

Reconfigurable Intelligent Surface Aided Position and Orientation Estimation Based on Joint Beamforming With Limited Feedback

KUNLUN LI¹, MOHAMMED EL-HAJJAR¹, (Senior Member, IEEE), AND LIE-LIANG YANG¹, (Fellow, IEEE)

School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K.

CORRESPONDING AUTHOR: L.-L. YANG (e-mail: lly@ecs.soton.ac.uk)

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Project under Grant EP/X01228X/1.

ABSTRACT The sparsity of millimeter wave (mmWave) channels in angular and temporal domains has been exploited for channel estimation, while the associated channel parameters can be utilized for localization. However, line-of-sight (LoS) blockage makes localization highly challenging, which may lead to big positioning inaccuracy. One promising solution is to employ reconfigurable intelligent surfaces (RIS) to generate virtual line-of-sight (VLoS) paths. Hence, it is essential to investigate the wireless positioning in RIS-aided mmWave systems. In this paper, an adaptive joint LoS and VLoS localization scheme is proposed, where the VLoS is constructed by a beamforming protocol operated between RIS and mobile station (MS). More specifically, to sense the location and orientation of a MS, a novel interlaced scanning beam sweeping algorithm is proposed to acquire the optimal beams. In this algorithm, LoS and VLoS paths are separately estimated and the optimal beams are selected according to the received signal strength so as to mitigate the LoS blockage problem. Then, based on the selected beams and received signal strength, angle of arrival (AoA), angle of departure (AoD), angle of reflection (AoR) and time of arrival (ToA) are estimated. Finally, with the aid of these estimated parameters, the location and orientation of the MS are estimated. We derive the Cramer-Rao lower bounds (CRLBs) for both location estimation and orientation estimation, and compare them with the corresponding results obtained from simulations. We compare the performance of our proposed scheme with that attained by three legacy schemes, when various aspects are considered. The performance results show the superiority of our proposed beam training algorithm, which is capable of achieving a localization error within 15 cm and an orientation error within 0.003 rads. Furthermore, the training overhead is 100 times less than that of the conventional exhaustive search algorithm, while obtaining a 13 dBm power gain when compared with the hierarchical codebook based search algorithm.

INDEX TERMS mmWave, localization, positioning, reconfigurable intelligent surfaces, beamforming.

I. INTRODUCTION

A. MOTIVATION AND BACKGROUND

LOCALIZATION has sparked a lot of research interest in recent years, owing to the numerous applications that benefit from localization information, such as intelligent transportation systems, unmanned aerial vehicles (UAVs), and so on [1]. However, the conventional localization systems, such as the global positioning system (GPS), are unable to provide high accuracy and also suffer from high latency [2], while they have poor coverage for indoor scenarios. On the other hand, millimeter wave (mmWave) based

localization techniques have attracted wide research interest in recent years, owing to the high beamspace resolution by employing large antenna arrays [3]. Nevertheless, mmWave communications have numerous challenges, including high path-loss, which can only be mitigated by employing beamforming using massive antenna array [3], [4].

Reconfigurable intelligent surfaces (RIS) have recently attracted intensive research attention to control the propagation environments [5], [6], [7]. RIS is a surface composed of a large number of controllable passive reflective elements, placed between the transmitter and receiver. Since

each element of RIS can independently change the phase (or/and) amplitude of the incident signal, RIS can be used to improve the communications energy efficiency, spectrum efficiency, positioning accuracy and even the communication security [8], [9]. However, the channel estimation of the cascaded channels in the RIS aided systems is a challenging task [10], [11], [12], [13], [14], but the concept of passive beamforming as well as joint active and passive beamforming considered in [10], [11] can be used in localization.

Benefiting from the high angular and temporal resolution, mmWaves have a great potential for accurate channel estimation and localization [15], [16], [17], [18], [19], [20], [21]. In [15], an adaptive codebook-based channel estimation algorithm was proposed, which uses multi-stage codebooks to achieve angular estimation, so as to attain a good trade-off between performance and implementation complexity in the conventional mmWave systems. However, when line-of-sight (LoS) path is blocked, the non-line-of-sight (NLoS) path based localization has low accuracy due to the significant reflection loss. For example, in [16], it was shown that the LoS relied localization outperforms significantly the NLoS base localization in the conventional mmWave systems. In [20], the LoS path was used to estimate the location of MS, and after the estimation of the MS's location, the locations of the scatterers are also estimated. The studies show that the performance of the LoS-based localization is much better than that of the NLoS based localization. Therefore, in the case of LoS blockage, the RIS aided localization may provide a promising alternative to improve the localization accuracy in mmWave systems. Furthermore, in [21], a successive localization and beamforming scheme was proposed to achieve the joint localization of user equipment (UE) and channel estimation in the mmWave MIMO systems without employing RIS.

In literature, there are a range of researches on the RIS-aided localization. In [22], the concept of continuous intelligent surface (CIS) was introduced, where the limits of the RIS-aided localization and communications system were presented. More specifically, in [22], a general signal model for the RIS-aided localization and communication system considering both far and near-field scenarios was studied, where both CIS and discrete intelligent surface (DIS) were introduced to improve the localization accuracy and spectral-efficiency via phase response optimization. In [23], the holographic network localization and navigation (NLN) was studied, where RISs were operated in the controlled electromagnetic environment attained by using polarization and specific antenna patterns. It was shown that the robustness of the holographic localization to obstructions can be improved with the aid of RISs. In [24], a localization scheme aided by the adaptive beamforming design using the hierarchical codebook algorithm was proposed for the RIS-aided joint localization and communication, when assuming that the LoS path was not available. In [25], a received signal strength (RSS) based positioning scheme was studied with a RIS-aided mmWave system. Moreover, in [26], a

multi-RIS aided multiple user beam training scheme was proposed for the Terahertz system, where a ternary-tree hierarchical codebook based algorithm was employed to reduce the complexity.

In the RIS-aided localization systems considered in literature, performance bounds have often been analyzed in order to demonstrate the achievable performance of the proposed localization schemes. Specifically in [27], the Cramer-Rao lower bound (CRLB) for the RIS-aided mmWave positioning was analyzed, where the impacts of the number of RIS elements and the design of phase shifters were studied. In this study, the optimal phases of RIS were assumed to be known and thus, the Fisher information matrix (FIM) does not contain the parameters of the reflection angle. Furthermore, the studies in [27] assumed the far-field communication environment. The authors of [28] analyzed the FIM and CRLB in a 3D scenario, where the RIS with unknown phase shift is assumed, and also compared the centralized and distributed deployments of RIS. In [29], a Bayesian framework for localization in the RIS-aided systems was developed, where the uncertainty was investigated in both the near- and far-field cases. In particular, the authors derived the Bayesian equivalent Fisher information matrix (EFIM), showing that, in the far-field case, the RIS orientation offset is unable to be corrected, when there is an unknown phase offset in the received signal. By contrast, the unknown phase offset has no effect in the near-field case. Furthermore, in [30], the Bayesian bounds for the localization of a UE with the aid of multiple RISs were derived, and it was demonstrated that the RIS related geometric channel parameters become lost, when the complex channel gains are not known to the UE. In [31], the authors analyzed the CRLBs of the active RIS-aided localization systems, where multiple signal transmissions and particle filtering algorithm were proposed to improve the signal power of the RIS-aided path. In [32], the CRLB of an indoor localization scheme in ultra-wide band (UWB) system was provided. In [33], the CRLB of the localization and orientation estimation was studied in association with a RIS-aided mmWave system. Furthermore, in [34], the downlink RIS-aided localization was studied, when assuming that LoS is obstructed. To improve the estimation performance, in [34], an alternative optimization and a gradient decent method were proposed to optimize the reflected beamforming and to minimize the CRLB.

In [35], the LoS blockage problem was considered and the effect of LoS blockage rate on the positioning was studied. In mmWave systems, the codebook based channel estimation has widely been studied [36], [37], [38]. However, these approaches cannot be directly applied to the RIS-aided systems, since RIS is regarded as the nearly passive metasurface.¹ For this sake, the codebook based techniques need to be revisited.

1. Note that, we use the term of 'nearly passive' instead of 'fully passive,' because a RIS also needs to consume a certain amount of power/overhead for configuration.

B. RESEARCH PROBLEM AND CONTRIBUTIONS

In this paper, we study the positioning and orientation estimation of a mobile station (MS) in the RIS-aided mmWave system with orthogonal frequency division multiplexing (OFDM) signaling. We set BS and MS to employ hybrid beamforming, while RIS to employ the nearly passive beamforming. Under these settings, the localization and orientation estimation problem can be treated as a downlink joint active and passive beamforming problem with limited feedback. Hence, our research objectives include to find the efficient methods to solve this problem, study the achievable performance and the associated training overhead, as well as to analyze the error bounds of the position and orientation estimation. More explicitly, our contributions can be summarized as follows.

- The position and orientation of a MS in the downlink OFDM mmWave system are estimated, when assuming that both BS and MS employ hybrid beamforming, while RIS is a nearly passive meta-surface, and the status of the LoS link between BS and MS is unknown. To estimate the location and orientation of the MS, the channel parameters of the LoS, and of the path reflected via RIS, which is referred to as the virtual line-of-sight (VLoS) path, are first estimated. In order to maximize the received signal power, we assume limited feedback from MS to RIS controller and from MS to BS for the sake of near optimal beam selection. After the estimation of the channel parameters, including angle of arrival (AoA), the angle of departure (AoD) of LoS path, the angle of reflection (AoR) of VLoS path, and time of arrival (ToA), the location and orientation of MS are then estimated.
- A novel beam sweeping algorithm is proposed and studied, which is capable of achieving a promising trade-off between localization accuracy and training overhead, when compared with the conventional beam sweeping algorithms, such as the exhaustive search (ES), hierarchical codebook (HC), and partial search (PS) algorithms.
- The CRLBs for both the location estimation and orientation estimation are derived, which are compared with the performance obtained from simulations for the proposed beam sweeping algorithm.
- In terms of the performance study and comparison, the effects of the codebook resolution, LoS blockage are investigated. Our studies and simulation results show that the proposed beam sweeping algorithm can achieve the accuracy of 15 cm in localization estimation and of 0.003 rads in orientation estimation. The training overhead is 100 times less than the conventional exhaustive search algorithm, while it can obtain 13 dBm of power gain when compared with the hierarchical codebook algorithm.

C. ORGANIZATION OF THE PAPER AND NOTATIONS

The rest of the paper is organized as follows. Section II introduces the system model for localization. Section III derives

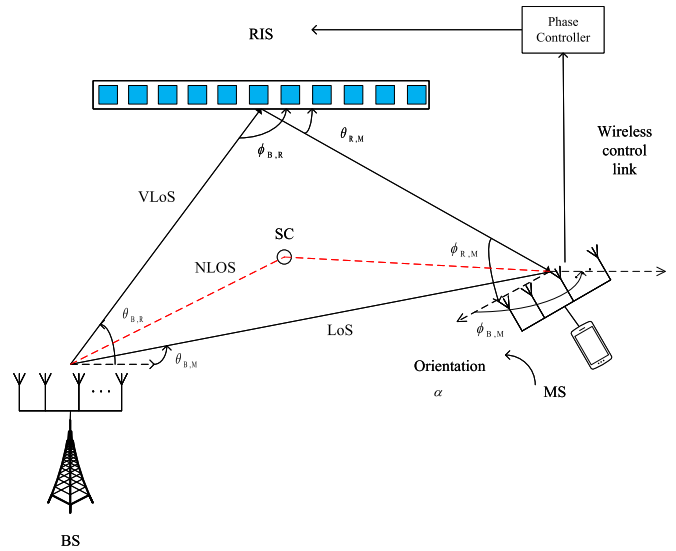


FIGURE 1. Illustration of the system model.

the CRLBs of the channel parameter estimation, where the position and orientation error bounds are also presented. The framework of the proposed joint beam alignment scheme is introduced in Section IV. The simulation setup and simulation results are detailed in Section V. Finally, we concisely conclude this paper in Section VI.

Notations: a , \mathbf{a} , \mathbf{A} stand for scalar, vector and matrix, respectively. \mathbf{A}^T , \mathbf{A}^H , \mathbf{A}^\dagger , $\|\mathbf{A}\|_2$ and $\|\mathbf{A}\|_F$ represent the transpose, Hermitian transpose, pseudoinverse, Euclidean norm and Frobenius norm of matrix \mathbf{A} , respectively. The (i, j) th entry of \mathbf{A} is $[\mathbf{A}]_{i,j}$, and $\text{diag}(\mathbf{A})$ is a diagonal matrix formed by the diagonal elements of \mathbf{A} .

II. SYSTEM MODEL

Fig. 1 shows the considered system model. We consider a downlink system, where BS transmit information to MS using OFDM based transmission aided by hybrid beamforming in the mmWave frequency band. We assume that there are N subcarriers. BS has N_B antennas and MS has N_M antennas. It is well-known that in mmWave channels, there is usually a dominant LoS path and several NLoS paths generated by random scatterers (SC). In most cases, the power of the time-varying NLoS path components is insignificant compared with that of the LoS path, which leads to poor performance if they are used for positioning [12]. Moreover, the LoS path based localization cannot be accomplished, when the LoS path is not available due to blockage in some communication scenarios. To overcome these problems, a nearly passive reconfigurable intelligent surface (RIS) with N_R discrete elements can be deployed between the BS and MS, as shown in Fig. 1, to help information transmission and also localization of the MS. Without any loss of generality, in this contribution, the uniform linear array structure is adopted by the BS, RIS as well as MS, but they can be extended to use other types of antenna arrays.

As shown in Fig. 1, we assume that BS is located at $\mathbf{b} = [b_x, b_y]^T$ and RIS is placed at $\mathbf{r} = [r_x, r_y]^T$, while MS is located at $\mathbf{m} = [m_x, m_y]^T$, which is to be estimated. We assume that BS transmits the positioning reference signals (PRS) either directly to the MS or through the reflection by the RIS to the MS. With the aid of these signals, the MS senses its own location and orientation based on the limited feedback assisted beam sweeping.

Therefore, in the considered mmWave system, there are possibly three types of propagation paths, which are the LoS path, the NLoS path and the virtual line-of-sight (VLoS) path, as explicitly depicted in Fig. 1. The MS will choose the path with the highest power to locate itself, when the LoS path is blocked. In the following subsections, we present the channel model, RIS control followed by the signal model.

A. CHANNEL MODEL

Without considering the RIS, the propagation channel between BS and MS in the frequency domain can be represented as [24]

$$\mathbf{H}_{\text{BM}}[n] = \rho_{\text{LoS}} \gamma_{\text{BM},0} e^{-j2\pi \tau_{\text{BM},0} \frac{nB}{N}} \mathbf{a}_{\text{r}}(\phi_{\text{BM},0}) \mathbf{a}_{\text{t}}^H(\theta_{\text{BM},0}) + \sum_{l=1}^L \gamma_{\text{BM},l} e^{-j2\pi \tau_{\text{BM},l} \frac{nB}{N}} \mathbf{a}_{\text{r}}(\phi_{\text{BM},l}) \mathbf{a}_{\text{t}}^H(\theta_{\text{BM},l}), \quad (1)$$

for the n -th, $n = 0, 1, \dots, N-1$, subcarrier. In (1), B is the signal bandwidth, and $\tau_{\text{BM},l}$, $l = 0, 1, \dots, L$, represents the time delay of the l -th path, where $\tau_{\text{BM},0}$ denotes the delay of the LoS path, while $l \neq 0$ correspond to the NLoS components. Furthermore, $\mathbf{a}_{\text{r}}(\phi_{\text{BM},l})$ and $\mathbf{a}_{\text{t}}(\theta_{\text{BM},l})$ denote the steering vectors of the l -th path at MS and BS, respectively, which are given by

$$\mathbf{a}_{\text{t}}(\theta_{\text{BM},l}) = \left[1, e^{j\frac{2\pi d}{\lambda} 1 \sin(\theta_{\text{BM},l})}, \dots, e^{j\frac{2\pi d}{\lambda} (N_{\text{B}}-1) \sin(\theta_{\text{BM},l})} \right]^T \quad (2)$$

$$\mathbf{a}_{\text{r}}(\phi_{\text{BM},l}) = \left[1, e^{j\frac{2\pi d}{\lambda} 1 \sin(\phi_{\text{BM},l})}, \dots, e^{j\frac{2\pi d}{\lambda} (N_{\text{M}}-1) \sin(\phi_{\text{BM},l})} \right]^T \quad (3)$$

where d and λ are the antenna spacing and signal wavelength, respectively, N_{B} and N_{M} are the numbers of antenna elements of the arrays at BS and MS, respectively, while $\theta_{\text{BM},l}$ and $\phi_{\text{BM},l}$ are the AoD and AoA of the l -th path between BS and MS, as seen in Fig. 1. Furthermore, in (1), $\rho_{\text{LoS}} \in \{0, 1\}$ is the LoS blockage indicator and $\gamma_{\text{BM},l}$ is the path-loss coefficient for the l -th path. Note that, in this paper, we divide the LoS component and VLoS component into two separate stages, while the NLoS components are treated as noise due to their significantly lower power in comparison with the LoS and VLoS paths. Thus, only the large-scale fading (path-loss) is needed to be considered with the LoS/VLoS paths [12], as shown in (1).

When RIS is employed, the VLoS path can be considered as a process of passive beamforming. Compared with the random scatterers, the reflected signals by RIS can be

coherently added to enhance the signal power of the desired users [39].

Specifically, when the RIS operated in the far-field scenario is considered, the channel can be treated as a two-hop channel, consisting of the first hop channel $\mathbf{H}_{\text{BR}}[n]$ and the second hop channel $\mathbf{H}_{\text{RM}}[n]$. Hence, the frequency domain cascaded RIS channel can be written as [24]

$$\mathbf{H}_{\text{BRM}}[n] = \mathbf{H}_{\text{RM}}[n] \Omega \mathbf{H}_{\text{BR}}[n] = \gamma_{\text{BRM}} e^{-j2\pi \tau_{\text{RM}} \frac{nB}{N}} \mathbf{a}_{\text{r}}(\phi_{\text{RM}}) \mathbf{a}_{\text{t}}^H(\theta_{\text{RM}}) \Omega \times e^{-j2\pi \tau_{\text{BR}} \frac{nB}{N}} \mathbf{a}_{\text{r}}(\phi_{\text{BR}}) \mathbf{a}_{\text{t}}^H(\theta_{\text{BR}}), \quad (4)$$

for the n -th subcarrier, where $\gamma_{\text{BRM}} = \frac{\sqrt{\zeta} \lambda}{4\pi(d_{\text{BR}} + d_{\text{RM}})}$ is the path-loss of the RIS channel [5], [12], where ζ denotes the reflection loss. while τ_{RM} and τ_{BR} are defined in the same way as that in (1), denoting the delay from RIS to MS and that from BS to RIS, respectively. In (4), the RIS's phase shift matrix is $\Omega = \frac{1}{\sqrt{N_{\text{R}}}} \text{diag}(e^{j\omega_1}, \dots, e^{j\omega_{N_{\text{R}}}})$, which is a diagonal matrix with constant-modulus entries according to [24].

Thus, when the LoS, NLoS as well as VLoS paths are considered, the entire downlink propagation channel can be written as

$$\mathbf{H}[n] = \mathbf{H}_{\text{BM}}[n] + \rho_{\text{VLoS}} \mathbf{H}_{\text{BRM}}[n], \quad (5)$$

where

$$\rho_{\text{VLoS}} = \begin{cases} 1, & \text{if RIS is activated;} \\ 0, & \text{if RIS is deactivated.} \end{cases} \quad (6)$$

Note that, only the large-scale fading is considered in (1) and (4), while the small-scale fading is ignored and treated as noise with its power proportional to the transmitted power, owing to the fact that in mmWave localization, only the strongest paths, i.e., the LoS or VLoS paths, are useful. Furthermore, since in our system, the localization scheme is separated into the LoS and VLoS path components, the spatial correlation between the LoS and VLoS paths can be neglected.

Let us now consider the RIS control.

B. RIS CONTROL

In this paper, the nearly passive RIS is employed, which consists of reflecting elements, copper backplane and control circuit board [8]. For the RIS control, the phase and magnitude of the reflected signal by a reflective element may be set by the controller, but the complexity is high and energy consuming. Therefore, in the proposed system, we assume that the RIS is only able to adjust the phases, while the reflected magnitude coefficient is 1, by following the widely used assumptions, such as, in [12], [24], [27]. Furthermore, we assume that the transmit beam by the BS towards the RIS is a sharp beam, as the result that the locations of both BS and RIS are usually fixed. After the reflection, the beam transmitted by the RIS should be directed towards the MS. This is achieved based on the assumption that the MS is capable of controlling the RIS via a lower frequency radio link, as done in [24].

On the other hand, as shown in Fig. 1, the reflected angle (phase) by RIS serves as a localization parameter. Hence, the localization goal is to estimate this phase based on the received signal and the power from the VLoS path. In the case that the LoS path is unavailable or too weak, the VLoS path will be utilized to sense the location of MS.

C. SIGNAL MODEL

To transmit information from BS to MS, at the BS, N_D data streams are firstly precoded by the digital beamformer $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{\text{RF}} \times N_D}$, where N_{RF} denotes the number of radio frequency (RF) chains. Then, on each RF chain, a N_{IFFT} -point inverse fast Fourier transform (IFFT) is implemented to transform the symbols from frequency-domain to time-domain, followed by adding a cyclic prefix (CP) before the RF level analog precoding. The CP length of $T_{\text{CP}} = L_{\text{CP}}T_s$ is assumed to be longer than the channel's delay spread, where T_s is the sampling interval, while L_{CP} denotes the CP length in samples. Then, an analog beamformer $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t \times N_{\text{RF}}}$ is applied. Therefore, the hybrid beamformer is $\mathbf{F} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$, which satisfies $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F = 1$. Specifically, when the n -th subcarrier is considered, after the positioning reference signal (PRS) $\mathbf{x}[n]$ is precoded by the hybrid beamformer F at the BS and sent to the channel, the received baseband signal at MS can be expressed as

$$\mathbf{y}[n] = \sqrt{P}\mathbf{H}[n]\mathbf{F}\mathbf{x}[n] + \mathbf{n}[n], n = 0, 1, \dots, N-1, \quad (7)$$

where P is the transmit power, $\mathbf{n}[n]$ is additive white Gaussian noise (AWGN) following the distribution of $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_M})$, where σ^2 is the noise variance, and $\mathbf{H}[n]$ is given by (5). Hence, the received signal of (7) contains the LoS, VLoS and NLoS components, which have the required localization parameters embedded in the LoS and VLoS components. Note that the NLoS path components usually vary fast and convey insignificant power in mmWave systems in comparison to the LoS path [12]. Therefore, the channel parameters governing the LoS and VLoS paths are more interested for localization purpose. To investigate the potentials of the joint LoS/VLoS paths based localization, below in Section III, we derive the CRLB as well as the corresponding positioning error bound and the orientation bound.

III. FIM AND CRLB FOR RIS-AIDED MMWAVE LOCALIZATION SYSTEM

To derive the lower bound of localization error, the FIM of the channel and the location information as well as the CRLB are derived from the aforementioned received signal model in this section. In this paper, a two-stage approach proposed in [16] is employed to jointly estimate the location and orientation of the MS. During the first stage, the FIM of the channel parameters is derived. Then, the FIM of the location/orientation parameters is derived during the second stage using transformation and the FIM of the channel parameters obtained from the first stage. Let us first derive the FIM of the channel parameters.

A. FISHER INFORMATION MATRIX FOR THE CHANNEL PARAMETERS

To obtain the error bound of the position and orientation estimations, the CRLB of the channel parameters should be first derived. Based on Fig. 1, the channel parameters are collected into a vector $\boldsymbol{\eta} = [\tau_{\text{BM}}, \theta_{\text{BM}}, \phi_{\text{BM}}, \gamma_{\text{BM}}, \tau_{\text{RM}}, \theta_{\text{RM}}, \phi_{\text{RM}}, \gamma_{\text{RM}}]^T$. By defining $\tilde{\boldsymbol{\eta}}$ as the unbiased estimator of $\boldsymbol{\eta}$, the mean squared error (MSE) of estimation is bounded by the inverse of the FIM, which is represented by [16]

$$\mathbb{E}_{\mathbf{y}|\boldsymbol{\eta}}\left[(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta})(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta})^T\right] \geq \mathbf{J}_{\boldsymbol{\eta}}^{-1}, \quad (8)$$

where the parameterized expectation $\mathbb{E}_{\mathbf{y}|\boldsymbol{\eta}}$ is taken with respect to the unknown channel parameters $\boldsymbol{\eta}$. Note that (8) means that $\mathbb{E}_{\mathbf{y}|\boldsymbol{\eta}} - \mathbf{J}_{\boldsymbol{\eta}}^{-1}$ is a semi-definite matrix. Since the channel parameters are closely related to the localization of BS, RIS, and MS, the lower bounds of the estimated channel parameters are first derived. Correspondingly, the FIM in terms of the channel parameter vector $\boldsymbol{\eta}$ can be expressed by a 8×8 matrix as

$$\mathbf{J}_{\boldsymbol{\eta}}[n] = \begin{bmatrix} \Psi_n(\eta_1, \eta_1) & \cdots & \Psi_n(\eta_1, \eta_8) \\ \vdots & \ddots & \vdots \\ \Psi_n(\eta_8, \eta_1) & \cdots & \Psi_n(\eta_8, \eta_8) \end{bmatrix}, \quad (9)$$

in which the (i, j) -th entry is given by

$$\Psi_n(\eta_i, \eta_j) = \frac{P}{\sigma^2} \Re \left\{ \frac{\partial \boldsymbol{\mu}^H[n]}{\partial \eta_i} \frac{\partial \boldsymbol{\mu}[n]}{\partial \eta_j} \right\}. \quad (10)$$

To compute (10), $\boldsymbol{\mu}[n]$ is assumed to be noiseless, which is the component embedded in $\sqrt{P}\mathbf{H}[n]\mathbf{F}\mathbf{x}[n]$ of (7), where each of the element in $\mathbf{f} = \mathbf{F}\mathbf{x}[n]$ is set to e^{jv} with v uniformly distributed in $(0, 2\pi]$. The details for the FIM of (9) are provided in Appendix A. Note that, the diagonal elements of $\mathbf{J}_{\boldsymbol{\eta}}[n]$ denote the CRLBs of the estimates of the corresponding channel parameters.

B. FISHER INFORMATION MATRIX FOR THE LOCATION PARAMETERS

In the previous subsection, the FIM and CRLBs of the channel parameters were derived. Since the channel parameters are related to the localization of communication terminals, the FIM for the channel parameters can be transformed to the FIM for the location parameters via a Jacobian transformation matrix [16]. Let the vector containing the localization parameters be expressed as $\boldsymbol{\zeta} = [m_x, m_y, \alpha]^T$, where α is the orientation of the array at MS. Then, the Jacobian transformation matrix T is obtained as [16]

$$\mathbf{T} = \frac{\partial \boldsymbol{\eta}^T}{\partial \boldsymbol{\zeta}}. \quad (11)$$

Furthermore, the FIM of the position parameters $\boldsymbol{\zeta}$ is related to the FIM for the channel parameters $\boldsymbol{\eta}$ via the formula [40] as

$$\mathbf{J}_{\boldsymbol{\zeta}}[n] = \mathbf{T}\mathbf{J}_{\boldsymbol{\eta}}[n]\mathbf{T}^H. \quad (12)$$

It can be shown that the entries of the transformation matrix T of (11) can be obtained from the geometric relationships as shown in Fig. 1, which are given in detail as

$$\tau_{\text{BM}} = \frac{\|\mathbf{b} - \mathbf{m}\|_2}{c}, \quad (13a)$$

$$\theta_{\text{BM}} = \arccos\left(\frac{(m_x - b_x)}{\|\mathbf{b} - \mathbf{m}\|_2}\right), \quad (13b)$$

$$\phi_{\text{BM}} = \pi + \theta_{\text{BM}} - \alpha, \quad (13c)$$

$$\tau_{\text{BR}} = \frac{\|\mathbf{b} - \mathbf{r}\|_2}{c}, \quad (13d)$$

$$\theta_{\text{BR}} = \arccos\left(\frac{(r_x - b_x)}{\|\mathbf{r} - \mathbf{b}\|_2}\right), \quad (13e)$$

$$\phi_{\text{BR}} = -\pi + \theta_{\text{BR}}, \quad (13f)$$

$$\tau_{\text{RM}} = \frac{\|\mathbf{m} - \mathbf{r}\|_2}{c}, \quad (13g)$$

$$\theta_{\text{RM}} = -\arccos\left(\frac{(m_x - r_x)}{\|\mathbf{r} - \mathbf{m}\|_2}\right), \quad (13h)$$

$$\phi_{\text{RM}} = \pi + \theta_{\text{RM}} - \alpha, \quad (13i)$$

where c is the speed of light. Furthermore, the channel gain of the LoS path can be expressed as [5]

$$\gamma_{\text{BM},1} = \frac{\lambda}{4\pi d_{\text{BM}}}, \quad (14)$$

where d_{BM} is the Euclidean distance between BS and MS given by $d_{\text{BM}} = \|\mathbf{b} - \mathbf{m}\|_2$, and λ is the wavelength. For the VLoS path, the channel gain γ_{BR} and γ_{RM} can be obtained similarly as (14). It is noteworthy that if the localization is only estimated by ToA and AoD (or AoR), the parameters reflecting the strength of received signal, i.e., γ_{BM} and γ_{RM} , can be neglected. On the other side, when the LoS link is blocked, all the parameters defining the LoS path can be discarded.

Consequently, the transformation matrix T can be obtained, which is in the form of (15), whose entries are detailed in Appendix B.

Finally, after T is obtained, the FIM for the location parameters of MS can be obtained by (12).

C. POSITION AND ORIENTATION ESTIMATION ERROR BOUNDS

After the FIM of the location parameters has been derived using the approach in the last section, the error bounds of the position and orientation estimation can be readily obtained by the superposition of the estimation errors from N subcarriers [16]. In detail, by exploiting the N subcarriers, the

FIM of the position parameters $\boldsymbol{\zeta}$ can be obtained as

$$\mathbf{J}_{\boldsymbol{\zeta}} = \sum_{n=-(N-1)/2}^{(N-1)/2} \mathbf{J}_{\boldsymbol{\zeta}}[n]. \quad (16)$$

Correspondingly, the error bound for position estimation (PEB) is given by

$$\text{PEB} = \sqrt{\text{trace}\left\{\left[\mathbf{J}_{\boldsymbol{\zeta}}^{-1}\right]_{1:2,1:2}\right\}}, \quad (17)$$

and the error bound for orientation estimation (OEB) is given by

$$\text{OEB} = \sqrt{\left[\mathbf{J}_{\boldsymbol{\zeta}}^{-1}\right]_{3,3}}, \quad (18)$$

where $[\cdot]_{1:2,1:2}$ and $[\cdot]_{3,3}$ denote the top-left 2×2 sub-matrix and the third diagonal entry, respectively. In Section V, the PEB and OEB will be compared with the proposed beam training based localization scheme that is to be detailed in the next section.

IV. JOINT BEAM TRAINING FOR LOCALIZATION

In [24], the LoS path is assumed to be fully obstructed, while in practice the LoS path may or may not exist. Moreover, in [12], both the LoS and VLoS paths are utilized for localization. However, the maximum likelihood algorithm employed introduces very high training overhead. Since mmWave channels are typical sparse channels, this property has been exploited for reducing the complexity and training overhead in channel estimation and localization [41]. However, unlike a conventional mmWave channel, the cascaded channel in the RIS embedded system is usually converted to a richly scattered channel, which contains an increased number of scatter paths due to the reflections of RIS [12]. It is well-known that the richly scattered channel makes it difficult to estimate the channel parameters for the purpose of localization. Hence, to simplify the joint beamforming problem in the RIS aided mmWave system, in this paper, we propose to divide the joint beamforming into an active beamforming, operated at BS and MS, and a passive beamforming, operated at RIS, for both the LoS and VLoS paths.

In our approach, the MS first determines the path with the highest power from the combined signal received from LoS and VLoS paths, and then use it to sense its own location. Explicitly, during the first stage, the RIS is de-activated and the beam sweeping algorithm is performed between BS and MS. Then, the RIS is activated during the second stage. During this stage, as the location of RIS is known *a priori* to BS, the AoD from BS to RIS can be assumed to be known,

$$\mathbf{T} = \begin{bmatrix} \frac{\partial \tau_{\text{BM}}}{\partial m_x} & \frac{\partial \theta_{\text{BM}}}{\partial m_x} & \frac{\partial \phi_{\text{BM}}}{\partial m_x} & \frac{\partial \gamma_{\text{BM}}}{\partial m_x} & \frac{\partial \tau_{\text{RM}}}{\partial m_x} & \frac{\partial \theta_{\text{RM}}}{\partial m_x} & \frac{\partial \phi_{\text{RM}}}{\partial m_x} & \frac{\partial \gamma_{\text{RM}}}{\partial m_x} \\ \frac{\partial \tau_{\text{BM}}}{\partial m_y} & \frac{\partial \theta_{\text{BM}}}{\partial m_y} & \frac{\partial \phi_{\text{BM}}}{\partial m_y} & \frac{\partial \gamma_{\text{BM}}}{\partial m_y} & \frac{\partial \tau_{\text{RM}}}{\partial m_y} & \frac{\partial \theta_{\text{RM}}}{\partial m_y} & \frac{\partial \phi_{\text{RM}}}{\partial m_y} & \frac{\partial \gamma_{\text{RM}}}{\partial m_y} \\ \frac{\partial \tau_{\text{BM}}}{\partial \alpha} & \frac{\partial \theta_{\text{BM}}}{\partial \alpha} & \frac{\partial \phi_{\text{BM}}}{\partial \alpha} & \frac{\partial \gamma_{\text{BM}}}{\partial \alpha} & \frac{\partial \tau_{\text{RM}}}{\partial \alpha} & \frac{\partial \theta_{\text{RM}}}{\partial \alpha} & \frac{\partial \phi_{\text{RM}}}{\partial \alpha} & \frac{\partial \gamma_{\text{RM}}}{\partial \alpha} \end{bmatrix} \quad (15)$$

which can be used by BS to design a narrow beam towards RIS. Furthermore, since BS and RIS are at different locations and the LoS component between BS and MS has been identified during the first stage, this LoS component can be removed from consideration in the received signal to get rid of its impact on the VLoS path. Hence, during the first stage, the AoD at BS, denoting θ_{BM} , can be estimated. During the second stage, the AoR at RIS, i.e., θ_{RM} , can be estimated. Moreover, we propose a novel beam sweeping method for beamformer design, so as to strike a good trade-off between the localization performance and the beam search time. This method will be compared with the conventional hierarchical codebook [24], partial beam sweeping [42], and exhaustive search approaches in Section IV.

Below in this section, different codebook-based schemes for beamformer design are considered. We start with the conventional hierarchical codebook based design proposed in [37]. The RIS codebook and BS/MS codebook designs are separated into two parts, where the RIS codebook contains only pure analog phase shifters, while the BS/MS codebook has a typical hybrid beamforming architecture. Therefore, in our codebook design, only analog codewords are considered for the RIS codebook, but hybrid codewords are considered for the BS and MS codebooks.

A. ACTIVE AND PASSIVE BEAMFORMING FOR LOS AND VLOS PATHS

In this paper, the location and orientation estimation based on the LoS or VLoS path is divided into two stages to address the LoS path blockage problem, which is the assumption similar to that used in [12]. However, in [12], random beamforming is employed, which causes a high complexity and also a high offline training overhead. To overcome these problems, in our hybrid LoS and VLoS paths based localization problem, we motivate to attain a good trade-off between the localization performance and training overhead, and investigate the impact of the different beam sweeping methods on the performance of localization.

To identify the LoS path or VLoS path for MS's location estimation, a two-stage strategy is introduced. In the first stage, the RIS is de-activated for the MS to sense if the LoS path is available. Then, in the second stage, the RIS is activated. Since the transmit beam from BS to MS is very narrow, the NLoS components between BS and RIS can be ignored. From the above-mentioned two stages, the MS identifies the strongest path for its location estimation. Below we detail this two-stage strategy.

Stage 1: As depicted in Fig. 1, when the RIS is de-activated, there are only LoS and NLoS channel components from BS to MS. In this case, $\rho_{\text{VLoS}} = 0$ in (7), and the received signals can be formulated as

$$\begin{aligned} \mathbf{y}_{\text{LoS}}[n] &= \sqrt{P} \mathbf{W}^H \mathbf{H}_{\text{BM}} \mathbf{f}_{\text{LoS}} x[n] + \mathbf{W}^H \mathbf{n}[n] \\ &= \sqrt{P} \rho_{\text{LoS}} \gamma_{\text{BM},0} e^{-j2\pi \tau_{\text{BM},0} \frac{nB}{N}} \\ &\quad \times \mathbf{W}^H \mathbf{a}_r(\phi_{\text{BM},0}) \mathbf{a}_t^H(\theta_{\text{BM},0}) \mathbf{f}_{\text{LoS}} x[n] \end{aligned}$$

$$\begin{aligned} &+ \sum_{l=1}^L \sqrt{P} \gamma_{\text{BM},l} e^{-j2\pi \tau_{\text{BM},l} \frac{nB}{N}} \\ &\quad \times \mathbf{W}^H \mathbf{a}_r(\phi_{\text{BM},l}) \mathbf{a}_t^H(\theta_{\text{BM},l}) \mathbf{f}_{\text{LoS}} x[n] \\ &+ \mathbf{W}^H \mathbf{n}[n], \end{aligned} \quad (19)$$

for $n = 0, 1, \dots, N-1$, where P is the transmit power, $\mathbf{n}[n]$ is additive white noise distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_M})$, while \mathbf{f}_{LoS} and \mathbf{W} are the hybrid precoder and combiner employed at BS and MS, respectively, which can be defined in the similar ways as that for (7). It is noteworthy that the NLoS paths in mmWave systems vary fast and their power is usually weak. Hence, the NLoS paths are highly unreliable for localization [12]. Thus, it is reasonable to treat the NLoS paths as extra noise, which we represent as $\mathbf{v}[n] = \sum_{l=1}^L \gamma_{\text{BM},l} e^{-j2\pi \tau_{\text{BM},l} \frac{nB}{N}} \mathbf{a}_r(\phi_{\text{BM},l}) \mathbf{a}_t^H(\theta_{\text{BM},l}) \mathbf{f}_{\text{LoS}} x[n]$. Since $\mathbf{v}[n]$ and $\mathbf{n}[n]$ are independent, the total noise $\tilde{\mathbf{n}}[n] = \mathbf{v}[n] + \mathbf{n}[n]$ in (19) follows the approximated Gaussian distribution of $\mathcal{CN}(\mathbf{0}, (\sigma^2 + \sigma_v^2) \mathbf{I}_{N_M})$ [12], where σ_v^2 is the power conveyed by the NLoS paths.

Consequently, by setting $x[n] = 1$, (19) can be written as

$$\begin{aligned} \mathbf{y}_{\text{LoS}}[n] &= \sqrt{P} \rho_{\text{LoS}} \gamma_{\text{BM},0} e^{-j2\pi \tau_{\text{BM},0} \frac{nB}{N}} \\ &\quad \times \mathbf{W}^H \mathbf{a}_r(\phi_{\text{BM},0}) \mathbf{a}_t^H(\theta_{\text{BM},0}) \mathbf{f}_{\text{LoS}} + \mathbf{W}^H \tilde{\mathbf{n}}[n]. \end{aligned} \quad (20)$$

Stage 2: In contrast to stage 1, when the RIS is activated by setting $\rho_{\text{VLoS}} = 1$, similar to (20), the received signals at MS can be represented as

$$\begin{aligned} \mathbf{y}_{\text{VLoS}}[n] &= \sqrt{P} \gamma_{\text{BRM}} e^{-j2\pi \tau_{\text{RM}} \frac{nB}{N}} \\ &\quad \times \mathbf{W}^H \mathbf{a}_r(\phi_{\text{RM}}) \mathbf{a}_t^H(\theta_{\text{RM}}) \mathbf{\Omega} e^{-j2\pi \tau_{\text{BR}} \frac{nB}{N}} \\ &\quad \times \mathbf{a}_r(\phi_{\text{BR}}) \mathbf{a}_t^H(\theta_{\text{BR}}) \mathbf{f}_{\text{VLoS}} + \mathbf{W}^H \mathbf{n}[n], \end{aligned} \quad (21)$$

for $n = 0, 1, \dots, N-1$, where γ_{BRM} denotes the large scale fading gain of the VLoS path, while \mathbf{f}_{VLoS} is the precoding vector operated by BS towards the RIS. Note that, since the locations of BS and RIS are fixed, the precoding vector for transmission from BS to RIS can be fixed to $\mathbf{f}_{\text{VLoS}} = \frac{1}{\sqrt{N_B}} \mathbf{a}_t(\theta_{\text{BR}})$. Furthermore, after the processing of RIS, the beam from RIS to MS can also be designed to be a narrow beam. Hence, with the *a priori* knowledge about the narrow beams transmitted from BS to RIS and from RIS to MS, the interference generated by the LoS path on the VLoS path can be ignored, as the result that the LoS component has already been estimated by MS during the first stage. This explains why the LoS component is not included in (21) for the sake of simplicity.

B. PROPOSED BEAM SWEEPING ALGORITHM

Since the phase shifts at RIS are controllable, the signal processing carried out by the RIS is the same as passive beamforming. Thus, the MS can adjust the reflected phases at RIS by the limited feedback transmitted through a wireless control link based on the downlink PRS reflected by the RIS to the MS, with the objective, such as, to maximize the received signal power [24], [43]. Fig. 2 shows some examples of beam alignment algorithms in beamspace.

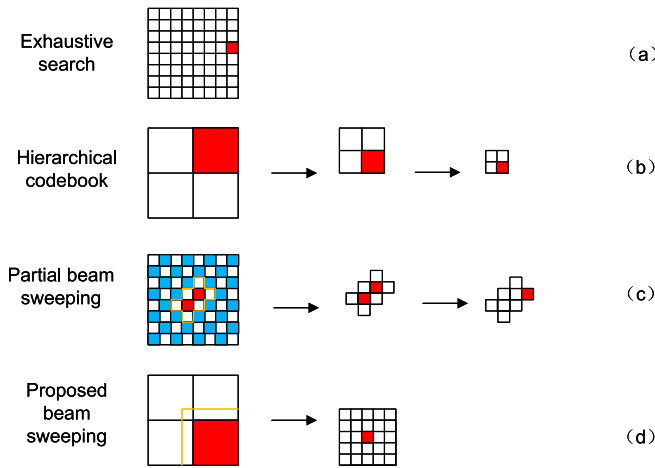


FIGURE 2. Illustration of different beam sweeping methods.

In more detail, as depicted in Fig. 2(a), the most basic beam training technique is to exhaustively test all possible pairs of codewords to determine the best pair that maximize the received signal power. Note that here that a pair of codewords contain one for the RIS to implement precoding and one for the MS to carry out combining. Hence, this scheme somewhat needs to sweep every possible angle [37], [42]. As a consequence, this method is known as the exhaustive search (ES), which demands extremely high training overhead, if the number of antenna elements at RIS or/and MS is high.

In the context of the hierarchical codebook (HC) proposed in [37], as shown in Fig. 2(b), the column and row on the grid correspond to the AoD at BS (or AoR at RIS) and the AoA at MS, respectively. This method follows a multi-stage search strategy. For instance, if at each side of the RIS and MS, the HC is used to search twice, it leads to a 2×2 grid at the first stage (level). With the HC based search, the HC has several levels, and the number of codewords increases as the level number increases. In other words, the lowest level uses a small number of low-resolution codewords that generate broad beams, while the higher level has a relatively higher number of high-resolution codewords that generate narrow beams with very high directional beam gain. Therefore, HC consists of multiple levels with the resolution increasing from lower level to higher level, which allows to significantly reduce the overhead for the codeword pair search.

The partial beam sweeping (PS) scheme proposed in [42] only uses a subset of the codewords for the beam sweeping so as to reduce the overhead. As shown in Fig. 2(c), the transmit and receive codewords are represented as the two-dimensional grids, where each grid corresponds to a pair of transmit and receive codewords directed to certain directions. The PS scheme typically divides the beam training into two stages of the initial test (IT) and the additional test (AT). Unlike the ES scheme that directly tests all the high resolution grids, in the PS scheme, only half or even fewer of the grids are searched during the IT stage, which aims to

TABLE 1. Training overhead.

	Training overhead
The HC algorithm	$SK_{\text{HC}}^2/N_{\text{RF}}$
The PS algorithm	$(K^2/2 + K_{\text{AT}}^2)/N_{\text{RF}}$
The ES algorithm	K^2/N_{RF}
The proposed algorithm	$[\text{ceil}(\sqrt{K})]^2 + [\text{ceil}(\sqrt{K}) + 2]^2/N_{\text{RF}}$

achieve the approximate estimations for the AoD (or AoR) and AoA. However, due to the power leakage of mmWave channel, the two neighboring columns and two neighboring rows in the grid with the largest power are chosen. As shown in Fig. 2(c), we assume that the two tested grids darkly filled have higher power, which yield a new grid, as shown in the middle of Fig. 2(c), for the testing at the AT stage. At the AT stage, the untested neighboring grids are tested, and the best pair are obtained from that yielding the highest power [42].

In contrast to the above three search schemes, our proposed beam sweeping scheme combines the benefits of all of them, to achieve the trade-off between performance and training overhead. As shown in Fig. 2(d), the proposed scheme uses two stages of search, called as the initial test (IT) and deep test (DT), respectively. Different from the PS method, during the IT stage, our scheme uses low resolution codewords instead of high resolution ones, and employs the HC scheme, in order to estimate the potential range of the beam directions on the grids. However, the possible power leakage of mmWave channel may result in miss identification and impact the performance of estimation. Thus, as shown in Fig. 2(d), one neighboring column and one neighboring row in the higher resolution grids (yellow grids) are included to the DT stage. Then, during the DT stage, the ES scheme is employed to search for the final pair of codewords. To be more specific, during the IT stage, our proposed beam sweeping method can be viewed as a higher resolution hierarchical codebook algorithm, which is more robust than the binary HC algorithm [24] and ternary-tree HC algorithm [26], while the complexity is lower than the PS algorithm [42] and ES algorithm [37]. Then, the second DT stage can benefit from the angle range estimated by the first IT stage, which provides a space with significantly reduced size for executing the exhaustive search. From above we can see that the proposed beam sweeping scheme has the potential to obtain the near-optimal codewords, while the search overhead is significantly lower when compared with the ES algorithm. In detail, the training overheads for the HC, ES, PS and the proposed algorithms are summarized in Table 1. Further details about the performance and training overhead of the different beam sweeping schemes will be provided in the next section.

Remark: For the first stage, the proposed beam sweeping algorithm conveys $[\text{ceil}(\sqrt{K})]^2$ potential optimal pair codewords for precoding and combining, which can be also termed as the beamspace grids. For example, as shown in Fig. 2(d), when $K_1 = 2$ codewords are employed for both precoding and combining, the complexity for the first stage is $O(K_1^2)$. Similarly, for the second stage, $K_2 \times K_2$

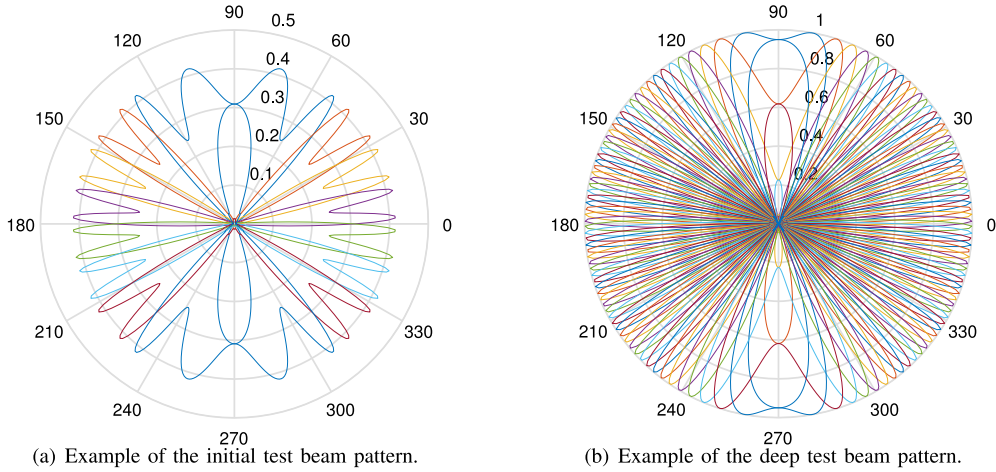


FIGURE 3. Illustration of the beam patterns at different stages with $N_R = 8$.

codewords are included, but only the specific range of grids is considered (as shown in the yellow grids in Fig. 2(d)). From Fig. 2(d) we can know that the related complexity is $O(K_2^2/4 + K_{\text{additional}})$, where $K_{\text{additional}}$ denotes the neighboring grids of the estimated range from the first stage. Thus, the total complexity of the proposed beam sweeping method is $O(K_1^2 + K_2^2/4 + K_{\text{additional}}) = O(29)$. Comparatively, for the HC algorithm, the complexity is $O(SK^2) = O(12)$, where $S = 3$ denotes the total number of layers of beam sweeping, and K is set to 2 in Fig. 2. For the ES algorithm, the complexity is $O(K_2^2) = O(64)$. Finally, the complexity of the PS algorithm is $O(K_2^2/2 + K_{\text{neighbor}}) = O(40)$, where K_{neighbor} denotes the neighboring grids in addition to the two grids selected by the first stage. We can deduce from Table 1 that the complexity of our proposed algorithm is lower than that of the PS and ES algorithms, as the result that the resolution sacrificed in the first stage can help to reduce the overall complexity of the algorithm. More explicitly, to attain a good performance and complexity trade-off, the proposed algorithm exploits the low-complexity of the HC algorithm, and the neighboring grids of the PS algorithm, which aids to avoid the beamspace power leakage.

C. DESIGN OF HYBRID AND PURE ANALOG CODEBOOK

The proposed beam sweeping codebook design requires two sets of codewords having different resolutions for operation in the two stages, namely the IT and DT stages, expressed as $s = 1, 2$, respectively. The objective of the IT stage employing a low resolution codebook is to estimate the approximate ranges of the beamspace AoD/AoR/AoA angles, while that of the DT stage with a high resolution codebook is to provide the final estimations with relatively high accuracy for the beamspace AoD/AoR/AoA angles.

In this paper, the method proposed in [37] is introduced to generate the codebook. Specifically, to obtain the codebook for using in the first s -th stage, the (u, v) -th element of the overcomplete quantized angle dictionary [24] can be

defined as

$$[\mathbf{A}]_{u,v} = e^{j2\pi \frac{d}{\lambda} (u-1)\bar{\theta}_v}, \quad (22)$$

where $u = 1, \dots, N_R(N_B)$, and $v = 1, \dots, G$ with $\bar{\theta}_v = -1 + \frac{2v-1}{G}$ for the G quantized angles in the interval $[-1, 1]$ [24]. Now let us define \mathbf{M}_s and \mathbf{W}_s as the codebooks of the RIS and MS, respectively, where each column of \mathbf{M}_s and \mathbf{W}_s denotes a codeword at the s -th stage, corresponding to a specific angle in the beamspace domain. As shown in [24], [37], the codebook \mathbf{M}_s (the same is \mathbf{W}_s) can be obtained by the operation of

$$\mathbf{M}_s = Z_s (\mathbf{A}^H)^\dagger \mathbf{G}_s \in \mathbb{C}^{N_R \times K_s}, \quad (23)$$

where $(C)^\dagger$ denotes the pseudoinverse operation, Z_s is the normalization constant and \mathbf{G}_s is a $G \times K_s$ matrix, which contains only binary elements of 1's and 0's, and K_s is the number of codewords used at stage s . More explicitly, \mathbf{G}_s has the structure of

$$\mathbf{G}_s = \begin{bmatrix} \mathbf{n}_{\text{non-zero}} & 0 & \cdots & 0 \\ 0 & \mathbf{n}_{\text{non-zero}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{n}_{\text{non-zero}} \end{bmatrix}, \quad (24)$$

where $\mathbf{n}_{\text{non-zero}}$ denotes a column vector containing G/K_s 1's. Hence, each column of \mathbf{G}_s has G/K_s 1's, while the remaining elements are 0's. In this way, the beam pattern of the codebook $\mathbf{M}_s \in \mathbb{C}^{N_R \times K_s}$ is constant over a specific angular range. Similarly, $\mathbf{W}_s \in \mathbb{C}^{N_M \times K_s}$ for MS can be obtained.

Each column of the codebook denotes a codeword corresponding to a specific angular beam. As an example, Fig. 3 shows the beam patterns of a RIS system, where the IT stage codebook consists of 8 codewords and the DT stage contains 64 codewords, and all the codewords are the same length of N_R . Explicitly, the DT stage codebook has much higher resolution than the IT stage codebook. Note that, as shown in Fig. 3, when the number of codewords exceeds N_R of the number of antennas, adjacent codewords will become

non-orthogonal. However, each of the non-orthogonal codewords points to a specific angle. This property can be used to find the angle with the accuracy higher than that obtained by only using the set of orthogonal codewords.

As mentioned previously, RIS is assumed to be only equipped with phase shifters but no radio frequency (RF) chains. Hence, each element in the phase shift matrix is required to have a constant modulus. To satisfy these requirements, once \mathbf{M}_s is obtained using (22)-(24), the codebook used for RIS at the s -th stage can then be obtained as

$$\begin{aligned} [\tilde{\mathbf{M}}_s]_{:,j} &= \arg \min \left\| [\hat{\mathbf{M}}_s]_{:,j} - [\mathbf{M}_s]_{:,j} \right\|_2, \\ \text{s.t. } \left| [\hat{\mathbf{M}}_s]_{i,j} \right| &= \frac{1}{\sqrt{N_R}}, \end{aligned} \quad (25)$$

where $[\cdot]_{:,j}$ denotes the j -th codeword, and $[\cdot]_{i,j}$ denotes the element located in the i -th row and j -th column.

On the other side, both MS and BS can employ hybrid beamforming. For the sake of simplicity of codebook generation, we assume that both MS and BS have the same number of N_{RF} RF chains. Simultaneously, we assume that the BS and MS can only access a maximal N_{RF} number of observations at a time. Then, consequently, let us assume that N_{RF} codewords are employed as the combiner at MS (or precoder at BS). Then, the matrices for the hybrid combining (or precoding) at the s -th stage can be obtained as [37]

$$\begin{aligned} \{\tilde{\mathbf{W}}_{RF,s,k}, \tilde{\mathbf{W}}_{BB,s,k}\} \\ = \arg \min \left\| \hat{\mathbf{W}}_{RF,s,k} \hat{\mathbf{W}}_{BB,s,k} - [\mathbf{W}_s]_{:, (k-1)N_{RF}+1:kN_{RF}} \right\|_F^2, \\ \text{s.t. } \left\| \hat{\mathbf{W}}_{RF,s,k} \hat{\mathbf{W}}_{BB,s,k} \right\|_F^2 = N_{RF}, [\hat{\mathbf{W}}_{RF,s,k}]_{i,j} = \frac{1}{\sqrt{N_M}}, \end{aligned} \quad (26)$$

for $k = 1, \dots, K_s/N_{RF}$ to consider all the codewords in \mathbf{W}_s . Note again that the number of codewords for the s -stage can be larger than the number of antennas, and for each beam sweeping time slot, only N_{RF} observations are processed.

Finally, the hybrid codebook for BS and MS is defined as $\tilde{\mathbf{W}} = [\mathbf{W}_1, \mathbf{W}_2]$, while the pure analogue codebook for the RIS is similarly expressed as $\tilde{\mathbf{M}} = [\mathbf{M}_1, \mathbf{M}_2]$, where 1 and 2 denote the IT and DT stage, respectively. For simplicity, we assume that BS and MS have the same number of antennas, and in this case, the same hybrid codebook $\tilde{\mathbf{W}}$ can be applied by both BS and MS.

D. PROPOSED DESIGN FOR JOINT BEAM TRAINING AND LOCALIZATION

Assume that LoS and VLoS paths are separated into two frames without interfering each other. Then, the impact from the small-scale fading can be ignored, while the weak NLoS paths can be treated as noise with its power proportional to the transmitted power [12]. In our proposed joint beam training and localization scheme, the LoS or VLoS path is first selected based on their power, and the one with higher power is selected for the joint training and localization. Specifically, in the first frame, the RIS is de-activated. In this case, the

localization problem can be treated as a conventional joint active beam alignment problem between BS and MS. By contrast, in the second frame when the RIS is activated, the localization can be achieved by the joint passive and active beamforming at the RIS and MS, respectively. Below we detail the operations under Frame 1 and Frame 2.

Frame 1: When operated under Frame 1, joint active beamforming between BS and MS is implemented to match to the LoS path, after the hybrid codewords are selected iteratively by the BS and the MS. As discussed in Section IV-C, there are two stages, i.e., IT and DT stages for the design. We assume that during the IT stage, the low resolution codebook employed by both BS and MS has the number of codewords equating the square root of that of the DT stage, i.e., $K_1 = \text{ceil}(\sqrt{K_2})$. The codewords used for precoding and combining are generated according to (26) in Section IV-C. Therefore, based on (20), the improved interlaced scanning (IS) based received signal for the LoS path can be represented as

$$\begin{aligned} \mathbf{y}_{LoS,s,k}[n] &= \sqrt{P} \rho_{LoS} \gamma_{BM,0} e^{-j2\pi\tau_{BM,0} \frac{nb}{N}} \\ &\times \mathbf{W}_s^H \mathbf{a}_r(\phi_{BM,0}) \mathbf{a}_t^H(\theta_{BM,0}) \mathbf{f}_{LoS,s,k} \\ &+ \mathbf{W}_s^H \tilde{\mathbf{n}}[n], s = 1, 2, \end{aligned} \quad (27)$$

where the columns of \mathbf{W}_s are selected from the codebook $\tilde{\mathbf{W}}$ for the s -th stage, while $\mathbf{f}_{VLoS,s,k}$ denotes the k -th column of \mathbf{W}_s , representing one codeword. Let us assume that K_s codewords are employed by both BS and MS for the training during stage s . Then, by collecting the measurements corresponding to all the $K_s \times K_s$ codewords, the received signals from the n -th subcarrier can be collected to form a $K_s \times K_s$ matrix

$$\mathbf{Y}_{LoS,s}[n] = \left[\mathbf{y}_{LoS,s}^{(1)}[n], \dots, \mathbf{y}_{LoS,s}^{(K_s)}[n] \right]. \quad (28)$$

Considering N subcarriers, the associated sum-power when the BS applies the i_{BS} -th codeword, while the MS applies j_{MS} -th codeword is given by

$$[\mathbf{P}_{LoS,s}]_{i_{BS},j_{MS}} = \sum_{n=1}^N \left| [\mathbf{Y}_{LoS,s}[n]]_{i_{BS},j_{MS}} \right|^2, \quad (29)$$

where $i_{BS}, j_{MS} \in [1, K_s]$. Then, $[\mathbf{P}_{LoS,s}]_{i_{BS},j_{MS}}$ are compared in terms of all the pairs of i_{BS} and j_{MS} , and the pair with the largest power are selected. To this point, if $s = 1$, the selected pair are set as the search range of the DT stage after adding the extra neighboring grids, as shown in Fig. 2. By contrast, if $s = 2$, the selected pair are used as the final codewords for designing the hybrid precoder and hybrid combiner.

Frame 2: During Frame 2, the joint active and passive beamforming between RIS and MS is implemented to match to the VLoS path. Following the LoS path scenario, a low resolution codebook \mathbf{M}_1 in the form of (25) is used for RIS passive beamforming during the IT stage, while a higher resolution codebook \mathbf{M}_2 is used during the DT stage. The combining codewords used by MS can be obtained according to (26). Then, following (21), the received signal from

subcarrier n at stage $s=1$ or 2 can be expressed as

$$\begin{aligned} \mathbf{y}_{\text{VLoS},s,k}[n] &= \sqrt{P} \rho_{\text{VLoS}} \gamma_{\text{BRM}} e^{-j2\pi \tau_{\text{RM}} \frac{nB}{N}} \\ &\quad \times \mathbf{W}_s^H \mathbf{a}_r(\phi_{\text{RM}}) \mathbf{a}_t^H(\theta_{\text{RM}}) \boldsymbol{\Omega}_{s,k} e^{-j2\pi \tau_{\text{BR}} \frac{nB}{N}} \\ &\quad \times \mathbf{a}_r(\phi_{\text{BR}}) \mathbf{a}_t^H(\theta_{\text{BR}}) \mathbf{f}_{\text{VLoS}} + \mathbf{W}_s^H \mathbf{n}[n], \\ n &= 0, \dots, N-1, \end{aligned} \quad (30)$$

when the precoding vector \mathbf{f}_{VLoS} employed by BS to steer the transmission signal towards RIS is assumed to be known. In (30), the K_s columns of \mathbf{W}_s are chosen from codebook $\tilde{\mathbf{W}}$ for hybrid beamforming, and $\boldsymbol{\Omega}_{s,k} = \text{diag}(\mathbf{m}_{s,k})$ is the phase shift matrix, with the diagonal elements given by the codeword $\mathbf{m}_{s,k}$ for the passive beamforming of RIS, where $\mathbf{m}_{s,k}$ is the k th column of \mathbf{M}_s .

Similar to (28) and (29), when K_s codewords are respectively applied by RIS, the MS can obtain $\mathbf{Y}_{\text{VLoS},s}[n]$ for subcarrier n in the form of (28). Then, to estimate the AoR at RIS and the AoA at MS, the sum power with respect to all $i_{\text{RIS}} \in [1, K_s]$, and $j_{\text{MS}} \in [1, K_s]$ is calculated as

$$[\mathbf{P}_{\text{VLoS},s}]_{i_{\text{RIS}},j_{\text{MS}}} = \sum_{n=1}^N \left| [\mathbf{Y}_{\text{VLoS},s}[n]]_{i_{\text{RIS}},j_{\text{MS}}} \right|^2, \quad (31)$$

Then, for the IT stage of $s=1$, the best pair of codewords are determined as the one having the highest power, which yield the search range for stage 2 after taking some neighbouring grids, as shown in Fig. 2, into account. Finally, after the DT stage of search by following the same procedure of the IT stage search, a pair of codewords can be obtained, RIS is used to generate the phase shift only. Then the codeword of precoding, and the codeword of MS is used to carried out hybrid combining. Note that, to inform the selection decision to RIS, MS needs $\log_2(K_s)$ feedback bits, which are sent to the RIS controller through wireless control link.

Path decision: From above discussion we know that there are possible LoS or/and VLoS paths between BS and MS. In the case that both LoS and VLoS paths present, the MS can choose a better path for positioning according to the received power of these two paths² Furthermore, after the selection of the pair of codewords at the DT stage, the channel parameters can be estimated, as detailed in the following discourses.

Blocked LoS case ($P_{\text{LoS},s} < P_{\text{VLoS},s}$): When the LoS path is blocked, the MS senses its location based on the parameters estimated from the VLoS path, which are

$$\begin{aligned} \hat{\theta}_{\text{RM}} &= \arg \max_{\theta_{\text{RM}}} \left| [\mathbf{U}_t(\theta_{\text{RM}}) \odot \mathbf{a}_t^*(\phi_{\text{BR}})]^H \mathbf{m}_2 \right|, \\ \hat{\phi}_{\text{RM}} &= \arg \max_{\phi_{\text{RM}}} \left| \mathbf{w}_2^H \mathbf{U}_r(\phi_{\text{RM}}) \right|, \\ \hat{\tau}_{\text{RM}} &= \arg \max_{\tau_{\text{RM}}} \left| \mathbf{y}_{\text{VLoS},2}^H \mathbf{t}(\tau_{\text{RM}}) \right| - \tau_{\text{BR}}, \end{aligned} \quad (32)$$

2. In RIS-assisted mmWave communications, the NLoS paths in practice vary fast and they are usually too weak to build an effective transmission link, although they are sometimes possible to have higher power than the VLoS path, as argued in [44]. In this paper, only the stable links, such as LoS and VLoS paths, are considered for positioning, while the NLoS paths are treated as interference.

where $\mathbf{m}_2 = [\mathbf{M}_2]_{:,i_{\text{RIS}}}$ and $\mathbf{w}_2 = [\mathbf{W}_2]_{:,i_{\text{MS}}}$ are the finally determined precoding and combining codewords of the RIS and MS obtained at the DT stage,

$\mathbf{t}(\tau_{\text{RM}}) = [e^{-j2\pi \frac{-B\tau_{\text{RM}}(N-1)/2}{N}}, \dots, e^{-j2\pi \frac{B\tau_{\text{RM}}(N-1)/2}{N}}]^T$ is the delay domain grid, while $\mathbf{U}_t(\theta_{\text{RM}}) \in \mathbb{C}^{N_{\text{R}} \times K_2}$ and $\mathbf{U}_r(\phi_{\text{RM}}) \in \mathbb{C}^{N_{\text{M}} \times K_2}$ are the search dictionaries of θ_{RM} and ϕ_{RM} , respectively. Finally, $\mathbf{y}_{\text{VLoS},2}$ is the received signal at the DT stage, which is a vector obtained stacking all the signals received from N subcarriers.

Based on the estimates to the channel parameters, the location and orientation of the MS can be estimated as [16]

$$\begin{aligned} \hat{\mathbf{m}} &= \mathbf{r} + c\hat{\tau}_{\text{RM}} [\cos(\hat{\theta}_{\text{RM}}), \sin(\hat{\theta}_{\text{RM}})], \\ \hat{\alpha} &= \pi + \hat{\theta}_{\text{RM}} - \hat{\phi}_{\text{RM}}. \end{aligned} \quad (33)$$

LoS case ($P_{\text{LoS},s} > P_{\text{VLoS},s}$): When the LoS path is available, the channel parameters of the LoS path can be estimated with the aid of the estimated hybrid codewords at the DT stage, which are represented as

$$\begin{aligned} \hat{\theta}_{\text{BM}} &= \arg \max_{\theta_{\text{BM}}} \left| \mathbf{f}_2^H \mathbf{U}_t(\theta_{\text{BM}}) \right|, \\ \hat{\phi}_{\text{BM}} &= \arg \max_{\phi_{\text{BM}}} \left| \mathbf{w}_2^H \mathbf{U}_r(\phi_{\text{BM}}) \right|, \\ \hat{\tau}_{\text{BM}} &= \arg \max_{\tau_{\text{BM}}} \left| \mathbf{y}_{\text{LoS},2}^H \mathbf{t}(\tau_{\text{BM}}) \right|, \end{aligned} \quad (34)$$

where $\mathbf{f}_2 = [\mathbf{W}_2]_{:,i_{\text{BS}}}$ and $\mathbf{w}_2 = [\mathbf{W}_2]_{:,i_{\text{MS}}}$ are the hybrid precoding codeword of BS and the hybrid combining codeword of MS determined during the DT stage. $\mathbf{U}_t(\theta_{\text{BM}}) \in \mathbb{C}^{N_{\text{B}} \times K_s}$ and $\mathbf{U}_r(\phi_{\text{BM}}) \in \mathbb{C}^{N_{\text{M}} \times K_s}$ are the search dictionaries corresponding to θ_{RM} and ϕ_{RM} , respectively, while $\mathbf{t}(\tau_{\text{BM}})$ is defined similarly to (32). Having obtained the estimates in (34), the location and orientation of MS can be obtained as

$$\begin{aligned} \hat{\mathbf{m}} &= \mathbf{b} + c\hat{\tau}_{\text{BM}} [\cos(\hat{\theta}_{\text{BM}}), \sin(\hat{\theta}_{\text{BM}})], \\ \hat{\alpha} &= \pi + \hat{\theta}_{\text{BM}} - \hat{\phi}_{\text{BM}}. \end{aligned} \quad (35)$$

Note that, as mentioned in Section IV-B, to mitigate the channel power leakage problem, one column and one row neighbouring to the grid identified at the IT stage are added to form the higher resolution search grids for the DT stage. Due to this, the total numbers of codewords searched are slightly different for the two search stages. To sum up, the proposed improved IS beam sweeping localization scheme is described by Algorithm 1 supported by the diagram-based in Fig. 4.

V. SIMULATION RESULTS AND ANALYSIS

In this section, we first compare the performance of the proposed improved IS beam sweeping algorithm with that of the conventional algorithms, such as, the exhaustive search (ES), hierarchical codebook (HC) and partial search (PS), when the impacts of the number of antennas, size of codebook, electrically large and small RIS are respectively considered. Then, we compare the training overhead required by the different beam sweeping algorithms, when the effect

Algorithm 1 Improved Adaptive Interlaced Scanning (IS) Beam Sweeping Based Localization

Inputs: Hybrid codebooks \tilde{W} for both BS and MS, pure analog codebook \tilde{M} for RIS, number of subcarriers N .

Objectives: estimation of \hat{m} , $\hat{\alpha}$.

```

1: Frame 1. Joint active beamforming at BS and MS.
2: for  $s = 1 : 2$  do
3:   for  $k = 1 : K_s$  do
4:     for  $n = -(N - 1)/2 : (N - 1)/2$  do
5:       Store the received signals for all-pair of codewords

$$y_{\text{LoS},s,k}[n] = \sqrt{P} \rho_{\text{LoS}} \gamma_{\text{BM},0} e^{-j2\pi \tau_{\text{BM},0} \frac{nB}{N}} W_s^H \times a_r(\phi_{\text{BM},0}) a_t^H(\theta_{\text{BM},0}) f_{\text{VLoS},s,k} + W_s^H \tilde{n}[n].$$

6:     end for
7:     Store the  $k$ -th received signal as (28).
8:   end for
9:   Calculate the sum-power matrix as (29), and find the desired pair of codewords, which are indexed by  $i_{\text{BS}}$  and  $j_{\text{MS}}$ .
10: end for

```

```

11: Frame 2. Joint active and passive beamforming at RIS and MS.
12: for  $s = 1 : 2$  do
13:   for  $k = 1 : K_s$  do
14:     for  $n = -(N - 1)/2 : (N - 1)/2$  do
15:       Store the received signals for all-pair of codewords according to (30).
16:     end for
17:     Store the  $k$ -th received signal similarly as (28).
18:   end for
19:   Calculate the sum-power matrix as (31), and find the desired pair of codewords, which are indexed by  $i_{\text{RIS}}$  and  $j_{\text{MS}}$ .
20: end for
21: Estimate the channel parameters by following (32) or (34) based on the selected pair of codewords.
22: Estimate the MS's location  $\hat{m}$  and orientation  $\hat{\alpha}$  using (33) or (35) based on the corresponding channel parameters estimated from the LoS or VLoS path.
23: return  $\hat{m}$ ,  $\hat{\alpha}$ .

```

of the LoS blockage rate is taken into account. Finally, we compare the CRLB in the scenarios, when LoS only, VLoS only and both LoS and VLoS present.

A. SIMULATION SETUP

Table 2 details the parameters used in simulations. A system with BS, MS and RIS of all employing 16 antennas is considered for comparing the performance of the different beam sweeping algorithms, and demonstrating the impact of the different number of codewords (beam-space resolutions). As shown in Table 2, the locations of BS, MS, and RIS are

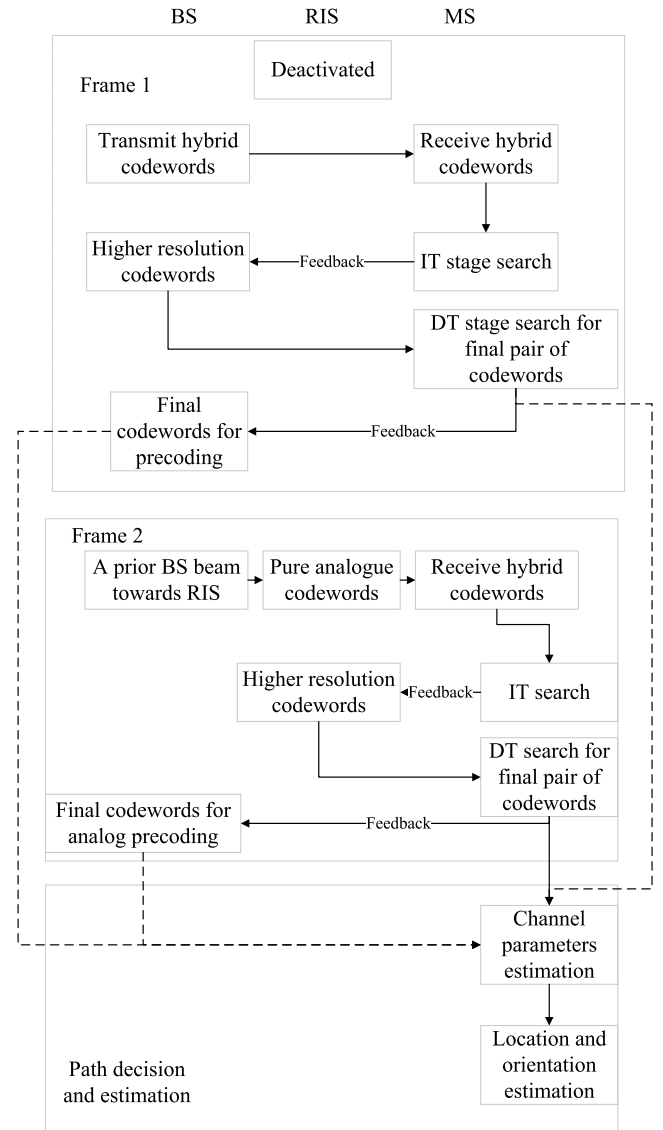


FIGURE 4. Flow chart to illustrate the proposed partial beam sweeping localization scheme.

TABLE 2. Simulation parameters.

Parameters	Value
No. of Antennas N_b, N_m, N_r	16
Speed of Light c	2.99792×10^8 m/s
Carrier Frequency f_c	60 GHz
Bandwidth B	100 MHz
No. of Total Codewords K_S	64 or 512
Path Loss γ	Eq. (14)
Reflection Loss ζ	-13 dB
No. of Subcarriers N	31
Location of RIS, \mathbf{p}	$[5, 30]^T$
Location of BS, \mathbf{b}	$[0, 0]^T$
Location of MS, \mathbf{s}	$[30, 0]^T$
Noise Power	-89 dBm

at $\mathbf{b} = [0, 0]^T$, $\mathbf{m} = [30, 0]^T$, and $\mathbf{r} = [5, 30]^T$, respectively. The carrier frequency f_c is set to 60 GHz, while the bandwidth is $B = 100$ MHz, and the noise power is set to -89 dBm. The reflection loss is set as $\zeta = -13$ dB [12].

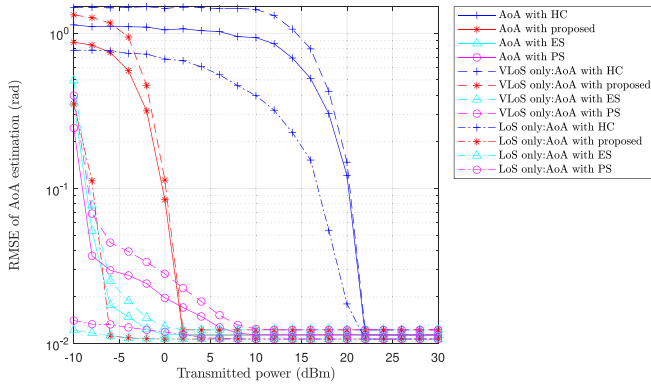


FIGURE 5. RMSE performance of the estimated AoA for $N_b = N_m = N_r = 16$ and 64 codewords at each of the terminals, when different beam sweeping algorithms are employed.

For the path blockage, we set the blockage rate according to that in [35], which is $\rho = \frac{2}{\pi} \lambda_{\text{blocker}} \nu \frac{(h_{\text{blocker}} - h_m)}{(h_b - h_m)} d_{\text{BM}}$, where $\lambda_{\text{blocker}} = 0.1$ denotes the blocker density, and $\nu = 1$ m/s is the blocker's moving speed, with $h_b = 5$ m, $h_m = 1.2$ m and $h_{\text{blocker}} = 2$ m denoting the heights of BS, MS and blocker, respectively. Finally, the estimation performance is measured by the root mean squared error (RMSE), defined as:

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_k \|\hat{\mathbf{x}} - \mathbf{x}\|^2}, \quad (36)$$

where K denotes the number of Monte Carlo trials, \mathbf{x} and $\hat{\mathbf{x}}$ are the true and estimated MS location or orientation.

B. COMPARISON OF DIFFERENT BEAM SWEEPING ALGORITHMS

In this subsection, the numbers of antennas at BS, MS and RIS are all set to 16. For the hierarchical codebook (HC) algorithm [24], the first layer contains 2 codewords at each side, and the total number of layers is $S = 6$. For the partial beam sweeping (PS) algorithm [42], two layers are assumed and both of them have the same resolution, and 32 of the 64 codewords are employed at either sides. After the first layer search, one neighbouring column and one neighbouring row are added to the second layer search, as illustrated in Fig. 2(c). For the ES algorithm, all the 64 codewords at both sides are trained for obtaining the optimal pair of codewords.

Finally, for the proposed improved interlaced scanning (IS) algorithm, the number of codewords for the ($s = 2$)nd layer is set to $K_2 = 64$ at the BS, RIS and MS, which are used during the DT stage of search. Thus, for the IT stage, we have $K_1 = \sqrt{64} = 8$ codewords.

In Fig. 5, we compare the RMSE performance of the AoA estimation by the different beam sweeping algorithms, including the ES, HC, PS and the proposed algorithm, when considering the different scenarios, including LoS only, VLoS only and the unknown LoS blockage. First, the simulation results show that, for all the four beam sweeping algorithms, the performance of the estimation based on the LoS path, if it is available, is better than that based on the

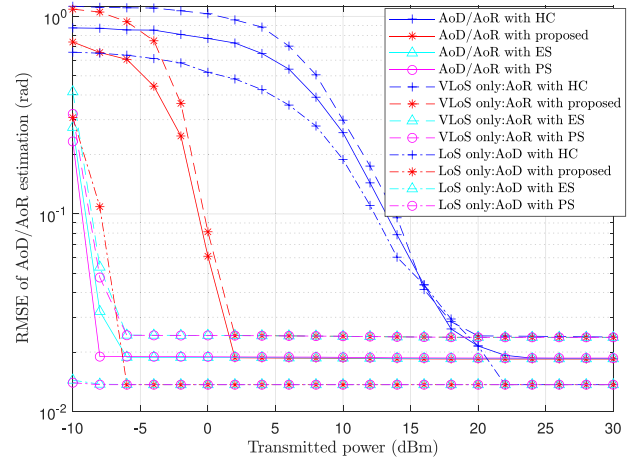


FIGURE 6. RMSE performance of the estimated AoD (AoR) for $N_b = N_m = N_r = 16$ and 64 codewords used at each terminal, when different beam sweeping algorithms are employed.

VLoS path, due to its larger path loss. Second, as expected, the ES algorithm outperforms all the other algorithms, since it does the exhaustive search and hence, demands the highest complexity. Third, it is known that the PS algorithm requires about half of the complexity of the ES algorithm, while its performance is much better than that of the HC codebook. However, we should note that, when the number of codewords is large, the first search stage of the PS algorithm will be of high complexity. Finally, as seen in Fig. 5, the proposed method converges at 2 dBm in the case of unknown LoS blockage, which is about 2 dBm worse than the ES algorithm. However, when taking into account the complexity, the proposed method can achieve a superior trade-off between complexity and performance, when compared with the ES algorithm. More explicitly, the complexity and training overhead trade-off will be analyzed in more detail later.

Fig. 6 compares the RMSE performance of AoD (or AoR) estimation, when different beam sweeping algorithms are respectively employed, and when assuming different scenarios, including LoS only, VLoS only and the unknown LoS blockage. Similar to Fig. 5, the LoS only scenario provides a higher estimation accuracy than the VLoS only scenario. When comparing Fig. 5 with Fig. 6, we can observe that there is difference between the estimation accuracy of the AoD (or AoR) and AoA of the same path. This can be explained by the system geometry, as the AOA to be estimated is closer to beamspace grid than the AoD (or AoR), yielding a higher accuracy of AoA estimation. Additionally, it can be seen that for a given beam sweeping algorithm, the RMSE floor presents, when SNR reaches a relatively high value. This floor is due to the beamspace grids, which unavoidably introduce quantization error. The other observations are similar as that obtained from Fig. 5.

The corresponding RMSE performance of the ToA estimation is shown in Fig. 7. As shown in Fig. 7, the RMSE of the LoS path based estimation converges to 2×10^{-9}

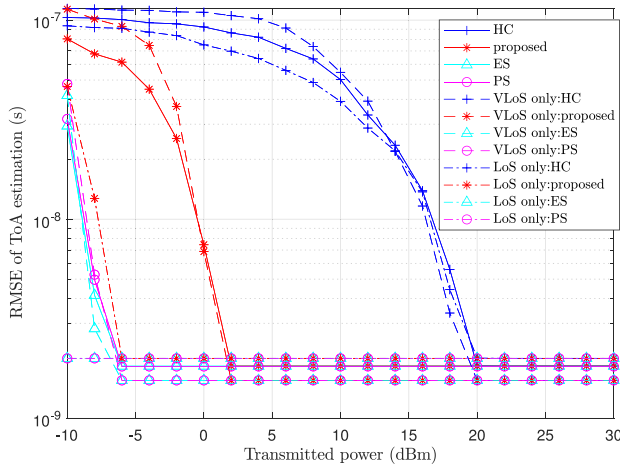


FIGURE 7. RMSE performance of the estimated ToA for $N_b = N_m = N_r = 16$ and 64 codewords at each of the terminals, when different beam sweeping algorithms are employed.

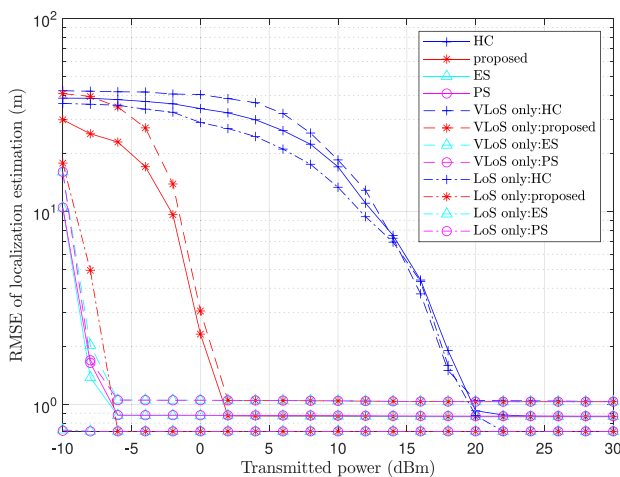


FIGURE 8. RMSE performance of the estimated location of the MS for $N_b = N_m = N_r = 16$ and 64 codewords at each of the terminals, when different beam sweeping algorithms are employed.

second(s), while the VLoS path based estimation converges to 1.55×10^{-9} s. Note that the ToA estimation is only related to the combined grids of the received signal power and delay. In this study, the step size of time delay is set to 10^{-8} s. Furthermore, for the pure LoS case, we can find that the ToA estimation can obtain a 4 dBm gain after convergence, which is provided by the dominated power of the LoS path.

Based on the estimated AoD (or AoR) and ToA as well as (33) and (35), the RMSE performance of positioning is demonstrated in Fig. 8, when the case of LoS only, VLoS only, and the unknown LoS blockage are respectively considered. Explicitly, for all the beam sweeping algorithms, the LoS case outperforms both the VLoS and the unknown LoS blockage cases, converging to the RMSE of 70 cm. For the VLoS case, the RMSE of localization is around 1 m, while for the unknown LoS blockage case, this error is about 85 cm. As shown in Fig. 8, our proposed algorithm needs about 9 dBm more transmit power to reach the convergent

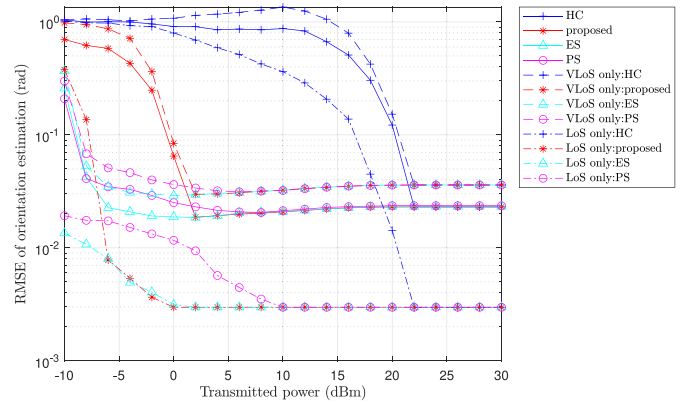


FIGURE 9. RMSE performance of the estimated orientation of the MS for $N_b = N_m = N_r = 16$ with 64 codewords at each of the terminals, when different beam sweeping algorithms are employed.

performance than the PS algorithm, which is still 18 dBm less than the HC algorithm required.

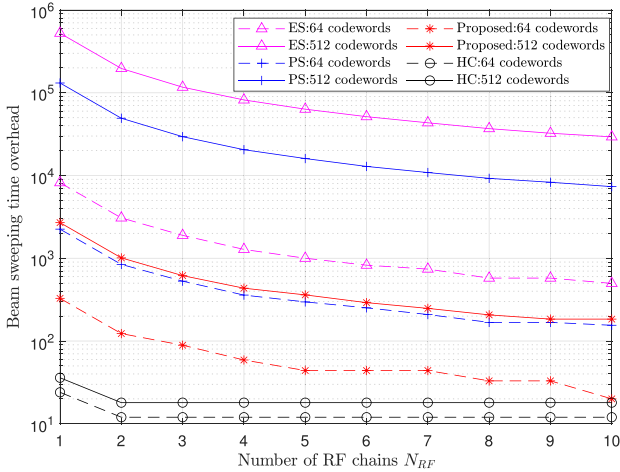
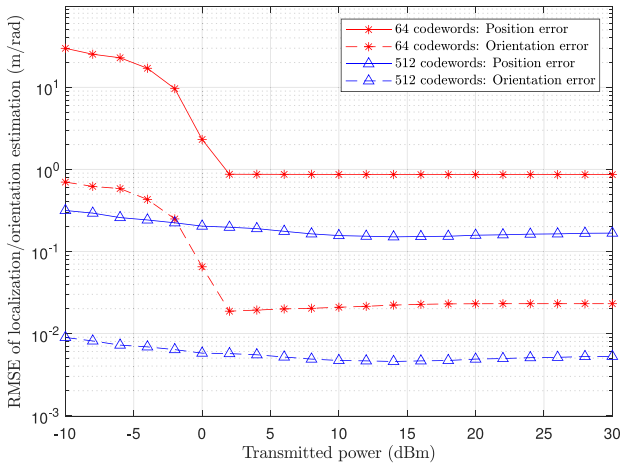
Finally, based on the estimated AoA, (33) and (35), the RMSE performance of the MS's orientation estimation is shown in Fig. 9, from which we can find that the orientation error has the same trend as the AoA estimation error shown in Fig. 5, as it is the AoA-based orientation estimation scheme. As shown in Fig. 9, in the LoS only case, the proposed algorithm is capable of achieving better performance than the PS and HC algorithms, when the transmitted power is less than 10 dBm. This is because the first level search in the proposed algorithm can provide an approximate orientation range, especially in low SNR regime, which is further enhanced by the second level search. More explicitly, although the PS algorithm has a higher resolution during the first level search, it only searches over half of the grids. By contrast, for the proposed algorithm, all the possible AoA ranges are searched in the low resolution domain, which is robust to noise. Hence, it outperforms the PS algorithm in low SNR region. For the VLoS only case, the PS algorithm attains better performance than the proposed algorithm. This is because in this case the received signal is very weak, and more codewords are required by the proposed algorithm at the IT stage. Therefore, the PS algorithm outperforms the proposed algorithm. However, the PS algorithm employs more codewords in the first level search than the proposed one, leading to much higher complexity, as in the first level search, the beamspace angle range is significantly wide, which requires more grids. The training overhead of all the algorithms is analyzed in the next subsection, to prove the superiority of our proposed algorithm.

C. TRAINING OVERHEAD

Since all the codebook-based algorithms converge to the same error floor of estimation, we investigate the training cost and performance trade-off. First, in Fig. 10, we compare the beam sweeping time overhead of the different algorithms, when the transmitted power $P_{TX} = 5$ dBm.

TABLE 3. Examples showing the relationship between codebook size and training overhead of different algorithms.

The codebook size of the proposed algorithm	4	8	16	32	64	128
The training overhead of proposed algorithm	10	17	26	50	82	170
The codebook size of the ES algorithm	2	4	8	16		
The training overhead of the ES algorithm	4	16	64	256		
The codebook size of the PS algorithm	4	8	16	32		
The training overhead of the PS algorithm	8	20	68	260		


FIGURE 10. Beam training time overhead comparison when different beam sweeping algorithms are invoked.

FIGURE 11. RMSE versus transmit power performance, when different codebook resolutions are employed.

Note that, since RIS is only capable of adjusting phases, no RF chains are employed. As shown in Fig. 10, when the number of RF chains $N_{RF} > 2$, the required training time becomes a constant for the HC algorithm, which is because the HC algorithm only considers 2 codewords per layer. Thus, $N_{RF} = 2$ is enough for the HC algorithm, which leads to the lowest hardware complexity and training overhead. With the increase of N_{RF} , the training overhead of the ES algorithm decreases significantly, as the result that it employs large number of codewords K_s for beam search. However, it still has a high training overhead, even when the number of RF chains is large. In terms of our proposed

algorithm, from Fig. 5 to Fig. 9, we know that it has a small performance loss in comparison with the ES algorithm. However, when compared with the ES and PS algorithms, the complexity and the training overhead of our proposed algorithm are reduced by about 25 times and 7 times, respectively. Moreover, according to Table 1, in Table 3, some codebook sizes and the corresponding training overheads of the ES, PS and proposed algorithms are provided. From the results we can see that the training overhead of the proposed algorithm is only linearly dependent on its codebook size. By contrast, the training overheads of the ES and PS algorithms are exponentially related to the codebook size. Hence, given a training overhead, the proposed algorithm is able to employ more codewords to achieve a higher resolution. Therefore, based on the above analysis, we can be inferred that the proposed algorithm is capable of providing a promising trade-off between the achievable performance and the training overhead, when compared with the HC, PS and ES algorithms.

D. EFFECT OF CODEBOOK RESOLUTION

In Fig. 11, the RMSE performance of both the positioning and orientation estimation of the MS is illustrated, when different codebook resolution is considered, where a bigger number of codewords yields a higher resolution. As shown in Fig. 11, when the number of codewords is increased from 64 to 512, the localization error is reduced from 87 cm to 15 cm in average, when the transmitted power is higher than 2 dBm. In this case, the orientation estimation accuracy is improved from 0.02 rads to 0.005 rads.

E. EFFECT OF LOS BLOCKAGE RATE AND FADING

In Fig. 12, we show the effect of the density of the blockers on the LoS path on the positioning performance, when the transmitted power is set to $P_{TX} = 20$ dBm. As shown in Fig. 12, first, it is explicitly shown that when the number of codewords is increased from 64 to 512, the RMSE of localization estimation is reduced from 96 cm to 26 cm. Second, according to [35], when the blocker density is $\lambda_{\text{blocker}} = 0.2$, the blockage rate of LoS path is $\rho = 0.8042$. As shown in Fig. 12, when $\lambda_{\text{blocker}} > 0.3$, the LoS path is basically not available, and hence the localization of MS is fully dependent on the VLoS path. Additionally, according to Fig. 8, when the transmitted power is $P_{TX} = 20$ dBm, the proposed beam sweeping algorithm has the same performance as the HC, PS and the ES algorithms. This can also be clearly seen in Fig. 12, the proposed algorithm has almost the same performance as the HC algorithm, when a given number of codewords is assumed.

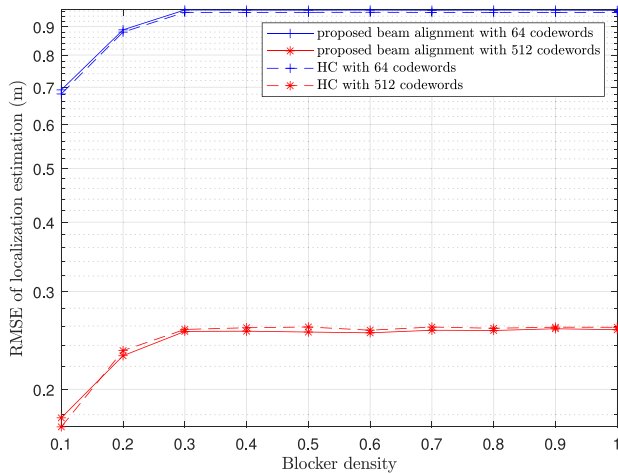


FIGURE 12. RMSE performance versus blocker density, when different number of codewords are invoked by the proposed and the HC algorithms.

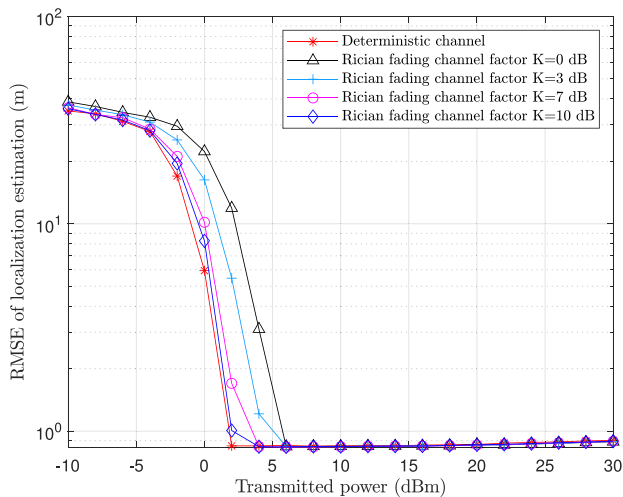


FIGURE 13. RMSE performance of the estimated location of MS under deterministic channel and fading channel for $N_b = N_M = N_R = 16$ with 64 codewords at each of the terminals, when the proposed algorithm is employed.

In the above simulations, the LoS/VLoS channels are assumed to be deterministic experiencing only the large-scale fading. In Fig. 13, we validate this assumption by assuming that the LoS/VLoS channels experience the small-scale Rician fading, in addition to the large-scale fading. In the Rician fading, the K-factor is defined as the ratio between the power of LoS component and that of NLoS components. Hence, a larger value of K implies that more power is delivered by the LoS component. As shown in Fig. 13, the positioning performance converges to that achieved in deterministic channel, as the value of K increases. Specifically, when $K \geq 7$ dB, the positioning performance achieved over Rician fading channel is close to that attained in deterministic channel. In practice, since mmWave systems utilize transmitter/receiver beamforming, the Rician K-factor is generally large, typically in the range of 9 to 15 dB [45]. Hence, it is reasonable to approximate the LoS/VLoS channels as the deterministic channels experiencing only large-scale fading.

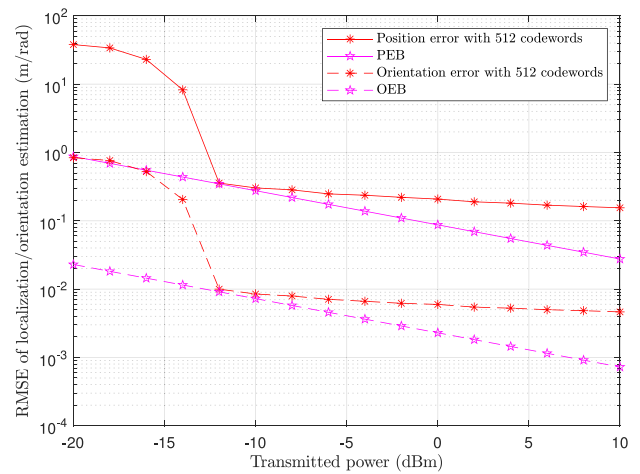


FIGURE 14. RMSE performance of positioning and orientation estimation, when 512 codewords are employed in the DT stage of the proposed algorithm.

F. PEB V.S. OEB

Finally, in Fig. 14, the PEB and OEB, as shown in (17) and (18), are compared with the proposed improved IS algorithm, when 32 antennas are respectively employed at BS, MS and RIS, and when 512 codewords are employed in the DT stage. From the results, we can see that the proposed method has 25 cm of localization error in average, while the PEB is about 10 cm, when the transmit power is -12 dBm. When the transmit power is higher than -12 dBm, the RMSE gap between the proposed algorithm and the PEB becomes larger, which is due to the number of codewords limited by the number of antennas. As shown in Fig. 14, the relationship between the orientation estimation error and the OEB is similar as that between the localization estimation error and the PEB.

VI. CONCLUSION

In this paper, the downlink joint active and passive beamforming based localization scheme with the aid of RIS was presented, where a limited feedback assisted adaptive phase shift scheme for localization was introduced. Specifically, a novel beam sweeping algorithm for both active and passive beamforming was proposed. Compared with the conventional beam sweeping algorithms, i.e., HC, PS and ES, introduced for estimating AoA, AoD (or AoR) and ToA for the position and orientation estimation, the proposed algorithm is capable of providing a superior trade-off between performance and training overhead and hence, it is feasible for practical application. In more detail, our studies showed that the training overhead of the proposed technique is 100 times less than the ES algorithm, and it requires 13 dBm less of transmit power than the HC algorithm. For the proposed algorithm alone, when the codebook resolution increases, our results showed that while the localization and orientation estimation accuracy improve, it can also attain the power gain for performance enhancement. This is especially important when LoS path has a high probability to be blocked. In this case,

employing a high resolution codebook is a promising way to enhance the positioning accuracy. Additionally, our studies showed that electrically large RIS is desirable for attaining relatively good positioning performance.

APPENDIX A ELEMENTS IN (10)

The elements obtained based on (10) are given by

$$\Psi_n(\tau_{\text{BM}}, \tau_{\text{BM}}) = \frac{PN_M \gamma_{\text{BM}}^2}{\sigma^2} \left(\frac{2\pi nB}{N} \right)^2 \|\mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n]\|_2^2, \quad (37)$$

$$\Psi_n(\tau_{\text{BM}}, \theta_{\text{BM}}) = \frac{P\gamma_{\text{BM}}^2}{\sigma^2} \times \Re \left\{ \frac{j2\pi nB}{N} \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{d}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (38)$$

$$\Psi_n(\tau_{\text{BM}}, \phi_{\text{BM}}) = \frac{P\gamma_{\text{BM}}^2}{\sigma^2} \times \Re \left\{ \frac{j2\pi nB}{N} \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \times \mathbf{d}_r(\phi_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (39)$$

$$\Psi_n(\tau_{\text{BM}}, \gamma_{\text{BM}}) = 0, \quad (40)$$

$$\Psi_n(\tau_{\text{BM}}, \tau_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \left(\frac{2\pi nB}{N} \right)^2 \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (41)$$

$$\Psi_n(\tau_{\text{BM}}, \phi_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \Re \left\{ \frac{j2\pi nB}{N} t[n] \alpha[n] \mathbf{x}^H[n] \times \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{d}_t(\theta_{\text{RM}}) \right\}, \quad (42)$$

$$\Psi_n(\tau_{\text{BM}}, \gamma_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BR}}}{\sigma^2} \Re \left\{ \frac{j2\pi nB}{N} t[n] \alpha[n] \mathbf{x}^H[n] \times \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (43)$$

$$\Psi_n(\tau_{\text{BM}}, \theta_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ \frac{j2\pi nB}{N} t[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \times \mathbf{d}_t^H(\theta_{\text{RM}}) \widehat{\boldsymbol{\Sigma}} \mathbf{a}_r(\phi_{\text{BR}}) \mathbf{a}_t^H(\theta_{\text{BR}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (44)$$

$$\Psi_n(\theta_{\text{BM}}, \theta_{\text{BM}}) = \frac{PN_M \gamma_{\text{BM}}^2}{\sigma^2} \|\mathbf{d}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n]\|_2^2, \quad (45)$$

$$\Psi_n(\theta_{\text{BM}}, \phi_{\text{BM}}) = \frac{P\gamma_{\text{BM}}^2}{\sigma^2} \times \Re \left\{ \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \times \mathbf{d}_r(\phi_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (46)$$

$$\Psi_n(\theta_{\text{BM}}, \gamma_{\text{BM}}) = \frac{PN_M \gamma_{\text{BM}}}{\sigma^2} \times \Re \left\{ \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (47)$$

$$\Psi_n(\theta_{\text{BM}}, \tau_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ \frac{-j2\pi nB}{N} t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \times \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (48)$$

$$\Psi_n(\theta_{\text{BM}}, \phi_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{d}_r(\phi_{\text{RM}}) \right\}, \quad (49)$$

$$\Psi_n(\theta_{\text{BM}}, \gamma_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BR}}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (50)$$

$$\Psi_n(\theta_{\text{BM}}, \theta_{\text{RM}}) = \frac{P\gamma_{\text{BM}}}{\gamma_{\text{BRM}} \sigma^2} \times \Re \left\{ t[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \times \mathbf{d}_t^H(\theta_{\text{RM}}) \widehat{\boldsymbol{\Sigma}} \mathbf{a}_r(\phi_{\text{BR}}) \mathbf{a}_t^H(\theta_{\text{BR}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (51)$$

$$\Psi_n(\phi_{\text{BM}}, \phi_{\text{BM}}) = \frac{P\gamma_{\text{BM}}^2}{\sigma^2} \|\mathbf{d}_r(\phi_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n]\|_2^2, \quad (52)$$

$$\Psi_n(\phi_{\text{BM}}, \gamma_{\text{BM}}) = \frac{P\gamma_{\text{BM}}}{\sigma^2} \times \Re \left\{ \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \times \mathbf{d}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (53)$$

$$\Psi_n(\phi_{\text{BM}}, \tau_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ \frac{-j2\pi nB}{N} t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \times \mathbf{d}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (54)$$

$$\Psi_n(\phi_{\text{BM}}, \phi_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{d}_r^H(\phi_{\text{BM}}) \mathbf{d}_r(\phi_{\text{RM}}) \right\}, \quad (55)$$

$$\Psi_n(\phi_{\text{BM}}, \gamma_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BR}}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{d}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (56)$$

$$\Psi_n(\phi_{\text{BM}}, \theta_{\text{RM}}) = \frac{P\gamma_{\text{BM}}\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ t[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{d}_r(\phi_{\text{BM}}) \mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \times \mathbf{d}_t^H(\theta_{\text{RM}}) \widehat{\boldsymbol{\Sigma}} \mathbf{a}_r(\phi_{\text{BR}}) \mathbf{a}_t^H(\theta_{\text{BR}}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (57)$$

$$\Psi_n(\gamma_{\text{BM}}, \gamma_{\text{BM}}) = \frac{PN_M}{\sigma^2} \|\mathbf{a}_t^H(\theta_{\text{BM}}) \mathbf{F} \mathbf{x}[n]\|_2^2, \quad (58)$$

$$\Psi_n(\gamma_{\text{BM}}, \tau_{\text{RM}}) = \frac{P\gamma_{\text{BRM}}}{\sigma^2} \times \Re \left\{ \frac{-j2\pi nB}{N} t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \times \mathbf{a}_t(\theta_{\text{BM}}) \mathbf{a}_r^H(\phi_{\text{BM}}) \mathbf{a}_r(\phi_{\text{RM}}) \right\}, \quad (59)$$

$$\Psi_n(\gamma_{BM}, \phi_{RM}) = \frac{P\gamma_{BRM}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{BM}) \mathbf{a}_r^H(\phi_{BM}) \mathbf{d}_r(\phi_{RM}) \right\}, \quad (60)$$

$$\Psi_n(\gamma_{BM}, \gamma_{RM}) = \frac{P\gamma_{BR}}{\sigma^2} \times \Re \left\{ t[n] \alpha[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{BM}) \mathbf{a}_r^H(\phi_{BM}) \mathbf{a}_r(\phi_{RM}) \right\}, \quad (61)$$

$$\Psi_n(\gamma_{BM}, \theta_{RM}) = \frac{P\gamma_{BRM}}{\sigma^2} \times \Re \left\{ t[n] \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{BM}) \mathbf{a}_r^H(\phi_{BM}) \mathbf{a}_r(\phi_{RM}) \times \mathbf{d}_t^H(\theta_{RM}) \widehat{\mathbf{\Omega}} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (62)$$

$$\Psi_n(\tau_{RM}, \tau_{RM}) = \frac{PN_M \gamma_{BRM}^2}{\sigma^2} \left(\frac{2\pi n B |\alpha[n]|}{N} \right)^2, \quad (63)$$

$$\Psi_n(\tau_{RM}, \phi_{RM}) = \frac{P\gamma_{BRM}^2}{\sigma^2} \times \Re \left\{ \frac{j2\pi n B}{N} \alpha^*[n] \mathbf{a}_r^H(\phi_{RM}) \mathbf{d}_t(\theta_{RM}) \times \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (64)$$

$$\Psi_n(\tau_{RM}, \gamma_{RM}) = 0, \quad (65)$$

$$\Psi_n(\tau_{RM}, \theta_{RM}) = \frac{PN_M \gamma_{BRM}^2}{\sigma^2} \times \Re \left\{ \frac{j2\pi n B}{N} \alpha^*[n] \mathbf{d}_t^H(\theta_{RM}) \widehat{\mathbf{\Omega}} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (66)$$

$$\Psi_n(\phi_{RM}, \phi_{RM}) = \frac{P\gamma_{BRM}^2}{\sigma^2} \|\alpha[n] \mathbf{d}_r(\phi_{RM})\|_2^2, \quad (67)$$

$$\Psi_n(\phi_{RM}, \gamma_{RM}) = \frac{P\gamma_{BR} \gamma_{BRM}}{\sigma^2} \times \Re \left\{ \alpha^*[n] \mathbf{d}_t^H(\theta_{RM}) \mathbf{a}_r(\phi_{RM}) \mathbf{a}_t^H(\theta_{RM}) \times \mathbf{\Omega} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (68)$$

$$\Psi_n(\phi_{RM}, \theta_{RM}) = \frac{P\gamma_{BRM}^2}{\sigma^2} \times \Re \left\{ \alpha^*[n] \mathbf{d}_r^H(\phi_{RM}) \mathbf{a}_r(\phi_{RM}) \mathbf{d}_t^H(\theta_{RM}) \times \widehat{\mathbf{\Omega}} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (69)$$

$$\Psi_n(\gamma_{RM}, \gamma_{RM}) = \frac{PN_M \gamma_{BR}^2}{\sigma^2} |\alpha[n]|^2, \quad (70)$$

$$\Psi_n(\gamma_{RM}, \theta_{RM}) = \frac{PN_M \gamma_{BR} \gamma_{BRM}}{\sigma^2} \times \Re \left\{ \alpha^*[n] \mathbf{d}_t^H(\theta_{RM}) \widehat{\mathbf{\Omega}} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}, \quad (71)$$

$$\Psi_n(\theta_{RM}, \theta_{RM}) = \frac{PN_M \gamma_{BRM}^2}{\sigma^2} \times \Re \left\{ \mathbf{x}^H[n] \mathbf{F}^H \mathbf{a}_t(\theta_{BR}) \mathbf{a}_r^H(\phi_{BR}) \times \widehat{\mathbf{\Omega}} \mathbf{d}_t(\theta_{RM}) \mathbf{d}_t^H(\theta_{RM}) \widehat{\mathbf{\Omega}} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n] \right\}. \quad (72)$$

$$\mathbf{d}_t(\theta_{BM}) = \partial \mathbf{a}_t(\theta_{BM}) / \partial \theta_{BM} = j2\pi \frac{d}{\lambda} \cos(\theta_{BM}) \text{diag}(0, \dots, N_B - 1), \quad (73)$$

$$\mathbf{d}_r(\theta_{BM}) = \partial \mathbf{a}_r(\theta_{RM}) / \partial \theta_{RM} = j2\pi \frac{d}{\lambda} \cos(\theta_{RM}) \text{diag}(0, \dots, N_R - 1), \quad (74)$$

$$\mathbf{d}_r(\phi_{BM}) = \partial \mathbf{a}_r(\phi_{BM}) / \partial \phi_{BM} = j2\pi \frac{d}{\lambda} \cos(\phi_{BM}) \text{diag}(0, \dots, N_M - 1), \quad (75)$$

$$\mathbf{d}_r(\phi_{RM}) = \partial \mathbf{a}_r(\phi_{RM}) / \partial \phi_{RM} = j2\pi \frac{d}{\lambda} \cos(\phi_{RM}) \text{diag}(0, \dots, N_M - 1). \quad (76)$$

Furthermore, the derivative of the diagonal phase shift matrix $\mathbf{\Omega}$ of the RIS can be represented as

$$\widehat{\mathbf{\Omega}} = \text{diag} \left(e^{j\pi(0, \dots, N_R - 1) \sin(\theta_{RM})} j\pi(0, \dots, N_R - 1) \cos(\theta_{\text{diff}}) \right), \quad (77)$$

where $\theta_{\text{diff}} = \theta_{RM} - \phi_{BR}$. Moreover, the exponential parameter related to the time delay is expressed as $t[n] = e^{-j2\pi n B (\tau_{BR} + \tau_{RM} - \tau_{BM}) / N}$, and $\alpha[n] = \mathbf{a}_t^H(\theta_{RM}) \mathbf{\Omega} \mathbf{a}_r(\phi_{BR}) \mathbf{a}_t^H(\theta_{BR}) \mathbf{F} \mathbf{x}[n]$.

APPENDIX B ENTRIES OF JACOBIAN MATRIX T IN (12)

$$\partial \tau_{BM} / \partial m_x = \cos(\theta_{BM}) / c, \quad (78)$$

$$\partial \theta_{BM} / \partial m_x = -\sin(\theta_{BM}) / \|\mathbf{b} - \mathbf{m}\|_2, \quad (79)$$

$$\partial \phi_{BM} / \partial m_x = -\sin(\theta_{BM}) / \|\mathbf{b} - \mathbf{m}\|_2, \quad (80)$$

$$\partial \gamma_{BM} / \partial m_x = \frac{\lambda \cos(\theta_{BM})}{4\pi \|\mathbf{b} - \mathbf{m}\|_2^2}, \quad (81)$$

$$\partial \tau_{RM} / \partial m_x = \cos(\theta_{RM}) / c, \quad (82)$$

$$\partial \theta_{RM} / \partial m_x = -\sin(\theta_{RM}) / \|\mathbf{r} - \mathbf{m}\|_2, \quad (83)$$

$$\partial \phi_{RM} / \partial m_x = -\sin(\theta_{RM}) / \|\mathbf{r} - \mathbf{m}\|_2, \quad (84)$$

$$\partial \gamma_{RM} / \partial m_x = \frac{\lambda \cos(\theta_{RM})}{4\pi \|\mathbf{r} - \mathbf{m}\|_2^2}, \quad (85)$$

$$\partial \tau_{BM} / \partial m_y = \sin(\theta_{BM}) / c, \quad (86)$$

$$\partial \theta_{BM} / \partial m_y = -\cos(\theta_{BM}) / \|\mathbf{b} - \mathbf{m}\|_2, \quad (87)$$

$$\partial \phi_{BM} / \partial m_y = -\cos(\theta_{BM}) / \|\mathbf{b} - \mathbf{m}\|_2, \quad (88)$$

$$\partial \gamma_{BM} / \partial m_y = \frac{-\lambda \sin(\theta_{BM})}{4\pi \|\mathbf{b} - \mathbf{m}\|_2^2}, \quad (89)$$

$$\partial \tau_{RM} / \partial m_y = \sin(\theta_{RM}) / c, \quad (90)$$

$$\partial \theta_{RM} / \partial m_y = \cos(\theta_{RM}) / \|\mathbf{r} - \mathbf{m}\|_2, \quad (91)$$

$$\partial \phi_{RM} / \partial m_y = \cos(\theta_{RM}) / \|\mathbf{r} - \mathbf{m}\|_2, \quad (92)$$

$$\partial \gamma_{RM} / \partial m_y = \frac{-\lambda \sin(\theta_{RM})}{4\pi \|\mathbf{r} - \mathbf{m}\|_2^2}, \quad (93)$$

$$\partial \tau_{BM} / \partial \alpha = 0, \quad (94)$$

$$\partial \theta_{BM} / \partial \alpha = 0, \quad (95)$$

$$\partial \phi_{BM} / \partial \alpha = -1, \quad (96)$$

$$\partial \gamma_{BM} / \partial \alpha = 0, \quad (97)$$

$$\partial \tau_{RM} / \partial \alpha = 0, \quad (98)$$

The derivatives of the steering vectors in above equations are defined as

$$\partial\theta_{RM}/\partial\alpha = 0, \quad (99)$$

$$\partial\phi_{RM}/\partial\alpha = -1, \quad (100)$$

$$\partial\gamma_{RM}/\partial\alpha = 0. \quad (101)$$

REFERENCES

- [1] J. del Peral-Rosado et al., "Whitepaper on new localization methods for 5G wireless systems and the Internet-of-Things," in *White Paper of the COST Action CA15104 (IRACON)*, COST Action CA15104, 2018, pp. 1–27.
- [2] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1124–1148, 2nd Quart., 2018.
- [3] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, 2nd Quart., 2018.
- [4] M. R. Akdeniz et al., "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.
- [5] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.
- [6] J. Zhao, "A survey of intelligent reflecting surfaces (IRSs): Towards 6G wireless communication networks," 2019, [arXiv:1907.04789](https://arxiv.org/abs/1907.04789).
- [7] Z. Chen, X. Ma, C. Han, and Q. Wen, "Towards intelligent reflecting surface empowered 6G terahertz communications: A survey," *China Commun.*, vol. 18, no. 5, pp. 93–119, 2021.
- [8] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
- [9] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, "Radio localization and mapping with reconfigurable intelligent surfaces: Challenges, opportunities, and research directions," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, Dec. 2020.
- [10] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
- [11] Q. Wu and R. Zhang, "Beamforming optimization for intelligent reflecting surface with discrete phase shifts," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2019, pp. 7830–7833.
- [12] W. Wang and W. Zhang, "Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6282–6297, Oct. 2021.
- [13] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, 2020.
- [14] J. He, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5786–5797, Sep. 2021.
- [15] Y. Zhang, M. El-Hajjar, and L.-L. Yang, "Adaptive codebook-based channel estimation in OFDM-aided hybrid beamforming mmWave systems," *IEEE Open J. Commun. Society*, vol. 3, pp. 1553–1562, 2022.
- [16] 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.
- [17] R. Koirala, B. Denis, B. Uguen, D. Dardari, and H. Wymeersch, "Localization and throughput trade-off in a multi-user multi-carrier mm-Wave system," *IEEE Access*, vol. 7, pp. 167099–167112, 2019.
- [18] Z. Lin, T. Lv, and P. T. Mathiopoulos, "3D indoor positioning for millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2472–2486, Jun. 2018.
- [19] D. Wang, M. Fattouche, and X. Zhan, "Pursuance of mm-level accuracy: Positioning in mmWave systems," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1169–1180, Jun. 2019.
- [20] K. Li, M. El-Hajjar, and L.-L. Yang, "Millimeter-wave based localization using a two-stage channel estimation relying on few-bit adcs," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1736–1752, 2021.
- [21] B. Zhou, A. Liu, and V. Lau, "Successive localization and beamforming in 5G mmWave MIMO communication systems," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1620–1635, Mar. 2019.
- [22] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, "Location awareness in beyond 5G networks via reconfigurable intelligent surfaces," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2011–2025, Jul. 2022.
- [23] M. Z. Win, Z. Wang, Z. Liu, Y. Shen, and A. Conti, "Location awareness via intelligent surfaces: A path toward holographic NLN," *IEEE Veh. Technol. Mag.*, vol. 17, no. 2, pp. 37–45, Jun. 2022.
- [24] J. He, H. Wymeersch, T. Sanguanpuak, O. Silvén, and M. Juntti, "Adaptive beamforming design for mmWave RIS-aided joint localization and communication," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, 2020, pp. 1–6.
- [25] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Towards ubiquitous positioning by leveraging reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 284–288, Jan. 2021.
- [26] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1376–1393, Feb. 2021.
- [27] J. He, H. Wymeersch, L. Kong, O. Silvén, and M. Juntti, "Large intelligent surface for positioning in millimeter wave MIMO systems," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, 2020, pp. 1–5.
- [28] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of positioning with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1761–1774, May 2018.
- [29] D.-R. Emenonye, H. S. Dhillon, and R. M. Buehrer, "RIS-aided localization under position and orientation offsets in the near and far field," 2022, [arXiv:2210.03599](https://arxiv.org/abs/2210.03599).
- [30] D.-R. Emenonye, H. S. Dhillon, and R. M. Buehrer, "Fundamentals of RIS-aided localization in the far-field," 2022, [arXiv:2206.01652](https://arxiv.org/abs/2206.01652).
- [31] G. Mylonopoulos, C. D'Andrea, and S. Buzzi, "Active reconfigurable intelligent surfaces for user localization in mmWave MIMO systems," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2022, pp. 1–5.
- [32] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, "Indoor localization with reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 161–165, Jan. 2021.
- [33] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," 2020, [arXiv:2009.02818](https://arxiv.org/abs/2009.02818).
- [34] Y. Liu, E. Liu, R. Wang, and Y. Geng, "Reconfigurable intelligent surface aided wireless localization," in *Proc. IEEE ICC Int. Conf. Commun.*, 2021, pp. 1–6.
- [35] I. K. Jain, R. Kumar, and S. Panwar, "Limited by capacity or blockage? A millimeter wave blockage analysis," 2018, [arXiv:1808.01228](https://arxiv.org/abs/1808.01228).
- [36] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.
- [37] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [38] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 141–153, Jan. 2017.
- [39] C. Pan et al., "An overview of signal processing techniques for RIS/IRS-aided wireless systems," 2021, [arXiv:2112.05989](https://arxiv.org/abs/2112.05989).
- [40] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
- [41] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

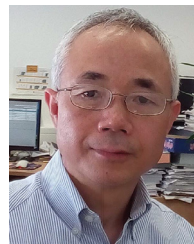
- [42] X. Sun, C. Qi, and G. Y. Li, "Beam training and allocation for multiuser millimeter wave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1041–1053, Feb. 2019.
- [43] Q. Wu, X. Guan, and R. Zhang, "Intelligent reflecting surface aided wireless energy and information transmission: An overview," 2021, *arXiv:2106.07997*.
- [44] A. Huang, L. Tian, T. Jiang, and J. Zhang, "NLOS identification for wideband mmWave systems at 28 GHz," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, 2019, pp. 1–6.
- [45] M. K. Samimi, G. R. MacCartney, S. Sun, and T. S. Rappaport, "28 GHz millimeter-wave ultrawideband small-scale fading models in wireless channels," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC-Spring)*, 2016, pp. 1–6.



KUNLUN LI received the B.Eng. degree in electronics engineering from Coventry University, U.K., in 2017, and the M.Sc. degree in wireless communications from the University of Southampton, U.K., in 2018, where he is currently pursuing the Ph.D. degree with the Next Generation Wireless Research Group. His research interests include millimeter-wave, massive MIMO, channel estimation, and reconfigurable intelligent surface for wireless positioning and communications.



MOHAMMED EL-HAJJAR (Senior Member, IEEE) received the Ph.D. degree in wireless communications from the University of Southampton, U.K., in 2008. Following the Ph.D., he joined Imagination Technologies as a Design Engineer, where he worked on designing and developing Imaginations multi-standard communications platform, which resulted in several patents. He is an Associate Professor with the School of Electronics and Computer Science, University of Southampton. He is the recipient of several academic awards and has published a Wiley-IEEE book and in excess of 80 journal and conference papers in addition to many patents. His research interests include the development of intelligent communications systems, energy-efficient transceiver design, millimeter wave communications, machine learning for wireless communications, and radio over fiber network design.



LIE-LIANG YANG (Fellow, IEEE) received the B.Eng. degree in communications engineering from Shanghai TieDao University, Shanghai, China, in 1988, and the M.Eng. and Ph.D. degrees in communications and electronics from Northern (Beijing) Jiaotong University, Beijing, China, in 1991 and 1997, respectively. He is a Professor of Wireless Communications with the School of Electronics and Computer Science, University of Southampton, U.K. He has published 400+ research papers in journals and conference proceedings, authored/coauthored three books and also published several book chapters. He has research interest in wireless communications, wireless networks and signal processing for wireless communications, as well as molecular communications and nano-networks. He also acted different roles for organization of conferences. He served as an associate editor to various journals and is currently a Senior Editor to the IEEE ACCESS and a Subject Editor to the *Electronics Letters*. He is a Fellow of IET and was a Distinguished Lecturer of the IEEE VTS. For more information, see <https://www.ecs.soton.ac.uk/people/llyang>.