IEEE Open Journal of Signal Processing

Received 5 December 2022; revised 15 February 2023; accepted 17 February 2023. Date of publication 24 February 2023; date of current version 10 March 2023. The review of this article was arranged by Associate Editor Fernando Cruz-Roldán.

Digital Object Identifier 10.1109/OJSP.2023.3249121

Joint Localization and Environment Sensing of Rigid Body With 5G Millimeter Wave MIMO

BIWEI LI [1] (Graduate Student Member, IEEE), XIANBIN WANG [1] (Fellow, IEEE), EDWARD AU [2] (Senior Member, IEEE), AND YAN XIN [2]

[1] Department of Electrical and Computer Engineering, Western University, London, ON N6A5B9, Canada
[2] Ottawa Research & Development Centre, Huawei Technologies Canada, Ottawa, ON K2K3J1, Canada

CORRESPONDING AUTHOR: XIANBIN WANG (e-mail: xianbin.wang@uwo.ca).

ABSTRACT Accurately localizing a target in three-dimensional (3D) space assisted with the fifth generation (5G) wireless systems in an indoor environment could enable a wide variety of new applications, including precise control for factory automation, self-maneuver of the vehicle and so on. However, controlling a target in 3D space relies on a rigid body modelling with six degrees of freedom, which dramatically increases the localization difficulty and complexity. Furthermore, for radio-based localization methods, the lack of line-of-sight (LOS) and the existence of reflection points in the environment will also influence the rigid body localization process. To improve the rigid body localization accuracy as well as unravel useful environmental information from the received signal, a novel rigid body joint active localization and environment sensing scheme is proposed in this article. Specifically, the multi-path effect of millimeter wave (mm-wave) signal with a single reflection can be exploited to enhance the rigid body localization accuracy, and it can also be utilized to locate the reflection points, which further enables a new way for environmental sensing. Hence, we first propose a two-step hierarchical compressive sensing algorithm to extract the angular and distance information of the LOS (if available) and single-bounce specular reflections. Then a particle swarm optimization (PSO) based method is derived to recover the posture of the rigid body and the location of reflection points. The Cramér-Rao lower bound (CRLB) on angle, rigid body posture and reflection points position uncertainty is also analyzed. The simulation results demonstrate that the proposed scheme can achieve high accuracy rigid body localization and locate the reflection points around the rigid body even under obstructed-line-of-sight (OLOS) conditions in an indoor scene.

INDEX TERMS Rigid body localization, mm-wave MIMO, CRLB, compressive sensing, NLOS, environment sensing.

I. INTRODUCTION

The rapid proliferation of localization-enabled applications including warehouse management, asset tracking and factory automation are significantly boosting the need for high-accuracy target positioning under challenging conditions [1]. Albeit there have been many investigations into real-time localization for many years, this area is attracting growing research interest due to new harsh localization environments and stringent accuracy requirements from many emerging applications. Traditionally, a target is modelled as a point target whose location is only defined by the 3D coordinates with three degrees of freedom while neglecting its orientations [2]. However, to achieve accurate maneuvering for tasks like cargo transportation, robotic control and others, not only the target's position but also its orientations are required. For this purpose, a target has to be modelled as a rigid body, which always maintains its shape and size. Apart from the 3D coordinates, a rigid body has the other three degrees of freedom including three Euler angles.

As a result, the rigid body localization (RBL) problem has been studied recently to determine a rigid body's centre position and orientation. In general, the rigid body localization problem for outdoor scenarios relies on the combined use of global navigation satellite systems (GNSS) and inertial measurement unit (IMU). However, GNSS suffers from limited signal coverage and poor position accuracy in

indoor environment, while IMU requires frequent calibration, which rules out most localization-based indoor applications. Meanwhile, some alternative indoor rigid body localization technologies, using radar [3], laser [4], infrared light (IR) [5], ultrasound [6] and cameras [7], are either very sensitive to environmental changes or dependent on expensive hardware.

Recently, the rapid advancements of 5G technologies such as multiple-input and multiple-output (MIMO) and mm-wave bring enormous improvements in network throughput and energy efficiency, as well as their potential for localization [8], [9]. Such technological advancements could be exploited to enhance localization performance. In particular, operating at carrier frequencies beyond 30 GHz with ultra-wide bandwidths, mm-wave based 5G MIMO systems provide extremely high data rates through dense spatial multiplexing by using a large number of antennas [10]. The high spatial resolution and large bandwidth provided by these 5G technologies are expected to bring in revolutionary impact on wireless localization [8]. In addition, 5G mm-wave MIMO not only enables the active localization of the rigid body but also provides the opportunity of positioning the surrounding physical objects serving as reflection points simultaneously to achieve environment sensing. It has been proved that the non-light-of-sight (NLOS) components in the channel estimation results can be turned into benefits to assist the radio map construction [11]. The precise rigid body posture and reflection points location information also in turn contribute to the design, operation and optimization of future 5G beyond and 6G wireless networks. For instance, location information of rigid body and reflection points can be utilized to construct location-specific channel state information (CSI), which can boost the spatial spectrum efficiency of future networks. Motivated by these considerations, joint localization and environment sensing of a rigid 5G mobile terminal equipped with MIMO has become one of the promising research directions.

While 5G mm-wave MIMO assisted rigid body localization and environment sensing have many advantages, several challenges of previous studies have yet to be overcome. On one hand, existing studies on angle information of a target only consider one [12] or two rotation angles [13] in 2D or 3D scenarios, which neglects part of rotation due to the inherent complexity. Solving the translation vector and rotation information of a rigid body from signals transmitted by fixed anchor nodes is a complicated nonlinear optimization problem. For rigid body localization, [14] proposed a range measurement method in a sensor network, in which four anchor nodes and five wireless sensors mounted on the rigid body are used. Jiang et al. [15], [16] presented a refined positioning algorithm using the range measurements between the anchors and several sensors in both 2D and 3D cases. In [17], an angle of arrival (AOA) based method is introduced by utilizing only one base station and four wireless sensors in the rigid body localization problem. However, the aforementioned methods need to guarantee that the topology of how the sensors are mounted on the rigid body is known and the LOS exists. In addition, these methods are based on passive localization which gathers the measurements from an incoming signal transmitted/reflected by sensors mounted on the rigid body and the robustness of distance estimation to the environment noises/shadowing is limited due to the fluctuations of practical path conditions. Therefore, active localization performed at the rigid body with received signals from only one anchor node is more controllable and efficient.

On the other hand, the derivation of NLOS components for reflection points localization has also been studied for years [18], [19], [20], [21], [22]. The reflection points localization problem can be solved by exploiting the AOA and angle of departure (AOD) with the sparsity of channel information. Among several popular algorithms, the distributed compressed sensing-simultaneous orthogonal matching pursuit (DCS-SOMP) [23], [24], [25] and its modifications like CoSaMP/OMP are useful tools by manipulating sparsity of mm-wave channels. [10] demonstrated a three-stage improved DCS-SOMP algorithm by linear antenna arrays with AOA and AOD to estimate the scatters (or reflection points) to improve the moving target localization accuracy only in the 2D scenario. However, the reflection points localization in the 3D scenario has not been investigated and its implementation at the rigid body end is yet to be studied. Thus, an integrated rigid body localization and environment sensing technique with accurate and robust performance in 3D is quite challenging yet demanding.

In this paper, the mm-wave MIMO assisted rigid body active localization and environment sensing scheme is proposed to tackle the aforementioned challenges. In the considered 3D scenario, the six degrees of freedom of rigid body are successfully resolved with only one base station equipped with uniform rectangular array (URA). Three different cases are investigated: with only line-of-sight (LOS), with both LOS and non-line-of-sight (NLOS), and with only the NLOS. Moreover, when there's NLOS, the position of reflection points in the environment is also estimated. Both the rigid body localization and reflection point estimation are achieved at the rigid body end with the one-way signals. To guarantee high rigid body/reflection points localization accuracy, a new strategy of the multi-step hierarchical design is developed. Specifically, the primary contributions of this paper comprise the following three aspects.

1) A novel rigid body active localization and environment sensing strategy is proposed with the assistance of one anchor node. To the best of our knowledge, there are no developed solutions for joint rigid body and reflection points localization estimation under 3D scenarios. Although some works were reported to detect the 2D reflection points, an optimal reflection points localization solution to deal with the 3D situation has never been studied using the mm-wave MIMO at the rigid body end. Our proposed strategy achieves the rigid body location estimation accuracy to the centimetre level and orientation estimation accuracy to 0.02 rd and improves the reflection points localization accuracy to decimeter level with reasonable computational complexity.

2) A novel hierarchical DCS-SOMP algorithm incorporated with iterative maximum likelihood (IML) is presented to resolve the channel information from received signals. In particular, by using the sparsity of the mm-wave channel in the angle domain, the AOA/AOD estimation accuracy is increased to 0.01 rd and distance estimation accuracy to the centimetre level. Compared with the original DCS-SOMP, the estimation precision is enhanced greatly.
3) The theoretical CRLB of angular, rigid body posture and reflection points location estimation is derived to evaluate the effectiveness of the proposed algorithms. Simulations show that our proposed joint rigid body and reflection points localization method approaches the theoretical bounds while the signal-to-noise ratio (SNR) is not less than 0 dB.

The rest of the paper is organized as follows. The system models including the environments, rigid body localization and reflection points localization are formulated in Section II. The novel compressive sensing method for channel parameter estimation is investigated in Section III-A. The PSO-based joint rigid body and reflection points localization strategy are proposed in Section III-B. The fundamental bounds are derived in Section IV. Section V presents the simulation results. Finally, Section VI summarizes and concludes the work.

II. SYSTEM MODEL

A. ENVIRONMENT MODEL

We consider a 3D indoor environment with an anchor node AN located at $\boldsymbol{s} = [s_x, s_y, s_z]^T \in \mathbb{R}^3$ and a rigid body with posture $[\boldsymbol{p}; \boldsymbol{q}] \in \mathbb{R}^6$. Note that $\boldsymbol{p} = [p_x, p_y, p_z]^T \in \mathbb{R}^3$ is the position and $\boldsymbol{q} = [q_1, q_2, q_3]^T \in \mathbb{R}^3$, $q_1, q_3 \in [0, 2\pi)$ and $q_2 \in [0, \pi]$, represents the orientation of rigid body. The objects in the environment are characterized by $L-1$ reflection points located at $\boldsymbol{\kappa}_l = [\kappa_{l,x}, \kappa_{l,y}, \kappa_{l,z}] \in \mathbb{R}^3$, $l = 1, 2, \ldots, L-1$. It is assumed that the anchor node as signal transmitter carries a uniform rectangular array (URA) consisting of N_T ($N_{Tx} \times N_{Ty}$) antennas while rigid body as receiver carries URA with N_R ($N_{Rx} \times N_{Ry}$) antennas. The value of $\boldsymbol{s}$ is assumed to be known, while $[\boldsymbol{p}; \boldsymbol{q}]$ and $\boldsymbol{\kappa}_l$ are unknown.

B. JOINT RIGID BODY AND REFLECTION POINTS LOCALIZATION MODEL

Fig. 1 illustrates the rigid body and reflection points localization scenario. In our mm-wave wireless system, the signal transmitted by AN is orthogonal frequency division multiplexing (OFDM) signal [10]. The mm-wave network operates at carrier frequency f_c and bandwidth B. In particular, G signals are transmitted sequentially for each subcarrier. The g-th signal is denoted as $\boldsymbol{x}^{(g)}[n] \in \mathbb{C}^{N_s \times 1}$ for n-th subcarrier $n = 0, 1, 2, \ldots, N-1$. $\boldsymbol{F}[n] = \boldsymbol{F}_{RF}\boldsymbol{F}_{BB}[n] \in \mathbb{C}^{N_T \times N_s}$ is the beamforming matrix where $\boldsymbol{F}_{RF} \in \mathbb{C}^{N_T \times N_T^{RF}}$ is analog precoding matrix and $\boldsymbol{F}_{BB} \in \mathbb{C}^{N_T^{RF} \times N_s}$ is the digital beamformer. N_T^{RF} is the number of RF chains at the transmitter. The transmitted power is denoted as P_T and satisfies $\sum_{n=0}^{N-1} \|\boldsymbol{F}[n]\|_F^2 = P_T$. In this work, without loss of generality, we don't designate the specific expression of $\boldsymbol{F}[n]$ but keep the general form during derivation. It is assumed there are L paths in the channel, where path index $l = 0$ is the LOS while the remaining paths are NLOS ($l = 1, 2, \ldots, L-1$) corresponding to $L-1$ reflection points. Each reflection point is assumed to provide one propagation path between AN and the rigid body. Under this model, the $N_R \times N_T$ channel matrix associated with n-th subcarrier is modeled as

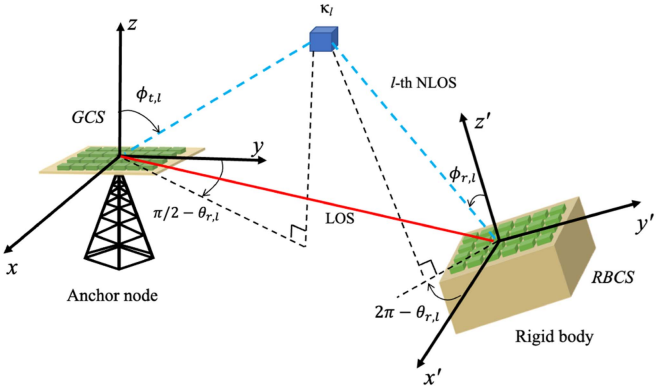

FIGURE 1. Schematic illustration of signal transmission between anchor node and rigid body.

$$\boldsymbol{H}[n] = \boldsymbol{A}_R[n]\boldsymbol{\Gamma}[n]\boldsymbol{A}_T^H[n], \tag{1}$$

where

$$\boldsymbol{\Gamma}[n] = \sqrt{N_R N_T}\ \mathrm{diag}\left\{\frac{h_0}{\sqrt{\rho_0}}e^{\frac{-j2\pi n\tau_0}{NT_s}}, \ldots, \frac{h_{L-1}}{\sqrt{\rho_{L-1}}}e^{\frac{-j2\pi n\tau_{(L-1)}}{NT_s}}\right\}. \tag{2}$$

In (2), h_l is complex channel gain and ρ_l here is the path loss of the l-th channel ($l = 0, 1, 2, \ldots, L-1$). $T_s = 1/B$ denotes the sampling period. The transmission time τ_0 at LOS is obtained by $\tau_0 = |\boldsymbol{p} - \boldsymbol{s}|/c$ (c is the speed of light) while τ_l ($l = 1, 2, \ldots, L-1$) is calculated by $\tau_l = (|\boldsymbol{p} - \boldsymbol{\kappa}_l| + |\boldsymbol{\kappa}_l - \boldsymbol{s}|)/c$. The response vectors at transmitter and receiver can be represented as

$$\begin{aligned}\boldsymbol{A}_T[n] &= [\boldsymbol{a}_{t,n}(\theta_{t,0}, \phi_{t,0}), \ldots, \boldsymbol{a}_{t,n}(\theta_{t,L-1}, \phi_{t,L-1})]^T,\\ \boldsymbol{A}_R[n] &= [\boldsymbol{a}_{r,n}(\theta_{r,0}, \phi_{r,0}), \ldots, \boldsymbol{a}_{r,n}(\theta_{r,L-1}, \phi_{r,L-1})]^T,\end{aligned} \tag{3}$$

where the sub-vector $\boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l}) = \boldsymbol{\chi}_x(\theta_{t,l}, \phi_{t,l}) \otimes \boldsymbol{\chi}_y(\theta_{t,l}, \phi_{t,l})$. And there are

$$\begin{aligned}\boldsymbol{\chi}_x(\theta_{t,l}, \phi_{t,l}) = \frac{1}{\sqrt{N_{Tx}}}\Big[&e^{-j\frac{N_{Tx}-1}{2}\frac{2\pi}{\lambda_n}d_o\cos(\theta_{t,l})\sin(\phi_{t,l})}\\ &, \ldots, e^{j\frac{N_{Tx}-1}{2}\frac{2\pi}{\lambda_n}d_o\cos(\theta_{t,l})\sin(\phi_{t,l})}\Big],\end{aligned} \tag{4}$$

$$\begin{aligned}\boldsymbol{\chi}_y(\theta_{t,l}, \phi_{t,l}) = \frac{1}{\sqrt{N_{Ty}}}\Big[&e^{-j\frac{N_{Ty}-1}{2}\frac{2\pi}{\lambda_n}d_o\sin(\theta_{t,l})\sin(\phi_{t,l})}\\ &, \ldots, e^{j\frac{N_{Ty}-1}{2}\frac{2\pi}{\lambda_n}d_o\sin(\theta_{t,l})\sin(\phi_{t,l})}\Big],\end{aligned} \tag{5}$$

where $[\theta_{t,l}, \phi_{t,l}]^T$ denotes the azimuth and elevation angles of AOD of l-th path at the anchor node side. $\lambda_n = c/(n/(NT_s) + f_c)$ is the signal wavelength at the n-th subcarrier and d_o denotes the distance between the antenna elements. The sub-vector $\boldsymbol{a}_{r,n}(\theta_{r,l}, \phi_{r,l})$ can also be derived with similar equations. $\theta_{r,l}$ and $\phi_{r,l}$ are azimuth and elevation angles of AOA of l-th path.

In this paper, the origin of the global coordinate system (GCS) is the position of the gravity center of URA antennas on the anchor node, and the z-axis of GCS is parallel to the antenna (as shown in Fig. 1). All of our calculations are under GCS. However, since the posture of URA antennas on the rigid body changes as the rigid body moves, $\phi_{r,l}$ and $\theta_{r,l}$ are measured at the rigid body coordinate system (RBCS). Thus, we need to calculate the transformation relationship from RBCS to GCS as shown

$$\begin{aligned}&[\cos(\theta_{r,l})\sin(\phi_{r,l}), \sin(\theta_{r,l})\sin(\phi_{r,l}), \cos(\phi_{r,l})]^T\\ &= \mathbf{R} \cdot [\cos(\theta_{r,l,g})\sin(\phi_{r,l,g}), \sin(\theta_{r,l,g})\sin(\phi_{r,l,g}), \cos(\phi_{r,l,g})]^T\\ &l = 0, 1, 2, \ldots, L-1. \end{aligned} \tag{6}$$

In (6), $\mathbf{R}$ is the rotation matrix related to the rigid body orientation $\boldsymbol{q}$. We use $\phi_{r,l,g}$ and $\theta_{r,l,g}$ to represent the elevation and azimuth angles of AOA at path l under the global coordinate system as

$$\begin{aligned}\phi_{r,l,g} &= \arccos\left(\frac{\kappa_{l,z} - p_z}{|\boldsymbol{\kappa}_l - \boldsymbol{p}|}\right),\\ \theta_{r,l,g} &= \arctan\left(\frac{\kappa_{l,y} - p_y}{\kappa_{l,x} - p_x}\right),\end{aligned} \tag{7}$$

$$\boldsymbol{R} = \begin{bmatrix} c_3c_2 & -s_3c_1 + c_3s_2s_1 & s_3s_1 + c_3s_2c_1 \\ s_3c_2 & c_3c_1 + s_3s_2s_1 & -c_3s_1 + s_3s_2c_1 \\ -s_2 & c_2s_1 & c_2c_1 \end{bmatrix}, \tag{8}$$

where $c_i = \cos(q_i)$ and $s_i = \sin(q_i)$, $i = 1, 2, 3$ for simplicity.

Therefore, the received signal at n-th subcarrier and g-th transmission can be expressed as (9).

$$\boldsymbol{y}^{(g)}[n] = \boldsymbol{H}[n]\boldsymbol{F}^{(g)}[n]\boldsymbol{x}^{(g)}[n] + \boldsymbol{n}^{(g)}[n], \tag{9}$$

where $\boldsymbol{n}^{(g)}[n] \in \mathbb{C}^{N_R}$ is a Gaussian noise vector with zero mean and variance $N_0/2$. $\boldsymbol{x}^{(g)}[n] = [x_1[n], \ldots, x_{N_s}[n]]^T$ is the g-th transmission. Our goal is to estimate rigid body posture $[\boldsymbol{p}; \boldsymbol{q}]$ and reflection points position $\boldsymbol{\kappa}_l$ from $\boldsymbol{y}$.

III. PROPOSED JOINT RIGID BODY AND REFLECTION POINTS LOCALIZATION SCHEME

In this section, a joint rigid body posture estimation and reflection points location estimation scheme is introduced. Firstly, channel parameters including AOD, AOA and transmission time are estimated by exploiting the sparsity of the mm-wave MIMO channel. Secondly, to recover the rigid body posture and reflection points location in both NLOS and OLOS scenarios, we use different expressions of the location of reflection points to build a minimization problem and solve it with the heuristic algorithm.

A. CHANNEL PARAMETERS ESTIMATION

Since the AOA and AOD estimation accuracy are limited to the size of the grid by using the original DCS-SOMP algorithm, we propose a novel two-step channel estimation algorithm. The coarse estimation is based on an improved DCS-SOMP algorithm called hierarchical DCS-SOMP (H-DCS-SOMP) while the further fine estimation is achieved by an iterative maximum likelihood (IML) method.

1) H-DCS-SOMP

In the coarse reflection points location estimation, assuming the received signal $\boldsymbol{y}$ from the anchor node is the compressive measurement, the reflection points localization can be formulated as compressive sensing (CS) problem. The angle, transmission time measurements and channel gain are solvable by constructing a hierarchical sensing dictionary.

Firstly, the mm-wave MIMO channel model is transformed into beamspace to reduce the complexity due to the sparsity of mm-wave signals in the angular domain [26]. In the 3D environment, we introduce the $N_T \times N_T$ transformation matrix, uniformly sampling the virtual spatial angles of AOD.

$$\boldsymbol{U}_T = \boldsymbol{U}_{Tx} \otimes \boldsymbol{U}_{Ty}, \tag{10}$$

$$\begin{aligned}\boldsymbol{U}_{Tx} &\triangleq [\boldsymbol{u}_{Tx}(-(N_{Tx}-1)/2), \ldots, \boldsymbol{u}_{Tx}((N_{Tx}-1)/2)],\\ \boldsymbol{U}_{Ty} &\triangleq [\boldsymbol{u}_{Ty}(-(N_{Ty}-1)/2), \ldots, \boldsymbol{u}_{Ty}((N_{Ty}-1)/2)],\end{aligned} \tag{11}$$

where

$$\begin{aligned}\boldsymbol{u}_{Tx}(b) &\triangleq \left[e^{-j2\pi\frac{N_{Tx}-1}{2}\frac{b}{N_{Tx}}}, \ldots, e^{j2\pi\frac{N_{Tx}-1}{2}\frac{b}{N_{Tx}}}\right]^T,\\ \boldsymbol{u}_{Ty}(b) &\triangleq \left[e^{-j2\pi\frac{N_{Ty}-1}{2}\frac{b}{N_{Ty}}}, \ldots, e^{j2\pi\frac{N_{Ty}-1}{2}\frac{b}{N_{Ty}}}\right]^T,\end{aligned} \tag{12}$$

where $-(N_{Tx}-1)/2 \le b \le (N_{Tx}-1)/2$. Similarly, we can define the $N_R \times N_R$ transformation matrix $\boldsymbol{U}_R$ with the same rule. The partial virtual representation of the channel concerning the angular domain can be written as (13). Furthermore, the received signal is expressed as

$$\tilde{\boldsymbol{H}}[n] = \boldsymbol{U}_R^H \boldsymbol{H}[n]\boldsymbol{U}_T, \tag{13}$$

$$\tilde{\boldsymbol{y}}[n] = \boldsymbol{\Omega}[n]\tilde{\boldsymbol{h}}[n] + \tilde{\boldsymbol{n}}[n], \tag{14}$$

where

$$\boldsymbol{\Omega}[n] = \begin{bmatrix} \boldsymbol{\Omega}^{(1)}[n] \\ \vdots \\ \boldsymbol{\Omega}^{(G)}[n] \end{bmatrix}, \tag{15}$$

$$\boldsymbol{\Omega}^{(g)}[n] = (\boldsymbol{U}_T^H \boldsymbol{F}^{(g)}[n]\boldsymbol{x}^{(g)}[n])^T \otimes \boldsymbol{U}_R, \tag{16}$$

$$\tilde{\boldsymbol{h}}[n] = vec(\tilde{\boldsymbol{H}}[n]). \tag{17}$$

In principle, this is a 3D CS problem, where $\boldsymbol{\Omega}[n]$ denotes the sparse dictionary (sensing matrix) and $\tilde{\boldsymbol{h}}[n] \in \mathbb{C}^{N_R N_T \times 1}$ is the vectorization of $\tilde{\boldsymbol{H}}[n]$ that corresponds to the coarse estimation of the AOA/AOD. With the received signal $\tilde{\boldsymbol{y}}[n]$,

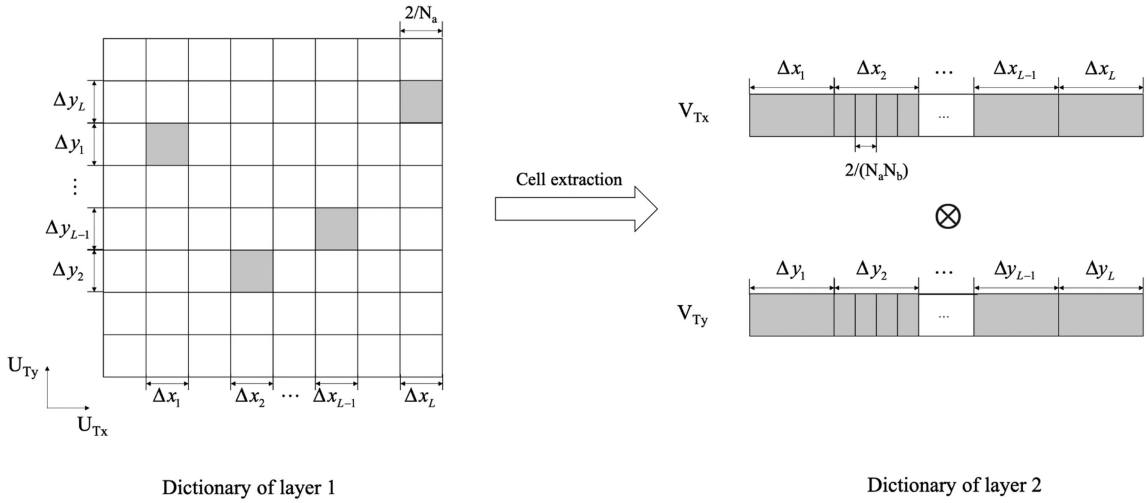

FIGURE 2. Dictionary construction in H-DCS-SOMP procedure.

$\boldsymbol{\Omega}[n]$ and the number of NLOS paths as input, the steps of DCS-SOMP to solve this problem can be summarized as

- For $n = 0, 1, \ldots, N-1$, initialize the residual vectors to $\boldsymbol{r}_{-1}[n] = 0$ and $\boldsymbol{r}_0[n] = \boldsymbol{y}_o[n]$. k is the iteration number and $\boldsymbol{\omega}_m[n]$ denotes the m-th column of measurement matrix.
- Find the largest projection of $r_{k-1}[n]$ on the columns by

$$\hat{n}_k = \arg\max_m \sum_{n=0}^{N-1} \frac{\left|\boldsymbol{\omega}_m^H[n]\boldsymbol{r}_{k-1}[n]\right|}{\|\boldsymbol{\omega}_m[n]\|_2}. \tag{18}$$

- Update indices according to (18).
- Calculate the k-th orthogonalized basis vector $\boldsymbol{\rho}_k[n]$ with Gram–Schmidt process using (19). When $k = 0$, $\boldsymbol{\rho}_0[n] = \boldsymbol{\omega}_0^H[n]$.

$$\boldsymbol{\rho}_k[n] = \boldsymbol{\omega}_k^H[n] - \sum_{k'=0}^{k-1} \frac{\boldsymbol{\omega}_k^H[n]\boldsymbol{\rho}_{k'}[n]}{\left\|\boldsymbol{\rho}_{k'}[n]\right\|_2^2}\boldsymbol{\rho}_{k'}[n]. \tag{19}$$

- Update the residual vector $\boldsymbol{r}_k[n]$.
- Repeat until all the L indices are found.

It is noted that the size of $\boldsymbol{\Omega}[n]$ will limit the upper bound on estimation accuracy. However, to improve the accuracy of parameter estimation, the size of the dictionary and the amount of computation will increase exponentially. In this paper, the channel parameter estimation is refined through a hierarchical procedure, which is initialized by the original DCS-SOMP. The structure of the dictionary matrix in H-DCS-SOMP is shown in Fig. 2. A new dictionary is constructed based on the subdivision of the L angular cells in DCS-SOMP. Since the dictionary resolution in DCS-SOMP depends on N_{Tx} and N_{Ty}, to make the second layer dictionary resolution the same in two dimensions, we set $N_a = \min\{\mathrm{N_{Tx}}, \mathrm{N_{Ty}}\}$ and consider the width between the adjacent virtual angle index as $2/N_a$ in both dimension of U_{Tx} and U_{Ty} for simplicity. Assume the indexes of L angular cells are $(\Delta x_1, \Delta y_1), \ldots, (\Delta x_L, \Delta y_L)$ in the first layer, the $(N_b \times L) \times (N_b \times L)$ transformation matrix $\boldsymbol{V}_T$ can be rebuilt, uniformly sampling the virtual spatial angles of multipath. In the other words, N_b depicts the sampling density of the second layer of the dictionary. In principle, the two-dimensional AOA/AOD estimation performance degrades greatly by using the original single layer DCS-SOMP due to the sparsity of antenna numbers. However, by using H-DCS-SOMP, increasing N_b will contribute to a higher upper bound on parameter estimation accuracy. We'll demonstrate the effectiveness of the hierarchical process in Section V.A. For the second layer of the dictionary, we have

$$\boldsymbol{V}_T = \boldsymbol{V}_{Tx} \otimes \boldsymbol{V}_{Ty}, \tag{20}$$

$$\boldsymbol{V}_{Tx}(\Delta x) \triangleq [\boldsymbol{v}_{Tx}(\Delta x_{1,c} - 1/N_a), \ldots, \boldsymbol{v}_{Tx}(\Delta x_{L,c} + 1/N_a)],$$

$$\boldsymbol{V}_{Ty}(\Delta y) \triangleq [\boldsymbol{v}_{Ty}(\Delta y_{1,c} - 1/N_a), \ldots, \boldsymbol{v}_{Ty}(\Delta y_{L,c} + 1/N_a)], \tag{21}$$

where $\Delta x_{l,c}$ represents the center of Δx_l and the sampling interval is $2/(N_a N_b)$ as demonstrated in Fig. 2. The construction of $\boldsymbol{v}_{Tx}(b)$ and $\boldsymbol{v}_{Ty}(b)$ are similar to (12). Also, we define the $N_R \times N_R$ matrix $\boldsymbol{V}_R$ in the same way. After conducting the DCS-SOMP algorithm again, indices with better precision can be obtained. If there are more than two layers, the construction rule is similar.

With the estimated $\hat{\tilde{\boldsymbol{h}}}$, $\hat{\boldsymbol{\Omega}}$, we can rewrite (17) as

$$\hat{\tilde{\boldsymbol{h}}}[n] = \Theta + \tilde{\boldsymbol{n}}[n], \tag{22}$$

where Θ can be separated into the multiplication of channel gain as

$$\Theta = \mathrm{diag}\{h_0, h_1, \ldots, h_{L-1}\}) \cdot f_1(\tau_0, \tau_1, \ldots, \tau_{L-1}) \cdot f_2(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{23}$$

and f_1, f_2 are determined by (1) and (2) [10]. $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are azimuth angles and elevation angles of all paths.

Therefore, h_l and τ_l can be resolved by minimizing the following least squares problem. For path l, there is

$$\underset{\tau_l, h_l}{\operatorname{argmin}} \sum_{n} \left\| \hat{\tilde{\boldsymbol{h}}}[n] - \Theta \right\|. \tag{24}$$

2) FINE ESTIMATION

In the following environment sensing process, the 3D position estimation error of reflection points is proportional to the square of AOA or AOD estimation. Therefore, to ensure satisfied angular estimation accuracy in applications as well as reduce the computation cost, an iterative approach is developed to further improve the channel estimation performance based on the maximum likelihood principle. The iterative procedure to refine the channel parameter estimation is initialized by the coarse estimation and the search for optimal angles is subject to $\pm\delta$ around the coarse estimation values.

Let $\boldsymbol{\eta} = [\eta_0, \eta_1, \ldots, \eta_{L-1}]^T$ be the vector of the parameters under estimation. $\boldsymbol{\eta}_l = [\theta_{t,l}, \phi_{t,l}, \theta_{r,l}, \phi_{r,l}, \tau_l, h_l]^T \in \mathbb{R}^6$ consists of the angle parameters, channel gain and transmission time for the l-th path. The likelihood function of the random vector $\boldsymbol{y}$ conditioned on $\boldsymbol{\eta}$ can be written as (25) [27],

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \propto \exp\left\{ \frac{2}{\mathrm{N}_0} \sum_{\mathrm{n}=0}^{\mathrm{N}-1} \Re\left\{ \boldsymbol{\mu}^{\mathrm{H}}[\mathrm{n}]\boldsymbol{y}[\mathrm{n}] \right\} - \frac{1}{\mathrm{N}_0} \sum_{\mathrm{n}=0}^{\mathrm{N}-1} \|\boldsymbol{\mu}[\mathrm{n}]\|^2 \right\}, \tag{25}$$

where $\boldsymbol{\mu}[n] = \boldsymbol{H}[n]\boldsymbol{F}[n]\boldsymbol{x}[n]$ and $\Re$ means taking the real part in $\{\cdot\}$.

Generally, the maximum of $f(\boldsymbol{y}|\boldsymbol{\eta})$ can be calculated directly using its Jacobian and Hesse matrix. However, due to the multi-dimensional ($\boldsymbol{\eta} \in \mathbb{R}^{6L}$) maximization of our problem, the complexity of using the analytical methods is high. The regularly utilized searching approach for ML estimator is not feasible well. Therefore, in this paper, a simplified method is employed by utilizing an iterative algorithm to update the variables one by one. Since there are L paths, we consider beginning the iteration with the LOS followed by the NLOS paths. For instance, considering the l-th path, we firstly update one of the angles by gradient descent algorithm with the remaining parameters fixed to maximize $f(\boldsymbol{y}|\boldsymbol{\eta})$, which stops when the variation of $f(\boldsymbol{y}|\boldsymbol{\eta})$ gets smaller than or equal to the threshold α. Since the estimation error of transmission time of H-DCS-SOMP is relatively small, in the fine estimation, we consider the refinement of angular parameters. Firstly, the four angles (AOA/AOD azimuth and elevation) in l-th path are updated iteratively until the variation of $f(\boldsymbol{y}|\boldsymbol{\eta})$ reaches threshold α. Then the L paths are updated in the loop to ensure the convergence of all the angular variables. The steps are summarized as follows.

- Initialize all the angles using the AOA/AOD measurements obtained from the coarse estimation.
- Calculate the likelihood function $f(\boldsymbol{y}|\boldsymbol{\eta})$ with only one angle as a variable while keeping all other angles fixed.
- Update the angle by using the gradient descent algorithm until the $f(\boldsymbol{y}|\boldsymbol{\eta})$ reaches the local maximum.
- Repeat the steps above for all the remaining angles until the variation of the likelihood function is less than or equal to threshold α.

During the iteration, there's no priority difference among the angular variables since the estimation precision of them cannot be distinguished theoretically. For instance, the iteration of l-th path is carried by using

$$\begin{aligned}
\theta''_{r,l} &= \underset{\theta_{r,l}}{\operatorname{argmax}} \left\{ f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta_{r,l}, \theta_{t,l}, {}'\phi_{t,l}, {}'\phi'_{r,l}} \right\}, \\
\theta''_{t,l} &= \underset{\theta_{t,l}}{\operatorname{argmax}} \left\{ f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta_{r,l}, {}''\theta_{t,l}, \phi_{t,l}, {}'\phi'_{r,l}} \right\}, \\
\phi''_{t,l} &= \underset{\phi_{t,l}}{\operatorname{argmax}} \left\{ f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta_{r,l}, {}''\theta_{t,l}, {}''\phi_{t,l}, \phi'_{r,l}} \right\}, \\
\phi''_{r,l} &= \underset{\phi_{r,l}}{\operatorname{argmax}} \left\{ f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta_{r,l}, {}''\theta_{t,l}, {}''\phi_{t,l}, {}''\phi_{r,l}} \right\},
\end{aligned} \tag{26}$$

where the results $\theta_{r,l}, {}''\theta_{t,l}, {}''\phi_{t,l}, {}''\phi''_{r,l}$ are calculated from the iteration results of last step $\theta_{r,l}, {}'\theta_{t,l}, {}'\phi_{t,l}, {}'\phi'_{r,l}$.

Here, the gradient descent method is used to obtain each maximum within a certain searching angular range. Taking the $\theta''_{r,l}$ as an example, with $\theta'_{r,l}$ as the original value, the steps are shown as follows.

1) Set $0 < \alpha < 1$ as the error bound and iteration number $k = 0$.
2) Calculate $\varphi^{(k)} = \nabla f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta^{(k)}_{r,l}, \theta_{t,l}, {}'\phi_{t,l}, {}'\phi'_{r,l}}$.
3) If $\|\varphi^{(k)}\| < \alpha$, break the iteration, return $\theta''_{r,l} = \theta^{(k)}_{r,l}$.
4) Find the value of $\varrho > 0$ from $\theta^{(k)}_{r,l}$ to get the $\underset{\varrho^k}{\operatorname{argmax}}\, f(\boldsymbol{y}|\boldsymbol{\eta})|_{\theta^{(k)}_{r,l} + \varrho \times \varphi^{(k)}, \theta_{t,l}, {}'\phi_{t,l}, {}'\phi'_{r,l}}$.
5) Make $\theta^{(k+1)}_{r,l} = \theta^{(k)}_{r,l} + \varrho \times \varphi^{(k)}$ and go to (2).

B. JOINT RIGID BODY AND REFLECTION POINTS LOCALIZATION

In this subsection, based on the estimation results of AOAs/AODs/transmission times, we demonstrate that the posture of the rigid body and the position of reflection points can be recovered from the parameters. Both two scenarios with and without LOS are discussed.

The case with LOS is first investigated. As shown in Fig. 1, the distances between AN and rigid body, AN and reflection point l, reflection point l and rigid body is denoted by d_0, $d_{l,1}$ and $d_{l,2}$ ($l = 1, 2, \ldots, L-1$), respectively. Their estimated values can be expressed with transmission times as

$$\begin{aligned}
\hat{d}_0 &= c \cdot \hat{\tau}_0, \\
\hat{d}_{l,1} + \hat{d}_{l,2} &= c \cdot \hat{\tau}_l.
\end{aligned} \tag{27}$$

Then, the rigid body position can be directly estimated with the AOD of LOS and $\hat{d}_0$ by using the geometric relationship as

$$\hat{\boldsymbol{p}} = \boldsymbol{s} + \hat{d}_0 \cdot [\sin\hat{\phi}_{t,0}\cos\hat{\theta}_{t,0}, \sin\hat{\phi}_{t,0}\sin\hat{\theta}_{t,0}, \cos\hat{\phi}_{t,0}]^T. \tag{28}$$

However, the orientation and reflection points cannot be resolved directly from channel parameters. Thus, we formulated

it as an optimization problem with the following steps. As can be seen in Fig. 1, the position of reflection point $\boldsymbol{\kappa}_l$ can be calculated by $\theta_{t,l}$, $\phi_{t,l}$ and $d_{l,1}$ as

$$\boldsymbol{\kappa}_l' = \boldsymbol{s} + d_{l,1} \cdot [\sin\hat{\phi}_{t,l}\cos\hat{\theta}_{t,l}, \sin\hat{\phi}_{t,l}\sin\hat{\theta}_{t,l}, \cos\hat{\phi}_{t,l}]^T, \quad (29)$$

and it can also be derived by $\theta_{r,l,g}$, $\phi_{r,l,g}$ and $d_{l,2}$ as

$$\boldsymbol{\kappa}_l'' = \hat{\boldsymbol{p}} + d_{l,2} \cdot [\sin\hat{\phi}_{r,l,g}\cos\hat{\theta}_{r,l,g}, \sin\hat{\phi}_{r,l,g}\sin\hat{\theta}_{r,l,g}, \cos\hat{\phi}_{r,l,g}]^T. \quad (30)$$

$\hat{\theta}_{r,l,g}$, $\hat{\phi}_{r,l,g}$ is expressed by $\hat{\theta}_{r,l}$, $\hat{\phi}_{r,l}$ and $\boldsymbol{q}$ as introduced in (6) and (8).

Hence, the problem is reformulated as the estimation of $\boldsymbol{q}$ and $d_{l,1}$, which is to minimize the difference between $\boldsymbol{\kappa}_l'$ and $\boldsymbol{\kappa}_l''$ subject to certain constraints as is shown in P.

$$P: \min_{\boldsymbol{q}, d_{l,1}} \sum_{l=1}^{L-1} ||\boldsymbol{\kappa}_l' - \boldsymbol{\kappa}_l''||, \quad (31)$$

$$s.t.$$

$$0 < d_{l,1} < d_l, \quad 31(a)$$

$$0 \leqslant q_1, q_3 < 2\pi, \quad 31(b)$$

$$0 \leqslant q_2 \leqslant \pi, \quad 31(c)$$

$$d_l = d_{l,1} + d_{l,2}. \quad 31(d)$$

To solve P, we propose to use the particle swarm algorithm (PSO) based approach [28]. PSO is a kind of population-based search algorithm. The advantage of PSO is that it can quickly find the optimal solution without searching the entire solution space, and it has low complexity to execute. In [29], the effectiveness of the PSO algorithm, especially for non-convex global optimization has been proved. Generally, it simulates the social behaviour of birds/fishes in nature. During the execution, the individuals' positions are changing with the social tendency of the group. In PSO, each individual, called a particle, benefits from the historical experience of its own and that of the other members when searching for food. In particular, each particle i records the best position it has experienced so far as $pbest_i$, and the best position of its neighbours or the global community as $gbest$. With the iteration, the particle i can update its velocity v_{ij} and position ϖ_{ij} (j-th dimension of optimization vector) through the personal best position and swarm's best position. In problem P, the position of particle i is denoted as $\boldsymbol{\varpi}_i = [q_1, q_2, q_3, d_{1,1}, d_{2,1}, \ldots, d_{L-1,1}] \in \mathbb{R}^{L+2}$. In Algorithm 1, w is the inertia weight, c_1 and c_2 are learning factors while $rand_1$ and $rand_2$ are random numbers between 0 and 1. The iteration is stopped when the $gbest$ is convergent or the maximum iteration number is reached.

Under the OLOS scenario, i.e., the LOS is not available due to obstruction, $\hat{d}_0$ cannot be obtained since (27) is not valid. In this case, we further revise Algorithm 1 to estimate the $\tilde{\boldsymbol{x}}_i = [p_1, p_2, p_3, q_1, q_2, q_3, d_{1,1}, d_{2,1}, \ldots, d_{L-1,1}] \in \mathbb{R}^{L+5}$ with NLOS information.

Algorithm 1: Joint Rigid Body and Reflection Points Localization.

Input: L, $\boldsymbol{s}$, $\hat{\tau}_l$, $\hat{\theta}_{t,l}$, $\hat{\phi}_{t,l}$, $\hat{\theta}_{r,l}$, $\hat{\phi}_{r,l}$
Output: $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{\kappa}_l$

Initializaiton:
Calculate the rigid body position $\hat{\boldsymbol{p}}$ and d_l using (27).
Generate particles $\mathcal{P}_1, \ldots, \mathcal{P}_S \in \mathbb{R}^{L+2}$ under (31a-c).
Generate velocities of each particle $v_{i,j}$.
Loop
Calculate the fitness score of each particle fit_i with (31).
Determine the $pbest_i$ and $gbest$ through the fitness score evaluation of the original swarm.
Update swarm at iteration number k with
(1) $v_{ij}(k+1) = \mathrm{w}v_{ij}(\mathrm{k}) + \mathrm{c}_1\mathrm{rand}_1(\mathrm{pbest}_{ij}(\mathrm{k}) - \varpi_{ij}(\mathrm{k})) + \mathrm{c}_2\mathrm{rand}_2(\mathrm{gbest}_{ij}(\mathrm{k}) - \varpi_{ij}(\mathrm{k}))$
(2) $\varpi_{ij}(k+1) = \varpi_{ij}(k) + v_{ij}(k+1)$.
Update the new fitness for each particle.
Update the $gbest$ through the fitness score evaluation.
End loop

C. COMPLEXITY ANALYSIS

In this subsection, we analyze the complexity of the channel parameter estimation algorithm and joint rigid body and reflection points localization algorithm.

1) CHANNEL PARAMETER ESTIMATION

The complexity in performing (18) is on the order of $O(N_R^2 N_T^2 GN)$ as same as [10]. During the coarse estimation, the coefficients derived from the second layer dictionary approximately take $O(N_b^2 GN)$. Eq. (24) requires $O(NL)$ operations. Consequently, the maximum complexity for H-DCS-SOMP is $L \times O(N_R^2 N_T^2 N_b^2 GN)$. In fine estimation, the complexity is mainly caused by iterations of the gradient descent algorithm. Since the gradient descent algorithm is an iterative process, the complexity is subject to initial values and step size factor ϱ, threshold α. Assume dimension of optimized vector is $\mathcal{N}_1$, the maximum complexity from fine estimation will be $\mathcal{N}_1 log(1/\alpha)/\varrho$.

2) JOINT POSITION ESTIMATION

The calculation of rigid body location in the LOS case is easy to implement since it involves only some basic operations. For rigid body posture estimation and reflection points position estimation, the complexity is mainly from the PSO algorithm. Assume the PSO algorithm iterates n_2 times and in each iteration of the algorithm $\mathcal{N}_3$ particles are updated. The (31) is invoked $\mathcal{N}_4$ times in each iteration. Hence the overall time complexity of the joint position estimation algorithm is $O(\mathcal{N}_2\mathcal{N}_3\mathcal{N}_4)$.

IV. FUNDAMENTAL BOUNDS

In this section, we first derive the Cramér-Rao lower bound (CRLB) of the channel parameters. Then the theoretical lower bounds of rigid body position, orientation, and reflection points position are calculated.

CRLB determines the lower limit of the variance of any unbiased estimator in the parameter estimation problems. The variance of the estimator can only approach the CRLB indefinitely, but cannot be less than it. Therefore, it provides important information for the performance evaluation of algorithms. Here, we derive the fisher information matrix (FIM) and the CRLB for the channel parameters and apply them to investigate rigid body posture and reflection points estimation bounds. To simplify the notation without loss of generality, we consider the case of $G = 1$ which means only one OFDM signal is transmitted. First, we discuss the situation where both LOS and NLOS exist in the channel.

Defining $\boldsymbol{\eta} = [\boldsymbol{\eta}_0, \boldsymbol{\eta}_1, \ldots \boldsymbol{\eta}_{L-1}]^T$ as the unbiased estimation of $\boldsymbol{\eta}_l = [\theta_{t,l}, \theta_{r,l}, \phi_{t,l}, \phi_{r,l}, h_l, \tau_l]^T \in \mathbb{R}^6$ $(l = 0, 1, 2, \ldots, L-1)$, the mean squared error (MSE) is bounded as

$$E_{\boldsymbol{y}|\boldsymbol{\eta}}[(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^T] \geq \boldsymbol{J}_{\boldsymbol{\eta}}^{-1}, \tag{32}$$

where $E[\cdot]$ represents the expectation of $[\cdot]$ and $6L \times 6L$ FIM $\boldsymbol{J}_{\boldsymbol{\eta}}$ is defined as

$$\boldsymbol{J}_{\boldsymbol{\eta}} = E_{\boldsymbol{y}|\boldsymbol{\eta}}\left[-\frac{\partial \ln f(\boldsymbol{y}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^T}\right]. \tag{33}$$

In (33), $f(\boldsymbol{y}|\boldsymbol{\eta})$ is the likelihood function of the vector $\boldsymbol{y}$ which is expressed as (25). Regarding each paths in the channel model, the FIM can be rewritten as (34) and the operator is defined as

$$\boldsymbol{J}_{\boldsymbol{\eta}} = \begin{bmatrix} \boldsymbol{\Psi}(\boldsymbol{\eta}_0, \boldsymbol{\eta}_0) & \cdots & \boldsymbol{\Psi}(\boldsymbol{\eta}_0, \boldsymbol{\eta}_{L-1}) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}(\boldsymbol{\eta}_{L-1}, \boldsymbol{\eta}_0) & \cdots & \boldsymbol{\Psi}(\boldsymbol{\eta}_{L-1}, \boldsymbol{\eta}_{L-1}) \end{bmatrix}, \tag{34}$$

$$\boldsymbol{\Psi}(\boldsymbol{\eta}_{l_1}, \boldsymbol{\eta}_{l_2}) = E_{\boldsymbol{y}|\boldsymbol{\eta}}\left[-\frac{\partial^2 \ln f(\boldsymbol{y}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}_{l_1} \partial \boldsymbol{\eta}_{l_2}^T}\right]. \tag{35}$$

In particular, the sub-matrix in $\boldsymbol{J}_{\boldsymbol{\eta}}$ is the sum of $\boldsymbol{\psi}_n(\boldsymbol{\eta}_{l_1}, \boldsymbol{\eta}_{l_2}) \in \mathbb{C}^{6\times 6}$ across all the subcarriers and

$$\boldsymbol{\Psi}(\boldsymbol{\eta}_{l_1}, \boldsymbol{\eta}_{l_2}) = \sum_{n=0}^{N-1} \boldsymbol{\psi}_n(\boldsymbol{\eta}_{l_1}, \boldsymbol{\eta}_{l_2}), \tag{36}$$

$$\boldsymbol{\psi}_n(\boldsymbol{\eta}_{l_1}, \boldsymbol{\eta}_{l_2}) = \psi_n \begin{bmatrix} (\theta_{t,l_1}, \theta_{t,l_2}) & (\theta_{t,l_1}, \theta_{r,l_2}) & \ldots & (\theta_{t,l_1}, \tau_{l_2}) \\ (\theta_{r,l_1}, \theta_{t,l_2}) & (\theta_{r,l_1}, \theta_{r,l_2}) & \ldots & (\theta_{r,l_1}, \tau_{l_2}) \\ \cdots & \cdots & \cdots & \cdots \\ (\tau_{l_1}, \theta_{t,l_2}) & (\tau_{l_1}, \theta_{r,l_2}) & \ldots & (\tau_{l_1}, \tau_{l_2}) \end{bmatrix}. \tag{37}$$

Taking $\boldsymbol{y}[n] = \boldsymbol{H}[n]\boldsymbol{F}[n]\boldsymbol{x}[n] + \boldsymbol{n}[n]$ into $\boldsymbol{J}_{\boldsymbol{\eta}}$ (note $E_{\boldsymbol{y}|\boldsymbol{\eta}}[\boldsymbol{n}[n]] = 0$), the scalar operator $\psi(x_{l_1}, x_{l_2})$ can be derived as

$$\psi(x_{l_1}, x_{l_2}) = \frac{2}{N_0} \Re\left\{\frac{\partial \boldsymbol{\mu}^H[n]}{\partial x_{l_1}} \frac{\partial \boldsymbol{\mu}[n]}{\partial x_{l_2}}\right\}, \tag{38}$$

where $x_{l_1}, x_{l_2} \in \{\theta_{t,l}, \theta_{r,l}, \phi_{t,l}, \phi_{r,l}, h_l, \tau_l, l = l_1, l_2\}$. A more detailed derivation can be found in APPENDIX A. With the FIM, the CRLB of channel parameters in LOS can be calculated by

$$\begin{aligned} \mathrm{CRLB}(\hat{\theta}_{t,0}) &= \sqrt{\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_1}, \\ \mathrm{CRLB}(\hat{\theta}_{t,0}) &= \sqrt{\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_2}, \\ \mathrm{CRLB}(\hat{\phi}_{r,0}) &= \sqrt{\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_3}, \\ \mathrm{CRLB}(\hat{\phi}_{r,0}) &= \sqrt{\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_4}. \end{aligned} \tag{39}$$

The averages of the CRLBs of AOD/AOA in NLOS are derived in (40) where $\mathrm{diag}[J]_{e_1,e_2,\ldots,e_{l_3}}$ denotes the diagonal matrix comprising of the $e_1, e_2, \ldots, e_{l_3}$-th diagonal elements in J. The CRLB of $\hat{h}_l$ and $\hat{\tau}_l$ can be deduced similarly as

$$\begin{aligned} \mathrm{CRLB}(\hat{\theta}_{t,l}) &= \sqrt{tr\left\{\mathrm{diag}\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_{7,13,\ldots,6L-5}\right\}/(L-1)}, \\ \mathrm{CRLB}(\hat{\theta}_{r,l}) &= \sqrt{tr\left\{\mathrm{diag}\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_{8,14,\ldots,6L-4}\right\}/(L-1)}, \\ \mathrm{CRLB}(\hat{\phi}_{t,l}) &= \sqrt{tr\left\{\mathrm{diag}\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_{9,15,\ldots,6L-3}\right\}/(L-1)}, \\ \mathrm{CRLB}(\hat{\phi}_{r,l}) &= \sqrt{tr\left\{\mathrm{diag}\left[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}\right]_{10,16,\ldots,6L-2}\right\}/(L-1)}. \end{aligned} \tag{40}$$

With the FIM of channel parameters, the CRLB of rigid body position $(\hat{\boldsymbol{p}})$, orientation $(\hat{\boldsymbol{q}})$ and reflection point position $(\hat{\boldsymbol{\kappa}}_l)$ can be derived through the variable transformation tensor T from $\boldsymbol{\eta}$ to $\boldsymbol{\xi} = [\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2, \ldots, \boldsymbol{\kappa}_{L-1}]^T$. Then the FIM of $\boldsymbol{\xi}$ can be expressed as

$$\boldsymbol{J}_{\boldsymbol{\xi}} = \mathbf{T}\boldsymbol{J}_{\boldsymbol{\eta}}\mathbf{T}^{\mathrm{T}}. \tag{41}$$

The transformation matrix T is calculated with

$$\mathrm{T} = \frac{\partial \boldsymbol{\eta}^{\mathrm{T}}}{\partial \boldsymbol{\xi}}. \tag{42}$$

The elements of T are calculated by using the geometric relationships in (43), (6), (7) and (8) as

$$\begin{aligned} \tau_0 &= |\boldsymbol{p} - \boldsymbol{s}|/c, \\ \tau_l &= (|\boldsymbol{p} - \boldsymbol{\kappa}_l| + |\boldsymbol{\kappa}_l - \boldsymbol{s}|)/c, \\ \phi_{t,0} &= \arccos[(p_z - s_z)/|\boldsymbol{p} - \boldsymbol{s}|)], \\ \theta_{t,0} &= \arctan[(p_y - s_y)/(p_x - s_x)], \end{aligned}$$

$$\phi_{t,l} = \arccos(\kappa_{3,l} - s_z)/|\boldsymbol{\kappa}_l - \boldsymbol{s}|),$$
$$\theta_{t,l} = \arctan[(\kappa_{2,l} - s_y)/(\kappa_{1,l} - s_x)]. \quad (43)$$

Here, we structured T analog to (34) by considering each path in the channel model ($[\boldsymbol{p}, \boldsymbol{q}] \subseteq \boldsymbol{\xi}$ is treated as path 0) as

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{1,0} & \cdots & \mathbf{T}_{0,L-1} \\ \vdots & \ddots & \vdots \\ \mathbf{T}_{L-1,0} & \cdots & \mathbf{T}_{L-1,L-1} \end{bmatrix}, \quad (44)$$

$$\mathbf{T}_{w,v} = \frac{\partial \boldsymbol{\eta}_w}{\partial \boldsymbol{\xi}_v}. \quad (45)$$

With the Jacobian matrix notation, it's clear that $T \in \mathbb{C}^{(3L+3)\times 6L}$, and the elements in $T_{w,v}$ can be abbreviated as shown as

$$T_{w,0} = \frac{\partial(\theta_{t,w}, \theta_{r,w}, \phi_{t,w}, \phi_{r,w}, h_w, \tau_w)}{\partial(p_x, p_y, p_z, q_1, q_2, q_3)},$$
$$T_{w,v,v\neq 0} = \frac{\partial(\theta_{t,w}, \theta_{r,w}, \phi_{t,w}, \phi_{r,w}, h_w, \tau_w)}{\partial(\kappa_{1l}, \kappa_{2l}, \kappa_{3l})}. \quad (46)$$

Consequently, the entries for the transmission time and AOD are derived as

$$\partial\tau_0/\partial p_i = (p_i - s_i)/(c|\boldsymbol{p} - \boldsymbol{s}|),\ i = x, y, z,$$
$$\partial\phi_{t,0}/\partial p_i = (p_i - s_i)(p_z - s_z)|\boldsymbol{p} - \boldsymbol{s}|^{-3}/|\sin\phi_{t,0}|,\ i = x, y,$$
$$\partial\phi_{t,0}/\partial p_z = \left[(p_z - s_z)^2|\boldsymbol{p} - \boldsymbol{s}|^{-3} - |\boldsymbol{p} - \boldsymbol{s}|^{-1}\right]/|\sin\phi_{t,0}|,$$
$$\partial\theta_{t,0}/\partial p_x = (p_y - s_y)/[(p_x - s_x)^2(1 + \tan^2\theta_{t,0})],$$
$$\partial\theta_{t,0}/\partial p_y = [(p_x - s_x)(1 + \tan^2\theta_{t,0})]^{-1}, \quad (47)$$

for LOS terms and

$$\partial\tau_l/\partial p_i = (p_i - \kappa_{l,i})/(c|\boldsymbol{p} - \boldsymbol{\kappa}_l|),\ i = x, y, z,$$
$$\partial\tau_l/\partial\kappa_{l,i} = -\partial\tau_l/\partial p_i,\ i = x, y, z,$$
$$\partial\phi_{t,l}/\partial\kappa_{l,i} = (\kappa_{l,i} - s_i)(\kappa_{l,z} - s_z)|\boldsymbol{\kappa}_l - \boldsymbol{s}|^{-3}/|\sin\phi_{t,0}|,$$
$$i = x, y,$$
$$\partial\phi_{t,l}/\partial\kappa_{l,z} = \left[(\kappa_{l,z} - s_z)^2|\boldsymbol{\kappa} - \boldsymbol{s}|^{-3} - |\boldsymbol{\kappa} - \boldsymbol{s}|^{-1}\right]/|\sin\phi_{t,l}|,$$
$$\partial\theta_{t,l}/\partial\kappa_{l,x} = (\kappa_{l,y} - s_y)/[(\kappa_{l,x} - s_x)^2(1 + \tan^2\theta_{t,l})],$$
$$\partial\theta_{t,l}/\partial\kappa_{l,y} = [(\kappa_{l,x} - s_x)(1 + \tan^2\theta_{t,l})]^{-1}, \quad (48)$$

for NLOS terms ($l \neq 0$) while the rest related to AOD and transmission time are zero.

Moreover, the entries regarding the AOA are more complex since the AOA is measured in RBCS. The expression of the same vector at RBCS and GCS is connected by rotation matrix R. Therefore, we introduce the notation $\boldsymbol{\zeta}_l(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{\kappa}_l) = [\cos(\theta_{r,l,g})\sin(\phi_{r,l,g}), \sin(\theta_{r,l,g})\sin(\phi_{r,l,g}), \cos(\phi_{r,l,g})]^T$, and there is

$$\phi_{r,l} = \arccos(R(3,:) \cdot \boldsymbol{\zeta}_l(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{\kappa}_l)),$$
$$\theta_{r,l} = \arctan\left(\frac{R(2,:) \cdot \boldsymbol{\zeta}_l(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{\kappa}_l)}{R(1,:) \cdot \boldsymbol{\zeta}_l(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{\kappa}_l)}\right). \quad (49)$$

The remaining entries in T containing AOA terms can be deduced by (50) (details of the final results are in APPENDIX B) while the rests entries in T are zero.

$$\partial\theta_{r,0}/\partial p_i = (R(3,:) \cdot \partial\boldsymbol{\zeta}_0/\partial p_i)/\left|\sin\theta_{r,0}\right|,$$
$$\partial\theta_{r,l}/\partial q_i = \frac{\partial R(3,:)}{\partial q_i} \cdot \frac{\boldsymbol{\zeta}_l}{\left|\sin\theta_{r,l}\right|},$$
$$\partial\theta_{r,l}/\partial\kappa_{l,i} = (R(3,:) \cdot \partial\boldsymbol{\zeta}_l/\partial\kappa_{l,i})/\left|\sin\theta_{r,l}\right|,$$
$$\partial\phi_{r,0}/\partial p_i = \frac{\partial\left[R(2,:) \cdot \boldsymbol{\zeta}_0/R(1,:) \cdot \boldsymbol{\zeta}_0\right]}{(1 + \tan^2\phi_{r,0})\partial p_i},$$
$$\partial\phi_{r,l}/\partial\kappa_{l,i} = \frac{\partial\left[R(2,:) \cdot \boldsymbol{\zeta}_l/R(1,:) \cdot \boldsymbol{\zeta}_l\right]}{(1 + \tan^2\phi_{r,l})\partial\kappa_{l,i}},$$
$$\partial\phi_{r,l}/\partial q_i = \frac{\partial\left[R(2,:) \cdot \boldsymbol{\zeta}_l/R(1,:) \cdot \boldsymbol{\zeta}_l\right]}{(1 + \tan^2\phi_{r,l})\partial q_i}. \quad (50)$$

Using the FIM $\boldsymbol{J}_{\boldsymbol{\xi}} \in \mathbb{C}^{(3L+3)\times(3L+3)}$, the CRLB of rigid body position and orientation estimation can be obtained as

$$\mathrm{CRLB}(\hat{\boldsymbol{p}}) = \sqrt{\mathrm{tr}\left\{[\boldsymbol{J}_{\boldsymbol{\xi}}^{-1}]_{1:3,1:3}\right\}},$$
$$\mathrm{CRLB}(\hat{\boldsymbol{q}}) = \sqrt{\mathrm{tr}\left\{[\boldsymbol{J}_{\boldsymbol{\xi}}^{-1}]_{4:6,4:6}\right\}}, \quad (51)$$

where $[\cdot]_{e_1:e_2,e_3:e_4}$, means the selection of the $(e_2 - e_1 + 1) \times (e_4 - e_3 + 1)$ submatrix and tr is the trace of the matrix. For the CRLB of reflection points estimation, the average bound of all the NLOS paths is taken as

$$\mathrm{CRLB}(\hat{\boldsymbol{\kappa}}_1) = \sqrt{\mathrm{tr}\left\{[\boldsymbol{J}_{\boldsymbol{\xi}}^{-1}]_{7:3\mathrm{L}+3,7:3\mathrm{L}+3}\right\}}. \quad (52)$$

In the case of OLOS, there is no AOA/AOD for LOS to be detected. Hence, the estimation of the channel becomes $\tilde{\boldsymbol{\eta}} = [\boldsymbol{\eta}_1, \ldots \boldsymbol{\eta}_{L-1}]^T$ while the $\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}$ since the NLOS channel model is still related to all components in $\boldsymbol{\xi}$. The following derivation under OLOS is similar to the case with LOS.

V. SIMULATION RESULTS

In this section, simulation results of the proposed 5G mm-wave rigid body active localization and environment sensing methods are discussed. First, the environmental setup and performance metric are introduced. Second, we present simulation results of the rigid body posture estimation for both NLOS and OLOS situations. Third, the reflection points localization results are presented including the performance comparison of different algorithms. The CRLB as the theoretical optimal solution is also investigated.

A. SIMULATION SETUP

We consider a 3D indoor scenario with the size of $5 \times 5 \times 5$ m^3 in MATLAB 2020b. The location of AN is $\boldsymbol{s} = [0, 0, 0]^T$. To simplify the problem without loss of generality, we randomly set the real position and orientation of rigid body within far-field regime $\boldsymbol{p} = [5, 5, 3]^T$ and

$\boldsymbol{q} = [\pi/6, 2\pi/5, \pi/5]^T$. Reflection points can be located anywhere in the cube indoor environment except the positions on the extension line between AN and the rigid body. The URAs with half-wavelength antennas inter-element spacing on both AN and the rigid body consists of $N_T = N_{Tx} \times N_{Ty} = N_R = N_{Rx} \times N_{Ry} = 8 \times 4$ antennas. We set $L = 5, N = 10, c = 0.3$ m/ns, $B = 100$ MHz, $f_c = 60$ GHz. The number of beams sent is 10 and the average reflection loss for the first order reflection is -10 dB similar to [10] introduced. The SNR represents the signal-to-noise ratio of the signal from $\boldsymbol{s}$. It is related to the channel information and can be defined as

$$\mathrm{SNR} \triangleq \sum_{\mathrm{n}=1}^{\mathrm{N}} \frac{\|\boldsymbol{H}[\mathrm{n}]\boldsymbol{F}[\mathrm{n}]\boldsymbol{x}[\mathrm{n}]\|_2^2}{\|\boldsymbol{n}[\mathrm{n}]\|_2^2}. \tag{53}$$

The root mean square error (RMSE) is used to evaluate the estimation accuracy of the joint rigid body and reflection points localization algorithm. Suppose $\hat{\boldsymbol{\Xi}}_i$ is the estimation of true value $\boldsymbol{\Xi}_i$ in the i-th simulation, there is

$$\mathrm{RMSE}_{\boldsymbol{\Xi}} = \sqrt{\frac{1}{\mathcal{M}} \sum_{\mathrm{i}=1}^{\mathcal{M}} \left| \boldsymbol{\Xi}_{\mathrm{i}} - \hat{\boldsymbol{\Xi}}_{\mathrm{i}} \right|^2}. \tag{54}$$

For the embedded PSO algorithm, the initial population is set to 300 and the inertia weight is 0.8 with both the self-learning factor and group learning factor being 0.5. The total iteration number is set to 50 to ensure convergence.

B. RIGID BODY POSTURE ESTIMATION

Firstly, we evaluate the CDF (Cumulative Distribution Function) of rigid body position and orientation estimation against different compressive sensing methods. Both the rigid body position and posture are denoted by the three dimension vector. The $\boldsymbol{\Xi}_i$ in (54) is the three components of the vector with $\mathcal{M} = 3$. The algorithms are repeated 200 times as is shown in Fig. 3. Here, the DCS-SOMP algorithm is taken from paper [30]. H-DCS-SOMP and our modified H-DCS-SOMP are introduced in Section III. A. In Fig. 3, we set $SNR = 20$ dB and $N_b = 4$. It can be observed that our proposed modified H-DCS-SOMP algorithm has narrower distribution width for both position and orientation estimation compared with the other two algorithms. Moreover, for rigid body position estimation in Fig. 3(a), the DCS-SOMP algorithm can achieve around 0.65 m RMSE with 90% probability, and with the same condition, H-DCS-SOMP can achieve around 0.25 m RMSE and our proposed method can achieve around 0.08 m RMSE. Thus, the performance of the DCS-SOMP algorithm is much more restricted compared with our method. In addition, for rigid body orientation estimation in Fig. 3(b), our proposed method achieves 0.1 rd RMSE with 90% probability, which is higher than the DCS-SOMP method.

In addition, the localization and posture error are investigated when SNR is 0 dB, indicating a low-to-medium SNR case. The CDF is shown in Fig. 4. When SNR is lower, the localization performance using the three algorithms degrades due to the increased noise level. However, it can be observed

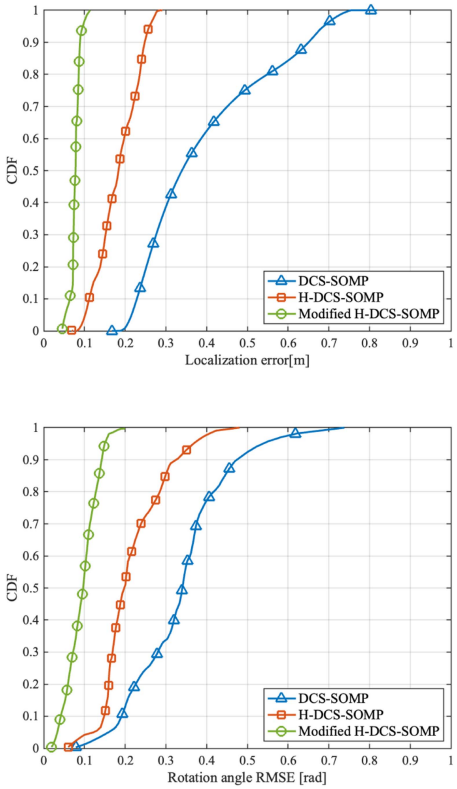

FIGURE 3. The rigid body localization RMSE (top) and rigid body rotation angle RMSE (bottom) of DCS-SOMP, H-DCS-SOMP and modified H-DCS-SOMP vs CDF when SNR is 20 dB.

that the performance degradation of modified H-DCS-SOMP is much smaller than the other two methods. The average error of position and posture decreases less than 0.1 m and 0.1 rd, with distribution less than 0.1 m and 0.25 rd, respectively. This demonstrates the great robustness of proposed method in rigid body position and posture estimation.

To further demonstrate the influence of SNR, we compare the variation of rigid body position and orientation RMSE of the three algorithms under different SNRs from -20 dB to 20 dB as shown in Fig. 5. Instead of using the CDF, all the performance is the average value of 200 simulations except for the CRLB. As can be seen in Fig. 5(a), when $SNR = 0$ dB our proposed method reaches the corresponding theoretical bound. When $SNR = 10$ dB, the RMSE of the rigid body position is less than 0.1 m. Similar to position estimation results, in Fig. 5(b), the orientation RMSE of our method reaches the corresponding theoretical bound when $SNR = 0$ dB. When $SNR = 20$ dB, our proposed algorithm's location and orientation precision could reach 0.04 m and 0.027 rd while the other two algorithms can only achieve 0.072 m and 0.035 rd, and 0.15 m and 0.107 rd. Both the position and orientation results show the effectiveness of our proposed method.

In addition, we also evaluate the performance of our proposed method when there is no LOS path (OLOS situation). It can be observed from Fig. 6, that when $SNR = -10$ dB,

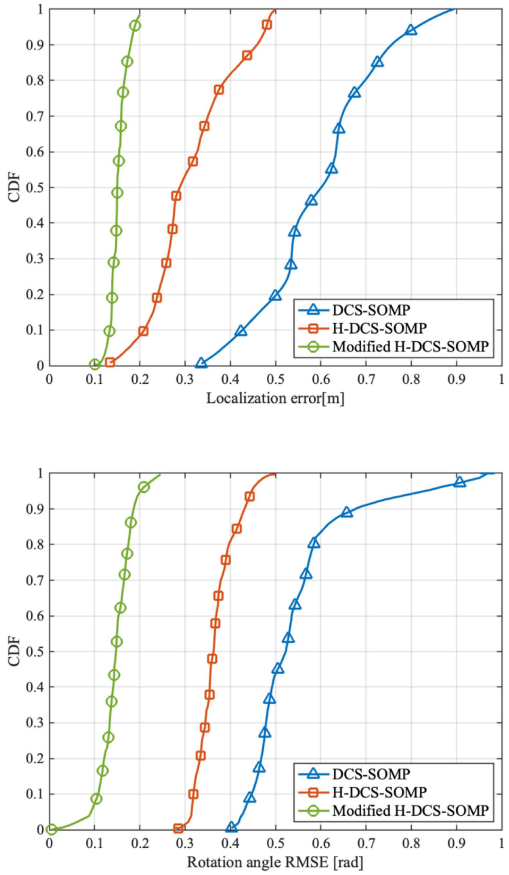

FIGURE 4. The rigid body localization RMSE (top) and rigid body rotation angle RMSE (bottom) of DCS-SOMP, H-DCS-SOMP and modified H-DCS-SOMP vs CDF when SNR is 0 dB.

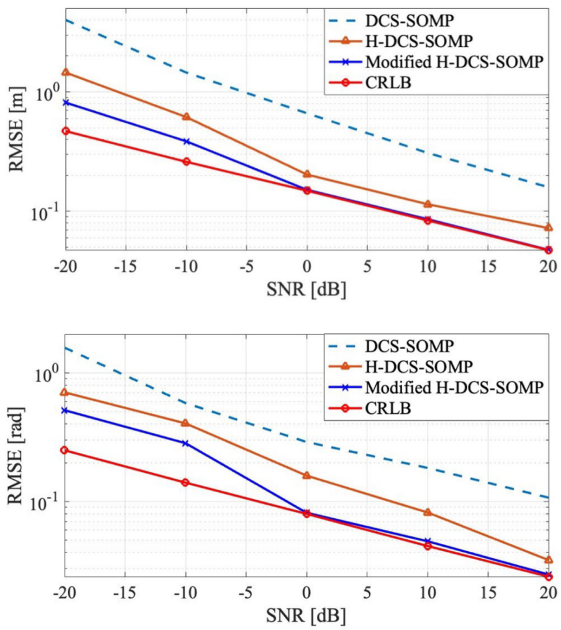

FIGURE 5. The rigid body localization RMSE (top) and rigid body rotation angle RMSE (bottom) of DCS-SOMP, H-DCS-SOMP, modified H-DCS-SOMP and corresponding theoretical bound vs different SNRs.

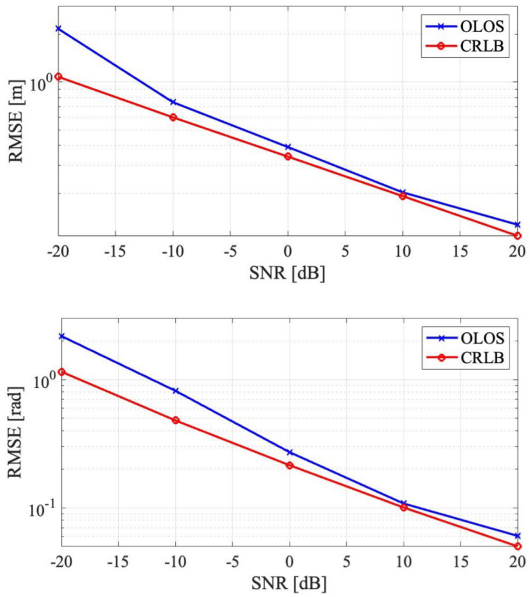

FIGURE 6. In the OLOS situation, the rigid body localization RMSE (top) and rigid body rotation angle RMSE (bottom) of modified H-DCS-SOMP and corresponding theoretical bound vs different SNRs.

position estimation RMSE reaches the corresponding theoretical bound and when $SNR = 0$ dB, orientation estimation RMSE reaches the corresponding theoretical bound. It is also noted that when the LOS is blocked, the theoretical bound of location and orientation estimation is worse than the previous situation due to the less effective information acquisition. In the absence of LOS, only angle and distance information from NLOS can be used to assist rigid body posture estimation. However, the information provided by NLOS has a relatively large error due to the influence of the path loss, which leads to the deterioration of the positioning performance. Compared with Fig. 5, the position and orientation estimation error are doubled. Even though, when the $SNR = 10$ dB, the rigid body localization and orientation estimation can reach decimeter level and 0.1 rd, respectively.

C. REFLECTION POINTS LOCATION ESTIMATION

In this subsection, we evaluate the performance of our proposed reflection points location estimation method. As introduced in Section III. B, the position of reflection points is estimated from the angle and distance of NLOS using 1. Thus, the NLOS channel parameter estimation results are first evaluated. Each NLOS path has four angle parameters (AOA/AOD azimuth and AOA/AOD elevation, $[\theta_{t,l}, \theta_{r,l}, \phi_{t,l}, \phi_{r,l}]^T$) and there are $L-1$ NLOS paths. We calculate the average RMSE and CRLB of each angle among different NLOS paths, respectively. Fig. 7 illustrates the average azimuth angle estimation RMSE of AODs and AOAs of NLOS paths. And the average CRLB is labelled with red lines. It is observed that the RMSE of both elevation angles and azimuth angles reach 0.1 rd accuracy when SNR is above 0 dB. While the SNR is large than -10 dB, the angular estimation converges to the

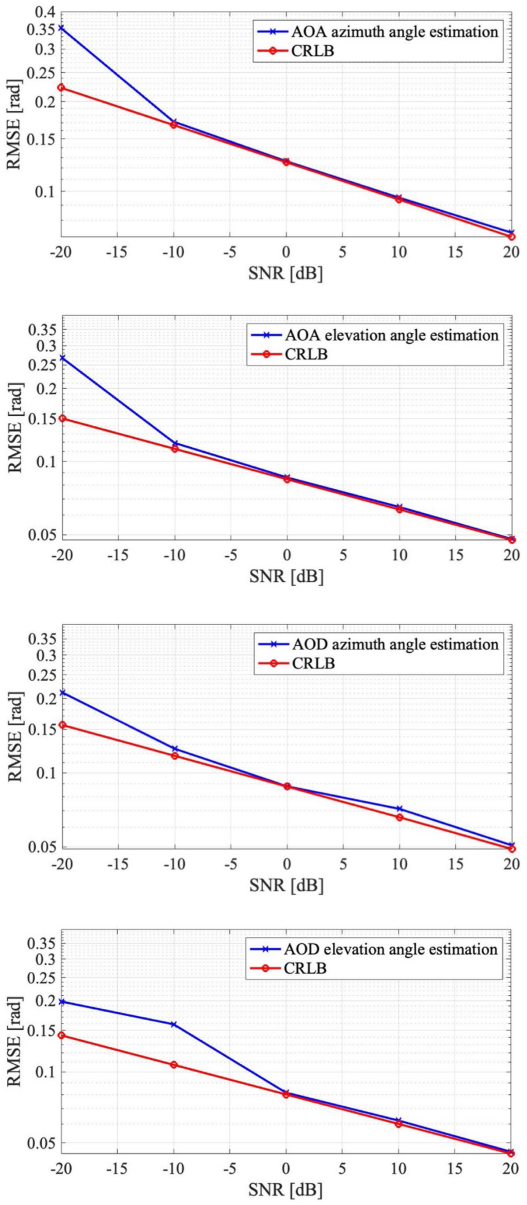

FIGURE 7. The average AOA/AOD RMSE of the NLOS paths and their corresponding theoretical bounds under different SNRs.

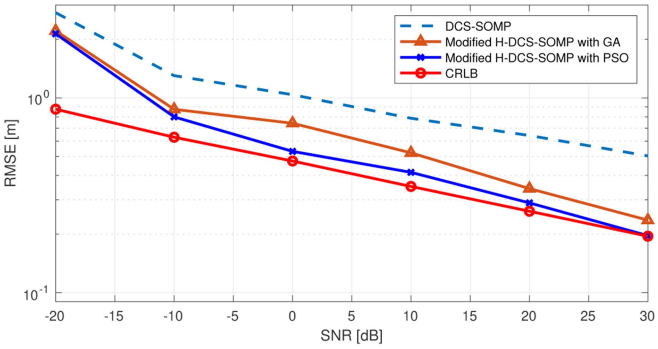

FIGURE 8. The reflection points localization RMSE of DCS-SOMP, modified H-DCS-SOMP with GA method, modified H-DCS-SOMP with PSO method and corresponding theoretical bound vs different SNRs.

theoretical bounds. The results demonstrate the validation of the proposed algorithm to resolve the NLOS channel for the following reflection points localization.

The theoretical lower bound and simulation results of RMSE for the different reflection points localization algorithms under various SNRs are shown in Fig. 8. The well-known genetic algorithm (GA) [31] has also been widely used for optimization problems. Thus, (31) can be potentially solved with the classic GA. Herein, if the optimal solution of (31) is obtained with GA, we denote it as Modified H-DCS-SOMP (GA-based). The cyan dash line is the RMSE of DCS-SOMP algorithm with the PSO algorithm, the orange line with a triangular symbol represents the RMSE of the modified H-DCS-SOMP algorithm with GA while the blue line with a cross symbol is the RMSE of the modified H-DCS-SOMP algorithm with PSO algorithm. The crossover rate and mutation rate are 0.5 and 0.1, respectively in the GA algorithm with an initial population of 500 and a maximal generation of 100. It is observed that in our optimization problem, under the same conditions, the PSO algorithm has better performance than GA and can converge faster to the theoretical lower bound. The performance of the PSO method is improved up to 40% compared with GA. The explanation is that, for the global optimization of continuous variables, the memory parameters *gbest* and $pbest_i$ of PSO make it easier to find the optimal solution while the GA-based method doesn't track the optimal solutions in the generations. Moreover, under restricted iteration number/generations, the PSO method costs less than the GA-based method to get converged. Thus, the PSO-based Algorithm 1 is much more preferred when the time resource is limited.

VI. CONCLUSION

In this paper, a rigid body active localization and environment sensing scheme with the assistance of 5G mm-wave MIMO is proposed. A novel hierarchical compressive sensing algorithm refined by an iterative maximum likelihood step for channel angle and distance estimation is presented. Furthermore, we joint consider the rigid body and reflection points localization problem and a PSO-based optimization algorithm is used to estimate the posture of the rigid body and position of reflection points. We also calculate the theoretical bounds (CRLB) on the rigid body posture, NLOS angular and reflection points position estimate uncertainties. Compared to the traditional rigid body localization approaches, our proposed method using only one anchor node can achieve centimeter-level rigid body posture estimation accuracy under both NLOS and OLOS situations. Simulation results reveal that the rigid body posture and reflection points position estimation of our proposed algorithm approach the corresponding bounds with an increase in SNR value. The results also demonstrate that the proposed scheme can achieve high performance for both rigid body localization and reflection points estimation when SNR reaches 20 dB.

APPENDIX A
THE FIM OF CHANNEL PARAMETERS IN (38)

The expansion of FIM is dependent on the partial derivatives of the channel matrix concerning the channel parameters. The elements in (38) are expressed as

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial \tau_l} = \frac{\boldsymbol{A}_R[n]\partial(\boldsymbol{\Gamma}[n])\boldsymbol{A}_T^H[n]}{\partial \tau_l}\boldsymbol{F}[n]\boldsymbol{x}[n], \tag{55}$$

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial \theta_{t,l}} = \frac{\boldsymbol{A}_R[n]\boldsymbol{\Gamma}[n]\partial(\boldsymbol{A}_T^H[n])}{\partial \theta_{t,l}}\boldsymbol{F}[n]\boldsymbol{x}[n], \tag{56}$$

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial \phi_{t,l}} = \frac{\boldsymbol{A}_R[n]\boldsymbol{\Gamma}[n]\partial(\boldsymbol{A}_T^H[n])}{\partial \phi_{t,l}}\boldsymbol{F}[n]\boldsymbol{x}[n], \tag{57}$$

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial \theta_{r,l}} = \frac{\partial(\boldsymbol{A}_R[n])\boldsymbol{\Gamma}[n]\boldsymbol{A}_T^H[n]}{\partial \theta_{r,l}}\boldsymbol{F}[n]\boldsymbol{x}[n], \tag{58}$$

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial \phi_{r,l}} = \frac{\partial(\boldsymbol{A}_R[n])\boldsymbol{\Gamma}[n]\boldsymbol{A}_T^H[n]}{\partial \phi_{r,l}}\boldsymbol{F}[n]\boldsymbol{x}[n], \tag{59}$$

$$\frac{\partial \boldsymbol{\mu}[n]}{\partial h_l} = \frac{\boldsymbol{A}_R[n]\partial(\boldsymbol{\Gamma}[n])\boldsymbol{A}_T^H[n]}{\partial h_l}\boldsymbol{F}[n]\boldsymbol{x}[n]. \tag{60}$$

Using the differential property of the Kronecker product, the derivative terms in (55–60) can be expanded with the following notations

$$\begin{aligned}
\mathcal{A}_x &= \left[-\frac{N_{Tx}-1}{2}, \ldots, \frac{N_{Tx}-1}{2}\right], \\
\mathcal{A}_y &= \left[-\frac{N_{Ty}-1}{2}, \ldots, \frac{N_{Ty}-1}{2}\right],
\end{aligned} \tag{61}$$

$$\frac{\partial \boldsymbol{\Gamma}[n]}{\partial \tau_l} = \mathrm{diag}(0, 0, \ldots, \sqrt{N_R N_T}\frac{-j2\pi n h_l}{NT_s\sqrt{\rho_l}} e^{\frac{-j2\pi n\tau_l}{NT_s}}, \ldots, 0), \tag{62}$$

$$\frac{\partial \boldsymbol{\Gamma}[n]}{\partial h_l} = \mathrm{diag}\left(0, 0, \ldots, \frac{1}{\sqrt{\rho_l}} e^{\frac{-j2\pi n\tau_l}{NT_s}}, \ldots, 0\right), \tag{63}$$

$$\frac{\partial(\boldsymbol{A}_T^H[n])}{\partial \theta_{t,l}} = \left[0, 0, \ldots, \frac{\partial \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l})}{\partial \theta_{t,l}}, \ldots, 0\right], \tag{64}$$

$$\frac{\partial \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l})}{\partial \theta_{t,l}} = j\frac{2\pi}{\lambda_n} d_o \sin(\phi_{t,l})\mathcal{D}_1 \cdot \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l}), \tag{65}$$

$$\mathcal{D}_1 = \begin{vmatrix} \cos\theta_{\mathrm{t},1} & \sin\theta_{t,l} \\ \mathcal{A}_x & \mathcal{A}_y \end{vmatrix}, \tag{66}$$

$$\frac{\partial(\boldsymbol{A}_T^H[n])}{\partial \phi_{t,l}} = \left[0, 0, \ldots, \frac{\partial \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l})}{\partial \phi_{t,l}}, \ldots, 0\right], \tag{67}$$

$$\frac{\partial \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l})}{\partial \phi_{t,l}} = j\frac{2\pi}{\lambda_n} d_o \cos(\phi_{t,l})\mathcal{D}_2 \cdot \boldsymbol{a}_{t,n}(\theta_{t,l}, \phi_{t,l}), \tag{68}$$

$$\mathcal{D}_2 = \begin{vmatrix} \sin\theta_{\mathrm{t},1} & -\cos\theta_{t,l} \\ \mathcal{A}_x & \mathcal{A}_y \end{vmatrix}. \tag{69}$$

The construction of $\frac{\partial(\boldsymbol{A}_R[n])}{\partial \theta_{r,l}}$ and $\frac{\partial(\boldsymbol{A}_R[n])}{\partial \phi_{r,l}}$ are similiar to (64) and (67), by changing the transmitter parameters into receiver parameters.

APPENDIX B
THE ENTRIES IN (50)

Using the derivative rules of the dot product, we can derive the following equations

$$\frac{\partial \boldsymbol{\zeta}_0}{\partial p_i} = \frac{-p_i\,|\boldsymbol{p}-\boldsymbol{s}| + (p_i - s_i)\,|\boldsymbol{p}-\boldsymbol{s}|^{-1}}{|\boldsymbol{p}-\boldsymbol{s}|^2}, \tag{70}$$

$$\frac{\partial \boldsymbol{\zeta}_l}{\partial \kappa_i} = \frac{-p_i\,|\boldsymbol{p}-\boldsymbol{\kappa}| + (\kappa_i - s_i)\,|\boldsymbol{p}-\boldsymbol{\kappa}|^{-1}}{|\boldsymbol{p}-\boldsymbol{\kappa}|^2}, \tag{71}$$

$$\frac{\partial\,[R(2,:)\cdot\boldsymbol{\zeta}_0/R(1,:)\cdot\boldsymbol{\zeta}_0]}{\partial p_i} = \frac{1}{(R(1,:)\cdot\boldsymbol{\zeta}_0)^2}\mathcal{D}_3 \cdot \frac{\partial \boldsymbol{\zeta}_0}{\partial p_i}, \tag{72}$$

$$\mathcal{D}_3 = \begin{vmatrix} R(1,:)\cdot\boldsymbol{\zeta}_0 & R(2,:)\cdot\boldsymbol{\zeta}_0 \\ R(1,:) & R(2,:) \end{vmatrix}, \tag{73}$$

$$\frac{\partial\,[R(2,:)\cdot\boldsymbol{\zeta}_l/R(1,:)\cdot\boldsymbol{\zeta}_l]}{\partial \kappa_i} = \frac{1}{(R(1,:)\cdot\boldsymbol{\zeta}_l)^2}\mathcal{D}_{4,l} \cdot \frac{\partial \boldsymbol{\zeta}_l}{\partial \kappa_i}, \tag{74}$$

$$\mathcal{D}_{4,l} = \begin{vmatrix} R(1,:)\cdot\boldsymbol{\zeta}_l & R(2,:)\cdot\boldsymbol{\zeta}_l \\ R(1,:) & R(2,:) \end{vmatrix}, \tag{75}$$

$$\frac{\partial\,[R(2,:)\cdot\boldsymbol{\zeta}_l/R(1,:)\cdot\boldsymbol{\zeta}_l]}{\partial q_i} = \frac{1}{(R(1,:)\cdot\boldsymbol{\zeta}_l)^2}\mathcal{D}_{5,l}\cdot\boldsymbol{\zeta}_l, \tag{76}$$

$$\mathcal{D}_{5,l} = \begin{vmatrix} R(1,:)\cdot\boldsymbol{\zeta}_l & R(2,:)\cdot\boldsymbol{\zeta}_l \\ \partial R(1,:)/\partial q_i & \partial R(2,:)/\partial q_i \end{vmatrix}, \tag{77}$$

$$\begin{aligned}
\frac{\partial R(1,:)}{\partial \boldsymbol{q}} &= \begin{bmatrix} 0 & -s_1 s_3 + c_3 s_2 c_1 & s_3 c_1 - c_3 s_2 s_1 \\ 0 & -c_3 s_1 + c_1 s_2 s_3 & -c_1 c_3 - s_1 s_2 s_3 \\ 0 & c_1 c_2 & -s_1 c_2 \end{bmatrix}^T, \\
\frac{\partial R(2,:)}{\partial \boldsymbol{q}} &= \begin{bmatrix} -s_2 c_3 & s_1 s_3 + c_3 c_2 s_1 & c_3 c_2 c_1 \\ -s_2 s_3 & s_3 c_2 s_1 & s_3 c_2 c_1 \\ -c_2 & -s_2 s_1 & -s_2 c_1 \end{bmatrix}^T, \\
\frac{\partial R(3,:)}{\partial \boldsymbol{q}} &= \begin{bmatrix} 0 & c_1 c_2 & -s_1 c_2 \\ -c_2 & -s_2 s_1 & -s_2 c_1 \\ 0 & 0 & 0 \end{bmatrix}^T.
\end{aligned} \tag{78}$$

and the notations in this part are the same as (8).

REFERENCES

[1] J. A. del Peral-Rosado et al., "Whitepaper on new localization methods for 5G wireless systems and the Internet-of-Things," COST Action CA15104, IRACON, White Paper, Apr. 2018. [Online]. Available: http://www.iracon.org/wp-content/uploads/2018/03/IRACON-WP2.pdf

[2] Y. Wang, G. Wang, S. Chen, K. C. Ho, and L. Huang, "An investigation and solution of angle based rigid body localization," *IEEE Trans. Signal Process.*, vol. 68, no. 9, pp. 5457–5472, Sep. 2020.

[3] X. Shen, H. Zheng, and X. Feng, "A novel FMCW radar-based scheme for indoor localization and trajectory tracking," in *Proc. IEEE 6th Int. Conf. Comput. Commun.*, 2020, pp. 298–303.

[4] J. Quenzel et al., "Autonomous MAV-based indoor chimney inspection with 3D laser localization and textured surface reconstruction," *J. Intell. Robot. Syst.*, vol. 93, no. 1–2, pp. 317–335, 2019.
[5] N. Q. Pham, K. Mekonnen, E. Tangdiongga, A. Mefleh, and T. Koonen, "User localization and upstream signalling for beam-steered infrared light communication system," *IEEE Photon. Technol. Lett.*, vol. 33, no. 11, pp. 545–548, Jun. 2021.
[6] J. Wang et al., "An ultra-low power, small size and high precision indoor localization system based on MEMS ultrasonic transducer chips," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 69, no. 4, pp. 1469–1477, Apr. 2022.
[7] G. Wang, W. Zhu, and N. Ansari, "Robust TDOA-based localization for IoT via joint source position and NLOS error estimation," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8529–8541, Oct. 2019.
[8] Y. Wang, Y. Wu, and Y. Shen, "Joint spatiotemporal multipath mitigation in large-scale array localization," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 783–797, Feb. 2019.
[9] F. Yin et al., "FedLoc: Federated learning framework for data-driven cooperative localization and location data processing," *IEEE Open J. Signal Process.*, vol. 1, no. 11, pp. 187–215, Nov. 2020.
[10] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.
[11] K. Witrisal et al., "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.
[12] J. Talvitie, M. Valkama, G. Destino, and H. Wymeersch, "Novel algorithms for high-accuracy joint position and orientation estimation in 5G mmwave systems," in *Proc. IEEE Globecom Workshops*, 2017, pp. 1–7.
[13] Z. Abu-Shaban, H. Wymeersch, T. Abhayapala, and G. Seco-Granados, "Single-anchor two-way localization bounds for 5G mmWave systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6388–6400, Jun. 2020.
[14] S. P. Chepuri, G. Leus, and A. van der Veen, "Rigid body localization using sensor networks," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4911–4924, Sep. 2014.
[15] J. Jiang, G. Wang, and K. C. Ho, "Accurate rigid body localization via semidefinite relaxation using range measurements," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 378–382, Mar. 2018.
[16] J. Jiang, G. Wang, and K. C. Ho, "Sensor network-based rigid body localization via semi-definite relaxation using arrival time and doppler measurements," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1011–1025, Feb. 2019.
[17] B. Zhou, M. Zhang, Y. -Q. Chen, N. Xiong, Y. Tian, and S. Ahmed, "Efficient AoA-based rigid body localization via single base station for Internet of Things applications," *IEEE Access*, vol. 7, pp. 171140–171152, 2019.
[18] S. W. Chen, C. K. Seow, and S. Y. Tan, "Elliptical lagrange-based NLOS tracking localization scheme," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3212–3225, May 2016.
[19] A. Abolfathi Momtaz, F. Behnia, R. Amiri, and F. Marvasti, "NLOS identification in range-based source localization: Statistical approach," *IEEE Sensors J.*, vol. 18, no. 9, pp. 3745–3751, May 2018.
[20] Y. Wang et al., "Single base station positioning based on multipath parameter clustering in NLOS environment," *EURASIP J. Adv. Signal Process*, vol. 20, p. 18, 2021, doi: https://doi.org/10.1186/s13634-021-00729-3.
[21] S. Al-Jazzar, J. Caffery, and H. -R. You, "Scattering-model-Based methods for TOA location in NLOS environments," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 583–593, Mar. 2007.
[22] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, "5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Jun. 2020.
[23] J. Talvitie, M. Koivisto, T. Levanen, M. Valkama, G. Destino, and H. Wymeersch, "High-accuracy joint position and orientation estimation in sparse 5G mmWave channel," in *Proc. IEEE Int. Conf. Commun.*, 2019, pp. 1–7.
[24] D. Needell and J. A. Tropp, "COSAMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.
[25] C. K. Anjinappa, Y. Zhou, Y. Yapici, D. Baron, and I. Guvenc, "Channel estimation in mmWave hybrid MIMO system via off-grid Dirichlet kernels," in *Proc. IEEE Glob. Commun. Conf.*, 2019, pp. 1–6.
[26] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3959–3973, Jun. 2018.
[27] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag, 1994.
[28] E. H. Houssein et al., "Major advances in particle swarm optimization: Theory, analysis, and application," *Swarm Evol. Comput.*, vol. 63, 2021, Art. no. 100868.
[29] S. Scott Hayward and E. Garcia-Palacios, "Multimedia resource allocation in mmwave 5G networks," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 240–247, Jun. 2015.
[30] M. F. Duarte et al., "Distributed compressed sensing of jointly sparse signals," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.*, 2005, pp. 1537–1541.
[31] S. Mirjalili, "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks*. Cham, Switzerland: Springer, 2019, pp. 43–55.

BIWEI LI (Graduate Student Member, IEEE) received the B.S. degree in 2017, and the M.Eng. degree in 2019, from the University of Science and Technology Beijing, China. She is currently working towards the Ph.D. degree in electrical and computer engineering with Western University, London, ON, Canada. Her research interests include wireless localization, resource allocation, and integrated localization, sensing and communication systems.

XIANBIN WANG (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2001. He is currently a Professor and a Tier-1 Canada Research Chair with Western University, London, Canada. Prior to joining Western University, he was with the Communications Research Centre Canada as a Research Scientist/Senior Research Scientist from 2002 to 2007. From 2001 to 2002, he was a System Designer with STMicroelectronics, Geneva, Switzerland. He has more than 500 highly cited journals and conference papers, in addition to 30 granted and pending patents and several standard contributions. His research interests include 5G/6G technologies, Internet of Things, communications security, machine learning, and intelligent communications. Dr. Wang is a Fellow of the Canadian Academy of Engineering, Ottawa, ON, Canada, and Engineering Institute of Canada, Ottawa. He was the recipient of the many prestigious awards and recognitions, including the IEEE Canada R. A. Fessenden Award, Canada Research Chair, Engineering Research Excellence Award at Western University, Canadian Federal Government Public Service Award, Ontario Early Researcher Award, and six IEEE best paper awards. He was involved in many IEEE conferences, including GLOBECOM, ICC, VTC, PIMRC, WCNC, CCECE, and CWIT, in different roles, such as General Chair, Symposium Chair, Tutorial Instructor, Track Chair, Session Chair, TPC Co-Chair, and Keynote Speaker. He is/was the Editor-in-Chief, Associate Editor-in-Chief, and Editor/Associate Editor for more than ten journals. He was the Chair of the IEEE ComSoc Signal Processing and Computing for Communications (SPCC) Technical Committee. He is currently the Central Area Chair of IEEE Canada.

EDWARD AU (Senior Member, IEEE) received the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2007. He is a Distinguished Engineer with Huawei Technologies Canada Company, Ltd., Markham, ON, Canada, where he works on product certification, test bed development, and standardization of Wi-Fi. He is currently the Chair of IEEE Technical Activities Board Committee on Standards that oversees the advancement and evangelization of technical standards activities across all IEEE Societies and Councils. He is also the Chair of Computer Society Standards Activities Board standards committee that manage working groups and standards development projects that are under the purview of the Society and for new lines of work or for standards for which a viable standards committee has not been identified. He also actively participates and contributes to IEEE 802 LAN/MAN Standards Committee. He is the Chair of IEEE 802.18 Radio Regulatory Technical Advisory Group supports the work of the IEEE 802 LAN/MAN Standards Committee by actively monitoring and participating in radio regulatory matters worldwide as an advocate for IEEE 802. He is also the Founding Chair of the IEEE 802.11ay, which is aimed at developing an amendment for next generation unlicensed millimeter wave technologies for Wi-Fi, a co-editor of the IEEE 802.11 maintenance task group, which is chartered to develop revisions for the IEEE 802.11 standards, and the Editor of the IEEE 802.11be, which focuses on developing a new major amendment for next generation Wi-Fi (a.k.a. Wi-Fi 7). He was a recipient of IEEE SA Standards Medallion For exceptional leadership and skill in driving the development of IEEE 802.11 Wireless LAN standards in 2021. He is a Senior Editor of IEEE Vehicular Technology Magazine, the Editor of IEEE Communications Standards Magazine, IEEE Open Journal of Vehicular Technology, and IEEE TRANSACTIONS ON COMMUNICATIONS. He was the recipient of the IEEE Transactions on Vehicular Technology Top Editor Award in recognition of outstanding performance for articles reviews in both quantity and quality for the journal, in 2013 and 2016.

YAN XIN received the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2002. He is currently a Principal Engineer with Huawei Technologies Canada Company Ltd., Ottawa, ON, Canada, where he works on research and standardization of wireless communications. Prior to joining Huawei Canada, he was a Senior Engineer with BlackBerry and Samsung Electronics. He has been extensively involved in 3GPP and IEEE 802.11 standardization activities, and was a Vice Chair of IEEE 802.11 TGay and the Technical Editor of IEEE 802.11 Full Duplex TIG. He has numerous granted U.S. patents and U.S. patent applications in pending. He has authored and coauthored more than 40 refereed international journals and conference papers. His research interests include physical layer development of wireless communication, signal processing, coding and modulation, MAC protocols and hardware implementation of communication systems.