Abstract—Wireless-based human activity recognition has become an essential technology that enables contact-free human–machine and human–environment interactions. In this article, we consider contact-free multitarget tracking (MTT) based on available communication systems. A radar-like prototype is built upon a sub-6-GHz distributed massive multiple-input and multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) communication system. Specifically, the raw channel state information (CSI) is calibrated in the frequency and antenna domain before being used for tracking. Then, the targeted CSIs reflected or scattered from the moving pedestrians are extracted. To evade the complex association problem of distributed massive MIMO-based MTT, we propose to use a complex Bayesian compressive sensing (CBCS) algorithm to estimate the targets’ locations based on the extracted target-of-interest CSI signal directly. The estimated locations from CBCS are fed to a Gaussian mixture probability hypothesis density (GM-PHD) filter for tracking. A multipedestrian tracking experiment is conducted in a room with a size of 6.5 m × 10 m to evaluate the performance of the proposed algorithm. According to the experimental results, we achieve 75th and 95th percentile accuracy of 12.7 and 18.2 cm for single-person tracking and 28.9 and 45.7 cm for multiperson tracking, respectively. Furthermore, the proposed algorithm achieves tracking purposes in real time, which is promising for practical MTT use cases.

Index Terms—Channel state information (CSI), indoor localization, integrated sensing and communication (ISAC), massive multiple-input and multiple-output (MIMO), multitarget tracking (MTT), radar.

I. INTRODUCTION

HUMAN activity recognition (HAR) using wireless devices has attracted considerable attention due to the advances in Internet of Things technology and radio frequency (RF) hardware circuits in the past decade [1]. It enables a wide variety of potential applications, such as elderly care, behavior analysis, virtual reality, location-based services (LBSs), etc. Generally, RF-based HAR techniques can be classified into two categories, namely, wearable device-based and device-free (also known as, contact-free or passive) solutions. Specifically, contact-free human sensing has become a more active research area as there is no need for user-attached sensors, which provides convenience and feasibility to specific use cases, such as infectious patient monitoring. Moreover, RF-based techniques are more privacy-preserving compared to image-based solutions. To date, a batch of academic and industrial efforts has been paid to promote human sensing techniques based on different standards, including RF identification (RFID), Bluetooth low energy (BLE), ultrawideband (UWB), Wi-Fi, millimeter wave, etc.

Among them, many efforts have been devoted to Wi-Fi-based human sensing because of the benefit of ubiquitous deployment. Wang et al. [2] proposed a channel state information (CSI)-based human activity monitoring algorithm by quantifying how the speed of human motion is related to the CSI power variations. Xu et al. [3] proposed to exploit the temporal CSI series to characterize the dynamic indoor motion event and realized high accuracy for real-time monitoring. Wi-Fi-based CSI signals have also been widely used for the single-target contact-free tracking [4], [5], [6], where the person is located via the geometrical metrics, such as Angle of Arrival (AoA), Doppler, Time-of-Flight (ToF), etc. Rather than single-target tracking, Karanam et al. [7] extended the work to multiple-person localization leveraging the AoA and Doppler estimates of the magnitude of CSI. Then, a particle filter together with joint probabilistic data association (JPDA) was used for tracking. Venkatnarayan et al. [8] proposed to
add horizontal and vertical polarization to perform multi-person tracking and achieved higher accuracy than the case without polarization diversity. However, this has a more rigorous demand for transceiver deployment. Besides Wi-Fi, UWB and millimeter-wave radios have also been widely used for contact-free human sensing due to the high-ranging resolution, for instance, vital sign monitoring [9], [10], pedestrian tracking [11], [12], etc.

As the standardization of fifth-generation (5G) wireless communication gradually solidifies, people are envisioning what the next-generation network will be. One of the great ambitions is to enable the wireless network to “see” or “understand” the physical world via sensing the targets or surroundings. Toward this end, the so-called integrated sensing and communication (ISAC) [13], [14] has been proposed which integrates the sensing functionality into wireless communication networks allowing the reuse of allocated spectrum, hardware, and even signaling resources. Rather than redesigning the whole hardware architecture, one of the feasible and promising schemes is to reutilize the existing and pervasive RF signals, e.g., Wi-Fi, long-term evolution (LTE), and 5G, for the sensing objectives [15]. As a milestone of the 5G wireless communication, the massive multiple-input and multiple-output (MIMO) technique not only improves communication in terms of channel capacity and spectral efficiency but also has the potential for accurate LBS applications resulting from the high spatial resolution [16], [17].

Despite the increasing discussions about enabling sensing via communication systems or the communication/radar co-design, little work is done on the experiment-based massive MIMO for ISAC except for our preliminary works in [18] and [19] for single-target sensing. To this end, in this article, we prototype a sub-6-GHz distributed massive MIMO communication system and investigate its radar-like functionality for human sensing. The established massive MIMO system adopts a standard cellular signal bandwidth and the orthogonal frequency-division multiplexing (OFDM) waveform, which is compatible with the available Wi-Fi, LTE, and sub-6-GHz 5G communication systems. Inspired by the user-centric concept of cell-free (or distributed) massive MIMO [20], which aims to surround the users with a large number of base station (BS) antennas, we implement a distributed massive MIMO radar-like system with a large antenna array separated in an indoor environment. Based on the prototype, we conduct the contact-free multiple pedestrians tracking experiment and analysis. The major contributions of this article are as follows.

1) We demonstrate a radar-like distributed massive MIMO system using an OFDM communication testbed. A standard cellular signal bandwidth, i.e., 18 MHz, is adopted. According to the best of our knowledge, it is the first bandwidth-limited distributed massive MIMO prototype for contact-free human sensing.

2) Based on the established prototype, we investigate the multitarget tracking (MTT) problem and propose a contact-free pedestrian tracking framework based on CSI data directly. Specifically, a serial interference cancellation and reconstruction (SICAR) algorithm is proposed for the targeted CSI extraction. The complex Bayesian compressive sensing (CBCS) and Gaussian mixture probability hypothesis density (GM-PHD) filter are introduced for positioning and tracking purposes. The proposed algorithm effectively evades complex association problems by exploiting the ideas of space sparsity and random finite set (RFS).

3) The MTT performance is evaluated comprehensively concerning the impact of bandwidth, the number of antennas, antenna deployment, the number of targets, etc. Besides, within the proposed MTT framework, each algorithm block is assembled organically, which helps to realize (quasi-) real-time MTT.

4) Furthermore, to evaluate contact-free MTT, we propose an active tracking algorithm built upon the idea of the synthetic aperture. A centimeter-level (even millimeter-level) accuracy is guaranteed based on the proposed active tracking algorithm, which is feasible as the benchmark of the contact-free MTT. The idea of the proposed method can also be generalized to other device-based tracking technologies, such as phase-based RFID tracking, phase-based BLE tracking, CSI-based Wi-Fi tracking, etc.

The remainder of this article is organized as follows. Section II introduces the established massive MIMO prototype, experimental campaign, and system overview. The practical challenges for MTT are also discussed in this section. Section III presents the CSI calibration and benchmark design for contact-free tracking. Section IV includes the Target-of-Interest (ToI) extraction, instantaneous locations estimation, and MTT algorithms design. The experimental results and analysis are presented in Section V. Section VI concludes this article.

II. Experiment and System Overview

A. Prototype and Experiment

We establish a radar-like prototype based on the massive MIMO communication testbed at KU Leuven ESAT-WAVECORE. The Massive MIMO system is built on time division duplex (TDD)-based OFDM signaling. In the system, the massive number of BS antennas are deployed as eight distributed uniform linear arrays (ULAs), each composing $1 \times 8$ patch antennas, as shown in Fig. 1. For the investigation in this
The measured CSI of a single transmission can be represented at the top of the participant's head, as shown in Fig.1. Therefore, active tracking to determine a “ground truth.” To avoid the burden of carrying an additional transmitter (i.e., UE1, UE2, or UE3) for monitoring movements. To handle this problem, we asked each participant not exactly follow the template tracks due to distinct individual characteristics. However, it should be noted that the participants can move. However, it should be noted that the participants cannot exactly follow the template tracks due to distinct individual movements. To handle this problem, we asked each participant to carry an additional transmitter (i.e., UE1, UE2, or UE3) for active tracking to determine a “ground truth.” To avoid the body along the time evolving was regarded as an assignment problem and solved via binary optimization and the Hungarian algorithm. As a single person walking in the scenario is a priori, the MPCs track with the highest power was selected. So for a single moving target, associating along the time domain is required for the tracking purpose. We name it evolving association. According to the best of our knowledge, there are also a few works on CSI-based multiperson tracking [7], [8], [23]. Besides the location- or velocity-related parameters extraction as in single-person tracking, we need to assign the estimated MPCs to the corresponding targets as we have no idea of which targets that MPCs are generated from. So there is an additional target association for MTT.

However, this is not the end of the challenges for our distributed massive MIMO-based MTT. For the contact-free tracking research mentioned above, most of them just used a single transceiver link or two links for accuracy enhancement. But in the case of distributed massive MIMO, there are a large number of propagation links. The MPCs of the multiple targets along the large antenna array are also varied for a given time slot, which means we also need to assign the massive and cluttered MPCs of all the links to the corresponding target, namely, link association. Even though many algorithms can be used to solve the assignment problem, such as JPDA [7], [23], multiple hypothesis tracking (MHT) [24], binary optimization [5], etc., joint association in these three domains is challenging and complex in computation, which is not feasible for many practical applications considering the computing resource and run-time issue.

Besides complex association, there are also like varied reflected power from the multiple targets, false alarms, and missed detection, challenging the distributed massive MIMO-based MTT problem. As presented in Section II-A, we have deployed the distributed massive MIMO as eight short subantenna arrays (ULA). As a planar wavefront assumption is satisfied for each short ULA. We adopt the consecutive CSI measurements within a given time window to estimate the Doppler. We set the time window as 0.1 s in this work, which is also widely adopted for person tracking [5], [25]. Then, a 2-D multiple signal classification (MUSIC) algorithm together with spatial–temporal smoothing [7] is implemented to generate the angular-Doppler spectrum. Fig. 2 shows the angular-Doppler profiles of three consecutive time windows in the case of three pedestrians. Three phenomena can be observed from Fig. 2, which bring in even greater challenges to association problems of distributed massive MIMO-based MTT.

### TABLE I

| Distributed Massive MIMO-OFDM Communication System Parameter Settings |
|---------------------------------------------------------------|
| **System Parameters** | **Value**       |
| Transmitter power        | 15 dBm          |
| Sampling rate            | 100 Hz          |
| Center frequency         | 2.61 GHz         |
| Wavelength               | 11.49 cm         |
| Bandwidth                | 18 MHz           |
| # of Sub-carrier         | 100              |
| # of antenna             | 64               |
| Modulation               | QPSK             |
| Patch antenna size       | 7 cm x 7 cm      |
| BS antenna altitude      | 1.205 m          |
| UE antenna altitude      | 0.8 m            |
| Size of experimental space | 6.5 m x 10 m    |

In recent years, there are increasing work on CSI-based contact-free tracking. However, most of them are focusing on single-target tracking (e.g., one walking person, a waving limb or finger) [4], [5], [6], [22]. The common idea behind these pioneer works is to extract the location- or velocity-related parameters from the CSI. For example, Qian et al. [5] adopted a space-alternating generalized expectation-maximization (SAGE) algorithm to estimate the multipath components (MPCs), i.e., amplitude, AoA, Doppler, and ToF. Then, the association of MPCs reflected by the body along the time evolving was regarded as an assignment problem and solved via binary optimization and the Hungarian algorithm. As a single person walking in the scenario is a priori, the MPCs track with the highest power was selected. So for a single moving target, associating along the time domain is required for the tracking purpose. We name it evolving association. According to the best of our knowledge, there are also a few works on CSI-based multiperson tracking [7], [8], [23]. Besides the location- or velocity-related parameters extraction as in single-person tracking, we need to assign the estimated MPCs to the corresponding targets as we have no idea of which targets that MPCs are generated from. So there is an additional target association for MTT.

For the contact-free MTT experiment, the communication links among the transmitter (i.e., UE0) and the distributed BS antennas form a sensing zone, where multiple moving pedestrians (up to three in this article) can be localized and tracked via characterizing the reflected/scattered signals on the human body. We have predefined the track templates for the participants to move. However, it should be noted that the participants cannot exactly follow the template tracks due to distinct individual movements. To handle this problem, we asked each participant to carry an additional transmitter (i.e., UE1, UE2, or UE3) for active tracking to determine a “ground truth.” To avoid the body along the time evolving was regarded as an assignment problem and solved via binary optimization and the Hungarian algorithm. As a single person walking in the scenario is a priori, the MPCs track with the highest power was selected. So for a single moving target, associating along the time domain is required for the tracking purpose. We name it evolving association. According to the best of our knowledge, there are also a few works on CSI-based multiperson tracking [7], [8], [23]. Besides the location- or velocity-related parameters extraction as in single-person tracking, we need to assign the estimated MPCs to the corresponding targets as we have no idea of which targets that MPCs are generated from. So there is an additional target association for MTT.

For the contact-free MTT experiment, the communication links among the transmitter (i.e., UE0) and the distributed BS antennas form a sensing zone, where multiple moving pedestrians (up to three in this article) can be localized and tracked via characterizing the reflected/scattered signals on the human body. We have predefined the track templates for the participants to move. However, it should be noted that the participants cannot exactly follow the template tracks due to distinct individual movements. To handle this problem, we asked each participant to carry an additional transmitter (i.e., UE1, UE2, or UE3) for active tracking to determine a “ground truth.” To avoid the body along the time evolving was regarded as an assignment problem and solved via binary optimization and the Hungarian algorithm. As a single person walking in the scenario is a priori, the MPCs track with the highest power was selected. So for a single moving target, associating along the time domain is required for the tracking purpose. We name it evolving association. According to the best of our knowledge, there are also a few works on CSI-based multiperson tracking [7], [8], [23]. Besides the location- or velocity-related parameters extraction as in single-person tracking, we need to assign the estimated MPCs to the corresponding targets as we have no idea of which targets that MPCs are generated from. So there is an additional target association for MTT.

For the contact-free MTT experiment, the communication links among the transmitter (i.e., UE0) and the distributed BS antennas form a sensing zone, where multiple moving pedestrians (up to three in this article) can be localized and tracked via characterizing the reflected/scattered signals on the human body. We have predefined the track templates for the participants to move. However, it should be noted that the participants cannot exactly follow the template tracks due to distinct individual movements. To handle this problem, we asked each participant to carry an additional transmitter (i.e., UE1, UE2, or UE3) for active tracking to determine a “ground truth.” To avoid the body along the time evolving was regarded as an assignment problem and solved via binary optimization and the Hungarian algorithm. As a single person walking in the scenario is a priori, the MPCs track with the highest power was selected. So for a single moving target, associating along the time domain is required for the tracking purpose. We name it evolving association. According to the best of our knowledge, there are also a few works on CSI-based multiperson tracking [7], [8], [23]. Besides the location- or velocity-related parameters extraction as in single-person tracking, we need to assign the estimated MPCs to the corresponding targets as we have no idea of which targets that MPCs are generated from. So there is an additional target association for MTT.
1) The reflected/scattered power from each target may be greatly varied. The major reasons are twofold: First, the different distances from the pedestrian to the transceivers. Second, the different radar cross-sections (RCSs) due to individual differences, body orientations, torso or limbs scattering, etc.

2) False alarms or ghost targets are prevalent. The possible reasons are threefold. The first one is the residual components due to not perfect background mitigation. Second, the walking pedestrians may partly block the reflections/scattering from the static objects causing the signal from the static objects to fluctuate. Third, the second- or higher-order signal bounces within human bodies or between the bodies and the surroundings.

3) Missed detection frequently happens. At some specific time slots or subantenna arrays, one or more targets cannot be detected, or several targets are merged from the observations. But it is possible to detect/separate them at other time slots or by other subantenna arrays.

C. Radar-Like System Overview

Different from conventional contact-free tracking solutions that exploit the estimated angle, distance, and Doppler estimates, in this article, we propose to achieve contact-free MTT based on CSI data directly. In the proposed framework, the CSI measured by each antenna element of the massive MIMO system is regarded as a whole for positioning and tracking, which is similar to the idea of a synthetic aperture. This kind of processing has two merits. First, it can decrease the probability of missed detection. Because if the target is not detected by part of the antennas, it is likely the target can be detected by other antennas due to the distributed deployment. Second, we treat the massive antenna array as a whole and estimate the locations of the targets, which avoids the link association problem in the MPCs-based solutions mentioned above. Given the fact that the targets’ states are sparse relative to the whole targeting space. So the locations can be estimated from the CSI matrix of the massive antennas via a compressive sensing algorithm. In this section, we summarize the whole procedures of the distributed massive MIMO prototype, including the CSI calibration, contact-free MTT algorithm design, and the benchmark for evaluation.

The architecture of the established radar-like system is given in Fig. 3. For a given scenario, several UEs are deployed, of which one UE (i.e., UE0) acts as a transmitter and is used for contact-free tracking, whereas the other UEs (UE1, UE2, and UE3) are used for benchmarking which are not necessary for practical MTT tasks. It is noteworthy that for a real-world deployment, UE0 (transmitter) can be part of the infrastructure (like BS) or a cooperative user. For the latter case, the location of the user should be obtained first via active localization.

In the established prototype, the BS receives the signals transmitted by all UEs, conducts the channel estimation, and obtains the CSI data. The procedures for achieving the contact-free MTT are as follows.

1) CSI Calibration: Before utilizing CSI for contact-free pedestrian tracking and benchmarking, CSI calibration is required along frequency and antenna domains [16], which will be introduced in Section III-A. In the CSI calibration, we can obtain a set of calibration parameters that are fixed for a given system setting (frequency, bandwidth, antenna, etc.). So we can use these parameters for the future CSI calibration directly without the need for redundant processing as in Section III-A, namely, one-fits-all CSI calibration.
sets can be effectively mitigated via a nonlinear regression radar experiment. According to [16], the mentioned phase off-poses. To this end, we need to calibrate the CSI before the makes the CSI intractable for localization and tracking pur-

**2) ToI Extraction:** First, we conduct background removal by subtracting the moving average CSI [18], [23], [26]. Then, we propose a SICAR algorithm to extract the CSI related to the multiple targets, namely, ToI, which will be discussed in Section IV-A.

**3) Instantaneous Locations Estimation:** We regard the CSI of the massive antenna array as a whole, and estimate the locations of the multiple targets directly. In Section IV-B, we will introduce the CBSC algorithm to estimate the instantaneous locations based on the extracted ToI. This can be achieved due to the fact that the locations of the targets are sparse relative to the entire monitoring area.

**4) MTT:** The estimated positions via the CBSC algorithm may be cluttered. In Section IV-C, the GM-PHD filter is adopted to track multiple targets simultaneously without the exact prior knowledge of the number of targets. GM-PHD filter is iterated within a Kalman recursion which ensures a real-time MTT.

**5) Benchmark and Evaluation:** In Section III-B, we propose to use active tracking results as the benchmark of the contact-free MTT. Stemming from the concept of synthetic aperture, the CSI of all distributed antennas is synthesized to generate the location likelihood. A particle filter is introduced to track the varying likelihood when the target is moving. The proposed tracking algorithm can guarantee centimeter-level, even millimeter-level accuracy, which is feasible for benchmarking. Furthermore, the body size compensation is considered for the accuracy evaluation, which will be explained in Section V-A.

### III. PRELIMINARY PREPARATIONS

#### A. CSI Calibration

Due to the imperfect synchronization and hardware signal processing, the measured raw CSI suffers from various frequency-dependent phase errors, including the sampling frequency offset (SFO) \( \phi_{SFO} \), symbol timing offsets (STOs) \( \phi_{STO} \), in-phase quadrature-phase (I-Q) imbalance \( \phi_{IQ} \), and the constant carrier phase offsets (CPOs) \( \phi_{CPO} \) [16], [27], which makes the CSI intractable for localization and tracking purposes. To this end, we need to calibrate the CSI before the radar experiment. According to [16], the mentioned phase offsets can be effectively mitigated via a nonlinear regression along the subcarriers (\( n_f = 1, \ldots, N_f = 100 \)), given by

\[
\arg \min_{\phi_{\text{ant}}} \sum_{n_f}(\Delta \Phi_{n_f} - \phi_{n_f}^{(n)})^2
\]  

where \( \Delta \Phi = [\varepsilon_g, \varepsilon_i, \phi_{SFO/STO}, \phi_{CPO}] \). \( \Delta \Phi \) is the residual phase after removing the calculated LoS based on ground truth. The phase error due to the I-Q imbalance is denoted as [27]

\[
\phi_{n_f}^{(n)} = \arctan \left( \frac{\sin(\eta_f \varepsilon_i + \varepsilon_p)}{\cos(\eta_f \varepsilon_i)} \right) 
\]  

where \( \varepsilon_g, \varepsilon_p, \) and \( \varepsilon_i \) represent the gain mismatch, phase mismatch, and unknown time offset, respectively. Moreover, due to the heterogeneity of hardware, there are also constant phase offsets among each element of the large antenna array. The phase offsets at \( n_f \)-th antenna element \( \phi_{\text{ant}}^{(n)} \) can be estimated in complex domain via

\[
\arg \min_{\phi_{\text{ant}}^{(n)}} \sum_{n_f} \sum_{n_f} (e^{j \Delta \Psi_{n_f}} - e^{j \phi_{\text{ant}}^{(n)}})^2
\]  

where \( \Delta \Psi \) is the residual phase after the frequency-domain calibration in (2) and \( j = \sqrt{-1} \). In the calibration, the UE is placed at different reference points (RPs), and \( n_f \) is the index of the RPs. We have used a computerized numerical control X-Y table (mm-level accuracy) to provide the ground truth of the RPs’ locations. After the processing in (2) and (4), we can obtain the calibration parameters \( \Phi \) and \( \phi_{\text{ant}}^{(n)} \). Note that we only need to estimate these parameters once, which can be used for future CSI calibration directly without the redundant processing again. Fig. 4 shows an example of the phases of the CSI before and after calibration. The calibrated phases in Fig. 4(c) have an excellent matching compared with the real phases in Fig. 4(a). Even though there are several phase jumps around \(-\pi \) and \( \pi \) rad, the errors in the complex angle domain are small. The mean phase error is only about 0.23 rad after the CSI calibration.

#### B. Benchmark: Active Tracking

As mentioned in Section II-C, we adopt the localization results of active tracking as the benchmark (“ground truth”) for contact-free tracking. For indoor massive MIMO localization, the planar wave-front assumption is not guaranteed for the whole BS antenna array. Therefore, the super-resolution
MPCs estimation algorithms, e.g., MUSIC, SAGE algorithm, cannot be implemented directly. In this section, we propose a tracking method exploiting the CSI signal directly. Define the ideal phase from the UE to the \( n \)th antenna as \( \phi_{n_t} = [2\pi / \lambda]\|P_{n_t} - P_{UE}\| \), where \( \lambda \) is the wavelength. \( P_{UE} \) and \( P_{n} \) are the coordinates of the UE and the \( n \)th antenna, respectively. Define the calibrated CSI of \( n \)th antenna as \( \tilde{H}_n \), so the corresponding phase is \( \tilde{\phi}_{n_t} = \angle \tilde{H}_n \) and amplitude \( \tilde{\alpha}_{n_t} = |\tilde{H}_n| \). Stemming from the concept of synthetic aperture radar (SAR), the location of the UE can be estimated via maximizing the matching function \( C(P_{UE}) \)

\[
\arg \max_{P_{UE}} C(P_{UE}) = \arg \max_{P_{UE}} \sum_{n_t=1}^{N_t} \tilde{\alpha}_{n_t} e^{j(\tilde{\phi}_{n_t} - \phi_{n_t})}
\]

where \( N_t = 64 \) is the number of antenna element at BS. For pedestrian tracking, the 2-D location is considered. So for the benchmarking, we can measure the height of the participant in advance and set it as a priori. As \( C(P_{UE}) \) in (5) is nonconvex, a common solution is conducting the grid search within the 2-D targeted area. However, the global grid search is time-consuming, especially for a large targeted area or a small grid size. To handle this problem, we propose to estimate the target’s location via tracking the changes of \( C(P_{UE}) \) based on a particle filter algorithm. For each location update, the candidate region is constrained by the pedestrian motion model, given by \( \tilde{x}_{UE}^{t+1} = Fx_{UE} + G\tilde{v}_{UE} \), where

\[
\begin{align*}
\tilde{x}_{UE} &= \begin{bmatrix} \tilde{P}_{UE}^x \ \tilde{P}_{UE}^y \end{bmatrix} \\
F &= \begin{bmatrix} I_2 & \Delta t \cdot I_2 \end{bmatrix} \\
G &= \begin{bmatrix} \Delta t \cdot I_{2 \times 1} \\
\Delta t \cdot I_{1 \times 2} \end{bmatrix}
\end{align*}
\]

where \( \tilde{v}_{UE} \) is the 2-D velocity of the moving target. \( \Delta t \) is the time difference between two consecutive timestamps. \( n_t \) is the random Gaussian velocity errors with the standard deviation \( \sigma_v = 0.3 \text{ m/s} \) [12]. In this way, we can avoid the global search for each update and speed up the computation. The \( i \)th (\( i = 1, \ldots, K \)) updated particle is weighted via \( \tilde{P}_w^{(i)} \), defined as [12]

\[
\ln \tilde{P}_w^{(i)} = \frac{1}{2\sigma^2} \left( \min_{1 \leq i \leq K} \left| \tilde{P}_w^{(i)} - P_{w}^{(i)} \right| \right)
\]

where \( P_w^{(i)} = (C^{(i)}(P_{UE}))^2 \), \( K \) is the number of particles, \( \sigma \) the standard deviation of \( C(P_{UE}) \), and \( C_{max} = 1 \) because we normalize \( C(P_{UE}) \) to \((0, 1)\). The normalized \( C(P_{UE}) \) can be regarded as the target’s location likelihood.

To validate the feasibility of active tracking as the benchmark, we evaluate the localization accuracy of active tracking using the collected data based on the computerized numerical control X-Y table and our open-access distributed massive MIMO data set [28]. The open-access data set was also collected via a similarly distributed antenna topology as shown in Fig. 1. In Fig. 5(a), the cumulative distribution functions (CDFs) of the tracking errors in these two data sets are presented. Fig. 5(b) shows an example of active tracking results, in which we formulate the tracking trajectory as a “cat shape.” We can see that the distributed massive MIMO system can guarantee millimeter-level mean tracking accuracy whereas the 90th percentile error is less than 0.78 cm. The tracking results indicate the feasibility of the proposed active tracking as the benchmark for the radar-like (contact-free) scheme. Regarding the active pedestrian tracking in the experiment, human walking is not perfectly controlled as the computerized numerical control table. But it is still reasonable to presume that the accuracy of active pedestrian tracking (e.g., centimeter-level is possible) is much higher than contact-free tracking in the case of distributed massive MIMO.

IV. CONTACT-FREE MULTITARGET TRACKING

A. Target-of-Interest

In case an unknown number of pedestrians are walking in an indoor scenario, the CSI at \( n \)th antenna, time \( t \), and frequency \( f \) (in equivalent baseband) can be given as

\[
H_n(t, f) = \tilde{H}_n(f) + \sum_{j \in D} \alpha_{n_t}^{(j)}(t, f) e^{-j2\pi \left( \frac{\nu_{n_t}^{(j)}(t)}{c} + w_{nj}^{(j)} \right)}
\]

where \( \tilde{H}_n(f) \) is the sum of all static signals. \( D \) denotes the set of dynamic targets. \( \alpha_{n_t}^{(j)}, \nu_{n_t}^{(j)}, \) and \( v_{nj}^{(j)} \) denote the complex gain, the link length, and Doppler shifts at the \( n \)th antenna element, respectively. In the framework of contact-free tracking, we are only interested in the MPCs reflected or scattered from the pedestrians. To remove the static clutters, a moving average method [18], [23], [26] is adopted to obtain the mean CSI and then subtracted from the instantaneous CSI. So the ToI can be
given as (note that we omit the indices of \( t \) and \( f \) for ease of expression)

\[
H_{{n_t}}^{\text{Tol}} = \sum_{j \in D} x_{{n_t}}^{(j)} e^{-j \frac{2\pi}{\lambda} (d_{{n_t}}^{(j)} + \Delta d_{{n_t}}^{(j)})} + n_{{n_t}}
\]

(9)

where \( n_{{n_t}} \) is the residual components after the background removal. \( \Delta d_{{n_t}}^{(j)} \) is the link length changes due to Doppler shift, given by [29]

\[
\Delta d_{{n_t}}^{(j)} = -\left( \frac{P_{{n_t}}^T - P_{{n_r}}}{\|P_{{n_t}}^T - P_{{n_r}}\|} + \frac{P_{{n_t}}^T - P_{\text{UE}0}}{\|P_{{n_t}}^T - P_{\text{UE}0}\|} \right)^T v_{{n_t}}^{(j)} \tau
\]

(10)

where \( P_{{n_t}}^T \) and \( v_{{n_t}}^{(j)} \) represent the location and velocity of the \( j \)-th target. For the pedestrians tracking, the 2-D coordinates and velocities, namely, \( P_{{n_t}} = [\chi \ \gamma \ \tau]^T \) and \( v_{{n_t}} = [\nu \ \omega \ \tau]^T \), are considered. Given the fact that the number of targets within the monitoring area is limited, it means that the targets’ states \( x_{{n_t}} = [P_{{n_t}}^T \ \tau]^T \) are sparsely distributed in the entire 4-D solution space \( X \), which reminds us of compressive sensing for location estimation.

Define the ToIs at all of the distributed antennas as \( H_{{n_t}}^{\text{Tol}} = [H_{{n_1}}^{\text{Tol}}; \ldots; H_{{n_T}}^{\text{Tol}}] \subseteq C^{N_\text{T} \times 1} \), so the estimation of targets’ states can be done by solving

\[
H_{{n_t}}^{\text{Tol}} = \Phi u + n
\]

(11)

where \( \Phi \in C^{N_\text{T} \times M} \) is the dictionary matrix whose \((n_{{n_t}}, m)\)th element is given as \( \Phi_{{n_{{n_t}}, m}} = e^{-j (2\pi/\lambda)(d_{{n_t}}^{(m)} + \Delta d_{{n_t}}^{(m)})} \). \( M = M_\text{L} \times M_\text{V} \times M_\text{V} \times M_\text{V} \) is the number of candidate states, where \( M_\text{L} \), \( M_\text{V} \), and \( M_\text{V} \) are the number of the partitioned coordinates and velocities in \( X \), respectively. \( u \in C^{M \times 1} \) indicating the sparse solutions of targets’ states, and \( n \in C^{N_\text{T} \times 1} \), \( n_{{n_t}} = n_{{n_t}} \). However, the second dimension of \( \Phi \) (i.e., \( M \)) is extremely large, resulting in a long computation time. For example, suppose that the size of the monitoring area is \( 2 \times 2 \) m, the walking speed ranges between \( \pm 1.5 \) m/s, and the grid sizes of coordinates and velocities are \( 0.05 \) m and \( 0.05 \) m/s, respectively, then the number of candidates \( M = 41^2 \times 61^2 = 6255001 > 10^6 \), which will cause the computing complexity curse and is unacceptable for practical usage. However, in the framework of pedestrian tracking, we are more concerned about the coordinates and the velocities that can be obtained via a tracking filter after obtaining the locations. Furthermore, in the TDD-LTE or 5G new radio network, the downlink or uplink subframes are not continuously transmitted due to transmission periodicity, which means the valid CSI sampling rate cannot be very high in practice, e.g., 100 Hz [15] (the setting in our experiment). In this case, the number of CSI samples \( N_\text{f} \) for Doppler estimation is limited, thus the Doppler estimation accuracy is not guaranteed. By these, we can rule out the Doppler impact and only retain the location-related part in (9). Then, the number of candidate states can decrease to \( M = M_\text{L} \times M_\text{V} = 41^2 \times 1681 \approx 6255001 \).

To handle this, we propose a ToI purification method based on the SICAR, as shown in Algorithm 1. Define the ToI of the \( n_{{n_t}} \)-th antenna element at the \( n_{{n_s}} \)-th subcarrier and \( n_{{n_t}} \)-th time slot as \( [H_{{n_t}}^{\text{Tol}} \in C^{N_\text{T} \times N_\text{f}}]_{{n_{{n_s}}, n_{{n_t}}}} \). Define the \( N_\text{D} \)-point phase-shifted discrete Fourier transform (DFT) matrix for Doppler as

\[
\begin{bmatrix}
D & [D_{{n_{{n_s}}, n_{{n_t}}}} (m, n_{{n_s}})]
\end{bmatrix} = \frac{1}{\sqrt{N_\text{D}}} \begin{bmatrix}
\text{e}^{-j \frac{2\pi}{\lambda} (n_{{n_s}} - N_\text{D}/2)}
\end{bmatrix}
\]

and the \( N_\text{T} \)-point phase-shifted DFT matrix for ranging as

\[
[T \subseteq C^{N_\text{T} \times N_\text{T}}]_{{n_{{n_s}}, n_{{n_t}}}} = \frac{1}{\sqrt{N_\text{T}}} \begin{bmatrix}
\text{e}^{-j \frac{2\pi}{\lambda} (n_{{n_t}} - N_\text{T}/2)}
\end{bmatrix}
\]

(12)

The idea of the proposed ToI purification method is built upon removing the Doppler of each MPC via the FT sequentially, which is similar to the expectation–maximization algorithm for channel parameters estimation in [30]. The iteration stops when the power of the current MPC is \( P_\text{th} \) dB (e.g., 30 dB) lower than the maximal MPC power. In this work, \( N_\text{D} \) and \( N_\text{T} \) are set as 128 and 100, respectively. Unlike many subspace methods, the 2-D DFT in Algorithm 1 does not require the exact number of targets and can be implemented efficiently. According to the experimental observations, the average number of iterations using the proposed SICAR algorithm is 4.6 with rare cases having more than 8 iterations. The computing time of ToI extraction will be investigated in Section V-C.

**Algorithm 1: SICAR-Based No-Doppler ToI Extraction**

**Input:** \( H_{{n_t}}^{\text{Tol}}, D, T, P_\text{th} 

**Output:** \( H_{{n_t}}^{\text{Tol}} 

1. Initialize: \( H_{{n_t}}^{\text{Tol}} = 0; \ X = H_{{n_t}}^{\text{Tol}}; \ l = 0; \)
2. while \( \text{TRUE} \) do
3. \( \ G = T^H X D; \ // \text{Range-Doppler profile} \)
4. \( (\hat{n}_f, \hat{n}_D) = \arg \max_{n_{{n_s}}, n_{{n_t}}} E[G \odot G^*]; \)
5. \( \phi_{{n_t}}^{\text{Tol}} = \angle G(\hat{n}_f, \hat{n}_D); \ // \text{Phase estimation} \)
6. \( /\Phi_1 = \phi_{{n_t}}^{\text{Tol}}; \ // \Phi_1 \)
7. \( /\Phi_2 = 2\pi(1 - \hat{n}_f); \)
8. \( /\Phi_3 = 2\pi(1 - \hat{n}_D); \)
9. \( /\Phi_4 = 2\pi(1 - \hat{n}_f - \hat{n}_D); \)
10. \( /\Phi_5 = 2\pi(1 - \hat{n}_f - \hat{n}_D); \)
11. \( \ X = X - a_{{n_t}}^{\text{Tol}} \Phi_{{n_t}}^{\text{Tol}}; \ // \text{Sequential extraction} \)
12. \( \ X = X - a_{{n_t}}^{\text{Tol}} \Phi_{{n_t}}^{\text{Tol}}; \ // \text{Reduction} \)
13. if \( l = l + 1 \) then
14. \( a_{{n_t}}^{\max} = a_{{n_t}}^{\text{Tol}}; \ // \text{Maximal amplitude} \)
15. if \( |a_{{n_t}}^{\text{Tol}}| < 10^{-2} P_\text{th}/20 \) then
16. \( \ X = X - a_{{n_t}}^{\text{Tol}} \Phi_{{n_t}}^{\text{Tol}}; \ // \text{Reconstruction} \)
17. \( \ \text{break}; \ // \text{Delete weak component and stop} \)
18. Return \( H_{{n_t}}^{\text{Tol}} \)

**B. Compressive Sensing-Based Positioning**

After ruling out the impact of Doppler, the dictionary matrix in (11) is simplified to \( \Phi_{{n_t}, m} = e^{-j (2\pi/\lambda) d_{{n_t}}^{(m)}} \) which can be generated by discretizing the targeted area. The corresponding localization problem can be solved via sparsity recovery.
techniques, e.g., basic pursuit (BP), orthogonal matching pursuit (OMP), Bayesian compressive sensing (BCS), etc. Among them, BCS provides a probabilistic inference method. It efficiently solves the sparse reconstruction problem by exploiting the partial common sparse support [31], which outperforms the conventional BP and OMP generally. However, the original BCS algorithm was designed for the real-valued problem and cannot be used directly for the complex problem in (11). Wu et al. [32] and Wang et al. [33] proposed to jointly model the real and imaginary components, presuming that they share the same hyperparameters \( \alpha_m, m \in [1, \ldots, M] \) (the inverse of the variances of intensity vector) and \( \beta_0 \) (the inverse of the variances of noise), and proposed complex BCS (CBCS) algorithm for the complex-value reconstruction problem. As the real and imaginary components share the same prior, the posterior density function for \( u = [u_1, \ldots, u_M]^\top \) can be given as [32], [33], [34]

\[
\Pr(u|H_{\text{Tot}}, \Phi, \alpha, \beta_0) = \prod_{m=1}^{M} \mathcal{N}(\Re(u_m)|0, \alpha_m^{-1} \beta_0^{-1}) \\
\times \prod_{m=1}^{M} \mathcal{N}(\Im(u_m)|0, \alpha_m^{-1} \beta_0^{-1}) \\
= C \mathcal{N}(u|\mu, \Sigma) \tag{12}
\]

where

\[
\mu = \beta_0 \Sigma \Phi^H H_{\text{Tot}} \\
\Sigma = (\text{diag}(\alpha) + \beta_0 \Phi^H \Phi)^{-1}. \tag{13}
\]

Based on the CBCS algorithm, we obtain the solution \( \hat{u} \) which has a limited number of nonzero elements. The nonzero values denote the complex amplitude and indicate the possible locations of the targets. Herein, we briefly introduce the CBCS modeling for the underlying localization problem. We refer to [33] and [35] for the detailed analysis and a fast-algorithm implementation which sequentially adds or deletes the candidate basis functions of the relevance vector machine (RVM) [31]. For the location estimation, instead of using the intermediate MPCs, we regard the CSI of all antennas as a whole and estimate the locations directly, so the link association challenge in the MBCs-based solutions can be evaded. For clarity, the essential procedures of the CBCS algorithm for instantaneous location estimates are referred to [33] and [35] and summarized in Algorithm 2.

C. Multitarget Tracking

For contact-free MTT, it is required to associate the multiple estimates to each target at every instantaneous time slot, namely, the so-called target and evolving association in Section II-B. There are several available algorithms to solve the association problem in MTT, such as JPDA [7], [23], MHT [24], and multitarget particle filter [36]. Besides the high computing complexity, they also require estimating the number of targets (or a priori) before tracking. Alternatively, the RFS-based MTT has attracted great attention as it evades the explicit associations between the targeting state and the measurement state and jointly estimates the number of targets and their corresponding trajectories.

A representative RFS-based MTT algorithm is the GM-PHD filter [37]. Under the assumption of independent and linear Gaussian multitarget models and Gaussian measurement models, the GM-PHD filter propagates the first moment of the multitarget posterior (as known as, the intensity function). It provides a computationally efficient closed-form alternative to the conventional PHD filter. GM-PHD filter is reasonable for our MTT objective because the sparse state in (12) of the CBCS algorithm has already been modeled as a Gaussian distribution and each pedestrian moves independently which is also indicated by the signal model in (9). Without considering the spawned targets, the targets’ states (i.e., coordinates and velocities) at time \( t \) can be modeled by the union of the surviving targets from previous time \( t - 1 \) and the spontaneous births at time \( t \). The predicted intensity \( u_{t|t-1} \) associated with the multitarget predicted density can be given by [37]

\[
u_{t|t-1}(x_T) = u_{S,t|t-1}(x_T) + \gamma_T(x_T) \tag{14}
\]

where

\[
u_{S,t|t-1}(x_T) = p_S \sum_{i=1}^{J_{S,t-1}} \omega_{i,t-1} \mathcal{N}(x_T; m_{s,t-1}^{(i)}, P_{s,t-1}^{(i)}) \tag{14a}
\]

\[
\gamma_T(x_T) = \sum_{i=1}^{J_{I,t}} \omega_{i,t} \mathcal{N}(x_T; m_{t}^{(i)}, P_{t}^{(i)}) \tag{14b}
\]

where \( p_S \) is the target survival probability. \( J_{(\cdot)} \) represents the number of Gaussian mixtures. \( \omega_{(\cdot)} \) is the weight of Gaussian
intensity. For contact-free tracking, we also adopt the constant velocity model as in active tracking in Section III-B. So we have the mean state update as \( \mathbf{m}_{s,t-1}^{(i)} = \mathbf{Fm}_{t-1}^{(i)} \) and the covariance update as \( \mathbf{P}_{s,t-1}^{(i)} = \mathbf{\sigma}^2_{P} \mathbf{G}^\top + \mathbf{FP}_{t-1}^{(i)} \mathbf{F}^\top \), where \( \sigma_{P} \), \( \mathbf{F} \), and \( \mathbf{G} \) define the pedestrian motion model as in (6). In (14b), the intensity of the new-born targets \( \gamma(x_T) \) is also assumed as a Gaussian mixture [37]. To handle the case that the newly emerging targets may appear anywhere and anytime, generally, the new-born targets are initialized to cover the entire target area [38]. However, it is not necessary for our MTT framework as we use the instantaneous locations estimated by the CBCS algorithm which covers the entire monitoring region naturally.

The measurement states \( Z_t \) are the instantaneous location estimates at time \( t \) based on the CBCS algorithm in Section IV-B. The estimates include the locations of the real targets and some false alarms. The posterior intensity under a Gaussian mixture model can be given by [37]

\[
u_{it}(x_T) = (1 - p_D)\nu_{it-1}(x_T) + \sum_{x \in Z_t} u_{D,t}(x_T; z) \quad (15)\]

where \( p_D = 0.95 \) is the detection probability. The target states are updated via the new CBCS estimates within the framework of Kalman recursion [37]

\[
u_{D,t}(x_T; z) = \sum_{i=1}^{J_{t-1}} \omega_{it}^{(i)}(z)\mathcal{N}(x_T; \mathbf{m}_{it}^{(i)}(z), \mathbf{P}_{it}^{(i)}) \quad (15a)\]

\[
\omega_{it}^{(i)}(z) = \frac{p_D p_{it}^{(i)} q_{it}^{(i)}(z)}{\kappa_{it}(z) + p_D \sum_{j=1}^{J_{t-1}} \omega_{jt}^{(j)} q_{jt}^{(j)}(z)} \quad (15b)\]

\[
\mathbf{m}_{it}^{(i)}(z) = \mathbf{m}_{it-1}^{(i)} + \mathbf{K}_{it}^{(i)}(z - \mathbf{Hm}_{it-1}^{(i)}) \quad (15c)\]

\[
\mathbf{P}_{it}^{(i)} = \left( \mathbf{I} - \mathbf{K}_{it}^{(i)} \mathbf{H} \right) \mathbf{P}_{it-1}^{(i)} \quad (15d)\]

\[
\mathbf{K}_{it}^{(i)} = \mathbf{P}_{it}^{(i)} \mathbf{H}^\top \left( \mathbf{H} \mathbf{P}_{it}^{(i)} \mathbf{H}^\top + \mathbf{R} \right)^{-1} \quad (15e)\]

\[
q_{it}^{(i)}(z) = \mathcal{N}(z; \mathbf{Hm}_{it-1}^{(i)} + \mathbf{R} + \mathbf{H_P}_{it-1}^{(i)} \mathbf{H}^\top) \quad (15f)\]

where \( \kappa_{it}(z) \) represents the clutter intensity which is generally assumed to follow a uniform distribution within the target space. \( \mathbf{H} = [\mathbf{I} \ 0] \) because we only have the location estimates from the CBCS algorithm for updating. \( \mathbf{R} \) is the observation noise covariance matrix. Moreover, in the context of the GM-PHD filter, pruning, merging, and capping are usually adopted to relieve the computing pressure. Pruning is to truncate the components having weaker weights than the defined threshold which is set as \( \omega_{\min} = 10^{-5} \) [37]. Merging is to combine the closely located components. In our work, we merge the components with a distance smaller than \( D_{\min} = 0.2 \) m. Capping is to disregard the weakest components when the number of components is larger than the given maximal number of Gaussian components, which is set as \( J_{\max} = 50 \) according to [37]. For clarity, the essential procedures of the GM-PHD filter for MTT are referred to [37] and summarized in Algorithm 3.

### D. Putting Things Together

This section summarizes the proposed contact-free MTT algorithm for the radar-like massive MIMO prototype. When the newly measured CSI is received, we conduct the calibration first using the parameters obtained in Section III-A. Then, the background static components are removed by subtracting the moving average CSI. In this work, we set the location update time window as 0.1 s, which is also widely adopted for pedestrian tracking [5], [25]. Within the time window, we presume the changes in pedestrians’ location are negligible and the Doppler velocity is constant. As the sampling rate of the established TDD massive MIMO system is 100 Hz, we conduct ToI extraction and MTT using 10 CSI data. In ToI extraction, Algorithm 1 is conducted for each antenna. Then, the ToI of all antennas is regarded as a whole and fed to the CBCS algorithm. GM-PHD filter will track the motion of the multitarget based on the instantaneous location estimates from CBCS. The number of predicted states in the GM-PHD filter indicates the number of targets [37]. The procedures of the proposed contact-free MTT algorithm are given in Algorithm 4.

Fig. 6 shows an example of the CBCS instantaneous location estimates and the GM-PHD filter-based tracking results for two-pedestrian tracking. In Fig. 6(a) and (b), we can see...
Algorithm 4: Proposed Contact-Free MTT Algorithm for the Distributed Massive MIMO System

```
1 foreach new CSI measurement do
2   CSI calibration using $\Upsilon$ and $\phi_{\text{ant}}(n_r)$, $n_r = 1, \ldots, N_r$;
3   Background removal;
4   foreach CSI data in 0.1-s time interval do
5     for $n_r = 1:N_r$ do
6       Tol extraction based on Algorithm 1;
7       Return $\hat{H}_{\text{Tol}}$;
8       CBCS for location estimates via Algorithm 2;
9       GM-PHD for tracking via Algorithm 3;
10      Return Updated states of the targets.
```

**Fig. 6.** CBCS positioning and the GM-PHD tracking results versus benchmark. (a) and (b) $x$ and $y$ coordinates. (c) Estimates of the number of targets.

that the CBCS algorithm estimates the targets’ locations correctly yet messily. The red circles represent the tracking results after feeding to the GM-PHD filter, which shows a smoother trajectory with a limited number of outliers. Fig. 6(c) shows the predicted number of targets based on the GM-PHD filter. In most cases, it obtains correct estimates but has some false alarms corresponding to the short lifetime trajectories and missed detection. This problem can be solved by jointly propagating the posterior intensity and the posterior cardinality distribution, such as the cardinalized PHD filter [39]. We would like to leave its adaption to our distributed massive MIMO radar-like system for future exploration.

**V. PERFORMANCE EVALUATION**

In this section, we evaluate the performance of contact-free human tracking using the sub-6 GHz distributed massive MIMO communication testbed. Three participants have been invited for the single-person tracking and multiperson tracking. Fig. 7 shows the room layout of the measurement campaign and the setup deployment for the three cases: single-person tracking, two-person tracking, and three-person tracking, respectively. For each scenario, the participant walks along the predefined track templates. As mentioned in Section II-A, we ask for each participant to carry an additional transmitter (i.e., UE1, UE2, and UE3 in Fig. 7) for benchmarking. The performance and the feasibility of active tracking have been validated in Section III-B. For the tracking performance evaluation, we calculate the distance between the locations of contact-free tracking and benchmark at each time instant.

**A. Single-Pedestrian Tracking**

1) **Accuracy Evaluation:** Fig. 8 shows the contact-free single-person tracking results using the proposed algorithm, in which the dot color depth indicates the tracking errors along the trajectory. In this scenario, each individual walks along a rectangular track as presented in Fig. 7(a). After reaching the start location for the first time, the person takes a U-turn and backtracks. The red arrows in Fig. 8(a)–(c) illustrate the walking directions. Fig. 8(d) summarizes the contact-free tracking errors for the three participants. The proposed algorithm achieves decimeter-level accuracy with the largest errors around 50 cm. Importantly, we can distinctly observe that the contact-free tracking trajectories are smaller rectangles compared with the benchmark results, which are closer to the transmitter UE0. This phenomenon can be explained by the truth that the human reflector is not a single point. In practice, the signal transmitted by the UE0 will be reflected/scattered by the curved body surface (torso or limbs) facing the UE0. This causes the reflection points to have an offset to the center of the human body, which roughly equals the radius of the torso (15–30 cm generally) under the cylindrical body hypothesis [40]. In our experiment, the antenna for benchmarking is placed on top of the head, which obtains the locations of the body centroid during movement. So the compensated tracking errors are calculated based on the distance from the estimates to the periphery of the cylindrical body given by $\| \hat{P}_T - P \|$, where $P$ is the cylindrical periphery, given by $P = (P_{BM} - \text{radius of torso})$, where $P_{BM}$ is the location of the benchmark. The median error is about 7.7 cm whereas the 95th percentile error is 16.9 cm. It should be noted that we adopt the tracking accuracy after the body size compensation is presented by the dashed plots in Fig. 8(d).

2) **Impact of Antenna Number:** Moreover, we evaluate the contact-free tracking performance concerning the impact of the number of antennas and the bandwidth, as shown in Fig. 9. For the impact of the antenna number evaluation, we first select the antenna elements sequentially as the labeled order shown in Fig. 7. For example, we use the sequential antenna elements from #1 to #24 for the case of 24-antenna evaluation. Fig. 9(b) presents the tracking errors of the varied antenna elements via the 75th percentile error bars with the low bound 50th percentile and upper bound 95th percentile errors, respectively. Increasing the number of antenna elements improves the tracking accuracy and reduces the outliers generally. Nevertheless, the accuracy saturates both in accuracy and precision when the number of antennas reaches 24, in which the 75th percentile...
errors are less than 13.6 cm and 95th percentile errors are less than 25.1 cm.

Furthermore, when considering the antenna deployment topology, a better tracking performance can be expected, especially for the cases of fewer antenna elements, e.g., 8 or 16. The proposed algorithm does not use the AoA but the CSI directly for positioning and tracking, so the antenna array construction is not necessary. For the 8-antenna evaluation, we select the antenna elements \{#4, #5, #20, #21, #36, #37, #52, #53\} labeled in Fig. 7. For the 16-antenna evaluation, besides the above selected 8 antennas, we include another 8 antenna elements \{#12, #13, #28, #29, #44, #45, #60, #61\}. Note that it cannot be expected to list all the possible selections. These two typical cases allowing antennas to be distributed as uniformly as possible are considered. As shown in Fig. 9(a), in the case of 8 antennas, the 75th and 95th percentile tracking errors using the distributed selection decrease from 1.38 and 2.64 m to 0.41 and 0.62 m, respectively. In the case of 16 antennas, the 75th and 95th percentile errors decrease by 0.14 and 0.68 m, respectively. The boosting performance can be explained by the geometrical diversity of the distributed antenna selection. It also indicates that antenna placement optimization is essential for practical implementation, which deserves further investigation and we would like to leave it for future work.

3) Impact of Bandwidth: In Fig. 9(b), the impact of bandwidth or number of subcarrier on tracking performance is presented. We can observe that when we decrease the number of adopted subcarriers, the tracking performance has no clear degradation even if a single subcarrier is adopted. The fluctuations of the 75th and 95th percentile errors are less than 1.8 and 4.3 cm, respectively. This is reasonable because the adopted massive MIMO communication system is with a standard cellular bandwidth of 18 MHz. The subcarrier interval is 180 kHz for 100 subcarriers. The limited bandwidth causes a coarse-grained ranging resolution. On the other hand, due to the limited contribution of bandwidth to tracking performance, the Fourier transform for ranging in Algorithm 1 can be disregarded to reduce the computing complexity. For example, we can average the CSI along the frequency domain before ToI extraction to speed up the location inference. The tracking accuracy using the average CSI is shown in Fig. 9(b).

B. Multiple Pedestrians Tracking

The measurement campaigns for two-person and three-person tracking are given in Fig. 7(b) and (c). In Figs. 10(a) and 11(a), the tracking results of two pedestrians and three pedestrians versus the benchmarks are presented, in which all 64 antennas are adopted for the tracking. The arrows with different colors represent the walking directions of different targets. Similar to the case of single-person tracking, the estimated trajectories are closer to the transmitter UE0 due to the reflections/scattering on the periphery of the
Fig. 10. Contact-free two-person tracking results. (a) Tracking trajectories versus benchmark (arrows with different colors indicate the walking directions of different pedestrians). (b) Impact of antenna numbers on tracking accuracy.

Fig. 11. Contact-free three-person tracking results. (a) Tracking trajectories versus benchmark (arrows with different colors indicate the walking directions of different pedestrians). (b) Impact of antenna numbers on tracking accuracy.

Table II

| Percentile Tracking Errors of the Proposed Contact-Free MTT Algorithm in Meter |
|-----------------------------------|----------------|----------------|----------------|
|                                   | 50th           | 75th           | 95th           |
| Single pedestrian                 | 0.073          | 0.127          | 0.182          |
| Two pedestrians                   | 0.148          | 0.234          | 0.457          |
| Three pedestrians                 | 0.191          | 0.289          | 0.422          |

torsos. Namely, body size compensation is adopted for the tracking accuracy evaluation. Besides, we observe that the tracking errors of the multipedestrian cases are larger than the single-pedestrian tracking. In Table II, we summarize the percentile errors concerning different numbers of targets. The median error for single-person tracking is only 7.3 cm and the 95th percentile error is 18.2 cm. But when the number of targets is three, the median error increases to 19.1 cm and the 95th percentile error to 42.2 cm. The possible reasons are twofold: First, when there are multiple targets, the signal reflected by one target may be partially blocked by another person. This results in the reflected signal cannot be detected by the receivers behind. This is equivalent to the case of fewer adopted antennas. Second, we use DFT for the ToI extraction in Section IV-A, which may have worse Doppler resolution for multiple targets due to the side-lobe power leakage of DFT. Moreover, the low sampling rate of our TDD-OFDM system also challenges the ToI extraction. In other words, the signal-to-noise ratio (SNR) of the ToI will decrease, which causes tracking performance degradation.

Moreover, the impact of antenna number on MTT accuracy is also investigated in Figs. 10(b) and 11(b). Note that the impact of bandwidth is not analyzed here because of its little influence on tracking accuracy as presented in Section V-A3. The average CSI along subcarriers is used for the MTT. The tracking accuracy improves as more antennas are utilized. However, the outliers in multipedestrian tracking increase distinctly compared with the case of a single pedestrian, which are indicated by the larger offsets from 75th to 95th percentile errors. For the smaller number of antennas, i.e., 8 and 16, we consider the distributed antenna selections as in Section V-A2. An accuracy improvement is observed more or less, but the overall performance is not promising concerning the size of the monitoring area in the experiment, because the 75th and 95th percentile errors are larger than 0.59 and 1.17 m, respectively. These observations indicate that a large number of antennas is necessary for the bandwidth-limited system to guarantee a satisfying contact-free MTT accuracy.

C. Inference Time

In this work, we adopt a 0.1-s time interval for CSI data accumulation and location updating, which means the real-time tracking is guaranteed if the location updating latency is less than 0.1 s. The running environment for evaluation is Dell OptiPlex 7050 with an Intel Core i7-7700 CPU@3.60 GHz and 16-GB RAM. The latency is calculated by the tic-toc commands in MATLAB 2021b. Fig. 12(a) shows the histograms of the inference time for different numbers of targets, in which the inference latency is less than 0.1 s mostly. The single-person tracking has the fastest updating in general with a 54.2-ms mean latency, whereas two- and three-person tracking has similar inference latency with mean values of 66.8 and 68.9 ms, respectively. These results testify the proposed contact-free tracking algorithm can achieve real-time processing for MTT purposes. We further evaluate the inference latency for the major algorithm blocks, namely, ToI extraction, CBCS location estimation, and GM-PHD filter, as shown in Fig. 12(b). Distinctly, the CBCS algorithm has the largest time consumption with a 28.9-ms mean value. This also shows the significance of the ToI purification in Section IV-A to avoid the computing complexity curse.

VI. CONCLUSION

In this article, a radar-like distributed massive MIMO prototype has been built upon an available sub-6-GHz OFDM
communication system. Based on the established prototype, we investigated the contact-free multipedestrian tracking problems. Instead of using the intermediate geometrical features (AoA, ToF, Doppler, etc.), we proposed to localize and track the multiple targets exploiting the CSI signal directly. Specifically, a recursive algorithm combining the CBCS and the GM-PHD filter was proposed for multiple pedestrians tracking in real time. According to the experimental results, we achieved the 75th and 95th percentile tracking accuracy of 12.7 and 18.2 cm for single-pedestrian tracking and 28.9 and 45.7 cm for multipedestrian tracking, respectively.

This article demonstrated our preliminary investigation of the sub-6-GHz distributed massive MIMO system for its radar-like functionality. It deserves further and extensive exploration. Future work will consist of both experimental analysis and algorithm development. For future experiments, a larger number of targets in various indoor scenarios will be investigated. The placement optimization for the massive antenna array in a specific scenario will be considered for the performance enhancement. Besides moving targets, how to localize multiple static targets and hybrid dynamic and static targets will be considered.

## References

[1] Z. Liu, X. Liu, J. Zhang, and K. Li, “Opportunities and challenges of wireless human sensing for the smart IoT world: A survey,” IEEE Netw., vol. 33, no. 5, pp. 104–110, Sep./Oct. 2019.

[2] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, “Device-free human activity recognition using commercial WiFi devices,” IEEE J. Sel. Areas Commun., vol. 35, no. 5, pp. 1118–1131, May 2017.

[3] Q. Xu, Y. Han, B. Wang, M. Wu, and K. J. R. Liu, “Indoor events monitoring using channel state information time series,” IEEE Internet Things J., vol. 6, no. 3, pp. 4977–4990, Jun. 2019.

[4] X. Li et al., “indoTrack: Device-free indoor human tracking with commodity Wi-Fi,” Proc. ACM Interact. Mobile Wearable Ubiquitous Comput., vol. 1, no. 3, pp. 1–22, Sep. 2017.

[5] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, “Widar2.0: Passive human tracking with a single Wi-Fi link,” in Proc. 16th Annu. Int. Conf. Mobile Syst., Appl. Services (MobiSys), Jun. 2018, pp. 350–361.

[6] K. Niu, X. Wang, F. Zhang, R. Zheng, Z. Yao, and D. Zhang, “Rethinking doppler effect for accurate velocity estimation with commodity WiFi devices,” IEEE J. Sel. Areas Commun., vol. 40, no. 7, pp. 2164–2178, Jul. 2022.

[7] C. R. Karanam, B. Korany, and Y. Mostofi, “Tracking from one side: Multi-person passive tracking with WiFi magnitude measurements,” in Proc. 18th Int. Conf. Inf. Process. Sensor Netw. (IPSN), Apr. 2019, pp. 181–192.

[8] R. H. Venkatnarayan, M. Shahzad, S. Yun, C. Vlachou, and K.-H. Kim, “Leveraging Polarization of WiFi signals to simultaneously track multiple people,” Proc. ACM Interact. Mobile Wearable Ubiquitous Technol., vol. 4, no. 2, pp. 1–24, Jun. 2020.

[9] T. Zheng, Z. Chen, C. Cai, J. Luo, and X. Zhang, “Vi2Fi: In-vehicle vital sign monitoring via compact RF sensing,” Proc. ACM Interact. Mobile Wearable Ubiquitous Technol., vol. 4, no. 2, pp. 1–27, Jun. 2020.

[10] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, “ViMo: Multiperson vital sign monitoring using commodity millimeter-wave,” IEEE Internet Things J., vol. 8, no. 3, pp. 1294–1307, Feb. 2021.

[11] C. Wu, F. Zhang, B. Wang, and K. J. Ray Liu, “mmTrack: Passive multi-person Localization using commodity millimeter wave radio,” in Proc. IEEE Conf. Comput. Commun. (INFOCOM), Jul. 2020, pp. 2400–2409.

[12] C. Li et al., “Device-free pedestrian tracking using low-cost ultra-wideband devices,” IEEE Trans. Instrum. Meas., vol. 71, pp. 1–4, Nov. 2021.

[13] C. De Lima et al., “Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges,” IEEE Access, vol. 9, pp. 26902–26925, Jan. 2021.
B.-T. Vo, B.-N. Vo, and A. Cantoni, “Analytic implementations of the Gaussian mixture probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo, “Adaptive target birth intensity for PHD and CPHD filters,” IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 2, pp. 1650–1668, Apr. 2012.

B.-T. Vo, B.-N. Vo, and A. Cantoni, “Analytic implementations of the cardinalized probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

M. Ghaddar, L. Talbi, T. A. Denidni, and A. Sebak, “A conducting cylinder for modeling human body presence in indoor propagation channel,” IEEE Trans. Antennas Propag., vol. 55, no. 11, pp. 3099–3103, Nov. 2007.

B.-N. Vo and W.-K. Ma, “The Gaussian mixture probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

Chenglong Li (Member, IEEE) was born in Nanchong, China, in May 1994. He received the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2022.

He was a Research Assistant with the Department of Information Technology, Ghent University-imec, Ghent, from 2018 to 2022. He is currently a Lecturer with the College of Electronic Science, National University of Defense Technology, Changsha, China. His current research interests include positioning and navigation, wireless sensing, wireless channel modeling, and mobile computing.

Dr. Li was a recipient of the International Union of Radio Science Young Scientist Award and the European Microwave Association Student Grant in 2022.

Sibren De Bast (Student Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from KU Leuven, Leuven, Belgium, in 2017 and 2022, respectively.

He is focusing on massive MIMO communications for aerial and ground robots, for which he created an experimental test and measurement environment. His current research interest is artificial intelligence for communications and localization.

Yang Miao (Member, IEEE) received the M.Sc. and Ph.D. degrees from Radio Propagation Laboratory, Mobile Communications Research Group, Tokyo Institute of Technology, Tokyo, Japan, in 2012 and 2015, respectively.

She was a Research Assistant with Takada Laboratory, Tokyo Institute of Technology from 2010 to 2015. From 2015 to 2018, she was a Postdoctoral Researcher with the Institute of Information and Communication Technologies, Electronics, and Applied Mathematics, Universite Catholique de Louvain, Louvain-la-Neuve, Belgium, and imec, Wireless, Acoustics, Environment, and the Expert Systems Laboratory, Ghent University, Ghent, Belgium. She was a part-time Senior Antenna Engineer with Jaguar Radio Wave Corporation, Shenzhen, China, from 2017 to 2018. She was a Research Assistant Professor with the Southern University of Science and Technology, Shenzhen, China, from 2018 to 2019. She has been an Assistant Professor with Radio Systems Group, University of Twente, Enschede, The Netherlands, since August 2019. She has been also affiliated in part-time with KU Leuven as a Marie Curie Individual Fellow since November 2021. Her current research focuses on joint communication and sensing, incorporating mobility, and human factors.

Sofie Pollin (Senior Member, IEEE) received the Ph.D. degree (Hons.) from KU Leuven, Leuven, Belgium, in 2006.

She is currently a Professor with KU Leuven focusing on wireless communication systems. Before that, she worked with imec, KU Leuven, and University of California at Berkeley, Berkeley, CA, USA. She pioneered a 5G testbed for distributed massive MIMO at KU Leuven, and is currently leading the way toward 6G tests in multiple large EU projects. Her research centers around wireless networks that require networks that are ever more dense, heterogeneous, battery powered, and spectrum constrained.

Prof. Pollin is a Belgian American Educational Foundation Fellow and a Marie Curie Fellow.

Emmerique Tanghe (Member, IEEE) was born in Tielt, Belgium, in August 1982. He received the M.Sc. and Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 2005 and 2011, respectively.

He was a Research Assistant with the Department of Information Technology, Ghent University (imec-Ugent/INTEC) from September 2005 to May 2011. His scientific work focused on the modeling of indoor and outdoor propagation through field measurements. Since May 2011, he has been a Postdoctoral Researcher with Ghent University and continues his work in propagation modeling. From October 2012 to September 2018, he was a Postdoctoral Fellow of the Research Foundation-Flanders, Brussels, Belgium. In October 2015, he became a part-time Professor in medical applications of electromagnetic fields in and around the human body.

Wout Joseph (Senior Member, IEEE) was born in Ostend, Belgium, in October 1977. He received the M.Sc. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2000, and the Ph.D. degree from the Department of Information Technology (INTEC), Ghent University, in March 2005.

He was a Research Assistant with INTEC, Ghent University from September 2000 to March 2005. During this period, his work was focused on electromagnetic exposure assessment. Since October 2007, he has been a Postdoctoral Fellow with the Research Foundation-Flanders, Brussels, Belgium. Since October 2009, he has been a Professor in the domain of experimental characterization of wireless communication systems. He has been an imec PI since 2017. His professional interests are electromagnetic field exposure assessment, propagation for wireless communication systems, antennas, and calibration. He specializes in wireless performance analysis and quality of experience. His research work dealt with measuring and modeling of electromagnetic fields around base stations for mobile communications related to the health effects of the exposure to electromagnetic radiation.