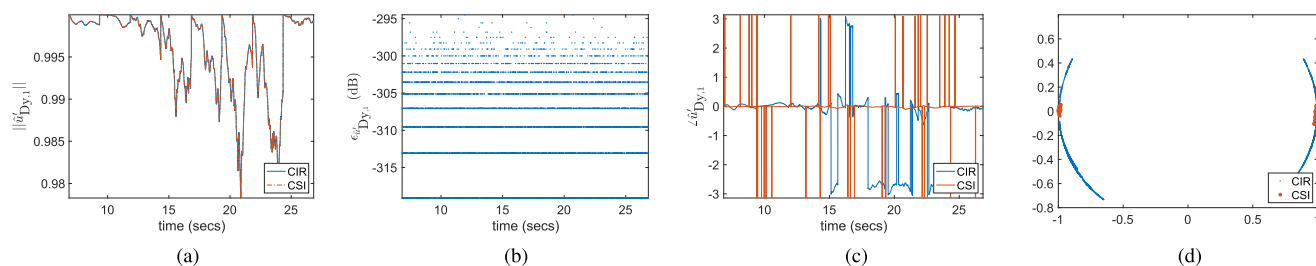
FIGURE 10. Waveforms for slope tracker corresponding to the running activity. differential metric is identical in magnitude, but subspace phase stabilities exhibit interesting variations that are different depending on whether subspace decomposition is performed in the time-domain or frequency-domain. (a) Magnitude. (b) Error power. (c) Scatter.

approach systematically fuses information contained in all subcarriers without the need to perform preconditioning. However, unlike the PCA-based approach [5], this fusion is principled, interpretable, and has its roots in formal wireless channel concepts [6]–[9].

As illustrated in Figure 7b, we use a robust sampling technique to obtain clean statistics from the differential unitarity measurements. In Figure 9, we illustrate the effect of this choice. Six single-user activities—standing up, sitting down, lying down, falling, walking, and running—for the slope tracker are depicted. Quick inspection of these plots corroborate the earlier findings of the spectrograms analysis; namely, that stochastic filters high-frequency channel perturbations compared to batch. That is, stochastic tracks the *envelope* of the activity rather than its and/or the channel's background high-frequency fluctuations. We have alluded to this *tunable* channel detail in the signal subspace, be it channel background- or activity-related, by the hat accent in Equation (9). The abrupt activity of falling has an impulse-like acceleration content, while running is the richest in such 2nd-order rate of change moments.

2) DUALITY

For completeness, we provide commentary on the pertinent issue of choosing a channel representation: time- versus frequency-domain. The structured model we introduced in Section II-C has been validated with empirical channel impulse response (CIR) measurements i.e. in the time-domain. Identical eigenspace formulation has been applied in the frequency-domain for CSI instead [18], and also validated with empirical capacity measurements. Since our subspace trackers are *differential* in nature, tracking is insensitive to the representation of the channel be it time- or frequency-domain. That said, a salient point in relation to the *phase* behavior of the trackers is worth making for completeness of treatment. The numerical perturbations experienced in the time-domain—as a function of human motion—differ to those experienced in the frequency-domain. Classic work on the stability of subspaces provides bounds on their trigonometric (i.e. angular) behavior as a function of technical mathematical issues ranging from eigenvalue spectral gap to numerical residuals [19].

To highlight this point, we revisit the waveform of the slope subspace tracker for the running activity depicted in Figure 9f.

We perform channel decomposition through to differential unitarity calculations both for the CIR and the CSI versions of measurements (i.e. time & frequency domains). The results are shown in Figure 10. As illustrated in Figures 10a & 10b, it is intuitive to note that the differential tracker performs identically in time and frequency domains. After all, a linear operator (i.e. [I]DFT) translates between one domain to another. The occasional polarity switch in the phase of the differential tracker (Figures 10c & 10d) can be explained by the effects studied in [19]. However, it is interesting to note the increased phase instabilities when running the differential metric on top of CIR measurements over those obtained from CSI measurements. This phenomenon can be readily seen in Figures 10c & 10d. We conjecture that the sparsity in the CIR measurements (i.e. impulse-like nature) compared to the smoother CSI measurements causes numerical instabilities which give rise to added phase instabilities in the subspace. The scatter plot of Figure 10d supports this hypothesis as can be seen by the tighter clustering in the CSI case. However, further investigations are needed to fully illuminate this issue before solid conclusions can be drawn.

V. EVALUATION
In what follows, we showcase how specialized occupancy and activity sensing can be built atop our featurization.

A. OCCUPANCY DETECTION
1) EXPERIMENTAL SETUP
We evaluate the performance of subspace tracking in terms of the robustness of occupancy detection. To evaluate the robustness, we investigate the accuracy of the classification model in new environments. More specifically, we trained the classification model using CSI data obtained from a certain placement and tested its accuracy on different placements.

a: DATA
We collected the CSI data in 8 places and on 41 placements in total. As depicted in figure 11, the places include six rooms, one lobby, and one lounge and have different characteristics such as room layout and furniture position. We collected the CSI data while varying the number of moving people from zero to 2 (P4, P5, P6) and to 3 (the rest). Each session lasted five minutes and participants were asked to freely move during the session. Figure 11 shows room layouts and device placements. The purpose of multiple placements are to
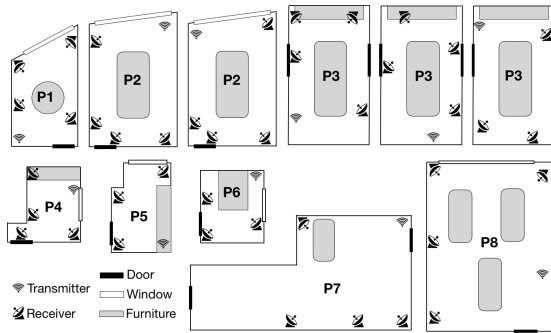
FIGURE 11. Device placements.

FIGURE 12. Occupancy performance. (a) Environment variations. (b) Number of classes.

investigate what is a realistic *upper-bound* on the classification performance of a *single* device under different training and testing conditions. MIMO CSI data were sampled at a nominal 500Hz rate. A stationarity period of 50ms was used and the subspace update was performed in a sliding window fashion with no overlap as in Equation (16).

b: PIPELINE

For the occupancy detection, we developed an inference pipeline using a long short-term memory (LSTM) classifier. We chose LSTM as a classifier to leverage spatio-temporal variation of our differential unitary features from subspace tracking. In the current implementation, we adopted two hidden LSTM layers, each of which has 50 nodes. Some prior presence detection work dwells on the signal much longer with distribution-based approach while using a diversity of frequency channels [20]. In contrast, we define a short 5 seconds inference window and with no channel frequency diversity. In this paper, our objective is to showcase how to specialize various subspace tracking-based applications rather than demonstrate best-in-class performance.

c: COMPARISON

For comparison, we implemented the baseline pipeline from [21]. It takes temporal variations of CSI data as feature values and uses linear discriminant analysis as a classifier.

d: TRAINING AND TEST

For training, we selected a receiver located at a diagonal position of the transmitter, thereby maximizing the RF coverage. Accordingly, we have 11 different models. For the evaluation, we considered three environment variations, *same*, *minor*, and *major*. *Same* refers where the data from the same receiver, i.e., same placement, is used both for training and test. *Minor* and *major* use the CSI data from different receivers placed in the same room and different room, respectively. *Same* represents the upper bound of the performance that the inference logic can achieve in a specific environment. *Minor* and *major* show how robust the inference pipeline is in unseen environments.

2) EXPERIMENTAL RESULTS

We investigate how the subspace tracking effectively mitigates the environmental effect of CSI on the occupancy
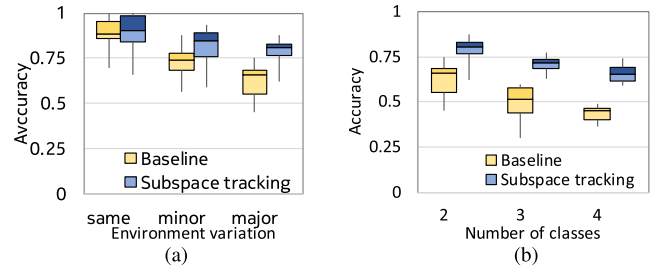
detection. Figure 12a shows the box plots of the accuracy of 11 models for different variations. Although the accuracy of both pipelines is similar in *same* variation, the subspace tracking retains more competitive accuracy as we introduce minor and major environmental changes compared to the baseline. The accuracy in *same* variation is 89% and 88% for the subspace tracking and baseline, respectively. However, in minor and major variations, the subspace tracking decreases to 82% and 78%, whereas the baseline does to 73% and 62%.

We further investigate the effect of the number of classes on the occupancy detection on major variation. Figure 12b shows the box plots of the accuracy while varying the number of classes. *2* classes represent presence detection, i.e., empty or occupied. *3* and *4* classes are for the number of people as [0, 1, 2+] and [0, 1, 2, 3], respectively.[5] The results show that the subspace tracking achieves reasonable performance even with higher number of classes. Our pipeline shows 85%, 70% and 65% for 2, 3, and 4 classes, respectively, whereas the baseline does 62%, 49%, and 43%.

B. PHYSICAL ACTIVITY

We use the activity recognition dataset available publicly by Yousefi *et al.* [1] to demonstrate the applicability of our subspace tracking technique on the problem domain of activity classification. The dataset is comprised of 6 single-user activities; namely, standing up, sitting down, lying down, falling, walking, and running. SIMO CSI data from three receiving multiple antennae is sampled at 1 ksps rate. We choose a stationarity period of 25ms and proceed to update the covariance matrix with 95% CSI overlap from previous stationarity period with $\lambda = 0.99$ for recursive subspace update as in Equation (15). This gives around 800 Hz subspace update rate. As illustrated previously in figure 9, recursive subspace tracking filters background channel activity and/or subspace noise. This unwanted channel activity has been alluded to in Equations (8) & (9).

In a preliminary evaluation, we build a simple classifier based around dynamic time warping (DTW) and K-nearest neighbors. This is applied to a single-dimensional Dy slope differential unitary (see figure 7b). We evaluate our classifier against the author's mid-range hidden Markov model (HMM) which uses a combination of PCA and the short-time

---

[5]By 2+ we mean 2 or greater than.
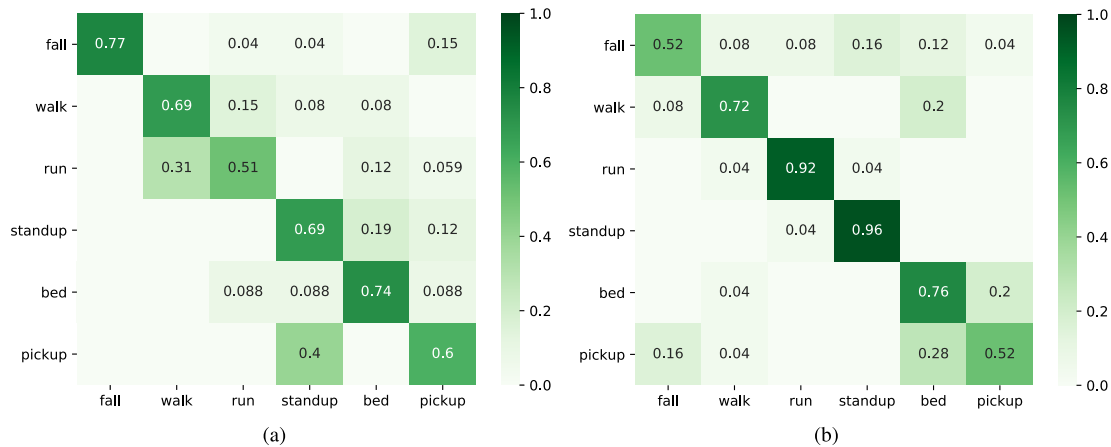
FIGURE 13. Activity recognition performance. (a) Ours: $\hat{u}_{\text{Dy},1}$ + DTW + K-nearest neighbours. (b) Yousef et al: PCA + STFT + HMM.

Fourier transform (STFT) time-frequency localization pre-processing. The results are shown in figure 13. Capability-wise, there is an asymmetry in that featurization based around 2D STFT + HMM is in principle far stronger than our 1D DTW + K-nearest. Nonetheless, on the whole, the performance of our simple classifier is not far from that reported by Yousef et al, albeit with different characteristics. For instance, while 2D STFT + HMM outperforms our 1D DTW + K-nearest in nearly all activities, our fall activity performance is substantially better. We attribute this to the high acceleration content of fall which our slope metric is able to capture easily as shown in figure 9a due to native acceleration sensing. Perhaps our pairwise metric with 2D time-frequency localization would perform much better. Since our focus in this paper is to only showcase a generic formal featurization suited for many applications, we leave improved classification for future work.

VI. DISCUSSION

In this section, we provide commentary on the limitations of our work and discuss relevance to other wireless systems, thereby exposing items of future research.

A. APPLICABILITY TO OTHER 802.11 STANDARDS

Physical propagation behavior will differ depending on the frequency band. Such behavior will be mirrored when viewed through the lens of the signal and noise sub-spaces. Our proposed featurization provides sensing primitives to track the variations in propagation dynamics that are induced by human motion. However, it is the role of the machine learning (ML) component to capture such behavior in a robust sensing model. Thus, when operating within different frequency bands, it is important to ensure that the back-end ML component is trained for the respective human-modulated propagation behavior corresponding to that specific band. Our experimental results in this paper are for the 5GHz Wi-Fi band with 40MHz bandwidth.

Nonetheless, other wireless standards—such as 802.11ah operating in the sub-1GHz band and 802.11ad/ay operating in the 60GHz band—could benefit from identical featurization, albeit after specializing the back-end ML component to capture their individual propagation characteristics as a function of human motion. Moreover, we have shown in Section IV that the magnitude of our differential subspace tracking behaves identically irrespective of the representation of the channel response, be it in time or frequency. This means that both the single-carrier and OFDM variants of WiGig would benefit from our subspace-based featurization. It is also worth pointing out that in relation to WiGig, 60GHz frequencies are quasi-optical and are less able to diffract around objects. The subspace will mirror this behavior; however, increased coverage of the environment may be possible by considering the beam training procedure that 802.11ad/ay implements. Specifically, recent work has shown that such beam training procedure from infrastructure access points can be used to localize a mobile user [22]. It would be interesting in this particular example to see if tracking the subspace would allow for inferring finer-grained details on the nature of the mobile node's movement. OFDMA systems such as 802.11ax can also benefit from the proposed subspace tracking; however, care should be taken to handle instances of transition in user-assigned subcarriers and their implications on the subspace.

B. ML MODEL COVERAGE AND VECTORS OF VARIATION

There are many variables that impact the robustness of the back-end ML model. We call these the vectors of variation of the ML model. Exhaustive training across these vectors of variation is needed for sufficient coverage of the sampling space in order to ensure the ML model generalizes in the real-world. One such vector of variation is that arising from the individualized way in which different users perform activities. Broadly, there are two methods in prior art for dealing with such variations: design-based and learning-based. In design-based methods, hand-crafted features by an

expert designer—such as careful frequency binning in [23] and coarser wavelet spectral bins in [5]—are engineered to absorb the expected variations in the real-world. In contrast, learning-based approaches rely on automatic coverage of these natural variations by the inference component through the sheer amount of empirical data used for training. In this paper, we focused on a formal and interpretable low-dimensional featurization of the wireless channel, with our evaluation (cf. figure 12) falling under the latter learning-based approach.

C. AXES OF RESOLUTION

The performance of sensing applications built atop channel tracking is fundamentally limited by the spatio-temporal resolutions of channel measurements. Specifically, the utilized bandwidth and number of antennae have a large bearing on what can be perceived unambiguously in the environment i.e. without *over-fitting* inference. To see this, consider the environmental imaging capability of the covariance $\mathbf{R}_x$ through its *beamspace* representation $\mathbf{F}\mathbf{R}_x\mathbf{F}^H$, where $\mathbf{F}$ is the Fourier transform matrix [6], [7], [24]. Clearly, for meaningful imaging, the number of antennae needs to be high in order to resolve environmental spatial scatterers. Similarly, bandwidth delivers the temporal resolution necessary for measuring the channel's delayspread (or frequency selectivity) more accurately. It is customary to see in related literature prolonged signal dwell times in order to compensate for the lack of spatio-temporal resolution as supplied by current research testbeds e.g. of the order of minutes dwell time to estimate occupancy in [3] and [20]. To put it in wireless terms, clearly the ''coherence'' time of crowd movement indoors is much shorter than 5 minutes. We, therefore, would argue that practical indoor channel sensing systems are likely to appear once we begin to see the roll-out of wireless infrastructure of enhanced spatio-temporal resolutions such as indoor massive MIMO in the millimeter-wave band.

VII. CONCLUSION

In this paper, we formalize the problem of Wi-Fi-based human sensing and cast it as a channel signal subspace tracking task. We demonstrate the equivalence of the two problems. We posit the optimality of such formulation citing prior established work from wireless literature. We conclude by providing evidence for the applicability of our subspace tracking across two usage scenarios: presence detection and activity recognition with promising early results. Future work will focus on machine learning classification using our subspace-based featurization.

ACKNOWLEDGMENT

REFERENCES

[1] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using WiFi channel state information," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 98–104, Oct. 2017. [Online]. Available: https://github.com/ermongroup/Wifi_Activity_Recognition

[2] W. Xi *et al.*, "Electronic frog eye: Counting crowd using WiFi," in *Proc. IEEE INFOCOM*, Apr./May 2014, pp. 361–369.

[3] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, Jul. 2015.

[4] J. Wang *et al.*, "Low human-effort, device-free localization with fine-grained subcarrier information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 11, pp. 2550–2563, Nov. 2018.

[5] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. MobiCom*. New York, NY, USA: ACM, 2015, pp. 65–76.

[6] N. Costa and S. Haykin, "A novel wideband MIMO channel model and experimental validation," *IEEE Trans. Antennas Propag.*, vol. 56, no. 2, pp. 550–562, Feb. 2008.

[7] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[8] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[9] F. J. Lopez-Martinez, E. Martos-Naya, J. F. Paris, and A. Goldsmith, "Eigenvalue dynamics of a central wishart matrix with application to MIMO systems," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2693–2707, May 2015.

[10] J. P. Delmas, "Subspace tracking for signal processing," *Adaptive Signal Processing*. Hoboken, NJ, USA: Wiley, 2010, ch. 4, pp. 211–270. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470575758.ch4. doi: 10.1002/9780470575758.ch4.

[11] *3rd Generation Partnership Project; Technical Specification Group GSM/EDGE Radio Access Network; Radio Transmission and Reception (Release 1999)*, document 3GPP v8.20.0, Annex C.3 Propagation models 1999.

[12] D. B. Smith and L. W. Hanlen, "Channel modeling for wireless body area networks," *Ultra-Low-Power Short-Range Radios*, P. P. Mercier and A. P. Chandrakasan, Eds. Cham, Switzerland: Springer, 2015, pp. 25–55. doi: 10.1007/978-3-319-14714-7_2.

[13] D. B. Smith, D. Miniutti, T. A. Lamahewa, and L. W. Hanlen, "Propagation models for body-area networks: A survey and new outlook," *IEEE Antennas Propag. Mag.*, vol. 55, no. 5, pp. 97–117, Oct. 2013.

[14] A. Fort, C. Desset, P. Wambacq, and L. V. Biesen, "Indoor body-area channel model for narrowband communications," *IET Microw., Antennas Propag.*, vol. 1, no. 6, pp. 1197–1203, Dec. 2007.

[15] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[16] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.

[17] C. Studer, S. Medjkouh, and E. Gönültaş, T. Goldstein, and O. Tirkkonen, "Channel charting: Locating users within the radio environment using channel state information," *IEEE Access*, vol. 6, pp. 47682–47698, 2018.

[18] Y. Zhang *et al.*, "A general coupling-based model framework for wideband MIMO channels," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 574–586, Feb. 2012.

[19] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation. III," *SIAM J. Numer. Anal.*, vol. 7, no. 1, pp. 1–46, 1970.

[20] X. Guo, B. Liu, C. Shi, H. Liu, Y. Chen, and M. C. Chuah, "WiFi-enabled smart human dynamics monitoring," in *Proc. SenSys*. New York, NY, USA: ACM, 2017, pp. 16:1–16:13.

[21] S. Di Domenico, M. De Sanctis, E. Cianca, and G. Bianchi, "A trained-once crowd counting method using differential WiFi channel state information," in *Proc. WPA*. New York, NY, USA: ACM, 2016, pp. 37–42.

[22] J. Palacios, P. Casari, and J. Widmer, "JADE: Zero-knowledge device localization and environment mapping for millimeter wave systems," in *Proc. IEEE INFOCOM*, May 2017, pp. 1–9.

[23] W. Wang, L. Yang, Q. Zhang, and T. Jiang, "Securing on-body IoT devices by exploiting creeping wave propagation," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 696–703, Apr. 2018.

[24] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

Authors' photographs and biographies not available at the time of publication.

• • •