

Toward Opportunistic Radar Sensing Using Millimeter-Wave Wi-Fi

Jian Wang^{1b}, Jack Chuang^{1b}, Samuel Berweger^{1b}, Camillo Gentile^{1b}, and Nada Golmie^{1b}

Abstract—Sensing with communication waveforms has drawn growing interest thanks to the ubiquitous availability of wireless networks. However, the required sensing resources may not always be available in a communication system. In addition, the communication system may have limited bandwidth, beamwidth, and transmit power, which could limit the sensing accuracy. To investigate such challenges, in this article, we study the feasibility of using the sector-level sweeping (SLS) procedure of IEEE 802.11ad to provide opportunistic indoor radar sensing service, which is vital to smart Internet of Things (IoT) applications. In particular, we design a framework to estimate the target’s spatial position with respect to delay and angle by employing the multiple signal classification (MUSIC) super-resolution algorithms. We conduct an extensive performance evaluation to understand the tradeoffs between sensing accuracy and required sensing resources in terms of system configurations (e.g., antenna array size and the overlapping of neighboring beams) and the impact of signal-to-noise ratio (SNR). Furthermore, based on the human multipath reflections captured from a real-world measurement campaign, we reconstruct the sensing channel, investigate the feasibility of monitoring the gesture behavior in a smart home environment, and discuss some findings and insights.

Index Terms—Edge processing, Internet of Things (IoT) applications, mmWave, opportunistic radar sensing, Wi-Fi sensing.

I. INTRODUCTION

PERFORMING sensing using communication waveforms has drawn growing interest recently, as sensing and communication can share common hardware modules, signal processing modules, as well as achieve efficient spectrum usage by using the same waveform for both purposes [1]. The same waveform used to deliver information to the intended receiver could also be used to illuminate the target to carry out sensing tasks by collecting and analyzing the scatterings from the target. To this end, channel estimation is essential to both sensing and communication components. In a communication system, channel estimation is the key to mitigating the impact of the frequency selective fading and Doppler effect to enable reliable communication. Meanwhile, channel estimation can be a viable method to detect and track sensing targets. With the ubiquitous availability of wireless networks, various smart-world Internet of Things (IoT) applications (smart cities, smart

homes, smart manufacturing, smart transportation, etc.) could be supported by the omnipresent sensing capability [1], [2], [3], [4], [5], [6], [7], [8].

Some existing research efforts have been conducted to investigate using commercial wireless standards to support various real-world IoT applications, such as vehicular range and speed detection [9], human activity recognition [10], [11], and environment monitoring [12], among others. These studies mainly focused on accomplishing a given sensing task, assuming the required sensing resource is always available without taking the communication performance into consideration. On the other hand, opportunistic sensing is a viable approach to enable simultaneous sensing and communication and have minimum impact on the performance of the existing communication system. The key idea of opportunistic sensing is to discover sensing possibilities with the existing communication protocols and procedures, aiming to obtain the sensing service “free” or at a very low cost in terms of resource usage.

However, designing effective opportunistic sensing needs to address several challenges. First, since the waveforms are designed with only communication performance in mind, the signal may have limited bandwidth, low signal power at the sensing receiver, and the nonideal range–Doppler ambiguity function [1], which could affect sensing performance. Second, to fulfill certain communication tasks, the sensing resources could be limited by availability, for example, the data flow is scheduled in an on-demand fashion. Thus, it is critical to not only consider the minimum sensing requirements to support a given application but also to design efficient and effective sensing algorithms to maximize sensing performance.

One essential sensing task is target localization, i.e., target range and angle estimation. The conventional approach to estimate range is to detect peaks in the delay profile of the matched filter output, each corresponding to paths scattered from distinct ambient objects. However, the delay resolution can be limited by the signal bandwidth B with $\Delta t \geq \frac{1}{B}$. Similarly, when identifying the signal direction, the angle resolution will be limited by the beamwidth if using the direction information of the highest power beam. Using IEEE 802.11ad single carrier (SC) signal (1.76-GHz bandwidth) as an example, the achievable 2-way delay resolution is 0.57 ns, corresponding to 8.5-cm range resolution. Also, with a 16-element uniform linear array (ULA), the achievable angle resolution is 6.4° , which may not be sufficient for some medium range (> 0.5 m) or large range (> 2 m) applications, such as gesture recognition in smart environments [1].

In this article, we have made the following contributions.

Manuscript received 15 May 2023; revised 24 June 2023; accepted 21 July 2023. Date of publication 2 August 2023; date of current version 25 December 2023. (Corresponding author: Jian Wang.)

The authors are with the Wireless Networks Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: jian.wang@nist.gov; jack.chuang@nist.gov; samuel.berweger@nist.gov; camillo.gentile@nist.gov; nada.golmie@nist.gov).

Digital Object Identifier 10.1109/JIOT.2023.3301006

U.S. Government work not protected by U.S. copyright.

- 1) To extend the detection resolution beyond the limitation of the system bandwidth and beamwidth, we design a framework to estimate the target's spatial position with respect to delay and angle by employing the multiple signal classification (MUSIC) super-resolution algorithms. Using IEEE 802.11ad beamforming training as a representative scenario, we leverage its pilot signal, including the synchronization and channel estimation fields (CEFs), to estimate the channel impulse response (CIR) for each beamforming direction. Considering one round of the sequentially switched beams as one observation, we estimate the target delay and angle by combining all the CIRs from each direction and then applying MUSIC super-resolution algorithms accordingly. To our best knowledge, our design is a new way of using the MUSIC algorithm on estimated CIRs to extract targets' delays and angles, providing the intelligent edge service for Wi-Fi sensing-based IoT applications.
- 2) We conduct an extensive performance evaluation to investigate performance tradeoffs by considering different beam widths (i.e., antenna array sizes), the level of overlapping of the neighboring beams, as well as the impact of the signal-to-noise ratio (SNR). Furthermore, we evaluate the efficacy of our approach in a gesture recognition application. Based on the target reflections captured from a real-world measurement campaign, we reconstruct the target propagation channel and evaluate the feasibility of monitoring and recognizing the gesture motion using the existing communication procedures and protocols.

The remainder of this article is organized as follows. In Section II, we provide the preliminary of opportunistic radar sensing. In Section III, we introduce our system model. In Section IV, we present our detection approach in detail. In Section V, we present the performance evaluation and results. In Section VI, we give the literature review. Finally, we conclude this article in Section VII.

Notations: A matrix, a vector, and a scalar are denoted as \mathbf{A} , \mathbf{a} , and a , respectively. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transpose, and conjugate transpose, respectively. $\text{diag}(\mathbf{a})$ is a diagonal matrix with the elements of \mathbf{a} as its diagonal elements. $\mathbf{A}[:, k]$ is the k th column of \mathbf{A} . $\|\mathbf{a}\|$ denotes the Forbenius norm. $\mathbf{A} \circ \mathbf{B}$ and $\mathbf{A} \otimes \mathbf{B}$ are Hadamard product and Kronecker product of matrices \mathbf{A} and \mathbf{B} , respectively. $\text{vec}(\mathbf{A})$ is the vectorization operation that stacks the columns of \mathbf{A} together to form a single column.

II. OPPORTUNISTIC RADAR SENSING

In this study, we consider opportunistic radar sensing that utilizes the IEEE 802.11-ad/ay (i.e., directional multigigabit (DMG) and enhanced DMG (EDMG) standards), operating at 60-GHz mmWave frequency. In order to meet the link budget requirement, beamforming, which applies an antenna weight vector (AWV) on the elements of a phased array antenna to concentrate the energy to a desired direction to boost the antenna gain, is necessary for achieving reliable communication at mmWave band. However, the communication

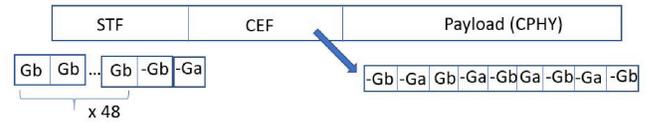


Fig. 1. SLS frame structure in IEEE 802.11ad. (The STF and CEF consist of 59 Golay sequences, with a total of 7552 symbols.)

performance can be greatly deteriorated due to the misalignment between the transmitting and receiving beams. Thus, beamforming training and beam refinement, as well as a periodic procedure to realign the transmitter and receiver beams, are required to maintain directional communication.

In an IEEE 802.11-ad/ay system, beamforming training and beam refinement are two typical procedures for beam alignment. The basic idea of these procedures is to sequentially transmit or receive a directional beamformed signal to find the best communication direction. In particular, sector-level sweeping (SLS) is an essential procedure in beamforming training. In SLS, to accommodate the unknown location of the receiver, the coverage area is divided into multiple sectors; access point (AP) or station (STA) sweeps from sector to sector sequentially, exchanging control PHY (CPHY) frames and identifying the sector that provides the best channel condition such as the highest receiving power. Suppose an object is located within the SLS-covered area. In that case, the reflections from the target could be exploited at the transmitter as a monostatic radar deployment, which can be further analyzed to estimate the target's position. In addition, with the periodic SLS procedure, the time evolving of the target reflection can be monitored to track the trajectory of the target movement and the Doppler shift caused by the motion.

Beacon interval (BI) is the super frame in the IEEE 802.11ad, defined as the time interval between the two consecutive beacon frames transmitted by the AP. Each BI starts with a beam transmission interval (BTI), during which, SLS is performed to send network announcements and perform beamforming training. The maximum value of BI is 1000 ms [13], and it is in the increment of one unit, with one unit equals 1/1024 s [14]. Thus, SLS is a periodic procedure at the AP, and the interval of the SLS can be configurable depending on system requirements. Shorter BI could incur more considerable overhead due to frequent beamforming training, while longer BI may increase the device initial attach time or even lead to connection loss. A typical BI value is set to 100 ms [13].

During SLS, the transmitter and receiver exchange a CPHY packet at each sector. The frame structure of the SLS frame is illustrated in Fig. 1. The SLS frame consists of the preamble and the payload. The preamble, composed of a short training field (STF) and a CEF, 59 length-128 Golay sequence in total, can be used for frame start detection and channel estimation. In this study, we use the preamble field to estimate the channel in each sector direction.

Beam refinement is another potential opportunity to be exploited to enable target sensing during the beam refinement phase. An optional training field (TRN) can be appended to the DMG data packet to perform additional beamform training to improve the SNR of the communication link. Each TRN

field consists of a number of TRN subfields. Up to 64 TRN subfields can be appended to a packet to refine a transmitter beam, with each TRN subfield transmitting to a distinct direction by applying a different AWW configuration. Compared to SLS, the TRN sequential beam sweeping can be completed in a single physical layer protocol data unit (PPDU), which can accommodate rapid changes in the environment. Similarly, a single PPDU could be used to extract the spatial position of the target, and multiple PPDUs transmitting over time can be leveraged for target tracking and velocity estimation.

In this study, we demonstrate the performance and feasibility of leveraging AP SLS at the beginning of each BI for opportunistic sensing. Since the TRN subfield also contains complementary Golay sequences, which can be used for channel estimation, the same target estimation techniques can be easily extended to the beam refinement phase.

III. SYSTEM MODEL

Our system model assumes that the AP operates in a full-duplex mode that incorporates both transmitter and receiver functionalities and can transmit and receive simultaneously. In order to perform monostatic sensing, the transmitter and receiver are both co-located at the AP. In particular, the transmitter (i.e., AP) sends out CPHY at each sector. In the meantime, the sensing receiver performs radar functionality by monitoring the reflection signals from the environment.

A. Channel Model

Generally speaking, the channel model that describes the reflection paths from the environment is

$$\mathbf{H}(t, \tau, \theta, \phi) = \sum_{l=1}^L g_l e^{j2\pi\nu(l)t} \delta(t - \tau_l) \mathbf{a}_r(\theta_l^r, \phi_l^r) \mathbf{a}_t(\theta_l^t, \phi_l^t)^* \quad (1)$$

where L is the total number of multipath component (MPC), g_l is the complex path gain of the l th path due to the signal propagation, and τ_l is the two-way path delay of the l th path. Also, $[\theta_l^r, \phi_l^r]$ and $[\theta_l^t, \phi_l^t]$ are the Angle of Arrival (AoA) and the Angle of Departure (AoD) of the l th MPC. Since the monostatic configuration is used in our study, we assume the same AoA and AoD angles. The symbols θ and ϕ represent the angles in the azimuth and elevation planes, respectively, and $\nu(l)$ is the Doppler shift of the l th path. Furthermore, $\mathbf{a}_r(\theta, \phi)$ and $\mathbf{a}_t(\theta, \phi)$ in (1) are the receive and transmit array response function vectors.

In one SLS, the AP sequentially steers toward a set of predefined directions to cover one area. Note that, in this study, we assume the AP only sweeps in the azimuth direction and keep the elevation steer angle fixed at 0° . Without loss of generality, we consider a linear phased array antenna with M elements at half-wavelength spacing, focusing on estimating the azimuth of the MPCs. Note that the algorithm can be extended to jointly estimate azimuth angle and elevation angle, when the beam is configured to steer in the elevation direction as well [15].

Denote M as the number of elements of the ULA. The array response function $\mathbf{a}(\theta)$ is given as

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}} \left[1, e^{j\pi \sin\theta}, e^{j2\pi \sin\theta}, \dots, e^{j(M-1)\pi \sin\theta} \right] \quad (2)$$

where θ is the azimuth angle of a propagation path.

In this study, we assume the phased array antenna at both the transmitter and the receiver has M elements. The analog beamforming matrices \mathbf{F} and \mathbf{W} are used at the transmitter and receiver, respectively, to steer the transmitter and receiver beams. The transmitter and receiver steer in the same direction, i.e., $\mathbf{F} = \mathbf{W}$. At the k th steer direction, beam steering vectors $\mathbf{F}[:, k] = \mathbf{a}_t(\theta_k) = \mathbf{a}_r(\theta_k) = \mathbf{W}[:, k] = \mathbf{a}(\theta_k)$ apply to the channel matrix $\mathbf{H}(t, \tau, \theta, \phi)$ and obtain a tap-delay channel response as follows:

$$\begin{aligned} \mathbf{h}_{\theta_k}(t, \tau) &= \sum_{l=1}^L \mathbf{a}_r(\theta_k)^* \mathbf{H}(t, \tau_l, \theta_l) \mathbf{a}_t(\theta_k) \\ &= \sum_{l=1}^L \alpha_l(\theta_k, \theta_l) e^{j2\pi\nu_l t} \delta(t - \tau_l) \end{aligned} \quad (3)$$

where $\alpha_l(\theta_k, \theta_l) = g_l \mathbf{a}_r(\theta_k)^* \mathbf{a}_r(\theta_l) \mathbf{a}_t(\theta_l) \mathbf{a}_t(\theta_k)$. We assume the delay, AoA, AoD, and the path gain stay the same during a single SLS.

B. Signal Model

The transmit signal consisting of the length L_s pilot sequence in an SLS data frame can be described as

$$s(t) = \sqrt{E_s} \sum_{m=0}^{L_s-1} b(m) p(t - mT_s) \quad (5)$$

where $p(\cdot)$ is the transmit pulse, E_s is the symbol energy, and $b(m)$ is the modulated symbols after $(\pi/2)$ -BPSK modulation with $b(m) \in \{1, -1, i, -i\}$.

At the receiver, with θ_k as the steering angle, when L copies of the transmit signal received through the L propagation paths, we have the received signal as follows:

$$r(t, \theta_k) = \sum_{l=0}^{L-1} \alpha_l(\theta_k, \theta_l) e^{j(2\pi\nu_l t)} s(t - \tau_l) + w(n) \quad (6)$$

where ν_l is the Doppler shift of path l , α_l is the complex gain of path l , τ_l is the propagation delay of the path l , and $w(n)$ is a complex circularly symmetric Gaussian random variable $CN(0, 2\sigma_w^2)$.

After the transmitter completes transmitting one CPHY data packet, it will switch to send a new CPHY data packet in the next direction. The same waveform as (6) will be received, but for a different steering angle θ_k . With the periodic SLS procedure, the AP could continuously monitor the environmental changes or the target's motion. In this study, we assume a slow-changing environment, i.e., very small ν_l , and $\alpha_l(\theta_k, \theta_l)$, θ_l , and τ_l are unchanged during one round of the SLS sweeping.

IV. ANGLE AND DELAY ESTIMATION

When SLS takes place, AP will scan through its coverage area using a directional beam formed through analog beamforming. A CPHY data frame is transmitted at each steering

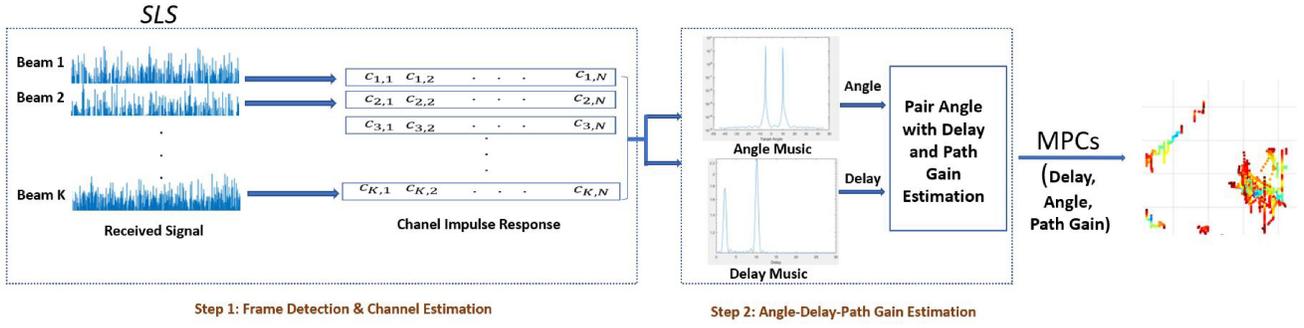


Fig. 2. Workflow for delay, angle, and path gain estimation.

angle, and the CIR will be estimated for each directional channel. The CIRs of each direction will be further combined to accomplish the radar functionality.

Fig. 2 describes the workflow of our approach. As seen in the figure, the SLS can be used to detect and track the target changes. In *Step 1: Frame Detection and Channel Estimation*, for each received CPHY data frame, our detection algorithm takes the received packets and computes the CIR based on the received pilot signal. Then, in *Step 2: Angle-Delay-Path Gain Estimation*, with the K CIRs, corresponding to K switched beams, the MUSIC algorithm is adopted to detect the number of propagation paths and their angles, delays, and path gain values. In particular, in our proposed approach, we first perform 1-D angle MUSIC estimation and 1-D delay MUSIC estimation separately. We then use maximum likelihood to pair the estimated angles and delays obtained from the 1-D MUSIC estimations. Finally, with periodic SLS, the detected paths can be tracked over time to detect the target motion pattern and support various applications. Our designed approach can be deployed at edge device to provide intelligent IoT applications based on Wi-Fi sensing. In the following, we describe the steps mentioned above in detail.

A. Step 1: Frame Detection and Channel Estimation

Frame detection and channel estimation are performed on each received beamformed waveform to obtain the CIR. Recall that communication receivers have traditionally used CIR to perform channel equalization to mitigate multipath effects such as frequency selective fading. More specifically, we carry out the following two key procedures in this step: 1) *Frame Detection Procedure*: We first perform frame detection to detect the frame and obtain a rough estimation of the start of the frame and 2) *CIR Estimation Procedure*: We leverage the Golay complementary sequences in the pilot to refine the frame timing and extract CIR.

1) *Frame Detection*: To detect the start of the frame, the cross correlation of the received waveform with the known pilot is first computed. Using the entire CPHY preamble can generate large peaks; however, it also introduces many high side peaks separated at the distance of the multiple of the Golay length [16], [17] due to the 48 repetitive length-128 Golay sequences in the STF, which is nonideal when SNR is low or for multitarget detection, as shown in Fig. 3. In addition, with Fig. 3(b), we can observe that cross correlation with

the last two Golay sequences in STF combined with the CEF field results in higher peak to sidelobe ratio. Thus, in this study, we choose the known correlation signal as the signal containing the 49th Golay sequence to the 59th Golay sequence in the CPHY preamble for frame detection.

The received signal is first detected by searching the start of the 49th Golay sequence \hat{l} that satisfies

$$\hat{l} = \arg \max_{l \in \mathbb{Z}} \frac{1}{L\sigma_w^2} \left| \sum_{i=0}^L r(l+i)b^*(i) \right|^2 > \gamma \quad (7)$$

where L is the symbol length of the correlation signal, \mathbb{Z} is the set of integer numbers, σ_w is the variance of the real and imaginary components of the complex white Gaussian noise $CN(0, 2\sigma_w^2)$, and $b(i)$ is the i th symbol of the correlation signal.

The γ value can be chosen to satisfy a given false alarm rate. When the received signal only contains the complex white Gaussian noise $CN(0, 2\sigma_w^2)$, the auto-correlation $\gamma(l)$ for a frame start value l can be computed as

$$\begin{aligned} \gamma(l) &= \frac{1}{L\sigma_w^2} \left| \sum_{m=0}^{L-1} w(l+m)b^*(m) \right|^2 \\ &= \frac{1}{L\sigma_w^2} \text{Re} \left[\sum_{m=0}^{L-1} w(l+m)b^*(m) \right]^2 \\ &\quad + \frac{1}{L\sigma_w^2} \text{Im} \left[\sum_{m=0}^{L-1} w(l+m)b^*(m) \right]^2. \end{aligned} \quad (8)$$

Both real and imaginary of $\gamma(l)$ follow a normal distribution with zero mean and unit variance. Thus, $\gamma(l)$ follows a central Chi-square distribution ($\chi^2(2)$) with degree of freedom of two.

False alarm rate $P_{FA} = \Pr(\gamma(l) > \gamma_0) = 1 - F_{\chi^2}(\gamma_0, 2)$. For example, to maintain a constant false positive rate of 0.1%, the γ should be larger than 13.8. In our detection algorithm, we set the threshold to be 13.8 for more reliable channel estimation. Note that with this threshold setting, we can detect a frame with SNR as low as -20 dB.

2) *CIR Estimation*: After detecting the frame, the next task is to locate the CEF field in the preamble to estimate the channel and obtain its CIR. The CEF field consists of two sets of 256-sample Golay complementary sequence pair \mathbf{G}_u and \mathbf{G}_v , where $\mathbf{G}_u = [a_{256,u} \ b_{256,u}] = [-b_{128} \ -a_{128} \ b_{128} \ -a_{128}]$ and $\mathbf{G}_v = [a_{256,v} \ b_{256,v}] = [-b_{128} \ a_{128} \ -b_{128} \ -a_{128}]$. The complementary Golay sequence has good correlation properties,

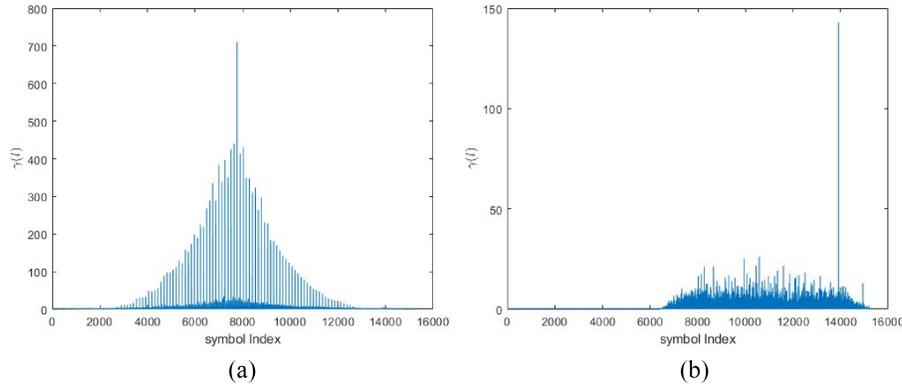


Fig. 3. Frame detection using (a) entire preamble and (b) last two Golay sequences in the STF plus the CEF field.

i.e., the sum of the autocorrelation function of the complementary Golay sequence has a value of zero at all lags except when lag equals zero, which can be mathematically expressed as $\sum_{i=0}^{L-1-l} a_{256}(i)a_{256}^*(i+l) + \sum_{i=0}^{L-1-l} b_{256}(i)b_{256}^*(i+l) = 2L\delta(l)$, where $L = 256$ is the length of the complementary sequence.

To estimate the channel, the received symbols are correlated with each Golay complementary sequence, and the correlations are summed as follows:

$$g(i) = \frac{1}{4L} \left(\sum_{l=0}^{L-1} r(l+i)a_{256,u}^*(l) + \sum_{l=0}^{L-1} r(l+i+L)b_{256,u}^*(l) + \sum_{l=0}^{L-1} r(l+i+2L)a_{256,v}^*(l) + \sum_{l=0}^{L-1} r(l+i+3L)b_{256,v}^*(l) \right). \quad (9)$$

Within $g(i)$, we search and locate a length 128 CIR that has the maximum sum power and store it as the estimated CIR for the steer direction k as $\hat{h}(k, n)$, $n \in [1, 128]$.

With CIR estimation for each steering angle, the delay estimate resolution is limited by the signal bandwidth, and the beamwidth of the switched analog beams. Cramer–Rao lower bound (CRLB) provides a lower bound of the estimation errors, which is a theoretical estimation limit [18]. Assuming additive white Gaussian noise (AWGN), CRLB of the range estimation can be computed as

$$\sigma_r^2 \geq \frac{c^2}{\frac{2}{3}\pi^2 B^2 L_s (\text{SNR})} \quad (10)$$

where c is the speed of light, B is the channel bandwidth, which is 1.76 GHz for IEEE 802.11ad CPHY, and L_s is the number of symbols in the detection sequence. According to (10), the detection error decreases with the increase of the SNR and the length of the pilot used for estimation. With $\text{SNR} = -10$ dB and 10 Golay sequence with $L_s = 128 \times 10$, the CRLB can be as low as 0.66 cm, which motivate us to explore super-resolution estimators to improve the detection resolution.

B. Step 2: Angle–Delay–Path Gain Estimation

Denote K as the number of sectors covering a given area. After the beam sweeps through K sectors, with one data packet transmitting in each direction, we have K estimated CIRs, denoted as $\hat{h}(1), \hat{h}(2), \dots, \hat{h}(K)$. In this step, we use the K estimated CIRs to calculate the arrival azimuth, the delay of the reflection paths, as well as their path gains. Note that the underlying assumption is that the channel is relatively static during each SLS. Thus, the channel parameters, including the AoAs, path delays, and path gains, remain unchanged. In the following, we describe the essential procedures: angle estimation, delay estimation, and pairing path delay and path angle with path gain estimation in detail.

1) *Angle Estimation:* For sector k , the transmit and receive steering vectors are denoted as $\mathbf{W}[:, k]$ and $\mathbf{F}[:, k]$, respectively. The estimated channel gain for the q th symbol can be described as

$$\begin{aligned} \hat{h}_k(q) &= \sum_{l=1}^L \mathbf{W}[:, k]^H g_l \mathbf{a}_r(\theta_l) \mathbf{a}_t^H(\theta_l) \mathbf{F}[:, k] p(qT_s - \tau_l) + n(q) \\ &= ((\mathbf{W}[:, k]^H \mathbf{A}_R) \circ (\mathbf{F}[:, k]^T \mathbf{A}_T^*)) \boldsymbol{\alpha}(q) + n(q) \end{aligned} \quad (11)$$

where

$$\begin{aligned} \boldsymbol{\alpha}(q) &= [g_1 p(qT_s - \tau_1), g_2 p(qT_s - \tau_2), \dots, g_L p(qT_s - \tau_L)]^T \\ &= [\alpha_1(q), \alpha_2(q), \dots, \alpha_L(q)]^T \end{aligned} \quad (12)$$

which is the path gain of the L MPCs at the q th tap, with $p(\cdot)$ as the transmit baseband pulse. And, \circ is the Hadamard product, $\mathbf{A}_T = [\mathbf{a}_t(\theta_1), \mathbf{a}_t(\theta_2), \dots, \mathbf{a}_t(\theta_L)]$ and $\mathbf{A}_R = [\mathbf{a}_r(\theta_1), \mathbf{a}_r(\theta_2), \dots, \mathbf{a}_r(\theta_L)]$. Note that $n(q)$ is the AWGN noise following $CN(0, \sigma^2)$. Due to the monostatic configuration, with the same phased array antenna configuration adopted for both transmitter and receiver, (11) can be further simplified as $\mathbf{W}[:, k] = \mathbf{F}[:, k]$ and $\mathbf{A}_T = \mathbf{A}_R = \mathbf{A}$.

The q th delay tap of the estimated CIRs at beams $1, 2, \dots, K$ can be expressed as

$$\begin{aligned} \hat{\mathbf{h}}(q) &= [\hat{h}_1(q) \hat{h}_2(q) \dots \hat{h}_K(q)]^T \\ &= ((\mathbf{W}^H \mathbf{A}_R) \circ (\mathbf{F}^T \mathbf{A}_T^*)) \boldsymbol{\alpha}(q) + \mathbf{n}(q) \\ &= ((\mathbf{W}^H \mathbf{A}) \circ (\mathbf{W}^T \mathbf{A}^*)) \boldsymbol{\alpha}(q) + \mathbf{n}(q) \\ &= \mathbf{B} \boldsymbol{\alpha}(q) + \mathbf{n}(q) \end{aligned} \quad (13)$$

where $\mathbf{B} = (\mathbf{W}^H \mathbf{A}) \circ (\mathbf{W}^T \mathbf{A}^*)$. The estimated CIR at each of the N delay taps can be used to estimate the covariance matrix by computing $\hat{\mathbf{R}}_h = \mathbb{E}(\hat{\mathbf{h}}(q)\hat{\mathbf{h}}(q)^H) = \frac{1}{N} \sum_{q=1}^N \hat{\mathbf{h}}(q)\hat{\mathbf{h}}(q)^H$ and N is the length of the estimated channel that consists of 128 samples as the 128 observations in our case. Then, we have $K \times K$ covariance matrix $\hat{\mathbf{R}}_h = \mathbf{B} \mathbb{E}(\boldsymbol{\alpha}(q)\boldsymbol{\alpha}(q)^T) \mathbf{B}^T + \sigma^2 \mathbf{I} = \mathbf{B} \boldsymbol{\Lambda}_\alpha \mathbf{B}^T + \sigma^2 \mathbf{I}$, where $\boldsymbol{\Lambda}_\alpha = \text{diag}\{\mathbb{E}(|\alpha_1(q)|^2), \mathbb{E}(|\alpha_2(q)|^2), \dots, \mathbb{E}(|\alpha_L(q)|^2)\}$, assuming path gain $\alpha_l(q)$ following an independent distribution with zero mean.

To perform the MUSIC algorithm, $\hat{\mathbf{R}}_{hh}$ is first separated into the signal subspace and the noise subspace through eigendecomposition. The eigenvectors corresponding to the largest L eigenvalues are the signal eigenvectors spanning the signal space. The eigenvectors corresponding to the smallest $K - L$ eigenvalues are the noise eigenvectors that construct the noise space \mathbf{E}_N . The angle spectrum can be computed as follows:

$$G(\theta) = \frac{1}{\mathbf{b}^*(\theta) \mathbf{E}_N \mathbf{E}_N^H \mathbf{b}(\theta)} \quad (14)$$

where $\mathbf{b}(\theta) = \mathbf{W}^H \mathbf{a}(\theta) \mathbf{a}(\theta)^H \mathbf{W}$. Since the noise space is orthogonal to the signal space, the θ values corresponding to the L largest peaks of the angle spectrum are the estimated azimuth angles of the L paths. Note that, in our approach, the number of signals is estimated by sorting the eigenvalues after covariance matrix eigendecomposition and locating the abrupt value changes of the eigenvalues to separate the signal space and noise space.

2) *Delay Estimation*: Similarly, the MUSIC algorithm can be used to estimate path delays. We rewrite (11) as

$$\hat{h}_k(q) = \sum_{l=1}^L \beta_{k,l} p(qT_s - \tau_l) + n_k(q) \quad (15)$$

where $\beta_{k,l}$ is the path gain at beam k for path l , and $\beta_{k,l} = g_l \mathbf{W}[:, k]^H \mathbf{a}_r(\theta_l) \mathbf{a}_t^H(\theta_l) \mathbf{F}[:, k]$. The CIR observed at beam k , i.e., $\hat{\mathbf{h}}_k = \mathbf{P} \boldsymbol{\beta}_k$, is a length-128 column vector, with $\mathbf{P} = [\mathbf{p}(\tau_1), \mathbf{p}(\tau_2), \dots, \mathbf{p}(\tau_L)]$ and $\boldsymbol{\beta}_k = [\beta_{k,1}, \beta_{k,2}, \dots, \beta_{k,L}]^T$. For the total K beams, we have K observations $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \dots, \hat{\mathbf{h}}_K]$, represented by a size $128 \times K$ matrix, which can be used to estimate path delays. To this end, we compute the estimated covariance matrix $\hat{\mathbf{R}}_H = \mathbb{E}(\hat{\mathbf{H}} \hat{\mathbf{H}}^H) = \mathbf{P} \mathbb{E}(\boldsymbol{\beta} \boldsymbol{\beta}^H) \mathbf{P}^H + \mathbf{N} \mathbf{N}^H$, where $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K]$. Assume, for a given beam k , $\beta_{k,l}$ follow the independent distribution with zero mean, we have $\hat{\mathbf{R}}_H = \mathbf{P} \boldsymbol{\Lambda}_\beta \mathbf{P}^H + \sigma^2 \mathbf{I}$, with $\boldsymbol{\Lambda}_\beta = \text{diag}\{\mathbb{E}(\|\boldsymbol{\beta}[:, 1]\|^2), \mathbb{E}(\|\boldsymbol{\beta}[:, 2]\|^2), \dots, \mathbb{E}(\|\boldsymbol{\beta}[:, L]\|^2)\}$, where $\boldsymbol{\beta}[:, l] = [\beta_{1,l}, \beta_{2,l}, \dots, \beta_{K,l}]^T$.

To estimate the delay, we take the eigenvalue decomposition of $\hat{\mathbf{R}}_H$. We partition the eigenvectors to have the eigenvectors corresponding to the $K - L$ smallest eigenvalues to form the noise subspace. Finally, we compute the delay spectrum

$$G(\tau) = \frac{1}{\mathbf{p}^H(\tau) \mathbf{U}_n \mathbf{U}_n^H \mathbf{p}(\tau)}. \quad (16)$$

By leveraging the fact that the signal subspace \mathbf{U}_s and noise subspace \mathbf{U}_n are orthogonal, the delay of the L path is estimated by locating the L peaks in the computed MUSIC delay spectrum.

3) *Pair Delay With Angle and Path Gain Estimation*: In order to pair the path delays detected by the delay MUSIC and the angles detected by the angle MUSIC, we perform the maximum-likelihood detection and search through the combinations of the m detected angles and n path delays. For example, if we have detected three AoAs, i.e., $[\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3]$ and two delays $[\hat{\tau}_1, \hat{\tau}_2]$. For a given detected angle $\hat{\theta}_l, l \in [1, 3]$, we can select the $\hat{\tau}_i$ by computing the projection of $\text{vec}(\hat{\mathbf{H}})$ on $\mathbf{r}(\hat{\theta}_l, \hat{\tau}_i)$ that satisfies

$$\max_{\hat{\tau}_i} \frac{|\mathbf{r}^H(\hat{\theta}_l, \hat{\tau}_i) \text{vec}(\hat{\mathbf{H}})|}{\|\mathbf{r}(\hat{\theta}_l, \hat{\tau}_i)\|^2} \quad (17)$$

and pair it with $\hat{\theta}_l$. Here, $\mathbf{r}(\hat{\theta}_l, \hat{\tau}_i) = \mathbf{b}(\hat{\theta}_l) \otimes \mathbf{p}(\hat{\tau}_i)$, which is the Kronecker product of $\mathbf{b}(\hat{\theta}_l)$ and $\mathbf{p}(\hat{\tau}_i)$.

In particular, $\mathbf{b}(\hat{\theta}_l)$ can be computed as

$$\mathbf{b}(\hat{\theta}_l) = \begin{pmatrix} \mathbf{W}[:, 1]^H \mathbf{A}(\hat{\theta}_l) \mathbf{W}[:, 1] \\ \mathbf{W}[:, 2]^H \mathbf{A}(\hat{\theta}_l) \mathbf{W}[:, 2] \\ \vdots \\ \mathbf{W}[:, K]^H \mathbf{A}(\hat{\theta}_l) \mathbf{W}[:, K] \end{pmatrix} \quad (18)$$

and delay vector $\mathbf{p}(\hat{\tau}_i) = [p(T_s - \hat{\tau}_i), p(2T_s - \hat{\tau}_i), \dots, p(NT_s - \hat{\tau}_i)]^T$.

After obtaining the delay $\hat{\tau}_i$ and AOA angle $\hat{\theta}_l$ for each path $l \in [1, 2, \dots, \hat{L}]$, we can apply the least-square method to estimate the complex path gains. Thus, (11) can be rewritten as

$$\text{vec}(\hat{\mathbf{H}}) = [\mathbf{r}(\hat{\theta}_1, \hat{\tau}_1), \mathbf{r}(\hat{\theta}_2, \hat{\tau}_2), \dots, \mathbf{r}(\hat{\theta}_L, \hat{\tau}_L)] [\hat{g}_1, \hat{g}_2, \dots, \hat{g}_L]^T. \quad (19)$$

To this end, the path gain can be computed as $\hat{\mathbf{g}} = (\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H \text{vec}(\hat{\mathbf{H}})$, where $\text{vec}(\cdot)$ is the vectorization function and $\mathbf{R} = [\mathbf{r}(\hat{\theta}_1, \hat{\tau}_1), \mathbf{r}(\hat{\theta}_2, \hat{\tau}_2), \dots, \mathbf{r}(\hat{\theta}_L, \hat{\tau}_L)]$.

The computational complexity depends on the number of sectors, the number of taps of the estimated channel, and the desired resolution for the parameter estimation (i.e., the grid resolution) to compute the MUSIC spectrum. The major computational complexity of estimation algorithms comes from the eigenvalue decomposition and the grid search. The eigenvalue decomposition of the angle and delay estimations are $O(K^3)$ and $O(N^3)$, respectively, where N is the length of the estimated channel, and K is the number of sweep sectors. The grid search for the angle MUSIC and delay MUSIC are $O(K^2|\theta|)$ and $O(N^2|\tau|)$, respectively, where $|\theta|$ and $|\tau|$ are the sizes of the search spaces for angle and delay. Note that $\mathbf{p}(\tau)$ and \mathbf{b} can be precomputed in practice. The computational complexity of the gain estimation is $O(L^2KN + LKN + L^3)$ for matrix multiplications and inversion, where L is the number of paths, and the complexity of gain estimation can be simplified as $O(L^2KN)$ to account for the dominant factor. In addition, the computation cost to pair the delay and angle is relatively small considering the mmWave channels' sparsity, which is $O(LKN)$.

In practice, several techniques can be adopted to reduce the computational complexity. For example, if the target falls into a specific range, we can truncate the estimate CIR accordingly

to reduce N . Furthermore, some MUSIC variations, such as root-MUSIC and ESPRIT [19], could be adopted to bypass the grid search phase. Since this article focuses on the feasibility of applying MUSIC to estimate angle and range/delay in an SLS scenario, these variations will be investigated in our future study.

V. PERFORMANCE EVALUATION

We design several simulations to illustrate and obtain the performance benchmark of the proposed approach. In the following, we introduce the evaluation methodology and then discuss the evaluation results.

A. Methodology

Our simulation assumes the system operates in a full duplex mode. The same ULAs are closely located at the AP for transmitting and receiving, with M elements spacing half wavelength apart. We assume the transmitter and receiver have sufficient isolation, and no direct propagation paths exist between the transmitter and receiver. In order to reduce side-lobe, a Kaiser window with a shape factor 3.4 is used as the beamforming taper.

In our study, we intend to cover directions between $[-45^\circ, 45^\circ]$, the beam is sequentially steered with a fixed value $\Delta\theta$ as the gap of the center of the neighboring beams. In our study, we introduce overlapped beams, as shown in Fig. 6(a), to cover the 90° area to achieve finer angle detection resolution. Note that, in order to cover a larger angle area, multiple phased-array antennas can be used. In this regard, four phased-array antennas can be stacked together to cover 360° area.

To evaluate the sensing performance, the IEEE 802.11ad compliant CPHY waveform, in particular the pilot signal, is generated first and analog beamformed at the transmitter. The IEEE 802.11ad CPHY operates in the SC mode, occupying a bandwidth of 1.76 GHz. The beamformed waveform propagates through the simulated channel and is received at the receiver. A sequence of algorithms, i.e., frame start detection, CIR estimation, and angle and delay estimation, are applied to the received waveform as described in Section IV. Note that both the angle estimation and delay estimation are 1-D grid searches with grid size 0.5° and $1/10$ sampling interval (0.057 ns), respectively.

We first perform a controlled experiment to evaluate how the transmission scheme and resource usage affect detection performance. We randomly generate L paths with their delays uniformly distributed between $[0, 30]$ samples and angles uniformly distributed between $[-45, 45]$ degrees. We run 500 simulations, i.e., 500 times sequential beam steering, to obtain average angle and delay detection performance for each SNR value. Note that SNR is defined as $\text{SNR} = [(E_s M^2)/\sigma^2]$, where E_s is the transmit symbol energy, M is the number of antenna elements of the phased-array antenna, and $\sigma^2 = 2\sigma_w^2$ is the noise energy of the complex Gaussian noise. In the simulations, the multipath channel is normalized to have unity gain. The metrics used for performance evaluation are the root-mean-square error (RMSE) between the

estimated delay, angle, and corresponding ground-truth values. The RMSE for angle and delay is computed separately as $\text{RMSE}_\theta = \sqrt{\mathbb{E}[(\theta_l - \hat{\theta}_l)^2]}$ and $\text{RMSE}_\tau = \sqrt{\mathbb{E}[(\tau_l - \hat{\tau}_l)^2]}$.

We further evaluate the detection performance based on real-world measurement data. At NIST, we carried out a human gesture measurement campaign with our mmWave phased-array channel sounder [20]. Our channel sounder operates at a center frequency of 28.5 GHz. A repeated M-ary pseudorandom-noise (PN) code of 2047 chips with 0.5-ns chip duration is used for sensing. PN codes generally have good ambiguity functions, such as sharp peaks and low side-lobes, and are robust to Doppler shift, which can improve sensing resolution. In addition, the 2047-chip PN code can provide 33-dB processing gain, significantly improving the receiver's SNR and increasing the link budget. The system samples at 16 Gsamples/s. The measurement campaign adopts a monostatic deployment, and the transmitter has a phased-array antenna with 64 elements, while the receiver employs a phased-array antenna with 256 elements to achieve fine angle resolution. In the measurement campaign, the channel was measured at a time interval of 2.6 ms. The human target is located at 2.5 m away from the transmitter and receiver, as shown in Fig. 7, and the target reflection angles are in the range of $[-10^\circ, 11^\circ]$. The MPCs were extracted from the received waveform using the CLEAN algorithm combined with least-square power estimation [21]. The average errors of the complex power, delay, and AoA of resolvable paths measured by the system are 0.48 dBm, 0.18 ns, and 0.47° , respectively [20]. We use the extracted MPCs as the ground truth of the target reflections to model the channel and verify the capability of gesture sensing using the mmWave WLAN signal.

B. Results

1) *Delay and Angle Estimation*: In Fig. 4, a basic $L = 3$ three-path scenario is demonstrated. In this basic configuration, the transmitter and receiver follow a monostatic deployment. The phased-array antennas at both the transmitter and the receiver consist of 16 antenna elements, and both beams sequentially steer at the 17 equally spaced directions. The azimuth angles of the three echoes are set to $[35.2864^\circ, 28.1517^\circ, 17.1803^\circ]$, with delays set to $[2.2532, 2.0707, 8.6339]$ ns, corresponding to $[3.96, 3.64, 15.20]T_s$, where $T_s = 0.5682$ ns is the symbol interval. The magnitudes of the three paths are set to roughly the same level, i.e., $[0.6285, 0.5500, 0.5500]$, with the overall power normalized to one unit and phases are randomly generated with rotation angle following uniform distribution between $[0, 2\pi]$ to focus on the capability of separating spatially closely located paths.

Fig. 4(a) illustrates the magnitude contour of the CIR estimation at the SNR level of 10 dB. In the figure, 17 rows correspond to the estimated CIRs for the 17 steer directions, and the color represents the relative power level. In this plot, we can observe that three paths are closely spaced in the upper left corner. Since path 1 and path 2 are very close in terms of delay, and due to side lobes existing in both delay

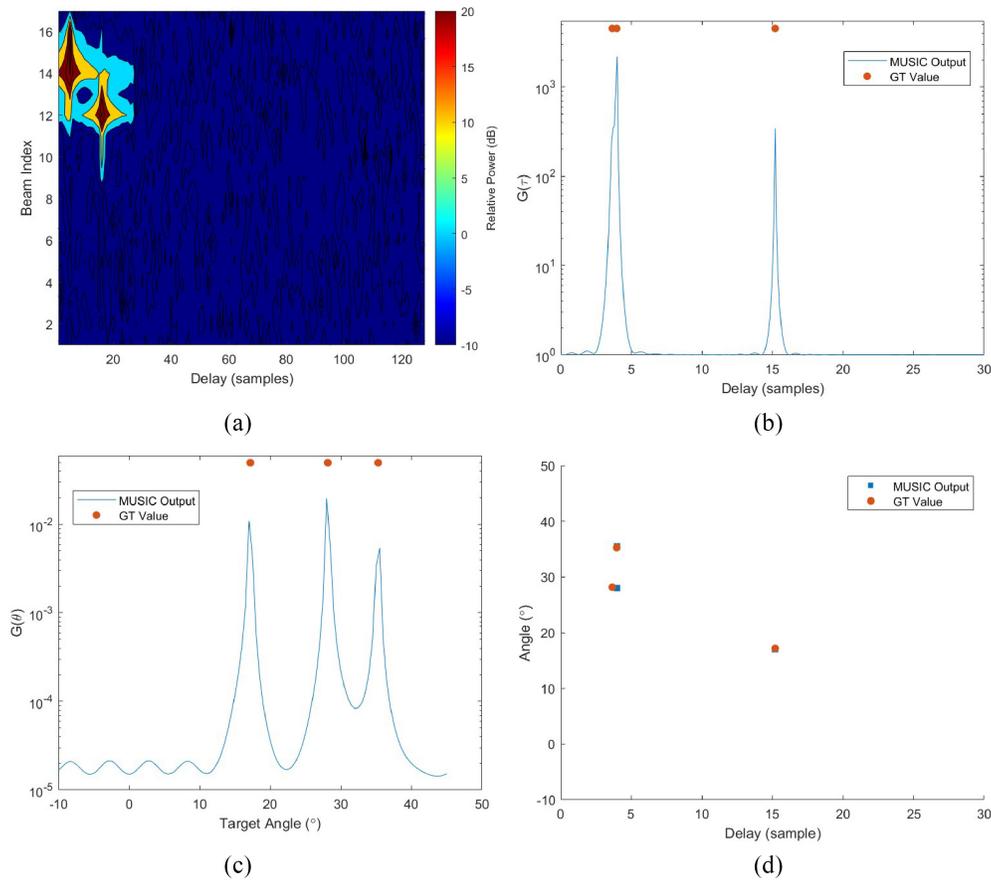


Fig. 4. Detected paths versus ground truth. The simulated three paths have delay at [2.2532, 2.0707, 8.5339] ns and angles at [35.2864, 28.1517, 17.1803] degree, respectively. SNR = 10 dB. (a) Contour of estimated CIRs. (b) Delay MUSIC spectrum. (c) Angle MUSIC spectrum. (d) Detected paths versus ground truth.

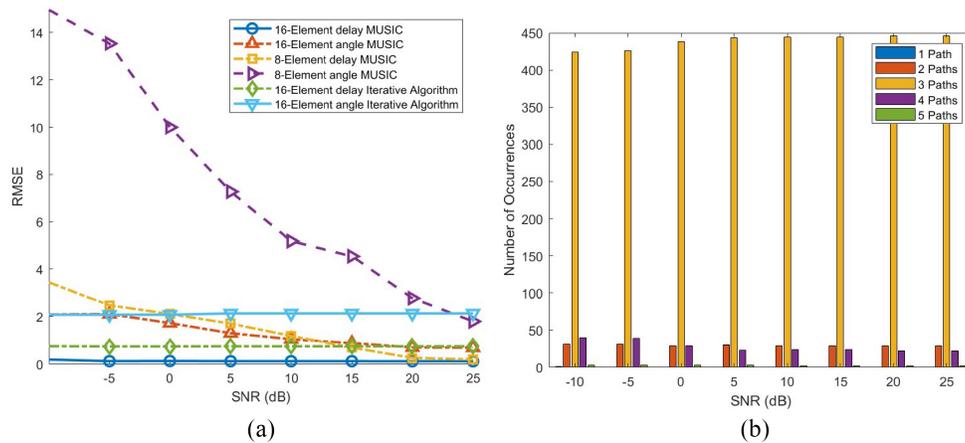


Fig. 5. 3-path detection performance. Three paths are randomly generated paths with their delays uniformly distributed between [0, 30] samples and angles uniformly distributed between [−45, 45] degrees. The units of delay RMSE are samples and the units of angle RMSE are degrees. (a) RMSE of the 3-path estimation. (b) Distribution of the number of detected paths (16-element phased array).

and beam directions, these two paths basically form a connected region. With the MUSIC super-resolution algorithm, in Fig. 4(b), two paths are detected using 1-D delay estimation, while in Fig. 4(c), three paths are detected using 1-D angle estimation. After pairing these two parameters, Fig. 4(d) demonstrates that the three paths are successfully separated in the angle–delay domain and are very close to the ground truth.

Next, we examine how the size of the phased array antenna can affect the delay and angle estimation performance. Fig. 5(a) compares the estimation RMSE versus SNR for ULAs with 8 and 16 elements, respectively. The plot shows the 16-element antenna outperforms the 8-element antenna in terms of detection accuracy since the 16-element antenna can generate narrower beams, resulting in more sweeps to cover

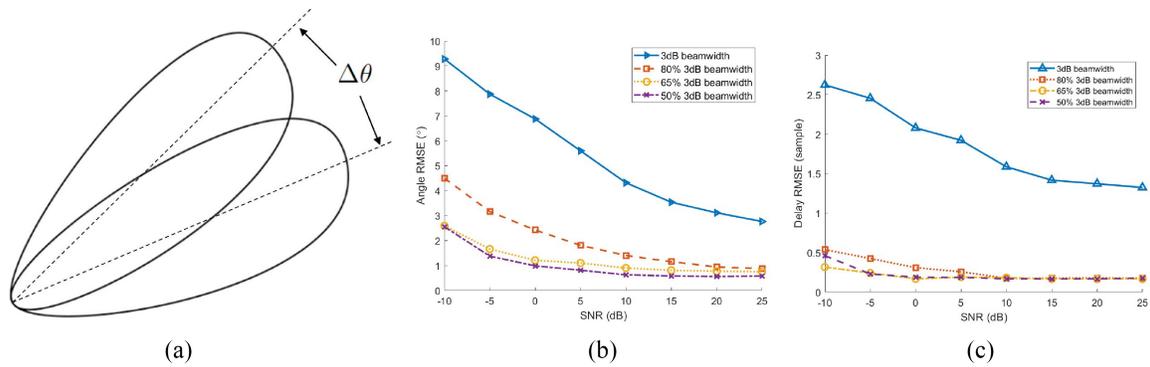


Fig. 6. Impact of beam overlapping on the detection performance with 16-element phased arrays. Four paths are randomly generated with their delays uniformly distributed between $[0, 30]$ samples and angles uniformly distributed between $[-45, 45]$ degrees. (a) Simplified beam overlapping (exclude the sidelobes). (b) Angle RMSE versus beam overlapping. (c) Delay RMSE versus beam overlapping.

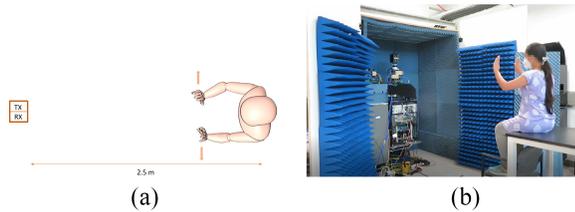


Fig. 7. Gesture measurement campaign. (a) Illustrate. (b) Measurement.

an area, improving the estimation performance. However, to cover a given area, the beam sweeping time will be longer for the 16-element ULA than for the 8-element ULA. In addition, the MUSIC angle RMSE decreases with the increase of the SNR value. In particular, the angle detection performance of the 8-element ULA degrades significantly for the low SNR scenario when the estimated CIRs are noisy. Fig. 5(a) also compares the performance of our approach with an iterative signal decomposition approach. With the iterative signal decomposition approach, the dominant signal is first estimated, and its contribution is then subtracted from the estimated CIRs for all sweep directions. The process is repeated until the stopping criterion is met. We assume the number of MPCs is known for the iterative signal decomposition approach, and the iteration stops when all the MPCs are identified. As seen in the figure, the detection performance with the iterative signal decomposition approach is limited by the system bandwidth and the beam width, demonstrating the benefit of our approach.

Fig. 5(b) illustrates the number of estimated propagation paths detected by our proposed estimation algorithm for 16-element ULA with different SNR conditions. When the echoes are closely located in terms of AoA angle and delay, they may not be reliably separated, and we observe fewer reflections than the ground truth (i.e., 3 paths). In other scenarios, the imperfect beam pattern and transmit pulse, introducing sidelobes in both the angle and delay domain, could cause false detection and lead to some extraneous reflections reported. Underdetection and overdetection become more severe when SNR is low and the beamwidth is wide.

Another factor that can affect angle–delay estimation is the beam overlapping ratio. The impact of the beam overlapping on the detection performance is demonstrated in Fig. 6. We randomly generate $L = 4$ paths with delay uniformly

distributed between $[0, 30]$ samples and angles distributed between $[-45, 45]$ degrees. Similarly, We run 500 simulations to obtain the average performance. Using the $\Delta\theta$ equals 3-dB beamwidth as the baseline, we increase the beam overlapping and reduce the gap of the steering angles of the neighboring beams to a fraction of the 3-dB beamwidth. We observe that in a general trend, the angle and delay detection performance improves when the beam overlapping increases. The figure shows angle RMSE and delay RMSE for four overlapping configurations, i.e., $\Delta\theta$ equals 3-dB beamwidth, 80%, 65%, and 50% of the 3-dB beamwidth, respectively. The figure shows that the estimation performance improves significantly when the beamwidth reduces from 100% to 80% of the 3-dB beamwidth. Also, from 80% to 65% and 50%, the trend of improvement continues; however, the amount of improvement decreases, especially for the high SNR situation. Note that the estimation performance depends on the beam pattern, including the main beam width and the main beam and side beam gains, the number of targets to be detected, and others. Although a smaller gap between steering angles can lead to better detection accuracy, it requires more beams to cover a given area.

2) *Application*: We now study the feasibility of leveraging SLS procedures to perform gesture recognition based on real-world human gesture measurement data. In our measurement campaign, the channel was acquired at a time interval of 2.6 ms, which means that MPCs were captured every 2.6 ms. The captured target-related MPCs (i.e., echoes from the human subject) are used to simulate the target-related signal propagation and reconstruct the channel model using (1). Note that nontarget-related echoes can be removed using clutter removal techniques such as background subtraction or delay-line canceller algorithms, among others [1]; therefore, we focus on target-related MPCs in this study. We first configure the SLS procedure to occur every 2.6 ms. With a 16-element ULA employed at the transmitter and the receiver, the angular spacing between adjacent sectors is set to be $0.65\theta_3$, i.e., 65% of the 3-dB beamwidth. We assume that the channel remains unchanged during each SLS, corresponding to a few tenths of 1 ms. For each beam direction, the minimum package length is 23 168 symbols [16], including 7552 pilot symbols. To cover an area spanning $[-45^\circ, 45^\circ]$ azimuth angle, the SLS accounts for at least 8.6% of the overall throughput. As the

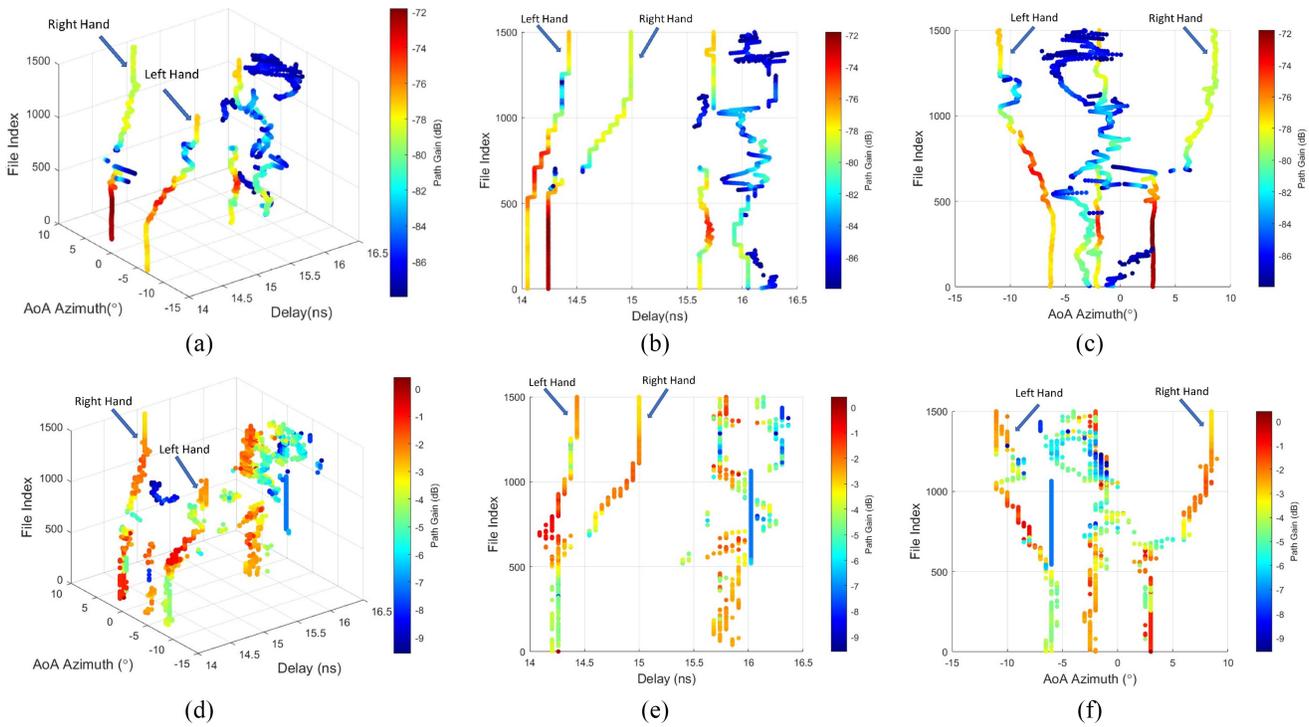


Fig. 8. Estimated MPCs versus ground-truth MPCs for a gesture case. (a)–(c) ground-truth MPCs over time (File Index), and (d)–(f) detected MPCs. SNR = 10 dB. (a) 3D View of the GT Multipath. (b) GT Delay versus Frame Idx. (c) GT Angle versus Frame Idx. (d) 3D View of the Detected Multipath. (e) Estimated Delay versus Frame Idx. (f) Estimated Angle versus Frame Idx.

beam sweeping interval increases, the SLS overhead decreases accordingly. In this section, we will examine the effect of the beam sweeping interval on the sensing results.

With the reconstructed propagation channel, we simulate the IEEE 802.11ad SLS procedure and estimate the target echoes using the approach described in Section IV. Fig. 8 displays the ground truth, i.e., the MPCs obtained through measurement campaigns versus the estimated MPCs, demonstrating the gesture motion seen through the RF signal. The scenario here is a person sitting on a table, facing the transmitter and receiver, with both hands moving outward away from the body in parallel to the ground, as illustrated in Fig. 7. The MPCs are displayed over time through the 1500 files indexed along the z -axis. Fig. 8(a)–(c) shows the ground-truth MPCs over time (File Index), and Fig. 8(d)–(f) shows the estimated MPCs. In particular, Fig. 8(a) and (d) are 3-D views in terms of angle, delay, and time; Fig. 8(b) and (e) displays the 2-D view of delay versus time; Fig. 8(c) and (f) illustrates the 2-D view of angle versus time. Comparing the ground-truth MPCs and the estimated MPCs, we observe that, in general, they closely follow the same pattern, except that some misdetection and false-alarm detection can be observed in the detected MPCs due to the spatial closeness of the reflections from the human body. Nevertheless, we can still clearly observe two trajectories of MPCs with increased azimuth angles over time in Fig. 8(f), representing the moving left and right hands. These patterns can provide unique signatures for gestures, facilitating gesture recognition applications. Note that the gesture pattern cannot be observed over time with the iterative signal decomposition approach, as shown in Fig. 9, demonstrating

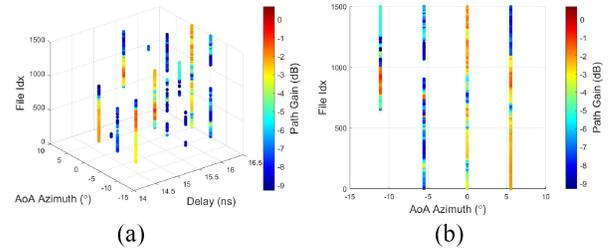


Fig. 9. Estimated MPCs using the iterative signal decomposition method for a gesture case. SNR = 10 dB. (a) 3-D view. (b) Angle versus frame Idx.

the benefits of our approach that apply the MUSIC algorithm to achieve super-resolution performance.

Next, we investigate the feasibility of gesture recognition by increasing the beam sweeping interval. To this end, we employ dynamic time warping (DTW) [22] as a metric to quantify the similarity between the detected MPCs and the ground truth. We focus our analysis on the angle–time domain since reflections from both hands may have very close delay values and delay patterns due to the symmetry of the deployment. As DTW can help compare the similarity between time series of different lengths, it can be used to assess whether motion patterns can be maintained with increased sensing intervals (i.e., beamforming training intervals in this study).

In order to extract the MPCs related to gesture motion, we first perform the clustering algorithm to identify the MPCs corresponding to the left and right hands. We leverage the density-based spatial clustering of applications with noise (DBSCAN)-based clustering algorithm [23] to filter out

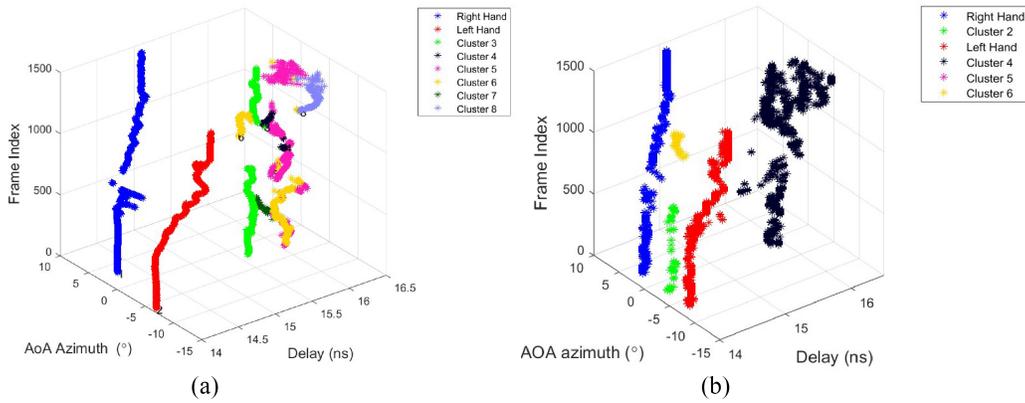


Fig. 10. MPCs clustering. (a) Clustering of the ground-truth MPCs. (b) Clustering of the detected MPCs.

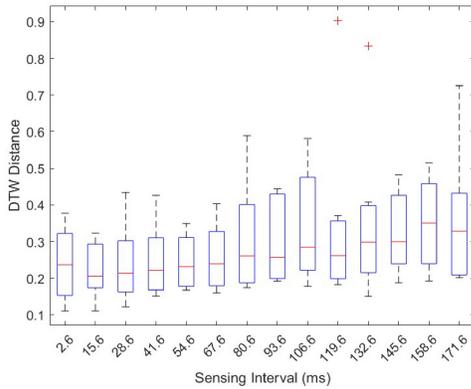


Fig. 11. DTW distance versus sensing interval.

the noise and cluster the MPCs in both spatial and time domains [24]. The two clusters with the smallest delay values are composed of echoes from the two hands, as shown in Fig. 10.

We then compute the cluster trajectory by averaging the angle and delay values of the MPCs in a cluster per time frame. After obtaining the hand trajectories from the estimated MPCs and ground-truth ones, we calculate the DTWs for the left and right hands separately and obtain DTW_L and DTW_R . To remove the effect of the sequence length on the DTW value, we define the normalized DTW as

$$DTW_n = \frac{DTW_L}{L_L} + \frac{DTW_R}{L_R} \quad (20)$$

where L_L and L_R are the stretched sequence lengths for left hand and right hand, respectively, after the DTW matching. This analysis considers two sets of opposite gestures: two hands moving outwards away from the human body and two hands moving inward toward the human body, i.e., the “open” and “close” hand gestures. We have four human subjects repeating each gesture motion. Fig. 11 shows the DTW statistics between the detected and the ground-truth trajectories when the sensing interval increases from 2.6 to 171.6 ms.

From Fig. 11, We discover that as the sensing interval increases, the DTW distance generally increases, with a fairly flat slope, and the average value is between 0.2 and 0.4. Note that the DTW distance between “open” and “close” hand trajectories is around 1.5. Thus, for the evaluated hand gestures,

the trajectory similarity stays consistent with the increased sensing interval, which is expected because the motion demonstrated here is only in one direction with a simple trend. It indicates that the sensing frequency can be as slow as 100 ms for the simple, straightforward gesture movement, which will dramatically reduce the sensing overhead compared to the initial sensing interval of 2.6 ms and have minimum impact on the existing Wi-Fi communication [25].

VI. RELATED WORK

In this section, we introduce the literature review closely related to our study.

Some existing research efforts use Wi-Fi signals for smart IoT applications, such as carrying out hand gesture recognition [26], [27] and human pose and seat occupancy classification [28], operating at mmWave band [26], [28] or sub-6-GHz frequency band [27]. In these efforts, machine-learning-based classifiers are incorporated to categorize human activities, and the data source for the classifiers include range–Doppler images (RDIs) [26], beam SNRs [28], and frequency–time Doppler profile [27]. Unlike these existing works, we consider the deployment scenario with large range and low Doppler shift. Our goal is to track the spatial location of the hands over time to identify gesture patterns with high resolution.

In the context of multipath parameters estimation using the Wi-Fi signal, the MUSIC algorithm has been adopted to extract Doppler frequency shift information [29] for device-free indoor human tracking using the Wi-Fi CSI samples collected using commercial Wi-Fi cards. Similarly, Liu et al. [30] extracted the range and velocity estimation of the vehicles in the intelligent transportation system. Their work proposed an auto-paired super-resolution algorithm to handle the orthogonal frequency-division multiplexing (OFDM) integrated radar and communications waveform. The performance of the super-resolution algorithms has inspired this work; however, this work concentrates on how to incorporate super-resolution in an estimation framework, which focuses on delay and angle estimation of the multipath by combining a set of CIRs collected via transmitting a sequence of switched analog beams. The designed algorithm can be deployed at an edge device to provide intelligent service for Wi-Fi sensing applications. Due

to the nature of the analog beamforming and no information available at the element level, the angle MUSIC algorithm is applied in the beam space [15].

Recently, some research has been carried out on opportunistic radar sensing, i.e., leveraging the existing wireless signal or wireless protocols to detect and track moving objects for area surveillance or activity monitoring, which are important to numerous IoT applications. The opportunistic radar sensing research involves advanced signal processing algorithms [16], [31] to improve detection accuracy and sensing with limited and irregular communication traffic data. In particular, IEEE 802.11ad SLS has been used to detect the presence of the objects and their position, radial velocity, and echo power strength [31] by assuming one target at each SLS direction. In that work, a generalized-likelihood ratio test (GLRT) and maximum-likelihood estimators were incorporated to estimate the range (delay), velocity (Doppler), and signal amplitude of a detected object. Unlike the existing study, we investigate the scenarios with multiple targets in the covered area by collectively analyzing the received waveforms from the sequentially switched beams.

VII. FINAL REMARKS

In this study, we investigated the feasibility of using the IEEE 802.11ad SLS procedure that provides opportunistic indoor radar sensing. In particular, we designed a framework to effectively estimate the target's spatial position through delay and angle. The switched beams in the IEEE 802.11ad beamforming training were used to estimate the CIR for each beamforming direction. Considering one round of the sequentially switched beams as one observation and combining all the CIRs from each direction, the target delay and angle could be estimated by engaging the MUSIC super-resolution algorithms. To validate the efficacy of our designed approach, we conducted an extensive performance study to understand the performance sensitivity with respect to parameters (e.g., beamwidth, the overlapping of neighboring beams, and SNR). Furthermore, using a gesture recognition application as an example, we conducted simulations to reconstruct the target propagation channel and discussed the feasibility of monitoring the gesture behavior based on existing communication procedures and protocols. In the future, we could extend our approach from the two perspectives. One direction is investigating other super-resolution algorithms, such as space-alternating generalized expectation-maximization (SAGE) and ESPRIT in WiFi sensing, and studying how radio resource availability can affect the estimation performance. The other direction is extending the CSI-based parameter estimation to include Doppler frequency or speed estimation to support more IoT applications.

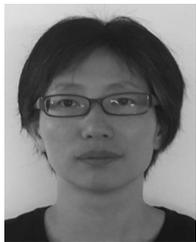
ACKNOWLEDGMENT

The authors thank their colleagues at NIST for allowing them to reuse the human gesture measurement data collected using NIST mmWave phased-array channel sounder. In particular, they would like to thank Chiehping Lai, Neha Pazare, Jelena Senic, and Neeraj Varshney.

REFERENCES

- [1] J. Wang, N. Varshney, C. Gentile, S. Blandino, J. Chuang, and N. Golmie, "Integrated sensing and communication: Enabling techniques, applications, tools and data sets, Standardization, and future directions," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23416–23440, Dec. 2022.
- [2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [3] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on Industrial Internet of Things: A cyber-physical systems perspective," *IEEE Access*, vol. 6, pp. 78238–78259, 2018.
- [4] F. Liang, W. Yu, X. Liu, D. Griffith, and N. Golmie, "Toward edge-based deep learning in Industrial Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4329–4341, May 2020.
- [5] Z. Yang, R. Wang, D. Wu, B. Yang, and P. Zhang, "Blockchain-enabled trust management model for the Internet of Vehicles," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 12044–12054, Jul. 2023.
- [6] J. A. Stankovic, "Research directions for the Internet of Things," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 3–9, Feb. 2014.
- [7] J. Yang, H. Zou, H. Jiang, and L. Xie, "Device-free occupant activity sensing using WiFi-enabled IoT devices for smart homes," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3991–4002, Oct. 2018.
- [8] S. Wu et al., "Survey on prediction algorithms in smart homes," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 636–644, Jun. 2017.
- [9] R. C. Daniels, E. R. Yeh, and R. W. Heath, "Forward collision vehicular radar with IEEE 802.11: Feasibility demonstration through measurements," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1404–1416, Feb. 2018.
- [10] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1629–1645, 3rd Quart., 2020.
- [11] Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: A survey," *ACM Comput. Surveys*, vol. 52, no. 3, p. 46, Jun. 2019. [Online]. Available: <https://doi.org/10.1145/3310194>
- [12] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, "Estimating crowd density in an RF-based dynamic environment," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3837–3845, Oct. 2013.
- [13] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11ad: Directional 60 GHz communication for multi-gigabit-per-second Wi-Fi [invited paper]," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 132–141, Dec. 2014, doi: [10.1109/MCOM.2014.6979964](https://doi.org/10.1109/MCOM.2014.6979964).
- [14] *IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band*, IEEE Standard 802.11ad-2012 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012 IEEE Std 802.11aa-2012), 2012.
- [15] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5384–5394, Aug. 2017.
- [16] E. Grossi, M. Lops, and L. Venturino, "Adaptive detection and localization exploiting the IEEE 802.11ad standard," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4394–4407, Jul. 2020.
- [17] L. Liu, H. Ju, X. Fang, Y. Long, and R. He, "Systematic design of radar detection under IEEE 802.11ad framework," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, 2021, pp. 1–5.
- [18] M. A. Richards, *Fundamentals of Radar Signal Processing*. New York, NY, USA: McGraw-Hill Prof., 2005. [Online]. Available: <https://mhebooklibrary.com/doi/book/10.1036/0071444742>
- [19] *Parameter Estimation II*. Hoboken, NJ, USA: Wiley, 2002, ch. 9, pp. 1139–1317. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471221104.ch9>
- [20] S. Y. Jun, J. Chuang, D. Caudill, C. Gentile, S. Blandino, and N. Golmie, "NIST mmWave phased-array channel sounder for human sensing," NIST, Gaithersburg, MD, USA, document IEEE 802.11-21/1675r0, 2021. [Online]. Available: <https://mentor.ieee.org/802.11/dcn/21/11-21-1675-00-00bf-mmwave-phased-array-channel-sounder-for-human-sensing.pptx>
- [21] J. A. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," *Astron. Astrophys. Supplement Series*, vol. 15, p. 417, Jun. 1974.

- [22] E. J. Keogh and M. J. Pazzani, "Derivative dynamic time warping," in *Proc. SIAM Int. Conf. Data Min.*, 2001, pp. 1–11. [Online]. Available: <https://epubs.siam.org/doi/abs/10.1137/1.9781611972719.1>
- [23] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Disc. Data Min.*, 1996, pp. 226–231.
- [24] N. Varshney, J. Wang, C. Lai, C. Gentile, R. Charbonnier, and Y. Corre, "Quasi-deterministic channel propagation model for an urban environment at 28 GHz," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, no. 7, pp. 1145–1149, Jul. 2021, doi: [10.1109/LAWP.2021.3073880](https://doi.org/10.1109/LAWP.2021.3073880).
- [25] J. Wang et al., "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.
- [26] Y. Ren et al., "Hand gesture recognition using 802.11ad mmWave sensor in the mobile device," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, 2021, pp. 1–6.
- [27] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38. [Online]. Available: <https://doi.org/10.1145/2500423.2500436>
- [28] J. Yu, P. Wang, T. Koike-Akino, Y. Wang, P. V. Orlik, and H. Sun, "Human pose and seat occupancy classification with commercial mmWave WiFi," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2020, pp. 1–6.
- [29] X. Li et al., "IndoTrack: Device-free indoor human tracking with commodity Wi-Fi," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 3, p. 72, 2017. [Online]. Available: <https://doi.org/10.1145/3130940>
- [30] Y. Liu, G. Liao, Y. Chen, J. Xu, and Y. Yin, "Super-resolution range and velocity estimations with OFDM integrated radar and communications waveform," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11659–11672, Oct. 2020.
- [31] E. Grossi, M. Lops, L. Venturino, and A. Zappone, "Opportunistic radar in IEEE 802.11ad networks," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2441–2454, May 2018.



Jian Wang received the B.S. degree in electrical engineering from Tongji University, Shanghai, China, and the M.S. degree in electrical engineering from Washington State University, Pullman, WA, USA, in 2000.

She has been an Electronics Engineer with the Communication Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA, since 2015. From 2000 to 2015, she worked in the wireless industry, including Nokia, Irving, TX, USA; Texas Instruments, Dallas, TX,

USA; and Digital Receiver Technology Inc., Germantown, MD, USA, on digital signal processing and wireless protocol research and development. Her research interests include next-generation wireless systems, channel modeling, and machine learning.



Jack Chuang received the Ph.D. degree from The Pennsylvania State University, State College, PA, USA, in 2008.

He was a Graduate Research Assistant with the Communications and Space Sciences Laboratory, The Pennsylvania State University. He then worked with BAE Systems, Merrimack, NH, USA, in electronic warfare and Cisco Systems, Richfield, OH, USA, in spectrum sharing. He is currently with the Communication Technology Laboratory, National Institute of Standards and Technology, Gaithersburg,

MD, USA, developing 5G mmWave channel sounders.



Samuel Berweger received the Ph.D. degree from the University of Washington, Seattle, WA, USA, in 2011.

He joined the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, in 2013. His past research interests were focused on scanning probe-based subdiffraction limit microscopy at optical and microwave frequencies. In 2022, he joined the Communication Technology Laboratory, NIST, to pursue interests in channel sounding and advanced free-space RF propagation measurements.



Camillo Gentile received the Ph.D. degree in electrical engineering from The Pennsylvania State University, State College, PA, USA, in 2001.

He joined the National Institute of Standards and Technology, Gaithersburg, MD, USA, in 2001, where he is currently leading the Radio Access and Propagation Metrology Group and the NextG Channel Measurement and Modeling Project in the Communications Technology Laboratory. He has authored over 90 peer-reviewed journal and conference papers, and a book on geolocation techniques and a book on millimeter-wave and subterahertz channel propagation modeling. His current interests include channel modeling and physical-layer modeling for 5G and 6G wireless systems.

techniques and a book on millimeter-wave and subterahertz channel propagation modeling. His current interests include channel modeling and physical-layer modeling for 5G and 6G wireless systems.



Nada Golmie received the Ph.D. degree in computer science from the University of Maryland at College Park, College Park, MD, USA.

Since 1993, she has been a Research Engineer with the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA. From 2014 until 2022, she served as the Chief for the Wireless Networks Division, NIST. She is an NIST Fellow with the Communications Technology Laboratory. Her research in media access control and protocols for wireless networks led to over 200 technical papers presented at professional conferences, journals, and contributed to international standard organizations and industry led consortia. She is the author of *Coexistence in Wireless Networks: Challenges and System-Level Solutions in the Unlicensed Bands* (Cambridge, U.K.: Cambridge University Press, 2006). She leads several projects related to the modeling and evaluation of future generation wireless systems and protocols and serves as the NextG Channel Model Alliance Chair.

technical papers presented at professional conferences, journals, and contributed to international standard organizations and industry led consortia. She is the author of *Coexistence in Wireless Networks: Challenges and System-Level Solutions in the Unlicensed Bands* (Cambridge, U.K.: Cambridge University Press, 2006). She leads several projects related to the modeling and evaluation of future generation wireless systems and protocols and serves as the NextG Channel Model Alliance Chair.