JUMP: Joint Communication and Sensing With Unsynchronized Transceivers Made Practical

Jacopo Pegoraro, *Member, IEEE*, Jesus O. Lacruz, Tommy Azzino, Marco Mezzavilla, *Senior Member, IEEE*, Michele Rossi, *Senior Member, IEEE*, Joerg Widmer, *Fellow, IEEE*, and Sundeep Rangan, *Fellow, IEEE*

***Abstract*— Wideband millimeter-wave communication systems can be extended to provide radar-like sensing capabilities on top of data communication, in a cost-effective manner. However, the development of *joint communication and sensing* technology is hindered by practical challenges, such as occlusions to the line-of-sight path and clock asynchrony between devices. The latter introduces *time-varying* timing and frequency offsets that prevent the estimation of sensing parameters and, in turn, the use of standard signal processing solutions. Existing approaches cannot be applied to commonly used phased-array receivers, as they build on stringent assumptions about the multipath environment, and are computationally complex. We present JUMP, the first system enabling *practical* bistatic and asynchronous joint communication and sensing, while achieving accurate target tracking and micro-Doppler extraction in realistic conditions. Our system compensates for the timing offset by exploiting the channel correlation across subsequent packets. Further, it tracks multipath reflections and eliminates frequency offsets by observing the phase of a dynamically-selected static reference path. JUMP has been implemented on a 60 GHz experimental platform, performing extensive evaluations of human motion sensing, including non-line-of-sight scenarios. In our results, JUMP attains comparable tracking performance to a full-duplex monostatic system and similar micro-Doppler quality with respect to a phase-locked bistatic receiver.**

Manuscript received 1 April 2023; revised 15 November 2023; accepted 9 February 2024. Date of publication 22 February 2024; date of current version 14 August 2024. This work was supported in part by the European Union through the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" under Grant PE0000001 - program "RESTART"; in part by Projects TSI-063000-2021-59 RISC-6G and TSI-063000-2021-63 MAP-6G funded by the Spanish Ministry of Economic Affairs and Digital Transformation through the European Union NextGeneration-EU Plan; and in part by MCIN/AEI/10.13039/501100011033/FEDER, UE, under Grant PID2022-136769NB-I00 ELSA6G. The work of Tommy Azzino, Marco Mezzavilla, and Sundeep Rangan was supported in part by NSF under Grant 1302336, Grant 1564142, Grant 1547332, Grant 1824434, Grant 2133662, and Grant 2148293; in part by SRC; and in part by the Industrial Affiliates of NYU WIRELESS. The associate editor coordinating the review of this article and approving it for publication was K. Adachi. *(Corresponding author: Jacopo Pegoraro.)*

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

Jacopo Pegoraro is with the Department of Information Engineering, University of Padova, 35122 Padua, Italy (e-mail: jacopo.pegoraro@unipd.it).

Jesus O. Lacruz and Joerg Widmer are with the IMDEA Networks Institute, Leganes, 28918 Madrid, Spain.

Tommy Azzino, Marco Mezzavilla, and Sundeep Rangan are with NYU Wireless Research Center, New York University (NYU), Brooklyn, NY 11201 USA.

Michele Rossi is with the Department of Information Engineering, University of Padova, 35131 Padua, Italy, and also with the Department of Mathematics "Tullio Levi-Civita," University of Padova, 35131 Padua, Italy.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2024.3365853.

Digital Object Identifier 10.1109/TWC.2024.3365853

***Index Terms*— Joint communication and sensing, wireless sensing, clock asynchrony, IEEE 802.11ay, human sensing, micro-Doppler.**

I. Introduction

JOINT Communication & Sensing (JCS) has emerged as a potential game changer for next-generation wireless networks, endowing communication systems with radar-like capabilities to perceive their surroundings [1], [2]. Joint Communication & Sensing (JCS) system designs are categorized in radar-centric and communication-centric, depending on which of the two functionalities is regarded as primary [3], [4]. Communication-centric JCS is the most promising approach for a cost effective solution, as it leverages the ubiquitous communication hardware and waveforms for sensing, adding minimal overhead and modifications to existing devices and protocols [5], [6]. To this end, communication-centric JCS should maintain the typical wireless network configuration of *separated* transmitter and receiver, with half-duplex capabilities (bistatic configuration) [7], as opposed to impractical full-duplex designs [8]. In this setting, sensing is performed by repurposing the Channel Impulse Response (CIR) or Channel State Information (CSI) estimation processes to extract information about the environment [9]. This allows estimating physical parameters such as the distance and velocity of nearby targets of interest, as these are related to signal reflection delays and Doppler shifts [10], [11]. However, a large-scale adoption of bistatic JCS is impaired by the fact that network nodes are *asynchronous*, i.e., they have different clock sources and oscillators for the Radio Frequency (RF) front-end [5]. This asynchrony causes a time-varying Timing Offset (TO) and Carrier Frequency Offset (CFO). The former appears as a common delay shift for all propagation paths in the CIR, preventing the correct estimation of actual path delays. The latter is a random frequency shift that destroys phase coherence across subsequent packets, hindering the estimation of the Doppler shift and of the micro-Doppler (μD) effect. The μD is a frequency modulation of the reflected signal around the main Doppler frequency induced by movements of a target or target components. It is the main signal feature used in fine-grained wireless movement sensing, and has a vast number of applications in target classification, human activity recognition, pervasive healthcare, and person identification, among others [12], [13], [14]. While existing communication systems use algorithms to compensate for TO and CFO, they treat the sensing parameters (delay and Doppler shift) as part of the undesired offsets and *remove* them [15]. This makes

such existing techniques unfit to our sensing purpose, as delay and Doppler are the channel parameters that we use to describe the physical environment. For these reasons, clock asynchrony is the main obstacle to large-scale practical implementations of JCS systems [5], [11], where TO and CFO due to clock asynchrony are to be removed while *retaining* the delay and the Doppler shift.

In this work, we design and implement JUMP, a JCS system that solves the problem of asynchrony and enables bistatic sensing in wideband communication systems, e.g., Millimeter-Wave (mmWave) networks, which are particularly problematic due to the high phase noise [16]. The system builds on two main insights. *First*, JUMP exploits the slow change of the CIR delay profile compared to the packet rate, and uses a fast correlation-based method to estimate the relative TO across subsequent packets. By compensating for the TO, it allows for target localization and tracking without ambiguity, even under frequent occlusions of the Line-of-Sight (LOS).

Second, it leverages the fact that the CFO is *constant* across the different signal propagation paths for the same packet. This property, paired with the accurate multipath resolution of wideband signals, allows identifying *static* reference multipath components, whose frequency shift is only due to CFO (being static, they do not contain the μD component). Hence, the signal collected from these static paths is used to remove the CFO from the sensing paths, which contain the μD from the targets of interest. This allows aggregating phase-coherent CIR estimates from subsequent packets, enabling μD estimation despite the clock asynchrony.

Existing bistatic sensing approaches use cross-antenna-based TO and CFO compensation, using the signal at one antenna of a Multiple Input Multiple Output (MIMO) receiver as a reference [17], [18], [19], and [20]. JUMP solves several practical problems associated with these methods: *(i)* it works with commonly used phased-array receivers, which do not allow access to the signal at each antenna; *(ii)* it preserves the linearity of the signal and does not introduce cross-terms that require complex estimation algorithms; *(iii)* it correctly operates in weak LOS and Non-Line-of-Sight (NLOS) conditions. Moreover, JUMP estimates the full μD spectrum of the targets, including the contributions of their different moving parts, not just their main Doppler frequency.

Unlike other JCS approaches, it is a *practical* solution that neither requires full-duplex hardware capabilities, as in [6], [8], [12], and [21], nor any synchronization between the nodes, nor any modification to the underlying communication standard. We stress that these qualities make JUMP appealing in more general settings than just JCS, e.g., in bistatic radar systems [5], [22], in case synchronization through fiber links or Global Positioning Systems (GPS) signals is not viable.

We provide analytical insights into JUMP's TO and CFO removal performance, complemented by numerical simulations. Moreover, we prototype it on a 60 GHz IEEE 802.11ay-based Software Defined Radio (SDR) platform. This implements independent transmitter and receiver pairs, thus serving as a realistic testbed for asynchronous JCS. In the implementation, we solve additional practical issues, such as symbol-level synchronization, which is usually neglected in JCS research [5]. To benchmark the system's target tracking and μD extraction capabilities, we conduct a vast experimental campaign in two indoor environments, performing people tracking and μD signature extraction. In the dataset collection, we augment the testbed with additional receivers in monostatic and bistatic phase-locked configurations, enabling a comparison with alternative JCS approaches from the literature.

The contributions of this work can be summarized as:

1. To the best of our knowledge, JUMP is the first system that enables practical JCS for asynchronous transceivers in realistic bistatic configurations, performing accurate target tracking and μD extraction.
2. We design a correlation-based algorithm to compensate for the relative TO across subsequent packets, attaining consistent bistatic target tracking even in challenging multitarget and NLOS scenarios.
3. We leverage the high multipath resolution of wideband transceivers to remove the CFO using a reference propagation path instead of a reference antenna, as done in the JCS literature. This largely simplifies the estimation of the Doppler spectrum, removing limiting assumptions on the multipath environment, or the necessity of a MIMO antenna array.
4. We build two JUMP prototypes and collect a large dataset of IEEE 802.11ay CIR measurements for human motion tracking and μD extraction.

Our work shows promising results for bistatic asynchronous JCS systems compared to monostatic full duplex and bistatic phase-locked ones.

The manuscript is organized as follows. In Section II we discuss the related work, while Section III introduces the system model. JUMP is presented in Section IV, along with a detailed explanation of each signal processing block. Section V contains an analysis of the TO and CFO compensation errors, complemented by numerical simulations. In Section VI we describe the implementation of JUMP on an experimental prototype. Section VII contains experimental results demonstrating JUMP's promising performance. Finally, in Section VIII we underline key aspects of our design and future research directions, while concluding remarks are given in Section IX.

II. Related Work

Full-duplex JCS. The use of full-duplex technology for JCS has been advocated in several works. In [8] and [21] the authors have investigated antenna arrays, beamforming, and waveform design to enable full-duplex JCS in mmWave 5G systems. In [6] and [12], full-duplex has been used as a practical solution to CFO for IEEE 802.11ay sensing, presenting algorithms to detect multiple targets and their μD signatures in indoor scenarios. However, full-duplex entails strong self-interference, which remains a challenging and unsolved problem in real communication systems, as it requires self-interference cancellation techniques [8], which are not yet a mature technology. As an alternative, the transmitted power can be reduced to avoid saturating the receiver, but this is only suitable for short-range indoor use [12].

Bistatic, asynchronous JCS. Some methods are available to provide accurate synchronization between communication network nodes [5], but require the use of GPS signals

or the cooperation of devices. In systems including a single transmitter-receiver pair or non-cooperating nodes, two main methods have been proposed: Cross-Antenna Cross-Correlation (CACC) [20] and Cross-Antenna Signal Ratio (CASR) [17], [18], [19]. Both use the signal collected at one of the receiver array's antennas as a reference to remove the TO and the CFO. These approaches have three main drawbacks that are solved by JUMP. First, they require a MIMO antenna array at the receiver, which is not available in commercial mmWave systems, as *phased-arrays* are preferred for their lower implementation cost and complexity. Second, they entail a higher complexity in the estimation of the sensing parameters, by doubling the number of parameters or introducing non-linearity [19], [20]. Third, for their correct operation, they require strong assumptions about the multipath environment, such as the continuous presence of a dominant LOS link between the transmitter and the receiver, or the presence of a single sensing target [19] in the monitored space. Contrarily, JUMP has no requirements for the receiver array and handles environments with multiple moving targets and LOS occlusions. Other techniques have been proposed that rely on partially overlapping subbands to remove TO and CFO, e.g., [23] and [24]. These require the availability of such subbands, which is rarely the case, and they can not work with Single Carrier (SC) systems.

Recently, a Kalman filter-based technique for removing the TO in bistatic asynchronous JCS has been proposed [25]. Differently from JUMP, however, [25] estimates the *joint Doppler plus CFO* frequency shift, which does not contain useful information about the target's movement. JUMP can instead remove the CFO and retain the Doppler shifts induced by the multiple moving parts of the target, enabling fine-grained sensing applications.

Bistatic sub-6 GHz Wi-Fi sensing. Several studies have proposed ways to perform activity recognition [14], [26], [27] using Wi-Fi Orthogonal Frequency Division Multiplexing (OFDM) CSI. Indeed, most of the CFO removal techniques described in the previous section have been originally proposed for Wi-Fi. The main drawback of sub-7 GHz sensing lies in its inherent low-ranging resolution, which is a direct consequence of the relatively low transmission bandwidth (40-80 MHz are typical values). There, reflected paths can only be resolved with an accuracy of a few meters, which causes major performance degradation with multiple concurrently moving subjects [28].

mmWave bistatic JCS. A few works have addressed bistatic JCS in mmWave systems, assuming synchronized transmitter and receiver [29], [30], [31]. Others have focused on tracking targets based on distance and Angle of Arrival (AoA), without addressing sensing tasks that require phase analysis or aggregation of signal samples over coherent processing intervals, such as Doppler speed estimation [32]. JUMP instead performs μD estimation in addition to tracking, and removes the unrealistic assumption of having time-synchronized nodes.

Radar systems. Bistatic radar systems have been widely studied and adopted [22], [33], [34]. In the so-called *active* bistatic radar setup, the TO and CFO are eliminated by connecting the transmitter and the receiver through high-speed fiber links, achieving phase locking, or exploiting GPS as the common clock source. In the *passive* radar case, the receiver exploits signals of opportunity from non-cooperative transmitters. Two receiving channels are used, one receiving the direct LOS signal and the other pointing towards the target of interest. Then, offsets are removed by cross-correlating the LOS signal with the reflected one [35]. Besides requiring complex implementation and being heavily dependent on the type of transmitted signal, these systems cannot handle occlusions of the LOS link [33].

III. Bistatic System Model

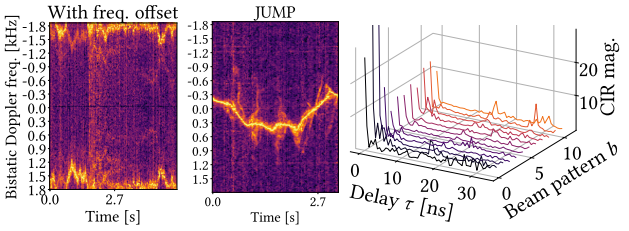
In this section, we introduce the bistatic system model.

A. CIR for Asynchronous Transceivers

The CIR estimation process consists of transmitting a pilot signal from a transmitter (TX) device that, after being reflected by nearby objects or humans, is collected at the receiver (RX). The latter correlates the known pilot waveform with the received one. The resulting CIR contains TO and CFO that prevent accurate estimation of the targets' parameters [5], [20]. In our model, we highlight *(i)* the role of TO and CFO, and *(ii)* the impact of discretizing the CIR for wideband systems. We use a SC CIR model to simplify the understanding of the implementation and experimental results, which are based on SC IEEE 802.11ay. However, JUMP is equally applicable to OFDM-based JCS systems. The model is based on, e.g., [12] and [36], but contains significant modifications to take into account TO, CFO, and the bistatic geometry.

1) Timing Offset: In wireless communications the TX and the RX clocks are not synchronized and thus exhibit an unstable relative clock drift over time [5], [11]. In the absence of an absolute timing reference, the RX performs packet detection and *coarse* synchronization. For example, in IEEE 802.11ad/ay packet detection is done with an auto-correlation method which returns a rough estimation of the start of the packet [37]. This operation does not yield a synchronization at the symbol-level, thus a residual TO is included into the channel estimate and removed via equalization. Note that these operations are performed separately on each packet, and the TO changes across subsequent packets depending on the selected synchronization point. In JCS we are interested in computing the exact delays of the signal propagation paths, without the residual synchronization error, as these are linked to the distance of the sensing targets. Hence, including the TO in the channel estimates results in a wrong estimation of the delays. Due to the time-varying nature of the TO, the cumulative estimation error across time can add up to tens of nanoseconds after just a few milliseconds [5]. This uncertainty in the delay easily leads to large errors in estimating the target's position (several meters), even within short time intervals [11].

2) Frequency Offset: Clock drifts can be typically ignored over the course of the preamble of a single communication packet, as their variation is negligible over short time periods (micro-second level). Therefore, in communication systems such as the IEEE 802.11ad/ay the CFO is estimated and compensated for each packet separately. This is done by computing the total *phase* error induced by the CFO



(a) μD with CFO. (b) JUMP μD. (c) CIR magnitude.

Fig. 1. μD spectrogram affected by the CFO (a) compared to JUMP's μD (b). (c) shows the CIR magnitude for different BPs.

using an autocorrelation technique [37] and removing it from the received samples. In JCS instead, we are interested in retrieving the Doppler shift caused by the target's movement speed. Therefore, compensating for the CFO for each packet separately is not a viable option, as this method removes the CFO *and* the Doppler shift, as they both appear in the phase of the received signal and they can not be easily separated. This is further discussed in Section IV-C. In addition, estimating the phase differences and Doppler shifts caused by moving targets requires processing the received samples *coherently across subsequent packets*. The CFO can not be considered constant within such extended time intervals (several milliseconds), and it makes the phase of signal samples across different packets *incoherent*, preventing the estimation of Doppler shifts.

CFO is especially severe in mmWave systems. As an example, it can reach hundreds of kHz with 60 GHz devices, while for a target moving at 10 m/s the Doppler shift is 4 kHz. This is shown in Fig. 1a, where we plot the CFO-affected normalized Doppler spectrum and that obtained by JUMP from a walking person. Without removing the CFO, the μD is completely corrupted and does not carry any useful information about the movement.

3) Transmitted Signal Model: We denote by $s[i]$ the discrete TX symbol sequence, by L the number of TX symbols, and by $p_{\text{tx}}(t)$ the pulse-shaping filter. Assuming unit energy per TX symbol and a sampling period equal to $\Delta\tau = 1/B$, with B the TX bandwidth, the SC TX signal is

$$x(t) = \sum_{i=0}^{L-1} s[i] p_{\text{tx}}\left(t - i\Delta\tau\right). \tag{1}$$

In a directional JCS system employing phased antenna arrays, different Beam Patterns (BPs) are used to steer the signal energy towards the intended direction. The TX signal is thus multiplied by the analog beamforming vector, $\mathbf{u}$, obtaining $\mathbf{x}(t) = \mathbf{u}x(t)$. In our simulations and implementation, $s[i]$ is a set of complementary Golay sequences used for channel estimation in IEEE 802.11ay, as detailed in Section VI. However, here we consider it to be a single sequence of pilot symbols for the sake of generality.

4) CIR Model: We denote by $N(t)$ the number of propagation paths at time t, by $\alpha_m(t)$ the complex signal attenuation coefficient of the m-th reflection, while and $\tau_m(t)$ is its delay. $f_{\text{D},m}(t)$ is the Doppler frequency due to the movement of the m-th reflector. Moreover, we call: $\mathbf{v}$ the vector of phase shifts used at the RX phased array (RX beamforming vector), θ_m, φ_m the Angle of Departure (AoD) and AoA of the m-th reflection, respectively, and $\mathbf{a}_{\text{tx}}(\theta_m), \mathbf{a}_{\text{rx}}(\varphi_m)$ the corresponding TX and RX array steering vectors. Considering uniform linear arrays at the TX and RX we have $\mathbf{a}_{\text{tx}}(\theta) = [1, \ldots, e^{-j(M_{\text{tx}}-1)\pi \sin\theta}]^{\mathsf{T}}$, $\mathbf{a}_{\text{rx}}(\theta) = [1, \ldots, e^{-j(M_{\text{rx}}-1)\pi \sin\theta}]^{\mathsf{T}}$, where M_{tx} and M_{rx} are the number of antennas at the TX and RX, respectively. Given these definitions, and calling $\delta(\cdot)$ the Dirac delta function, the continuous-time CIR model is

$$h(t,\tau) = \sum_{m=1}^{N(t)} A_m(t) e^{j2\pi f_{\text{D},m}(t)t} \delta(\tau - \tau_m(t)), \tag{2}$$

with $A_m(t) = \alpha_m(t)\mathbf{v}^{\mathsf{H}}\mathbf{a}_{\text{rx}}(\varphi_m)\left(\mathbf{a}_{\text{tx}}(\theta_m)\right)^{\mathsf{H}}\mathbf{u}$, where superscript $(\cdot)^{\mathsf{H}}$ indicates the complex conjugate transpose. Note that we include in the CIR model also the effect of TX and RX beamforming, that are combined in the complex-valued coefficients $A_m(t)$. $\alpha_m(t)$ instead accounts for the combined effect of the propagation loss and the target's Radar Cross Section (RCS) [34].

5) Received Signal Model: The RX signal is the result of a convolution between the TX signal and the CIR. In addition, it is also affected by TO and CFO, which we denote by $\tau_{\text{o}}(t)$ and $f_{\text{o}}(t)$, respectively. The RX applies a receive filter, $p_{\text{rx}}(t)$, and we define $p(t) = p_{\text{tx}}(t) * p_{\text{rx}}(t)$ to be the convolution between the TX and RX filters. The noise-free RX signal is

$$y(t) = \sum_{m=1}^{N(t)} A_m(t) e^{j2\pi(f_{\text{D},m}(t)+f_{\text{o}}(t))t} x'(t - \tau_m(t) - \tau_{\text{o}}(t)), \tag{3}$$

with $x'(t) = \sum_{i=0}^{L-1} s[i]p(t - i/B)$. Due to the presence of TO, the TX signal is further delayed with respect to the sole propagation time. The CFO instead causes an additional frequency shift that is added to the Doppler effect. We remark that TO and CFO are *constant* across the signal paths, and hence they are independent of index m.

The RX signal is sampled with period $\Delta\tau$, obtaining a discrete sequence of samples for each packet. We denote by T the time between two subsequent packets. In deriving the expression of the sampled RX signal we make the following common assumptions: (i) the combined filter $p(t)$ satisfies the Nyquist condition, i.e., $p(i\Delta\tau) = \delta[i]$ (Kronecker delta function), and (ii) the number of propagation paths and their parameters (amplitudes, Doppler shift, and delay) can be considered constant within a short processing interval of K packets, provided that T is sufficiently small that the propagation environment changes slowly compared to it [11]. We stress that the TO and CFO instead vary in each packet, hence they retain their dependence on time. The delay resolution of the system after discretization is determined by the sampling period $\Delta\tau$, thus being inversely proportional to the TX bandwidth. Due to the finite resolution, reflections with delay difference below $\Delta\tau$ are not *resolvable* at the RX. Hence, out of the N true signal propagation paths, only $N_{\text{r}} \leq N$ resolvable paths appear in the RX signal. We index such paths by variable n to distinguish them from the true propagation paths. Due to sampling, the n-th path delay and the TO are approximated as multiples of the sampling period, $\tau_n = \rho_n\Delta\tau$ and $\tau_{\text{o}}(kT) = \rho_{\text{o}}(kT)\Delta\tau$.

Although they cannot be distinguished in the delay domain, unresolvable paths could be discerned based on the Doppler shift. The n-th resolvable path is the superposition of N_n non-resolvable paths, with $N = \sum_{n=1}^{N_\mathrm{r}} N_n$. The i-th RX symbol in packet $k \in [0, \dots, K-1]$ is

$$y[k,i] = \sum_{n=1}^{N} \tilde{h}_n(kT) s[i - \rho_n - \rho_\mathrm{o}(kT)], \tag{4}$$

$$\tilde{h}_n(kT) = \sum_{\nu=1}^{N_n} A_{n,\nu} e^{j2\pi(f_{\mathrm{D},n,\nu} + f_\mathrm{o}(kT))kT}, \tag{5}$$

where $A_{n,\nu}$ and $f_{\mathrm{D},n,\nu}$ are the complex gain and the Doppler frequency of the ν-th superimposed path within resolvable path n, respectively. In Eq. (4) we neglect fractional components of the n-th path delay and the TO, since these cause sensing errors which are below the resolution of the system. Eq. (4) represents the *noiseless* RX symbols. In our model, we consider $y[k,i]$ to be affected by additive complex white Gaussian noise distributed as $\mathcal{CN}(0, \sigma_w^2)$, where σ_w^2 is the noise variance. Recalling we assumed unit TX symbol energy, the Signal-to-Noise Ratio (SNR) is denoted by $\Gamma = |\tilde{h}_1(kT)|^2/\sigma_w^2$ where $\tilde{h}_1(kT)$ is the gain of the LOS propagation path.

6) Estimated CIR Model: CIR estimation is carried out by cross-correlating the received samples with the known transmitted pilot sequence $s[i]$. This yields

$$h[k,\ell] = \sum_{n=1}^{N_\mathrm{r}} \tilde{h}_n(kT) \psi_n[\ell - \rho_n - \rho_\mathrm{o}(kT)], \tag{6}$$

for $k \in [0, \dots, K-1]$, where $\psi_n[\ell]$ is the cross-correlation of $s[i]$ with its delayed and frequency-shifted version (due to Doppler and CFO) for resolvable path n [34], [36]. In our implementation, we use complementary Golay sequences which exhibit perfect autocorrelation property when the frequency shift is zero [38]. This condition does not hold in Eq. (6), but such non-ideality can be neglected if the product of the Doppler plus CFO frequency and $L\Delta\tau$ (the duration of a pilot sequence) is much smaller than 1, as shown in [36]. As an example, with a CFO of 100 kHz, a Doppler shift of 4 kHz, $\Delta\tau = 0.568$ ns, and $L = 128$ we get $0.008 \ll 1$. It follows that the estimated CIR at a specific time is represented as a vector of complex channel gains, or *taps*, indexed by $\ell = 0, \dots, L-1$. The ℓ-th tap is related to a corresponding distance, $d_\ell = c\ell\Delta\tau$, with c being the speed of light.

Note that, as a result of Eq. (6), the SNR on the peaks of the estimated CIR is increased with respect to Γ as a result of the coherent integration of the signal. We denote the coherent integration gain by G. The exact gain depends on the frequency shift of the received signal (due to Doppler and CFO), and is upper bounded by $G \leq L$, which is the coherent gain in case of no frequency shift [34].

In the following, the TX is considered to be directional, whereas the RX is assumed to use a quasi-omnidirectional BP. Note that this is coherent with how existing mmWave commercial devices equipped with phased-arrays operate. This is not restrictive, as our method also works in the symmetric case where the RX listens to the signal using directional BPs, and the TX is omnidirectional.

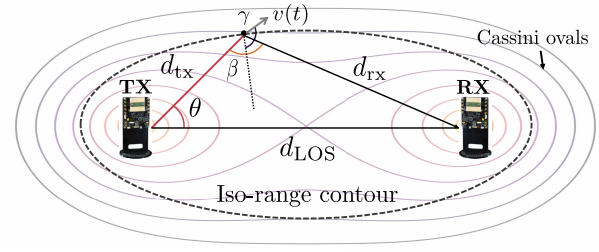


Fig. 2. Schematic representation of the bistatic geometry.

To account for the effect of analog TX beamforming, we denote by $h_b[k,\ell]$ the ℓ-th tap of the CIR at time kT obtained using BP b, i.e., when the TX uses beamforming vector $\mathbf{u}_b$. The complex gain of a path is denoted by $A_{\nu,n,b}$, which contains the antenna gain given by BP b along the direction that points to reflector n, together with the propagation loss and the RCS [34]. A visual representation of the magnitude of the CIR for different BPs is shown in Fig. 1c.

B. Bistatic Sensing Configurations

Next, we detail how the bistatic geometry affects the multipath reflection delays and Doppler frequencies.

1) Bistatic Reflection Delay: We denote by τ the *relative* delay of a generic multipath reflection, measured relative to the LOS propagation path along which the first signal copy arrives at the RX. Also, we denote by d_LOS the LOS distance between the TX and the RX. d_tx and d_rx are the distances from the TX to the target and from the target to the RX, respectively, as shown in Fig. 2. The following relation holds $c\tau = d_\mathrm{tx} + d_\mathrm{rx} - d_\mathrm{LOS}$, which states that the relative delay is the time needed for the signal to propagate along the excess length of the reflected path ($d_\mathrm{tx} + d_\mathrm{rx}$) with respect to the LOS path (d_LOS). Note that all reflectors located on the ellipse with focii coinciding with the TX and the RX will yield the same measured relative delay due to having the same $d_\mathrm{tx} + d_\mathrm{rx}$. This elliptical region is usually termed *iso-range* contour. In Fig. 2, we also represented the AoD of the reflection, called θ, as the angle formed by d_tx and d_LOS. Exploiting the geometric relations shown in Fig. 2, we can relate d_LOS, d_tx, and d_rx through the law of cosines

$$d_\mathrm{rx}^2 = d_\mathrm{tx}^2 + d_\mathrm{LOS}^2 - 2 d_\mathrm{tx} d_\mathrm{LOS} \cos\theta. \tag{7}$$

The distance of the reflector with respect to the TX, for a given delay τ, is obtained as

$$d_\mathrm{tx}(\tau) = \frac{(c\tau + d_\mathrm{LOS})^2 - d_\mathrm{LOS}^2}{2(c\tau + d_\mathrm{LOS} - d_\mathrm{LOS}\cos\theta)}. \tag{8}$$

The capability of a bistatic system to distinguish two reflectors located at different distances (hence causing different reflection delays) is termed *range resolution*. As described next, this value depends on the CIR delay resolution $\Delta\tau$. We define β as the *bistatic angle*, i.e., the angle formed by the segments connecting the TX to the target and the target to the RX. Due to the elliptical shape of the iso-range contours, the range resolution depends on the bistatic

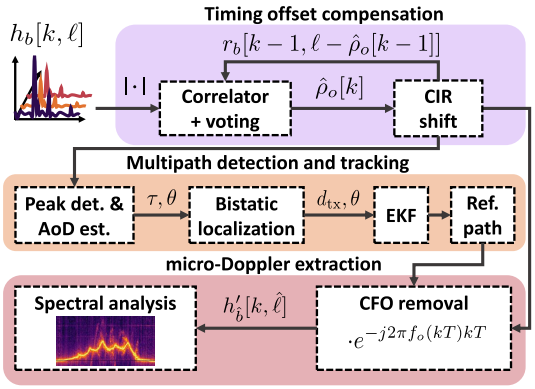


Fig. 3. Block diagram of the proposed method.

angle as $\Delta d \approx c\Delta\tau/(2\cos(\beta/2))$ [33]. This is a crucial aspect of bistatic systems. In practice, it prevents accurate localization of the reflectors when they approach the LOS segment, as $\Delta d \to +\infty$ when $\beta \to \pi$. The elliptical area around the LOS segment in which the resolution degrades is typically termed *forward scattering* region.

2) Bistatic SNR: In bistatic systems, for scattering-type reflections, the expected SNR at the RX is inversely proportional to the square of the product $d_{\text{tx}}d_{\text{rx}}$ [33]. For this reason, the locii of points along which the SNR is constant follow the so-called *Cassini ovals*, as shown in Fig. 2. As such, compared to a monostatic system, reflectors far from the TX device can still yield high SNR if they are sufficiently close to the RX.

3) Bistatic μD Spectrum: Consider the bistatic μD spectrum of the n-th resolvable path of the CIR. According to Eq. (6), it contains the contribution of N_n reflectors, each having a possibly different speed, together with the CFO. Moreover, the reflectors' movement speed maps to a different Doppler shift depending on the motion's direction and the bistatic angle. Denoting by $\gamma_{n,\nu}$ the angle between the bisector line of the bistatic angle $\beta_{n,\nu}$ and the velocity vector of the ν-th reflector superimposed resolvable path n, with magnitude $v_{n,\nu}$, we have [33]

$$f_{\text{D},n,\nu} = \frac{2v_{n,\nu}}{\lambda}\cos\gamma_{n,\nu}\cos\frac{\beta_{n,\nu}}{2}, \tag{9}$$

where $\lambda = c/f_c$ is the wavelength of the carrier f_c of the TX signal. As a consequence, in a bistatic scenario, the Doppler shift depends on the bistatic angle in addition to the direction of motion. The standard method for extracting the μD spectrum applies a Short Time Fourier Transform (STFT) to windows of subsequent CIR estimates [12]. Taking a processing window of K packets, the spectrum will present peaks at frequencies $f_{\text{D},n,\nu}$. However, in our asynchronous JCS setting, this approach is not directly applicable as the Doppler frequency is added to the time-varying CFO. Therefore, we cannot directly estimate the Doppler shift without removing the CFO first.

IV. Methodology

This section presents JUMP, whose block diagram is shown in Fig. 3. It includes three main steps, as described next.

1. *Timing offset compensation:* After obtaining the CIR in the current timestep, we leverage the previous CIR to estimate the relative random TO between the two, as detailed in Section IV-A. The random TO is estimated by computing the correlation between the CIR magnitude profiles with different BPs, resulting in the candidate TOs, and then combining them through a majority voting scheme (Section IV-A.1). Subsequently, the current CIR is shifted in the delay domain, compensating for the random offset (Section IV-A.2).

2. *Multipath detection and tracking:* Once the TO is compensated for, we perform detection and tracking of the multipath reflections in the environment. This is achieved by extracting the peaks in the CIR magnitude, computing their AoD, and smoothing the measurements with an Extended Kalman Filter (EKF) tracking algorithm. Section IV-B provides a detailed description of this step. Finally, we select the strongest *static* reflection from the set of tracked multipath components by measuring the variance of its 2D location estimated by the EKF (Section IV-B.5). We use this path to remove the CFO from the other reflections.

3. *μD extraction:* To extract the μD signature of the sensing targets, we compensate for the CFO by multiplying the CIR by a complex exponential with the phase of the previously selected static path (see Section IV-C). The key insight behind this operation is that the CFO is the same on all the multipath reflections. Subsequently, we apply STFT to the CIR to extract the μD.

A. Timing Offset Compensation

Here, we propose a method to compensate for the TO between two subsequent estimates of the CIR. We (reasonably) assume that, besides the reflections on the sensing targets, the background contains multipath reflections on *static* objects, which are *slowly* time-varying compared to the packet transmission rate. In addition, we consider that the packet transmission rate is high with respect to the movement speed of the sensing targets, which is a common assumption for radar and JCS. As a consequence, occlusions to the LOS lead to a progressive and "slow" disappearance of the corresponding path from the CIR, which happens over the course of several CIR estimates.

Consider two subsequent CIR estimates obtained from packets k and $k-1$ using the same BP b, and denote the magnitude of the CIR by $r_b[k,\ell] = |h_b[k,\ell]|$. If the two estimates are obtained at a sufficiently high sampling rate T, the propagation paths due to reflections on static objects do not change significantly, producing similar CIR magnitude profiles (envelopes). TO is expressed as $\tau_{\text{o}}(kT) = \rho_{\text{o}}(kT)\Delta\tau$ with $\rho_{\text{o}}[k] \in \{0,1,\dots,L-1\}$, as in Section III-A.5, defining $\rho_{\text{o}}[k] \triangleq \rho_{\text{o}}(kT)$. Therefore, our assumption about the slowly varying multipath environment can be written as $r_b[k,\ell+\rho_{\text{o}}[k]] \approx r_b[k-1,\ell]$. This equality means that the CIR envelope is preserved between subsequent time steps. It is just shifted by $\rho_{\text{o}}[k]$ in the path delay dimension, and this shift $\rho_{\text{o}}[k]$ is constant across all BPs but varies on a per packet basis (with index k).

In the following, we propose our method to estimate $\rho_o[k]$ and then use it to compensate for the TO, and shift the CIR back to a standard timing reference.

1) Estimation of the Timing Offset: Due to the similarity of the CIR envelopes in subsequent packets, a good estimate of the shift $\rho_o[k]$ can be obtained by maximizing the cross-correlation between the two CIRs magnitudes as follows

$$\hat{\rho}_b[k] = \arg\max_{\rho} \sum_{\ell=0}^{L-1} r_b[k, \ell+\rho] r_b[k-1, \ell]. \quad (10)$$

In Eq. (10), the estimate $\hat{\rho}_b[k]$ depends on the BP used in the estimation of the CIR, although $\rho_o[k]$ is the same across all BP. As such, a more robust estimate considers the P CIR estimates obtained for each packet, one per BP. We turn $\hat{\rho}_b[k]$ onto a one-hot vector representation $\hat{\mathbf{\Lambda}}_b[k] \in \{0,1\}^L$, which equals 1 at index $\hat{\rho}_b[k]$ and 0 elsewhere. Then, we use majority voting over the BPs $\hat{\rho}_b[k], b = 1, \ldots, P$, i.e.,

$$\hat{\rho}_o[k] = \arg\max_{i} \sum_{b=1}^{P} \hat{\Lambda}_{b,i}[k], \quad (11)$$

where $\hat{\Lambda}_{b,i}[k]$ is the i-th component of $\hat{\mathbf{\Lambda}}_b[k]$. Eq. (11) combines the estimated offsets from all the directions illuminated by the TX BPs. This procedure improves the robustness of the final decision, ignoring *outlying* $\hat{\rho}_b[k]$ that can be produced by BPs pointing in directions without static reflectors.

When transitioning from a LOS situation to a NLOS one, due to occlusion from a moving obstacle, $\rho_o[k]$ can be large. This is caused by the fact that the first packet copy to be detected at the RX will not travel along the (blocked) LOS. Still, the remaining static multipath reflections in the CIR are expected to remain constant, especially when considering a combination of all the BPs like in Eq. (11). Therefore, our method reliably estimates the relative TO even in such cases, as proven by our NLOS results in Section V and in Section VII.

2) CIR Shift: Upon obtaining an estimate of the TO, its effect is compensated for by shifting the CIR in the delay domain by $\hat{\tau}_o(kT) = \hat{\rho}_o[k]\Delta\tau$. In practice, this is implemented as a shift of the CIR along indices $\ell = 0, \ldots, L-1$ by $\hat{\rho}_o[k]$ positions. The corrected CIR is obtained as

$$h_b^c[k, \ell] = h_b[k, \ell - \hat{\rho}_o[k]], \quad \forall \ell = 0, 1, \ldots, L'-1, \quad (12)$$

where $L' < L$ is the length of the CIR portion of interest, which depends on the size of the monitored environment. The CIR taps having negative index after Eq. (12) are removed. The CIR estimated from the previous packet and $h_b^c[k, \ell]$ are aligned, thus allowing the computation of delays relative to the same reference point.

3) Initialization: Suppose that the sensing operation starts in a LOS condition. The direct path between the TX and RX is the *first* path to be detected at the RX. In this case, we initialize the CIR alignment process with respect to the LOS. Subsequent CIR estimates are shifted to compensate for the TO as previously described, allowing consistent tracking of the locations of the reflectors. If, instead, the LOS is unavailable when the sensing process starts, the CIR alignment is initialized with the first reflection that arrives at the RX. This introduces an offset in the obtained delays, which is corrected *a-posteriori* when the first packet traveling along the LOS path is detected. Therefore, having visible LOS during at least one timestep enables consistent tracking of the reflections.

B. Multipath Detection and Tracking

All the CIR estimates get the same timing reference using the TO compensation method from Section IV-A, which is represented by the first CIR tap in $h_b^c[k, \ell]$. As commonly done in JCS systems, we assume that the LOS distance, d_{LOS}, and the relative orientation between the TX and the RX, α, are known a priori. d_{LOS} can be estimated by applying *localization* methods to the RX node, for example, [39], while α is typically obtained from a beam alignment protocol [38]. Unlike existing approaches, we are interested in tracking both the dynamic *and* the static multipath components in the CIR, as the latter are used to remove the CFO. This makes the detection and tracking much more challenging as we want to preserve the information about the static objects in the environment, so clutter mitigation cannot be applied.

1) Delay Measurement by Peak Detection: The computation of the position of the peaks in the CIR magnitude after the alignment step of Section IV-A yields the relative delay of the reflection with respect to the LOS, τ. To reliably detect CIR peaks in the presence of clutter and background noise, we apply the Cell-Averaging Constant False-Alarm Rate (CA-CFAR) algorithm to the CIR magnitude after the TO compensation, denoted by $r_b^c[k, \ell]$ [34]. This operation consists of computing a dynamic threshold using a moving window. We call p the index of a generic peak returned by CA-CFAR at time kT. The delay of the corresponding reflection is then obtained as $\tau = p\Delta\tau$.

2) Angle-of-Departure Estimation: Upon receiving the signal reflected from an object in the environment, the RX computes the corresponding AoD relative to the TX, θ, using a modified version of the algorithm proposed in [40]. The latter computes the correlation between each BP shape and the CIR profile for a specific channel tap and requires that the TX BP shapes be known or estimated in advance. Unlike [40], however, in our system model more than one reflection can overlap in the same peak in the CIR magnitude, allowing multiple AoD to be detected for a single peak p in the delay domain, and collecting them into set Θ_p. In the following, the time index k is omitted to simplify the notation. We call $g_b(\theta)$ the gain of BP b along direction θ, and $\mathbf{r}[\ell] = [r_1^c[\ell], \ldots, r_P^c[\ell]]^\mathsf{T}$ the vector of CIR magnitudes for a specific ℓ.

Our AoD estimation algorithm computes

$$\Theta_p = \left\{ \theta \in [0, \pi] \text{ s. t. } \sum_{b=1}^{P} g_b(\theta) \frac{r_b^c[p]^2}{||\mathbf{r}[p]||_2^2} > \vartheta \right\}, \quad (13)$$

where ϑ is a threshold value used to adjust the sensitivity to multiple peaks in the correlation profile of the AoD.

3) Distance Estimation: To track the position of the targets with respect to the TX, we compute $d_{\text{tx}}(\tau)$ in Eq. (8) for $\tau = p\Delta\tau$ for each peak. Note that d_{tx} depends on p (already estimated) and d_{LOS} (assumed to be known) and θ, which belongs to set Θ_p. Hence, a set of candidate target locations

is obtained for each CIR peak p, by collecting all pairs (d_{tx}, θ), with $\theta \in \Theta_p$. A set of measurement vectors $\mathbf{z} = [d_{\text{tx}}, \theta]^{\mathsf{T}}$ collects all such pairs, *for all* the identified peaks p, and represents the input of an EKF [41] that tracks the Cartesian position of the targets relative to the TX.

4) Target Tracking (EKF): The EKF takes the measurement vectors $\mathbf{z}$ as input and, based on their evolution, obtains estimates of the states of the targets. The state is defined as the x, y position of a target and its velocity components along the two axes, $\mathbf{x} = [x, y, \dot{x}, \dot{y}]^{\mathsf{T}}$. A constant-velocity model is adopted to approximate the targets' movement, where we denote by $\mathbf{F}$ the state transition matrix and by $\mathbf{w}, \mathbf{v}$ the state and measurement noises, respectively. The tracking process operates following the EKF model equations $\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k$ and $\mathbf{z}_k = g(\mathbf{x}_k) + \mathbf{v}_k$, where $g(\mathbf{x}) = [\sqrt{x^2+y^2}, \arctan(y/x)]^{\mathsf{T}}$ is the non-linear function relating the current state to the current measurements. For measurements-to-tracks association, we adopt the cheap-nearest-neighbors joint probabilistic data association filter [42], [43], while to handle the initialization and termination of tracks, we use the method described in [44]. Such track management strategy filters out most false targets by taking into account the tracks' evolution across time, and allows quick initialization of new tracks whenever new targets enter the monitored area.

5) Static Path Selection: As described in Section IV-B.4, the state of each tracked multipath component contains an estimate of the current location of the reflector, i.e., the state components x and y. To identify a reliable static path, we first compute the mean location across the last K tracking steps for all the tracked reflectors. Then, we obtain the variance of the locations in the K frames with respect to this mean value. We select the reflectors with a variance under a certain threshold σ^2_{thr} for all the K frames. These are candidate reference paths, as their location is stable around a fixed position. The *strongest* path among them is the one providing the most reliable CIR phase value, having higher SNR. This path serves as a reference in the subsequent CFO removal.

C. μD Spectrogram Extraction

Here, we detail our CFO removal strategy and subsequent μD spectrum extraction. To compute the μD spectrograms of the tracked reflectors, we identify the CIR taps associated with the corresponding multipath components. This requires mapping the EKF states of the tracked components to the corresponding element of the CIR in the delay and angle domains. First, an estimate of the distance between the target and TX is $\hat{d}_{\text{tx}} = \sqrt{x^2+y^2}$. Then, denoting by α the angle between the TX and the RX, in the Cartesian reference system of the TX, an estimate of the AoD is given by $\hat{\theta} = |\arctan(y/x) - \alpha|$. Using $\hat{d}_{\text{tx}}$ and $\hat{\theta}$, we obtain an estimate of the distance between the target and the RX

$$\hat{d}_{\text{rx}} = \sqrt{\hat{d}_{\text{tx}} + d^2_{\text{LOS}} - 2 d_{\text{LOS}} \hat{d}_{\text{tx}} \cos\hat{\theta}}. \tag{14}$$

Once $\hat{d}_{\text{tx}}$ and $\hat{d}_{\text{rx}}$ are known, we estimate the CIR tap containing the reflection of the target through the estimated delay associated with a path of length $\hat{d}_{\text{tx}} + \hat{d}_{\text{rx}}$, as $\hat{\tau} = (\hat{d}_{\text{tx}} + \hat{d}_{\text{rx}} - d_{\text{LOS}})/c$. From this estimate, the CIR tap corresponding to each target is obtained as the one minimizing the delay difference with $\hat{\tau}$, i.e., $\hat{\ell} = \arg\min_\ell |\hat{\tau} - \ell\Delta\tau|$. Similarly, we select the BP pointing in the direction of the target, $\hat{b}$, as the one having the strongest gain along the direction $\hat{\theta}$.

The key idea behind our CFO correction method is that the term $f_{\text{o}}(kT)$ is constant in all multipath components of the CIR. Therefore, we can isolate the CFO component from the reference *static* path and remove it from the reflections on the sensing targets. The static reference path, which could either be the LOS path or a reflection on a stationary object (e.g., a wall), is identified as detailed in Section IV-B.5. Its corresponding CIR tap, $\hat{\ell}_{\text{s}}$, is obtained as detailed above. Neglecting the BP index, the expression of the $\hat{\ell}_{\text{s}}$-th CIR tap, corresponding to the static path, is

$$h^{\text{c}}[k, \hat{\ell}_{\text{s}}] = \tilde{h}_{n_{\text{s}}}(kT) = A_{n_{\text{s}}} e^{j2\pi f_{\text{o}}(kT)kT}, \tag{15}$$

where n_{s} is the index of the static path in $1, \ldots, N_{\text{r}}$. Note that since the path is static, even if it is a result of the superposition of multiple unresolvable reflections, these cannot be distinguished in the Doppler domain since their Doppler shift is equal to zero. As a result, Eq. (15) shows that the $\hat{\ell}_{\text{s}}$-th tap contains a single complex exponential component, whose phase only contains the CFO, without additional Doppler frequencies. Therefore, the offset can be removed from any other CIR tap corresponding to a target, with indices $\hat{\ell}, \hat{b}$, by computing

$$h'_{\hat{b}}[k, \hat{\ell}] = \frac{h^{\text{c}}_{\hat{b}}[k, \hat{\ell}]}{e^{j\angle h^{\text{c}}[k, \hat{\ell}_{\text{s}}]}} = \sum_{\nu=1}^{N_{n_{\text{t}}}} A_{n_{\text{t}},\nu,\hat{b}} e^{j2\pi f_{\text{D},n_{\text{t}},\nu} kT}, \tag{16}$$

where $\angle h^{\text{c}}[k, \hat{\ell}_{\text{s}}] = \arctan\left(h^{\text{c}}_Q[k, \hat{\ell}_{\text{s}}]/h^{\text{c}}_I[k, \hat{\ell}_{\text{s}}]\right)$ represents the phase of the reference path, letters I and Q denote the in-phase and quadrature components of $h^{\text{c}}[k, \hat{\ell}_{\text{s}}]$, and n_{t} is the index of the resolvable path corresponding to the target.

The μD spectrum of $h'_{\hat{b}}[k, \hat{\ell}]$ contains the time-varying contribution of the different reflections superimposing in CIR tap $\hat{\ell}$, as specified by Eq. (9). To compute it, we take the squared magnitude of the STFT of the cleaned CIR h', obtaining $S_{\hat{b}}[w, q, \hat{\ell}] = |\text{STFT}\{h'_{\hat{b}}[k, \hat{\ell}]\}|^2$, where w and q are the discrete frequency and STFT time-frame indices, respectively. Finally, for a tracked target, the μD spectrum is obtained by combining Q CIR taps preceding and following $\hat{\ell}$, as $\mu\text{D}[w, q] = \sum_{\ell=\hat{\ell}-Q}^{\hat{\ell}+Q} S_{\hat{b}}[w, q, \ell]$. This accounts for the extension of the target, which may exceed the system's ranging resolution.

We stress that, through Eq. (16), JUMP effectively removes the CFO *without affecting* $f_{\text{D},n_{\text{t}},\nu}(kT)$ (the Doppler term). Conversely, standard CFO removal methods used in communication systems compensate for the cumulative phase error caused by the CFO. This is done by following, e.g., the technique in [37]. Considering a single packet, we drop index k and denote by f_{o} the CFO, and by f_{D} the Doppler shift from a reflector. This causes a phase error of $\phi_{\text{o}} = 2\pi(f_{\text{o}} + f_{\text{D}})m\Delta\tau$, which increases *linearly* with time across subsequent samples in the packet preamble (here indexed by m). Then, the phase of the autocorrelation of the received signal at lag M amounts to $M\phi_{\text{o}}$, where M is the length of the pilot sequence used

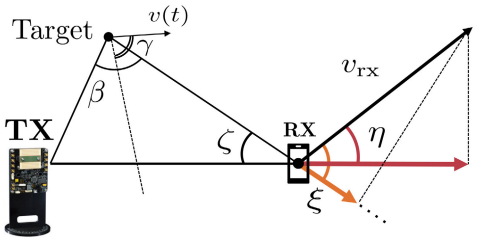


Fig. 4. Schematic representation of the bistatic geometry with a moving RX.

in the preamble (e.g., a Golay sequence in IEEE 802.11ay). This means that the cumulative phase error can be estimated by computing the phase of such autocorrelation, and dividing it by M. However, the estimated $\phi_{\rm o}$ contains $f_{\rm D}$, and it is therefore useless from a JCS perspective. For JCS, we instead aim at removing $f_{\rm o}$, while retaining $f_{\rm D}$.

D. Impact of RX Mobility

One of the key assumptions we made so far is that TX and RX are static, as only in this case static reference paths can be identified and used to remove the CFO. In this section, we discuss the impact of movements of the RX on JUMP. We show that if: (i) the RX is capable of estimating the AoA of the received multipath components, and (ii) an onboard sensor (e.g., an accelerometer on a mobile device) is available that can estimate the RX velocity, then the effect of RX motion can be compensated for. A full analysis of TX and RX motion is however beyond the scope of this paper, and constitutes a primary future research direction.

Consider the scenario in Fig. 4, where the RX device is moving with velocity $v_{\rm rx}$ and angles: η, with respect to the extension of the segment connecting TX and RX, and ξ, with respect to the extension of the segment connecting the sensing target and the RX. We assume the RX and the target velocities to be constant within a short processing time interval.

As a result of the RX motion, the phase of CIR along the propagation path caused by the target (*sensing path*) is given by $\phi_{\rm t}(kT) = 2\pi \left[f_{{\rm D},n_{\rm t},\nu} + f^{\rm rx}_{\rm D,t} + f_{\rm o}(kT) \right] kT$, where $f^{\rm rx}_{\rm D,t} = \frac{v_{\rm rx}}{\lambda} \cos\xi$ is the Doppler shift caused by the movement of the RX on the sensing path, indexed by $n_{\rm t}$. The phase of the CIR along the direct TX-RX path is $\phi_{\rm s}(kT) = 2\pi[f^{\rm rx}_{\rm D,s} + f_{\rm o}(kT)]kT$, where $f^{\rm rx}_{\rm D,s} = \frac{v_{\rm rx}}{\lambda} \cos\eta$ is the Doppler frequency caused by the RX along the TX-RX path. The expressions of $\phi_{\rm t}(kT)$ and $\phi_{\rm s}(kT)$ highlight that the Doppler shift caused by the RX movement is *different* along the TX-RX path and the target-RX path, due to the different angles η and ξ. This prevents aggregating the RX movement Doppler to the CFO and removing it by using the phase of the static path. Indeed, if one uses Eq. (16) directly, the resulting phase of the sensing path, $\phi'_{\rm t}$, contains a residual frequency offset that depends on $v_{\rm rx}$, η, and ξ, i.e.,

$$\phi'_{\rm t}(kT) = 2\pi \left[f_{{\rm D},n_{\rm t},\nu} + \frac{v_{\rm rx}}{\lambda} (\cos\xi - \cos\eta) \right] kT. \qquad (17)$$

Note that from the geometry in Fig. 4, by calling the AoA of the sensing path ζ, we have $\xi = \zeta + \eta$. It follows that if the RX can estimate the AoA, $\hat{\zeta}$, and its own movement speed vector (e.g., from an onboard accelerometer), $\hat{v}_{\rm rx}, \hat{\eta}$, the additional offset in Eq. (17) can be compensated for by multiplying the CIR by a complex exponential with phase $-2\pi \frac{\hat{v}_{\rm rx}}{\lambda} [\cos(\hat{\zeta}+\hat{\eta}) - \cos\hat{\eta}]kT$. An accurate estimate of $v_{\rm rx}$, ζ, and η allows JUMP to recover the correct Doppler frequency of the target. Errors on the estimation of $v_{\rm rx}$, ζ, and η can lead to a residual phase error, which is amplified by the presence of the wavelength λ in the denominator of Eq. (17). At mmWave frequencies, due to the short wavelength, JUMP's robustness to RX movements heavily depends on the quality of the external velocity estimate from the RX accelerometer.

V. Analysis and Numerical Simulation

In this section, we analyze the TO and CFO compensation capabilities of JUMP by providing insights on the residual timing and phase errors, as well as numerical simulations to validate our claims. We take IEEE 802.11ay as a reference, as it is also used in our testbed implementation. In this context, channel estimation fields appended to the packets are called beam-training (TRN) fields [38]. A TRN field includes 6 TRN *units*, each made of complementary Golay sequences of $L = 128$ samples modulated with Binary Phase-Shift Keying (BPSK), which can be transmitted using different antenna BPs.

In the simulations, we set $B = 1.76$ GHz (as in a standard IEEE 802.11ay channel), hence $\Delta\tau = 0.57$ ns and $\Delta d = 17$ cm. The inter-packet transmission time is $T = 0.27$ ms. We generate 10^4 CIR realizations with a random number of scatterers between 2 and 10. Each scatterer is located at a random distance from the TX, chosen uniformly at random in $[1.5, 10]$ m. The CIR taps' amplitudes are computed using the bistatic radar equation [34], with a random RCS for the scatterers in the interval $[-20, 10]$ dBsm (dB per square meter). The transmitted signal is a single IEEE 802.11ay TRN field, for a total length of 768 symbols. To evaluate JUMP's performance in intermittent LOS/NLOS conditions, in some of the simulations we modulate the CIR amplitude for the LOS path with an exponential profile which decreases to zero. This simulates the blockage due to scatterers passing between the TX and the RX.

The TO and the phase shift due to CFO are obtained as uniform random variables in the intervals $[0, 20\Delta\tau]$ s and $[0, 2\pi]$, respectively.

A. TO Compensation Error Analysis

We consider errors in the estimation of the relative TO from Eq. (10). This is a conservative approach, in that we do not account for the subsequent robustness improvement brought by the majority voting scheme in Eq. (11). For simplicity, we neglect the time index k, following our assumption that the channel profile can be considered static in the short time between two subsequent packets.

The accuracy of the relative TO estimate in Eq. (10) depends on (i) the relative location of the CIR peaks, and (ii) their sharpness. Indeed, cross terms due to the multiplication of CIR peaks caused by different propagation paths may result in strong, secondary cross-correlation peaks that cause ambiguity in the TO estimation. Moreover, despite the perfect autocorrelation properties of Golay sequences, in practice the CIR peaks

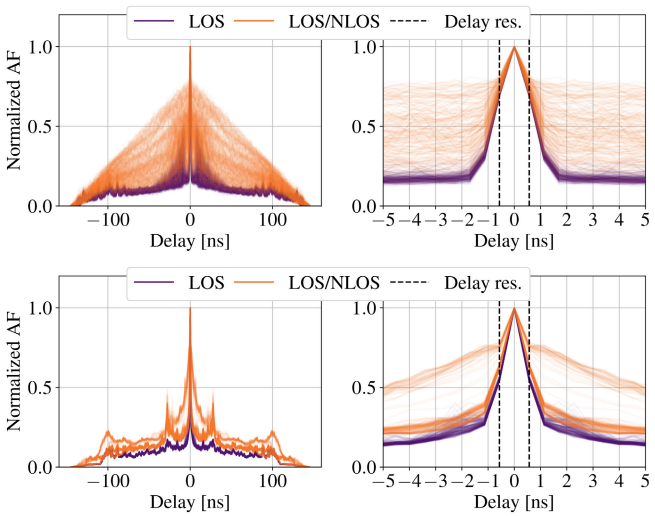


Fig. 5. Ambiguity function of the CIR amplitude profiles obtained from real data for LOS and intermittent LOS/NLOS (a zoom around ±5 ns is shown on the right). Different curves represent different channel realizations, for a total of 1000.

are not exact Kronecker delta functions, but have a non-zero *width*. This is caused by multiple factors, among which are the imperfect symbol synchronization (that causes a small error in the sampling point at the RX) and the fact that extended targets cause multiple reflections that overlap in the same or adjacent delay bins. However, because the CIR profiles depend on the underlying physical environment, which can be very different depending on the scenario, TO errors are difficult to model analytically without making restrictive assumptions.

For this reason, we take an alternative approach using the concept of Ambiguity Function (AF) in radar signal processing [34]. The AF is defined as the cross-correlation of a waveform with its delayed and frequency shifted versions and it measures the sharpness of the main correlation peak and the secondary peaks level. In our setting, we compute the AF of the CIR amplitude profiles used in Eq. (10), hence Doppler shift has no impact and the AF reduces to the autocorrelation of each CIR amplitude profile. The ideal AF in this situation is a Kronecker delta centered in zero. We obtain the AF of the CIR profiles in simulation and using *real* measurements from our testbed implementation and setup described in Section VI and Section VII. The resulting normalized AF for 1000 CIR profiles is shown in Fig. 5 for LOS and intermittent LOS/NLOS scenarios. In both real and simulated CIRs, the LOS AF exhibits a sharp peak at 0 ns, which is well contained within the delay resolution of the system, shown by the black dashed lines. In LOS/NLOS, the AF shows a higher noise floor, which is due to the lower amplitude of the NLOS peaks. Nevertheless, the main peak is still clear and sharp around 0 ns, showing the robustness of our approach. Note that the difference between simulated and real measurements is due to the much higher channel variability that can be obtained in simulation. The simulated AF is a more general assessment of JUMP's robustness for diverse channel realizations. Fig. 6a and Fig. 6b show the TO estimation Root Mean Squared Error (RMSE) varying the SNR and the system bandwidth, respectively. JUMP achieves an RMSE lower than the delay resolution $\Delta\tau$ (black dashed line) for an SNR of -5 dB in LOS and 0 dB in LOS/NLOS. This is sufficient to correctly estimate the integer part of the TO. Note that this is exactly the purpose of JUMP's TO compensation method: we are not interested in estimating the TO itself, but only in computing the correct integer shift to be applied to consecutive CIR profiles to obtain consistent range measurements. At higher SNR values, the error converges to the standard deviation of the fractional part of the true TO used in the simulations (black dotted line). This is expected since the fractional part of the TO is neglected in Eq. (10).

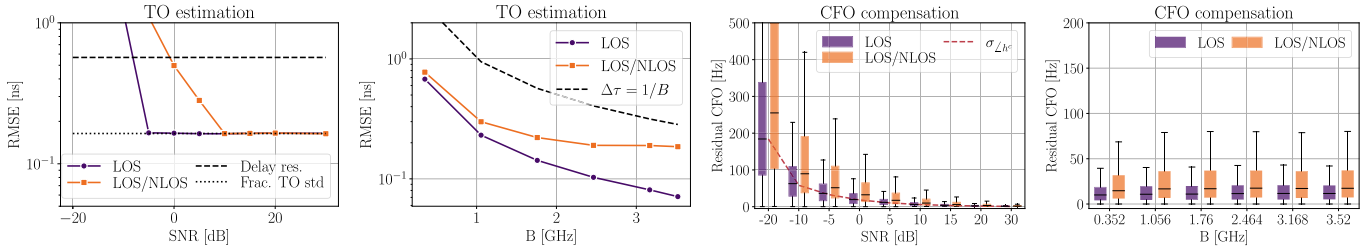
Fig. 6b shows the TO compensation RMSE varying the TX bandwidth, B, for an SNR of 5 dB. In LOS, the gain from using a larger B scales similarly to the delay resolution. Conversely, in LOS/NLOS conditions JUMP cannot make full use of larger bands due to the worse actual SNR, and thus the error does not improve with the delay resolution beyond 2 GHz.

B. CFO Removal Error Analysis

JUMP removes the CFO using the phase of the static reference path. Such phase is affected by noise on the RX signal, which causes a residual phase error on the sensing path. We use expression $\angle h^{\mathrm{c}}[k,\hat{\ell}_{\mathrm{s}}] = \arctan\left(h^{\mathrm{c}}_Q[k,\hat{\ell}_{\mathrm{s}}]/h^{\mathrm{c}}_I[k,\hat{\ell}_{\mathrm{s}}]\right)$ introduced in Section IV-C to compute the residual phase error variance, $\sigma^2_{\angle h^{\mathrm{c}}}$. The variance of $h^{\mathrm{c}}_I[k,\hat{\ell}_{\mathrm{s}}]$ and $h^{\mathrm{c}}_Q[k,\hat{\ell}_{\mathrm{s}}]$ is $\sigma^2_w/(2G)$, and $\sigma^2_{\angle h^{\mathrm{c}}}$ can be obtained by propagating the error induced by noise through the expression of $\angle h^{\mathrm{c}}[k,\hat{\ell}_{\mathrm{s}}]$, as shown, e.g., in [45]. For low $\sigma^2_w/(2G)$, this gives

$$\sigma^2_{\angle h^{\mathrm{c}}} = \frac{\sigma^2_w}{2G|h^{\mathrm{c}}[k,\hat{\ell}_{\mathrm{s}}]|^2} = \frac{1}{2G\Gamma}\frac{|\tilde{h}_1(kT)|^2}{|h^{\mathrm{c}}[k,\hat{\ell}_{\mathrm{s}}]|^2}, \tag{18}$$

where we recall that $\tilde{h}_1(kT)$ is the gain of the LOS path. Eq. (18) shows the dependency of the residual phase error on the power of the CIR tap corresponding to the reference path used for the CFO removal. This is validated by the simulation result in Fig. 6c. We show the residual CFO for different values of the SNR on the received signal, in both LOS and intermittent LOS/NLOS conditions. We also plot the theoretical error standard deviation from Eq. (18) in the LOS case for $G = L$ (rescaled by $2\pi T$) with a black dashed line, which matches well with the simulation result for higher values of the SNR. The best performance is obtained using the LOS path as the reference since it is typically much stronger than any first order reflection. Nevertheless, JUMP shows excellent CFO estimation even in intermittent LOS/NLOS, where a static first order reflection is used as described in Section IV-B.5. Finally, in Fig. 6d we report the residual CFO distributions obtained fixing the SNR to 5 dB and changing the TX bandwidth B between 350 MHz and 3.52 GHz. The results demonstrate that the resulting error is independent of the bandwidth, as expressed by Eq. (18). This makes our CFO removal approach also applicable to communication systems with lower bandwidth (e.g., 4G-LTE and 5G-NR), provided that a reference path can be identified.



(a) $B = 1.76$ GHz. (b) SNR = 5 dB. (c) $B = 1.76$ GHz. (d) SNR = 5 dB.

Fig. 6. Simulation results for TO estimation RMSE (a-b) and CFO compensation error distribution (c-d), varying the SNR and the TX signal bandwidth. We report the results for LOS and intermittent LOS/NLOS conditions.

C. Impact on Communication: Overhead

To perform sensing, JUMP requires obtaining: (i) CIR estimates covering the full angular space in order to detect and track the targets, and (ii) CIR estimates obtained with TX BPs pointing towards the targets of interest, with sufficient granularity to achieve a sufficient maximum Doppler frequency $f_{\mathrm{D}}^{\max} = 1/(2T)$. Condition (i) is easily satisfied by beam training operations that are commonly performed in mmWave systems to align the TX and RX beams for communication. As an example, in IEEE 802.11ay a full beam sweep is done periodically (e.g., every 100 ms) by appending TRN units to the packets. Due to typical human movement speeds being in the order of $1-4$ meters per second, the beam training frequency is sufficient to fulfill JUMP's requirements without adding any overhead to the communication protocol. Condition (ii) instead demands that additional channel estimation fields are appended to communication packets during normal traffic. Specifically, JUMP needs *one* additional field *per target*, and *one* field for the reference static path, transmitted using a BP that illuminates the corresponding direction. This adds overhead to the communication protocol, as shown in Fig. 7. We plot the overhead introduced by adding 1 to 12 TRN units, which corresponds to sensing up to 11 targets concurrently, plus the reference path. Different colors correspond to different Physical layer Service Data Unit (PSDU) sizes used in the standard [38], while solid and dashed lines refer to Modulation and Coding Schemes (MCSs) 8 and 12, respectively. The overhead is computed as the ratio between the length of the added PHY layer symbols in the TRN units and the total number of symbols in the packet, considering the payload plus PHY and MAC headers [38]. For PSDU sizes 66 kB to 4194 kB the overhead is less than 3%, with 11 targets, which is a negligible impact on communication. Note that in practical scenarios, given the limited TX range at mmWave frequencies, it is unlikely to sense such a high number of targets concurrently. For PSDU 4 kB, the overhead can reach 10-30%. However, note that this PSDU size should be avoided anyways in IEEE 802.11ay deployments, as due to the high rates available in mmWave communication it yields a very low MAC layer efficiency. Specifically, with a PSDU of 4 kB, the MAC efficiency is 74% solely because of the MAC and PHY layers overheads.

Finally, we underline that in this paper we do not consider the irregularity of the inter-packet times in communication systems. We address this in a separate work, [6], proposing a method to reconstruct μD signatures from the irregular CIR estimates obtained from communication packets.

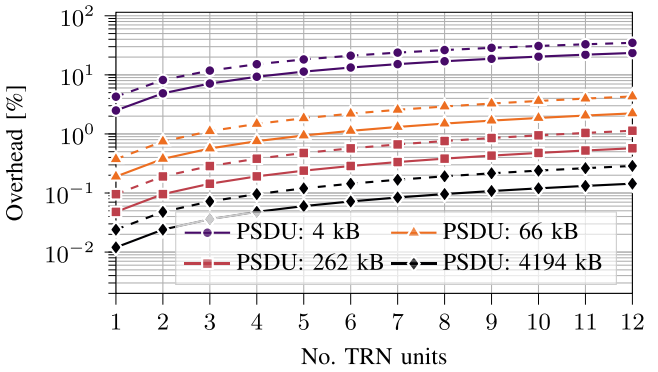


Fig. 7. Overhead introduced by JUMP on the IEEE 802.11ay PHY layer packet, as a function of the number of added TRN units, for different PSDU sizes. Solid lines correspond to MCS 8, while dashed lines represent MCS 12.

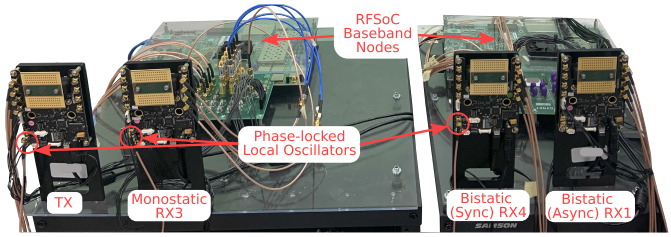


Fig. 8. Testbed components.

VI. Implementation

JUMP's prototype is based on the 60 GHz SC IEEE 802.11ay standard [38], maintaining its packet structure. CIR estimation is performed by using trailing TRN fields.

1) Testbed Design: Our implementation is a customization of MIMORPH [46], an open-source project for mmWave experimentation inclusing a Xilinx Radio Frequency System on a Chip (RFSoC) board and Sivers' mmWave front-end [47]. Fig. 8 shows the main components of our testbed. The MIMORPH Field Programmable Gate Array (FPGA) logic was modified to allow its operation as a TX, RX, or both functionalities simultaneously. In this way, we can emulate a monostatic JCS system, which we use as a baseline for comparison. The Analog-to-Digital Converters (ADCs) and Digital-to-Analog Converters (DACs) on the board are configured to operate at 3.52 GHz sampling frequency,

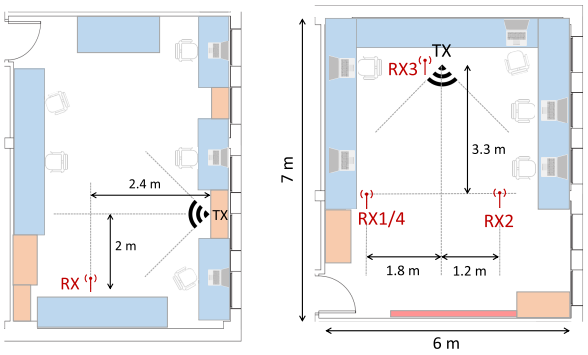
fulfilling the requirements of mmWave Wi-Fi standards (IEEE 802.11ad/ay), with 1.76 GHz of RF bandwidth.

The TX implements a loopback memory that feeds the DACs with the I/Q symbols to be transmitted, which are loaded from an external processor. To enable AoD estimation (Section IV-B.2), we implement a real-time antenna reconfiguration mechanism that allows to sweep through different BPs in the TRN fields of a single packet, as in [46].

At the RX, after downconversion and sampling, packet detection is performed by searching for peaks in the autocorrelation of the received signal [40], [48]. In communication between asynchronous devices, in addition to TO and CFO, the non-ideal sampling point at the RX causes a symbol timing offset. This introduces Inter-Symbol Interference (ISI). Although this is generally neglected in the JCS literature [5], [11], in our experiments we found that it affects μD quality. To avoid this, we implement a symbol synchronization block including a configurable fractional delay filter, based on a Farrow structure [49], followed by a fast Golay correlator [48]. After proper sampling point selection and subsequent downsampling, the I/Q symbols are fed to the onboard RAM.

2) Different Testbed Configurations: Our implementation includes a TX, two JUMP receivers, RX1-2, and two receivers, RX3-4, which we use to assess and compare JUMP's performance against monostatic and phase-locked configurations. JUMP's RXs are completely independent of the TX, as they do not share any clock source or oscillators. The monostatic RX, denoted by RX3, is located in the same location as the TX, as shown in Fig. 8. For the monostatic case, the RFSoC board concurrently operates as TX and RX (i.e., as a full-duplex system). To remove the CFO, the local oscillator is fed from the TX antenna to the RX3 antenna (Fig. 8). The direct LOS from TX to RX3 provides a reliable reference to align the CIR estimates, removing the TO. With RX4, a phase-locked version of the bistatic JCS system was implemented. This is similar to the monostatic RX3, but in this case, RX4 is co-located with JUMP's receiver, RX1. Similarly to the monostatic case, we phase-locked the mmWave front-ends by sharing the TX local oscillator signal through a cable, thus eliminating the CFO. RX4 is used as a reference to assess the quality of the JUMP reconstructed μD spectrum.

3) System Parameters: We use the same values of B and T used in the simulation in Section V. For μD extraction, we apply STFT with a Hanning window of size $W = 256$. This yields a maximum resolvable Doppler frequency of $f_{\mathrm{D}}^{\max} = 1/(2T) = 1851.85$ Hz and a resolution of $\Delta f_D = 1/(WT) = 14.47$ Hz. The velocity corresponding to such bistatic Doppler frequency depends on γ and β as per Eq. (9). For the AoD measurements, we set ϑ to 0.7 multiplied by the maximum correlation value in Eq. (13). This allows capturing the AoD of different reflections in the same CIR tap while rejecting peaks due to sidelobes of the BPs. $K = 100$ frames are used for the selection of the static reference path, with a variance threshold $\sigma^2_{\mathrm{thr}} = 0.005$ m^2 (see Section IV-B.5). For the μD spectrum, we use $Q = 2$, considering channel taps in a (bistatic) range of $\pm Q\Delta d = \pm 34$ cm from the target, which we empirically found to be suitable for human sensing, given typical human body sizes.



(a) Env1. (b) Env2.

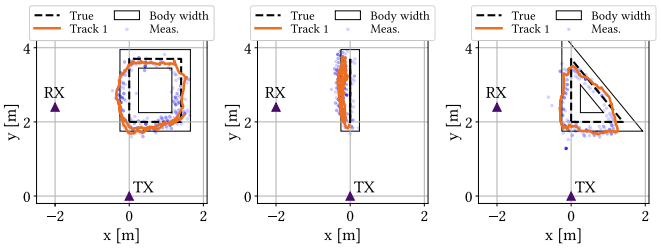
Fig. 9. Schematic representation of the two environments.

VII. Experimental Results

JUMP was tested on the tasks of people tracking and μD extraction, in two different indoor environments to verify the system effectiveness under different conditions. The test areas are research laboratories, denoted by Env1, of dimensions 5×8 m, and Env2, of dimensions 6×7 m. Fig. 9 shows them, along with the position of the TX and RX nodes. The placement of the RXs with respect to the TX, in terms of relative distance and orientation, is shown in Fig. 9 for Env1 and Env2. We use Env1 for baseline evaluations, using a single JUMP receiver, denoted by RX. In Env2 instead, we perform comparisons with the monostatic and phase-locked systems, using RX1 and RX3-4, and we evaluate the performance of a *multistatic* JUMP deployment using RX1-2 concurrently. To reconstruct the ground truth movement trajectory of the subjects, we mark the floor at specific locations in the two environments and instruct them to move across the markers. To perform measurements under NLOS conditions, we use rectangular panels of absorbing material interposed between the TX and RX to block the LOS signal. This is used to simulate common occlusions of the LOS that can happen in dynamic environments, caused, e.g., by other subjects moving around. We collect over 60 CIR sequences in Env1 and Env2, each with duration of about 12 seconds (40k packets).

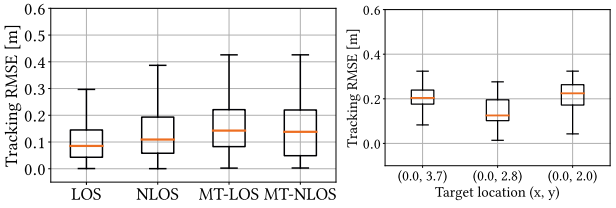
A. Bistatic Tracking Accuracy

To evaluate JUMP's tracking accuracy, we compute the RMSE of the estimated movement trajectory of the subjects with respect to the ground truth. Fig. 10 shows the tracks outputted by the EKF in a LOS setting in Env1 for three different movement trajectories of a single person. Along with the tracks, we plot the ground truth with a solid black line and the measurement vectors, $\mathbf{z}$, with gray dots. Moreover, we also represent the body width of the subject during the movement with a black solid rectangle. The body width is non-negligible due to the high delay resolution of our system. Although it is not possible to know the exact reflection point of the signal on the subject's body, we expect the collected measurements to be biased towards the body side that is facing the TX/RX pair. This is clearly visible in Fig. 10. To account for the bias introduced by the body width, in the following analysis we compute the tracking error with respect to the trajectory of the



(a) Square. (b) Line. (c) Triangle.

Fig. 10. Example estimated trajectories.



(a) LOS/NLOS tracking in the single and multi-target (MT) cases. (b) Tracking error depending on the target location.

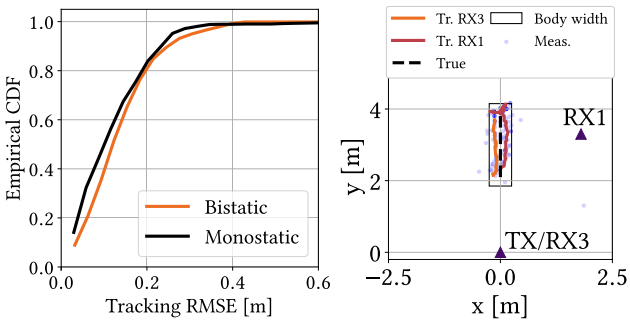
Fig. 11. Boxplots of the normalized tracking RMSE.

body side facing the TX/RX, as this is more representative of the true tracking accuracy of the system. The body size was measured for each subject involved in the experiments.

1) LOS Tracking: As a baseline result, we compute the average RMSE for different measurement sequences obtained in `Env1`, obtaining a median RMSE of 8.5 cm, as shown in the first boxplot of Fig. 11a. This result provides a first assessment of the capability of our system to compensate for the TO due to clock asynchrony between TX and the RX. In Fig. 11b, we show the RMSE obtained when tracking a subject sitting down and standing up at three different locations: 3.7, 2.8, and 2 m in front of the TX. We observe that, on the one hand, distant targets cause reflections with high $d_{\text{tx}}d_{\text{rx}}$ product, which are harder to detect and track due to low SNR. On the other hand, targets too close to the LOS link have the bistatic angle β close to π, yielding a lower-ranging resolution. For these reasons, the median tracking errors in the first and third cases are higher than in the second one.

2) Impact of NLOS: In the second boxplot in Fig. 11a we show the RMSE distribution obtained in *intermittent* LOS/NLOS conditions. In collecting these CIR measurements, we block the LOS link using the absorbing panel intermittently for time intervals of approximately $1-2$ s. This test is very challenging in terms of TO compensation, as the received packet can be detected from one of the reflections on surrounding objects, thus presenting a possibly large relative shift between subsequent packets. Still, our approach can successfully remove the TO, obtaining a median RMSE of 10.9 cm, with the third quartile of the error distribution being less than 20 cm.

3) Comparison With a Monostatic System: In Fig. 12 we compare our system with a *monostatic* sensing configuration (see Fig. 12b). This evaluation provides a comparison with existing approaches in the literature that adopt full-duplex monostatic sensing systems, e.g., [8] and [12]. In addition,



(a) RMSE CDF. (b) Example trajectory.

Fig. 12. Monostatic vs. bistatic tracking RMSE comparison.

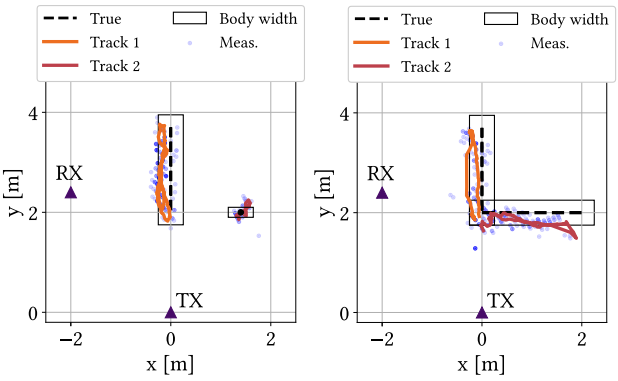
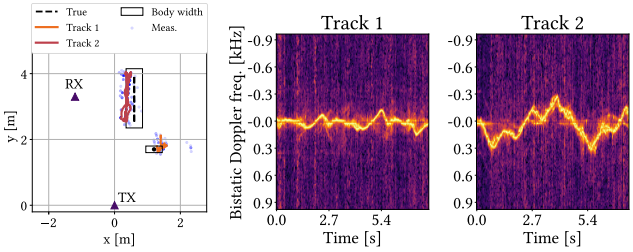


Fig. 13. Multitarget tracking in `Env1`.

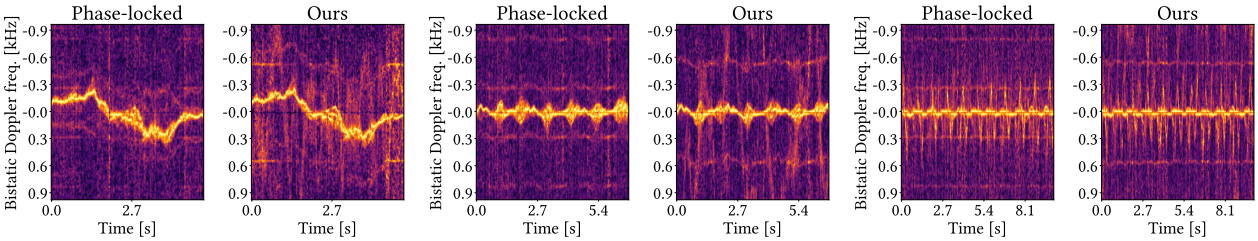


(a) Tracking. (b) Extracted μD.

Fig. 14. μD extracted in a multitarget scenario.

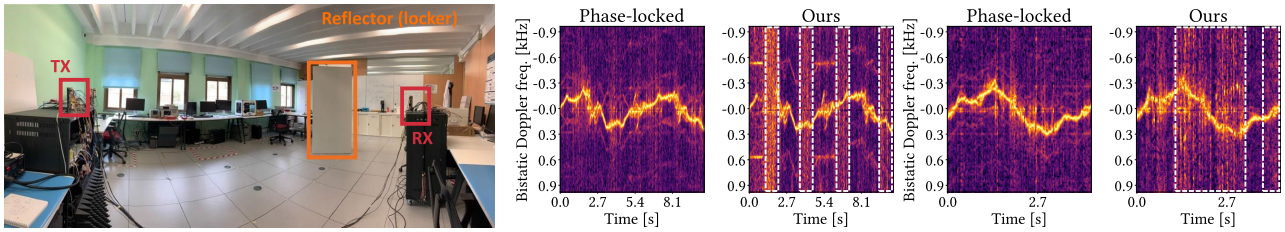
it allows evaluating the impact of the resolution as a function of the bistatic angle (see Section III-B). For this, we concurrently collect CIR sequences using RX3 and RX1, having the subject walk along different linear trajectories. Fig. 12a shows the empirical RMSE Cumulative Distribution Function (CDF) obtained with the monostatic and bistatic systems. Although less accurate, our system performs similarly to the monostatic configuration. This is because bistatic systems produce better sensing SNR in the region around RX1, which instead is a low SNR region for RX3. When multiple trajectories and locations are considered, our results show that this compensates for the degradation in the bistatic range resolution.

4) Multitarget Scenario: Finally, we evaluate the tracking accuracy when multiple subjects concurrently move in the environment. Fig. 13 shows two example results obtained with two subjects when one is walking and the other is sitting down/standing up repeatedly (left), and when both are walking on different trajectories (right). The system can accurately



(a) Walking. (b) Sitting. (c) Hand gestures.

Fig. 15. μD signatures for three different activities obtained with a phase-locked system and with JUMP.



(a) Env2 with reflector. (b) W/o. reflector. (c) W. reflector (4.8 m).

Fig. 16. μD reconstruction with intermittent LOS/NLOS. NLOS regions are enclosed in white dashed rectangles.

track multiple targets obtaining the RMSE distributions shown in the last two boxplots of Fig. 11a. We report the results for both the LOS and the intermittent LOS-NLOS cases. The median RMSE slightly degrades compared to the single subject case (14 cm), due to the more frequent outliers caused by the more challenging multi-target tracking task.

B. Micro-Doppler Quality

Next, we evaluate the quality of the μD signatures obtained with our CFO removal technique in Env2. The μDs shown here and in the rest of this section are obtained as $\log \mu\mathrm{D}[w, q]$, and normalized frame-wise in the interval $[0, 1]$. Darker colors correspond to lower energy levels, while lighter ones represent high energy. We show an example result in Fig. 14, with two subjects, one sitting down and standing up (Track 1) and one walking (Track 2). Notice that the μD signatures after the removal of the CFO contain a strong center frequency due to the torso's movement and sidebands due to the limbs. To provide a ground-truth reference for the μD spectrograms, we co-locate JUMP's receiver RX1 and the phase-locked RX4 at coordinates $(1.8, 3.3)$ m.

1) Comparison With a Phase-Locked System: In Fig. 15, the μD of our system is compared with the phase-locked system for three different activities: (a) walking, (b) sitting/standing, and (c) hand gestures. These tests are performed in a LOS setting. The cleaned μD reflects the phase-locked one very accurately, even for fine-grained Doppler shifts involved in hand gestures. To quantitatively evaluate the difference between the phase-locked spectra and JUMP's ones, we compute the frame-wise, *masked* RMSE of the spectrograms, after temporally aligning them. A direct computation of the RMSE would not give a clear indication of the quality of the μD, as it would equally weigh differences in the background noise floor and the body μD contribution. For this reason, we apply a 2D Gaussian filter, with parameter $\sigma = 2$, to the phase-locked μD

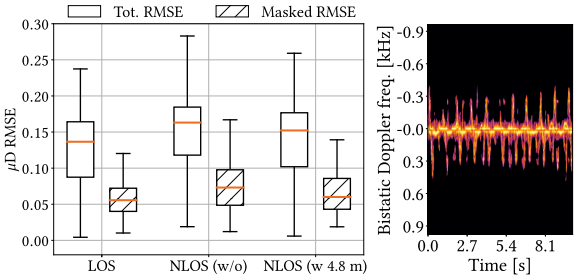
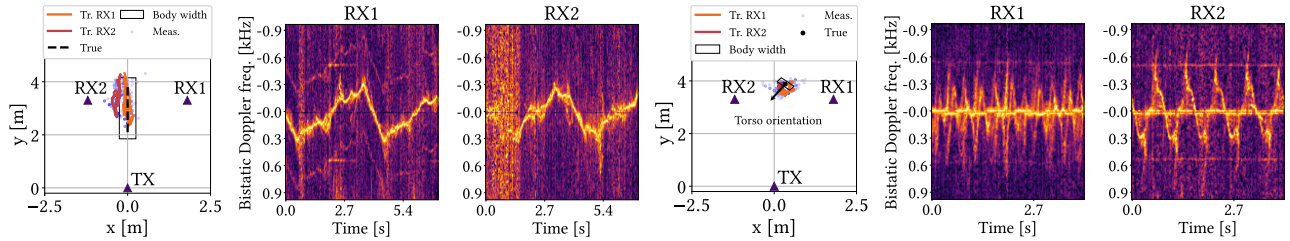


Fig. 17. (Left) Normalized RMSE distributions on the μD spectrograms, using the phase-locked ones as reference, in: LOS, NLOS with a strong reflector (metal locker) placed 4.8 m from the TX, and in NLOS with no strong reflectors. (Right) Example masked locked μD obtained from Fig. 15c.

to filter out background noise, as shown in Fig. 17 (right). Then we restrict the RMSE computation to the spectrogram elements in which the filtered spectrogram lies over a threshold of 0.45. The RMSE distribution over all the collected frames is reported in Fig. 17 (LOS). For completeness, we also report the standard RMSE, calculated over the entire spectrogram.

2) Impact of NLOS: In NLOS situations, our algorithm removes the CFO using the strongest static multipath reflection as a reference. Depending on the corresponding signal's strength and reflector's location, the resulting μD spectrum can be of higher or lower quality. In Fig. 16, we show two μD spectrogram examples obtained in Env2 using the same setup described above. We intermittently block this evaluation's LOS link using the absorbing panel. The NLOS regions in the μD are highlighted using white dashed rectangles. In Fig. 16b, we used the standard Env2 setup, without reflectors close to the testbed devices, while in Fig. 16c, we placed a metal locker at 4.8 m in front of the TX. In the first case, the μD reconstruction is particularly challenging, as the algorithm's reference multipath reflections are weak due to the large distance of the reflectors. Despite this, our method successfully



(a) Walking. (b) Walking μD. (c) Sit/stand. (d) Sit/stand μD.

Fig. 18. Independent multistatic tracking and μD extraction using multiple RXs.

recovers the μD in 3 out of 4 LOS occlusion events, while the spectrum appears corrupted during the first one. We verified, as exemplified in Fig. 16c, that when the metal locker is present on the scene, the μD corruption no longer occurs, as the algorithm can always find a reliable and strong reflection for removing the CFO. Quantitative results for this scenario are reported in Fig. 17, without (w/o) and with (w) the reflector. In the worst case of NLOS without reflectors, JUMP obtains a normalized RMSE of 0.07.

C. Multistatic Scenario

The proposed method easily scales to multiple RXs, allowing thorough and robust sensing of the environment. In Fig. 18, we show two measurement sequences obtained by concurrently operating RX1 and RX2 in `Env2`, with a subject walking (a-b) and performing hand gestures (c-d). In the first case, the subject starts moving close to the TX-RX2 LOS, thus yielding a noisy μD. However, the spectrum obtained from RX1 does not show noisy regions thanks to the better position of RX1 to capture the movement. In addition, the shape of the μD is slightly different due to the different viewpoints for RX1-2. This effect is captured even more clearly in Fig. 18b, where the subject is waving his arms in front of RX1 and RX2, with the torso oriented midway between TX and RX2. In this position, the full movement of the arms is not clearly visible to RX1, as it is blocked by the subject's back. This results in very different μD patterns that provide richer information about the movement.

VIII. Discussion and Future Work

1) Multistatic Data Fusion: The results provided in Section VII-C pave the way for multistatic data fusion across multiple receivers. As shown in Fig. 18, each JUMP receiver returns an independent view of the physical environment to be monitored. Leveraging multiple receivers would then provide a means to improve tracking accuracy, e.g., by combining their estimated tracks to solve possible occlusion events. Moreover, multiple receivers usually have different observation angles for the same targets. The fusion of μD spectrum features coming from such diverse viewpoints allows for enhanced diversity and, overall, for an improved understanding of the underlying movement, especially for extended targets with many moving parts. Target identification or motion recognition algorithms can exploit such enhanced feature representations to boost their performance.

2) NLOS Operation: Our target tracking accuracy results in the presence of NLOS events show that, for indoor human sensing, our assumption that most background reflections remain constant in between subsequent tracking frames holds. Moreover, when blocked, the LOS path gradually disappears across subsequent CIR estimations, yielding correct and clear correlation peaks. In general, however, the tradeoff between inter-packet time and the level of dynamicity in the environment plays a key role. As shown in [12], a minimum packet rate has to be ensured to accurately capture the Doppler spectrum of the physical targets, to avoid aliasing and low resolution. In the case of very challenging dynamic environments, with few static reflections, the JUMP correlation approach could be replaced by optimization-based association mechanisms between multipath reflections appearing in different CIR estimates. This aspect, and the possible improvements that it may bring, are left to future investigations.

3) Mobility: The analysis in Section IV-D shows that JUMP's CFO removal step can be extended to also compensate for RX movement. However, this requires estimating the RX velocity vector using external sensors (e.g., an onboard accelerometer). The resulting compensation is highly sensitive to errors in such estimation, which may degrade performance in practical settings. Therefore, we identify mobility as a key challenge for future JCS research that needs further attention. Interesting solutions may involve enhancing the accelerometer estimate using information from the wireless channel.

IX. Concluding Remark

The problem of integrating sensing functionalities in bistatic asynchronous communication systems is addressed in this paper. For that, we designed and prototyped JUMP, a JCS scheme that overcomes timing and frequency offsets due to clock asynchrony following two main pathways. First, it leverages the correlation in the channel propagation paths across different packets, second, it identifies a static reference path whose phase is used to remove the frequency offset from the reflections on sensing targets. An extensive CIR measurement campaign has been carried out targeting indoor human movement sensing, using a 60 GHz IEEE 802.11ay-based implementation of the system. Results show that JUMP is competitive with synchronous full-duplex and phase-locked solutions. It achieves a median worst-case multitarget tracking error of 14 cm and a human μD spectrum normalized reconstruction error of 0.07 in intermittent NLOS conditions.

REFERENCES

[1] J. A. Zhang et al., "Enabling joint communication and radar sensing in mobile networks—A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 306–345, 1st Quart., 2022.

[2] F. Liu et al., "Seventy years of radar and communications: The road from separation to integration," *IEEE Signal Process. Mag.*, vol. 40, no. 5, pp. 106–121, Jul. 2023.

[3] A. Zhang, L. Rahman, X. Huang, Y. J. Guo, S. Chen, and R. W. Heath, "Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing," *IEEE Veh. Technol. Mag.*, vol. 16, no. 2, pp. 20–30, Jun. 2021.

[4] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Netw.*, vol. 35, no. 5, pp. 158–167, Sep. 2021.

[5] J. A. Zhang, K. Wu, X. Huang, Y. J. Guo, D. Zhang, and R. W. Heath, "Integration of radar sensing into communications with asynchronous transceivers," *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 106–112, Nov. 2022.

[6] J. Pegoraro, J. O. Lacruz, M. Rossi, and J. Widmer, "SPARCS: A sparse recovery approach for integrated communication and human sensing in mmWave systems," in *Proc. 21st ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, May 2022, pp. 79–91.

[7] S. Lu et al., "Integrated sensing and communications: Recent advances and ten open challenges," 2023, *arXiv:2305.00179*.

[8] C. B. Barneto, S. D. Liyanaarachchi, M. Heino, T. Riihonen, and M. Valkama, "Full duplex radio/radar technology: The enabler for advanced joint communication and sensing," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 82–88, Feb. 2021.

[9] Y. Cui, F. Liu, C. Masouros, J. Xu, T. X. Han, and Y. C. Eldar, "Integrated sensing and communications: Background and applications," in *Integrated Sensing and Communications*. Singapore: Springer, 2023, pp. 3–21.

[10] S. Mazahir, S. Ahmed, and M.-S. Alouini, "A survey on joint communication-radar systems," *Frontiers Commun. Netw.*, vol. 1, Feb. 2021, Art. no. 619483.

[11] J. A. Zhang et al., "An overview of signal processing techniques for joint communication and radar sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1295–1315, Nov. 2021.

[12] J. Pegoraro, J. O. Lacruz, F. Meneghello, E. Bashirov, M. Rossi, and J. Widmer, "RAPID: Retrofitting IEEE 802.11ay access points for indoor human detection and sensing," *IEEE Trans. Mobile Comput.*, early access, 2023, doi: 10.1109/TMC.2023.3291882.

[13] X. Li, Y. Cui, J. A. Zhang, F. Liu, D. Zhang, and L. Hanzo, "Integrated human activity sensing and communications," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 90–96, May 2023.

[14] F. Meneghello, D. Garlisi, N. D. Fabbro, I. Tinnirello, and M. Rossi, "SHARP: Environment and person independent activity recognition with commodity IEEE 802.11 access points," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 6160–6175, Oct. 2023.

[15] Y. Chen, X. Su, Y. Hu, and B. Zeng, "Residual carrier frequency offset estimation and compensation for commodity WiFi," *IEEE Trans. Mobile Comput.*, vol. 19, no. 12, pp. 2891–2902, Dec. 2020.

[16] M. E. Rasekh, M. Abdelghany, U. Madhowz, and M. Rodwell, "Phase noise analysis for mmWave massive MIMO: A design framework for scaling via tiled architectures," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–6.

[17] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–26, Sep. 2019.

[18] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "MultiSense: Enabling multi-person respiration sensing with commodity WiFi," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–29, Sep. 2020.

[19] X. Li, J. A. Zhang, K. Wu, Y. Cui, and X. Jing, "CSI-ratio-based Doppler frequency estimation in integrated sensing and communications," *IEEE Sensors J.*, vol. 22, no. 21, pp. 20886–20895, Nov. 2022.

[20] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, "Uplink sensing in perceptive mobile networks with asynchronous transceivers," *IEEE Trans. Signal Process.*, vol. 69, pp. 1287–1300, 2021.

[21] M. Heino, C. B. Barneto, T. Riihonen, and M. Valkama, "Design of phased array architectures for full-duplex joint communications and sensing," in *Proc. 15th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2021, pp. 1–5.

[22] M. Malanowski and K. Kulpa, "Two methods for target localization in multistatic passive radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 572–580, Jan. 2012.

[23] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity Wi-Fi," *IEEE Trans. Mobile Comput.*, vol. 18, no. 6, pp. 1342–1355, Jun. 2019.

[24] H. Zhu, Y. Zhuo, Q. Liu, and S. Chang, "ΠI-splicer: Perceiving accurate CSI phases with commodity WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 17, no. 9, pp. 2155–2165, Sep. 2018.

[25] X. Chen, Z. Feng, J. A. Zhang, X. Yuan, and P. Zhang, "Kalman filter-based sensing in communication systems with clock asynchronism," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 403–417, Jan. 2024.

[26] H. Li, X. He, X. Chen, Y. Fang, and Q. Fang, "Wi-Motion: A robust human activity recognition using WiFi signals," *IEEE Access*, vol. 7, pp. 153287–153299, 2019.

[27] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "WiFi CSI based passive human activity recognition using attention based BLSTM," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2714–2724, Nov. 2019.

[28] B. Korany, H. Cai, and Y. Mostofi, "Multiple people identification through walls using off-the-shelf WiFi," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6963–6974, Apr. 2020.

[29] L. Pucci, E. Matricardi, E. Paolini, W. Xu, and A. Giorgetti, "Performance analysis of a bistatic joint sensing and communication system," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 73–78.

[30] O. Kanhere, S. Goyal, M. Beluri, and T. S. Rappaport, "Target localization using bistatic and multistatic radar with 5G NR waveform," in *Proc. IEEE 93rd Veh. Technol. Conf.*, Apr. 2021, pp. 1–7.

[31] Z. Gao et al., "Integrated sensing and communication with mmWave massive MIMO: A compressed sampling perspective," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1745–1762, Mar. 2023.

[32] D. Garcia, J. O. Lacruz, P. Jiménez Mateo, and J. Widmer, "POLAR: Passive object localization with IEEE 802.11ad using phased antenna arrays," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Jul. 2020, pp. 1838–1847.

[33] H. Kuschel, D. Cristallini, and K. E. Olsen, "Tutorial: Passive radar tutorial," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 2, pp. 2–19, Feb. 2019.

[34] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of Modern Radar*. Raleigh, NC, USA: Scitech, 2010.

[35] P. Samczyński et al., "5G network-based passive radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5108209.

[36] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, "IEEE 802.11ad-based radar: An approach to joint vehicular communication-radar system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, Apr. 2018.

[37] W.-C. Liu, T.-C. Wei, Y.-S. Huang, C.-D. Chan, and S.-J. Jou, "All-digital synchronization for SC/OFDM mode of IEEE 802.15.3c and IEEE 802.11ad," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 2, pp. 545–553, Feb. 2015.

[38] *IEEE Draft 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: Enhanced Throughput for Operation in License-Exempt Bands Above 45 GHz*, IEEE Standard P802.11ay/D3.0, 2019.

[39] A. Shastri et al., "A review of millimeter wave device-based localization and device-free sensing technologies and applications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1708–1749, 3rd Quart., 2022.

[40] J. O. Lacruz, D. Garcia, P. J. Mateo, J. Palacios, and J. Widmer, "mmFLEX: An open platform for millimeter-wave mobile full-bandwidth experimentation," in *Proc. 18th Int. Conf. Mobile Syst., Appl., Services*, New York, NY, USA, 2020, pp. 1–13.

[41] M. I. Ribeiro, "Kalman and extended Kalman filters: Concept, derivation and properties," *Inst. Syst. Robot.*, vol. 43, no. 46, pp. 3736–3741, 2004.

[42] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Syst. Mag.*, vol. 29, no. 6, pp. 82–100, Dec. 2009.

[43] R. J. Fitzgerald, "Development of practical PDA logic for multitarget tracking by microprocessor," in *Proc. Amer. Control Conf.*, Seattle, WA, USA, Jun. 1986, pp. 889–898.

[44] T. Wagner, R. Feger, and A. Stelzer, "Radar signal processing for jointly estimating tracks and micro-Doppler signatures," *IEEE Access*, vol. 5, pp. 1220–1238, 2017.

[45] H. H. Ku, "Notes on the use of propagation of error formulas," *J. Res. Nat. Bureau Standards*, vol. 70, no. 4, pp. 263–273, 1966.

[46] J. O. Lacruz, R. Ruiz, and J. Widmer, "A real-time experimentation platform for sub-6 GHz and millimeter-wave MIMO systems," in *Proc. ACM MobiSys*, 2021, pp. 427–439.

[47] Sivers Semiconductors. (2022). *Evaluation Kits*. [Online]. Available: https://www.sivers-semiconductors.com/sivers-wireless/evaluation-kits//

[48] C.-Y. Liu et al., "Dual-mode all-digital baseband receiver with a feed-forward and shared-memory architecture for dual-standard over 60 GHz NLOS channel," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 3, pp. 608–618, Mar. 2017.
[49] C. Farrow, "A continuously variable digital delay element," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 3, Jun. 1988, pp. 2641–2645.

Jacopo Pegoraro (Member, IEEE) received the Ph.D. degree in information engineering from the University of Padova, Padua, Italy, in 2023. He was a Visiting Research Scholar with New York University Tandon School of Engineering, Brooklyn, NY, USA, in 2022, working on joint communication and sensing. He is currently a Post-Doctoral Researcher with the SIGNET Research Group, Department of Information Engineering, University of Padova. His research interests include signal processing, sensor fusion, and machine learning with applications to mmWave sensing.

Jesus O. Lacruz received the bachelor's degree in electrical engineering from Universidad de Los Andes, Venezuela, in 2009, and the Ph.D. degree in electronic engineering from Universidad Politècnica de València, Spain, in 2016. Since 2017, he has been a Research Engineer with the IMDEA Networks Institute, Spain. His research interests include the design and implementation of fast signal-processing algorithms for digital communication systems in FPGA devices.

Tommy Azzino received the B.Sc. degree in information engineering and the M.Sc. degree in telecommunications engineering from the University of Padova, Italy, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering with New York University Tandon School of Engineering, Brooklyn, NY, USA, under the supervision of Prof. S. Rangan. He was a Visiting Researcher with the Department of Commerce, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, in 2018, and Nokia Bell Labs France, Paris, France, in 2019. He took part in the 2018 Edition of the Seeds for the Future Project promoted by Huawei Technologies, Shenzhen, China, in 2018. His research interests include 5G mobile cloud/edge computing, 5G network simulation and prototyping, MAC and network-layer resource allocation, wireless communications, and machine learning for wireless communications.

Marco Mezzavilla (Senior Member, IEEE) received the B.A.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Padova, Italy. He was a Visiting Researcher with NEC Network Laboratories, Heidelberg, Germany, in 2009; the Centre Tecnológic Telecomunicacions Catalunya (CTTC), Barcelona, Spain, 2010; and Qualcomm Research, San Diego, CA, USA, in 2012. In 2014, he joined New York University (NYU) Tandon School of Engineering, USA, where he currently leads several research projects that focus on mmWave and sub-THz radio access technologies for 5G and beyond. He is also a Research Faculty with NYU Tandon School of Engineering and the Co-Founder of Pi-Radio, a spin-off company of NYU that develops mmWave/sub-THz software-defined radios (SDRs). His research interests include communications, cybersecurity, and robotics. He has served as the co-chair and a reviewer for several IEEE and ACM conferences and as a guest editor for various journals and magazines.

Michele Rossi (Senior Member, IEEE) is currently a Professor of wireless networks with the Department of Information Engineering (DEI), University of Padova, Italy, where he is the Head of the master's degree in ICT for internet and multimedia (http://mime.dei.unipd.it/). He also teaches human data analysis at the data science master's degree with the Department of Mathematics (DM), University of Padova (https://datascience.math.unipd.it/). Since 2017, he has been the Director of the DEI/IEEE Summer School of Information Engineering (http://ssie.dei.unipd.it/). Over the years, he has been involved in several EU projects on wireless sensing and the IoT. He has collaborated with major companies, such as Ericsson, DOCOMO, Samsung, and INTEL. His research is currently supported by the European Commission through the H2020 projects MINTS on "mmWave networking and sensing" and GREENEDGE on "green edge computing for mobile networks" (project coordinator). His research interests include wireless sensing systems, green mobile networks, edge, and wearable computing. He was a recipient of seven best paper awards from the IEEE. He serves on the editorial boards of IEEE TRANSACTIONS ON MOBILE COMPUTING and IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.

Joerg Widmer (Fellow, IEEE) is currently a Research Professor and the Research Director of the IMDEA Networks Institute, Madrid, Spain. Before, he held positions at DOCOMO Euro-Labs, Munich, Germany; and EPFL, Switzerland. He was a Visiting Researcher with the International Computer Science Institute, Berkeley, CA, USA; University College London, U.K.; and TU Darmstadt, Germany. He has authored more than 150 conference papers and journal articles and three IETF RFCs, and holds 13 patents. His research interests include wireless networks, ranging from extremely high frequency millimeter-wave communication, and MAC layer design to mobile network architectures. He was awarded an ERC Consolidator Grant, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation, a Mercator Fellowship of the German Research Foundation, a Spanish Ramon y Cajal Grant, as well as eight best paper awards. He is a Distinguished Member of the ACM.

Sundeep Rangan (Fellow, IEEE) received the B.A.Sc. degree in electrical engineering from the University of Waterloo, Canada, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of California at Berkeley, Berkeley, CA, USA. He is currently a Professor in electrical and computer engineering with New York University (NYU) and the Associate Director of NYU Wireless Research Centre. He has held postdoctoral appointments with the University of Michigan, Ann Arbor, and Bell Laboratories. In 2000, he co-founded (with four others) Flarion Technologies, a spin-off of Bell Laboratories, that developed Flash OFDM, the first cellular OFDM data system and pre-cursor to 4G cellular systems, including LTE and WiMAX. In 2006, Flarion was acquired by Qualcomm Technologies. He was the Senior Director of Engineering with Qualcomm involved in OFDM infrastructure products. In 2010, he joined NYU Tandon (formerly NYU Polytechnic) where he is currently a Professor of electrical and computer engineering. He is the Associate Director of NYU WIRELESS, an industry-academic research center.

Open Access funding provided by 'Università degli Studi di Padova' within the CRUI-CARE Agreement