

# Channel Estimation and User Localization for IRS-Assisted MIMO-OFDM Systems

Yuxing Lin<sup>1</sup>, *Student Member, IEEE*, Shi Jin<sup>2</sup>, *Senior Member, IEEE*,  
Michail Matthaiou<sup>3</sup>, *Senior Member, IEEE*, and Xiaohu You<sup>4</sup>, *Fellow, IEEE*

**Abstract**—We consider the channel estimation problem and the channel-based wireless applications in multiple-input multiple-output orthogonal frequency division multiplexing systems assisted by intelligent reconfigurable surfaces (IRSs). To obtain the necessary channel parameters, i.e., angles, delays and gains, for environment mapping and user localization, we propose a novel twin-IRS structure consisting of two IRS planes with a relative spatial rotation. We model the training signal from the user equipment to the base station via IRSs as a third-order canonical polyadic tensor with a maximal tensor rank equal to the number of IRS unit cells. We present four designs of IRS training coefficients, i.e., random, structured, grouping and sparse patterns, and analyze the corresponding uniqueness conditions of channel estimation. We extract the cascaded channel parameters by leveraging array signal processing and atomic norm denoising techniques. Based on the characteristics of the twin-IRS structures, we formulate a nonlinear equation system to exactly recover the multipath parameters by two efficient decoupling modes. We realize environment mapping and user localization based on the estimated channel parameters. Simulation results indicate that the proposed twin-IRS structure and estimation schemes can recover the channel state information with remarkable accuracy, thereby offering a centimeter-level resolution of user positioning.

**Index Terms**—Channel estimation, parameter decoupling, tensor factorization, twin-IRS structure, user localization.

## I. INTRODUCTION

MILLIMETER wave (mmWave) (30–300 GHz) technologies have been identified as a promising candidate for tackling the data traffic deluge and frequency resource shortage in the fifth generation era [1]. Exploiting higher-frequency

spectrum, e.g., terahertz (0.1–10 THz), has also been regarded as one of the potential development directions of the future sixth generation wireless communications [2]. The short signal wavelength enables miniaturized implementation of massive multiple-input multiple-output (MIMO) arrays, providing considerable directional beamforming gains [3]. In order to address the limited coverage and line-of-sight (LoS) obstruction of high-frequency systems, intelligent reconfigurable surfaces (IRSs) have been recently investigated to artificially establish controllable non-line-of-sight (NLoS) links [4].

IRS technologies can contribute to the full coverage and broadband connectivity of the future wireless networks [5]. An IRS, typically a programmable metasurface, is composed of a massive number of unit cells independently interacting with incident waves [6]. The reflection coefficients of IRS elements can be predefined by digital controllers to perform real-time manipulation of electromagnetic responses [7]. IRSs can help realizing a smart wireless propagation environment, providing additional degrees of freedom for transceiver design and network optimization [8]. Some works have already integrated IRSs into wireless applications, e.g., MIMO detection, non-orthogonal multiple access, radio localization and mapping, etc. [9]–[11]. Nevertheless, these services require precise knowledge of the propagation channels and multipath parameters, thereby entailing fundamental challenges to channel estimation involving fully passive IRS modules.

We will now delineate the following relevant works: An efficient scheme that developed parallel and sequential training designs was proposed in [12]. A compressed sensing (CS) method that converted the channel estimation into a sparse recovery problem was proposed in [13]. A tensor-based approach that factorized the cascaded channel by iterative decompositions was presented in [14]. A message-passing algorithm that solved a dictionary learning problem was presented in [15]. A two-timescale training protocol that individually estimated the one-hop channels was developed in [16]. A matrix-calibration scheme that developed a joint calibration and estimation algorithm was developed in [17]. Most of these works achieve separated channels from the base station (BS) or user equipment (UE) to the IRS plane, which, however, induces inherent estimation ambiguities that hinder the exact recovery of multipath parameters. These ambiguities limit the integration of IRSs into environment-dependent applications [11], which can only be removed with strong or unrealistic assumptions, e.g., normalized power, channel

Manuscript received February 5, 2021; revised July 13, 2021; accepted August 28, 2021. Date of publication September 16, 2021; date of current version April 11, 2022. The work of Shi Jin was supported in part by the National Science Foundation of China (NSFC) for Distinguished Young Scholars under Grant 61625106 and in part by the NSFC under Grant 61941104. The work of Michail Matthaiou was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program under Agreement 101001331 and in part by the Research Grant from the Department for the Economy Northern Ireland under the US-Ireland Research and Development Partnership Program. The associate editor coordinating the review of this article and approving it for publication was W. Ni. (*Corresponding authors: Xiaohu You; Shi Jin.*)

Yuxing Lin, Shi Jin, and Xiaohu You are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: yxlin@seu.edu.cn; jinshi@seu.edu.cn; xhyu@seu.edu.cn).

Michail Matthaiou is with the Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, Belfast BT3 9DT, U.K. (e-mail: m.matthaiou@qub.ac.uk).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2021.3111176>.

Digital Object Identifier 10.1109/TWC.2021.3111176

reciprocity, quasi-static states, *a priori* long-term information, etc [14]–[17].

In this paper, we consider the channel estimation of IRS-assisted MIMO orthogonal frequency division multiplexing (OFDM) systems. By leveraging the concept of multi-IRS networks [18], we propose a novel twin-IRS structure composed of two IRS planes with a relative spatial rotation. This three-dimensional (3-D) structure can provide channel information along three orthogonal spatial dimensions, which enables us to accurately extract the channel parameters in 3-D propagation spaces. By leveraging the geometric relationship of cascaded parameters obtained from the twin-IRS structures, we can precisely recover the channel parameters, i.e., angles of arrival (AoAs), angles of departure (AoDs) and time delays, by solving systems of nonlinear equations. Unlike some existing works that rely on strong assumptions, e.g., BS-IRS LoS components with *a priori* information [19], [20], our work is able to exactly decouple the cascaded parameters of common channels, supporting precise environment mapping and localization applications. The main contributions of this paper are summarized as follows:

- We model the training signal as a third-order canonical polyadic (CP) tensor, transforming the channel estimation problem into a tensor factorization problem [21], [22]. By leveraging the assignment flexibility and distribution regularity of IRS unit cells, we apply the MIMO antenna designs and array processing techniques to the IRS configuration. We introduce four training designs, i.e., random, structured, grouping and sparse patterns, as well as, the corresponding uniqueness conditions of channel estimates. The structural information artificially attached to IRS coefficients can help to efficiently reduce the training overhead.
- We combine the concept of sparse/grouping arrays and atomic norm denoising techniques with the tensor operations to realize the cascaded parameter recovery [23]–[26]. Moreover, by leveraging the physical relationship of the novel twin-IRS structures, i.e., the adjacent position and relative rotation, we introduce additional information of channel paths into the phases/amplitudes of the IRS array response vectors. We further formulate a nonlinear equation system of cascaded parameters and propose two efficient decoupling modes to exactly recover the channel parameters, i.e., AoA/AoDs and time delays. Note that in most prior works (e.g. [14]–[17], [19], [20]), these parameters cannot be obtained without strong assumptions or estimation ambiguities.
- By leveraging the geometric characteristics of mmWave propagation, we implement a straightforward 3-D propagation environment mapping, including scatterer positioning of NLoS paths, user localization and orientation determination, based on the recovered channel path parameters. Dedicated wideband and narrowband training schemes with single and multiple twin-IRS structures are respectively proposed.

Simulation results indicate that the proposed schemes can achieve considerable channel estimation accuracy with reduced training overhead for large-scale IRSs with massive

TABLE I  
SYMBOLS OF IMPORTANT VARIABLES AND PARAMETERS

Symbol	Definition
$N_{\text{B}(U)}$	BS (UE) antennas
$M_{\text{B}(U)}$	BS (UE) RF chains
$N_{\text{I}}$	reflectors in one IRS plane
$N_{\text{h}(v)}$	horizontal (vertical) IRS reflectors
$K$	OFDM subcarriers
$K_{\text{tr}}$	training subcarriers
$P_{\text{tr}}$	training frames
$T_{\text{tr}}$	training time slots per frame
$M_{\text{tr}}$	BS training streams
$L_{\text{BI}(U),r}$	BS-IRS (IRS-UE) channel paths
$\alpha_{\text{BI}(U),r,\ell}$	BS-IRS (IRS-UE) path gain
$\tau_{\text{BI}(U),r,\ell}$	BS-IRS (IRS-UE) path delay
$\phi_{\text{B}(U),r,\ell}$	horizontal AoA at BS (AoD at UE)
$\theta_{\text{B}(U),r,\ell}$	vertical AoA at BS (AoD at UE)
$\phi_{\text{A}(D),r,\ell}$	horizontal AoA (AoD) at IRS
$\theta_{\text{A}(D),r,\ell}$	vertical AoA (AoD) at IRS
$G$	grouping pattern: total sub-groups
$N_{\text{gh}(gv)}$	grouping pattern: sub-group size
$N_{\text{hd}(vd)}$	sparse pattern: dense subarray size
$N_{\text{hs}(vs)}$	sparse pattern: sparse subarray size

number of elements. More importantly, the proposed twin-IRS structure can help decouple the multipath parameters, leading to accurate environment mapping with a centimeter-level resolution.

The rest of the paper is organized as follows: Section II introduces the IRS-assisted MIMO-OFDM system, as well as, the proposed twin-IRS structure. Section III presents the training pattern designs of IRS coefficients and the tensor-based channel estimation schemes. Section IV presents the recovery and decoupling procedures of cascaded channel parameters. Section V discusses the application of user localization based on the estimated channel information. Section VI presents the numerical results of the channel estimation schemes and environment mapping applications. Section VI draws the most important conclusions.

*Notations:*  $\mathbf{a}$ ,  $\mathbf{A}$ ,  $\mathcal{A}$  and  $\mathcal{A}$  denote vectors, matrices, tensors and element sets, respectively;  $(\cdot)^T$ ,  $(\cdot)^*$ ,  $(\cdot)^H$  and  $(\cdot)^\dagger$  denote the transpose, conjugate, Hermitian transpose and pseudo-inverse, respectively;  $\mathbf{A}^{(N)}$ ,  $\mathbf{A}^{[N]}$  denote the columns and rows of  $\mathbf{A}$  indexed by  $\mathcal{N}$ , respectively;  $\mathbf{a} \bullet \mathbf{b}$  and  $\mathbf{a} \times \mathbf{b}$  denote the dot and cross products, respectively;  $\|\mathbf{a}\|$ ,  $\|\mathbf{A}\|_{\text{F}}$  denote the 2-norm and Frobenius-norm, respectively;  $\otimes$ ,  $\odot$  and  $\circ$  denote the Kronecker, Khatri-Rao and tensor outer products, respectively;  $\text{Diag}(\mathbf{a})$  denotes the diagonal matrix formed by  $\mathbf{a}$ ;  $\text{Tr}(\mathbf{A})$  denotes the matrix trace;  $\mathcal{I}(n) \triangleq \{1, \dots, n\}$ ;  $\text{r}(\mathbf{A})$ ,  $\text{k}(\mathbf{A})$  denote the general rank and Kruskal-rank, respectively;  $\mathbf{0}_n$ ,  $\mathbf{1}_n$  and  $\mathbf{I}_n$  denote all-zeros, all-ones vectors and identity matrices, respectively;  $\text{Matr}(\mathcal{X}; [k_1, \dots, k_P], [k_{P+1}, \dots, k_N])$  *matricizes*  $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_N}$  into  $\mathbf{X} \in \mathbb{C}^{\prod_{p=1}^P I_{k_p} \times \prod_{q=P+1}^N I_{k_q}}$ ;  $\text{Tens}(\mathbf{X}; [I_1, \dots, I_N], [k_1, \dots, k_P], [k_{P+1}, \dots, k_N])$  *tensorizes*  $\mathbf{X}$  to  $\mathcal{X}$ ;  $\mathcal{X} \times_n \mathbf{U}$  denotes the mode- $n$  tensor-matrix product [27]. The symbols of important variables and parameters are summarized in Table I.

## II. SYSTEM MODEL

We consider an IRS-assisted MIMO-OFDM system as illustrated in Fig. 1(a), where one base station (BS) with  $N_{\text{B}}$  antennas and  $M_{\text{B}}$  radio frequency (RF) chains communicates

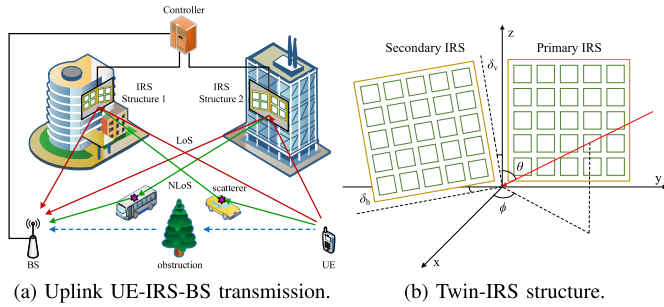


Fig. 1. An IRS-assisted MIMO-OFDM system. (a) One BS communicates with one UE via multiple IRSs with reflection coefficients configured by a controller. (b) A twin-IRS structure composed of two IRS planes with a relative spatial rotation.

with one user equipment (UE) with  $N_U$  antennas and  $M_U$  RF chains. The system occupies  $K$  subcarriers with a central carrier frequency  $f_c$  and a bandwidth  $f_s$ , where  $K_{tr}$  subcarriers are allocated for training. Multiple two-dimensional (2-D) IRSs with  $N_I$  elements are employed [5]–[7]. We propose to arrange every two IRS planes closely together, constructing  $R$  *twin-IRS* structures as shown in Fig. 1(b). Without loss of generality, we assume that the *primary IRS* is configured on the  $yz$ -plane (in the local coordinate system), while the *secondary IRS* is rotated with horizontal and vertical angles, i.e.,  $\delta_h$  and  $\delta_v$ , relative to the primary one. Note that  $\delta_h$ ,  $\delta_v$  can be viewed as the yaw and pitch angles of the rotation about the  $z$ - and  $y$ -axes, respectively. Other definitions of rotation, e.g., Euler angles or quaternion [28], can also be applied.

We consider an uplink communication scenario, where the direct BS-UE channel is assumed to be obstructed.<sup>1</sup> The channel estimation procedure occupies  $P_{tr}$  training frames, each containing  $T_{tr}$  training time slots. The training signal within the  $t$ th time slot of the  $p$ th frame at the  $k$ th subcarrier reflected via the primary IRS of the  $r$ th twin-IRS structure can be represented as<sup>2</sup>

$$\mathbf{y}_{r,k,p,t} = \mathbf{W}_{r,k}^H \mathbf{H}_{BI,r,k} \text{Diag}(\boldsymbol{\psi}_{r,k,p}) \times \mathbf{H}_{IU,r,k} \mathbf{F}_{r,k,t} \mathbf{x}_{r,k,t} + \mathbf{n}_{r,k,p,t}, \quad (1)$$

where  $\mathbf{x}_{r,k,t} \in \mathbb{C}^{M_U}$  and  $\mathbf{n}_{r,k,p,t} \in \mathbb{C}^{M_B}$  denote the transmitted pilots and additive noise, respectively;  $\mathbf{F}_{r,k,t} \in \mathbb{C}^{N_U \times M_U}$  denotes the precoding beamformer within the  $t$ th time slot of each frame;  $\mathbf{W}_{r,k} \in \mathbb{C}^{N_B \times M_{tr}}$  denotes the combining beamformer with  $M_{tr} \leq M_B$  data streams processed in parallel;  $\mathbf{H}_{BI,r,k} \in \mathbb{C}^{N_B \times N_I}$  and  $\mathbf{H}_{IU,r,k} \in \mathbb{C}^{N_I \times N_U}$  denote the one-hop IRS-BS and UE-IRS channels, respectively;

<sup>1</sup>Actually, the BS-UE channel is expected to remain static with a large coherence time. It can be efficiently removed by considering two received signals with different IRS training coefficients and mutually subtracting them. The IRS training pattern designs presented in this work are still applicable.

<sup>2</sup>The adjacent primary/secondary planes of each twin-IRS structure are expected to share identical propagation conditions, which can be equivalently regarded as a  $2N_I$ -element 3-D IRS. The proposed training designs can be directly applied to this 3-D topology. Moreover, multiple twin-IRS structures are distributed far apart from each other, whose propagation channels tend to be weakly correlated. By leveraging the spatial orthogonality with directional beamforming, or applying interference elimination methods with orthogonal reflection coefficients [29], one can identify the signals via different twin-IRS structures. For simplicity of notation and illustration, we here only represent the signal via the primary IRS of one twin-IRS structure.

$\boldsymbol{\psi}_{r,k,p} \in \mathbb{C}^{N_I}$  denotes the dynamic reflection coefficient vector during the  $p$ th training frame, which is sensitive to signal frequencies [30].

We consider IRS-assisted systems working at mmWave frequencies, where the uplink spatially sparse channels  $\mathbf{H}_{BI,r,k}$  and  $\mathbf{H}_{IU,r,k}$  via the  $r$ th primary IRS can be expressed as

$$\begin{aligned} \mathbf{H}_{BI,r,k} &= \sum_{\ell=1}^{L_{BI,r}} \underbrace{\alpha_{BI,r,\ell} e^{-j\frac{2\pi k f_s}{K} \tau_{BI,r,\ell}} F(\phi_{D,r,\ell}, \theta_{D,r,\ell})}_{\beta_{BI,r,k,\ell}} \\ &\quad \times \mathbf{a}_B(\phi_{B,r,\ell}, \theta_{B,r,\ell}) \mathbf{a}_I^H(\phi_{D,r,\ell}, \theta_{D,r,\ell}) \\ &= \mathbf{A}_{B,r} \text{Diag}(\boldsymbol{\beta}_{BI,r,k}) \mathbf{A}_{D,r}^H, \end{aligned} \quad (2a)$$

$$\begin{aligned} \mathbf{H}_{IU,r,k} &= \sum_{\ell=1}^{L_{IU,r}} \underbrace{\alpha_{IU,r,\ell} e^{-j\frac{2\pi k f_s}{K} \tau_{IU,r,\ell}} F(\phi_{A,r,\ell}, \theta_{A,r,\ell})}_{\beta_{IU,r,k,\ell}} \\ &\quad \times \mathbf{a}_I(\phi_{A,r,\ell}, \theta_{A,r,\ell}) \mathbf{a}_U^H(\phi_{U,r,\ell}, \theta_{U,r,\ell}) \\ &= \mathbf{A}_{A,r} \text{Diag}(\boldsymbol{\beta}_{IU,r,k}) \mathbf{A}_{U,r}^H, \end{aligned} \quad (2b)$$

where  $L_{BI,r}$ ,  $L_{IU,r}$  denote the numbers of propagation paths;  $\alpha_{BI,r,\ell}$ ,  $\alpha_{IU,r,\ell}$  denote the random complex gains;  $\tau_{BI,r,\ell}$ ,  $\tau_{IU,r,\ell}$  denote the time delays;  $\phi_{B(U),r,\ell}$ ,  $\theta_{B(U),r,\ell}$  denote the horizontal and vertical AoAs at the BS (AoDs at the UE), respectively;  $\phi_{D(A),r,\ell}$ ,  $\theta_{D(A),r,\ell}$  denote the horizontal and vertical AoDs (AoAs) at the primary IRS of the  $r$ th twin-IRS structure, respectively;  $F(\phi, \theta)$  denotes the angle-dependent IRS power radiation pattern defined as [31]

$$F(\phi, \theta) \triangleq (\sin \theta \cos \phi)^{\frac{q}{2}}, \quad (3)$$

where  $q \geq 0$  is the power radiation coefficient;  $q = 0$  corresponds to omnidirectional IRS elements;  $\mathbf{A}_{B,r} \in \mathbb{C}^{N_B \times L_{BI,r}}$ ,  $\mathbf{A}_{U,r} \in \mathbb{C}^{N_U \times L_{IU,r}}$  concatenate the antenna steering vectors formed by  $\{\phi_{B,r,\ell}, \theta_{B,r,\ell}\}$ ,  $\{\phi_{U,r,\ell}, \theta_{U,r,\ell}\}$ , respectively;  $\mathbf{A}_{D,r} \in \mathbb{C}^{N_I \times L_{BI,r}}$ ,  $\mathbf{A}_{A,r} \in \mathbb{C}^{N_I \times L_{IU,r}}$  concatenate the IRS response vectors formed by  $\{\phi_{D,r,\ell}, \theta_{D,r,\ell}\}$ ,  $\{\phi_{A,r,\ell}, \theta_{A,r,\ell}\}$ , respectively;  $\boldsymbol{\beta}_{BI,r,k} \in \mathbb{C}^{L_{BI,r}}$ ,  $\boldsymbol{\beta}_{IU,r,k} \in \mathbb{C}^{L_{IU,r}}$  concatenate the equivalent path gains  $\{\beta_{BI,r,k,\ell}\}$ ,  $\{\beta_{IU,r,k,\ell}\}$ , respectively. Note that  $\mathbf{H}_{BI(IU),r,k}$  may contain both the LoS and NLoS components or only one kind of them, where the Rician  $K$ -factor is denoted by  $K_{BI(IU),r}$ . In this paper, we consider only the first-order NLoS components scattered/reflected once by the environment objects, as illustrated in Fig. 1(a).<sup>3</sup>

The reflectors of an IRS plane form an  $(N_h \times N_v)$ -uniform planar array (UPA) topology with  $N_I = N_h N_v$ . The array response vector of the primary IRS can be represented in a form of a Kronecker product  $\mathbf{a}_I(\phi, \theta) \triangleq \mathbf{a}_h(\phi, \theta) \otimes \mathbf{a}_v(\phi, \theta)$ , where

$$\mathbf{a}_h(\phi, \theta) \triangleq \left[ 1, e^{j\frac{2\pi d_h}{\lambda_c} \sin \theta \sin \phi}, \dots, e^{j\frac{2\pi d_h}{\lambda_c} (N_h-1) \sin \theta \sin \phi} \right]^T, \quad (4a)$$

$$\mathbf{a}_v(\phi, \theta) \triangleq \left[ 1, e^{j\frac{2\pi d_v}{\lambda_c} \cos \theta}, \dots, e^{j\frac{2\pi d_v}{\lambda_c} (N_v-1) \cos \theta} \right]^T, \quad (4b)$$

<sup>3</sup>It has been shown in [32], [33] that the higher-order rays with more than two bounces suffer severe path loss and attenuation; as such, their contribution to the total received energy is about 2%–10% and can be ignored. Moreover, the second-order rays could have a power degradation of 10–20 dB with respect to the first-order ones, which can be identified and filtered out along with the diffuse scattering components, to enhance the accuracy of environment mapping and positioning (see for instance [34]).

with  $d_{h(v)}$  and  $\lambda_c$  denoting the azimuth (elevation) inter-element spacing of IRS arrays and central carrier wavelength, respectively. The BS/UE antennas can form either a UPA or a uniform linear array (ULA) topology, whose array steering vectors can be similarly defined as in (4).

Recall now that the secondary IRS is placed close to the primary one with rotation angles  $\delta_h$ ,  $\delta_v$ . In the far-field scenario, the radiation pattern of the secondary IRS can be expressed as

$$F(\phi, \theta; \delta_h, \delta_v) \triangleq (\sin \theta \cos \delta_v \cos(\phi - \delta_h) - \cos \theta \sin \delta_v)^{\frac{q}{2}}. \quad (5)$$

Similarly, the array response elements of the secondary IRS can be expressed as

$$[\mathbf{a}_h(\phi, \theta; \delta_h, \delta_v)]_{n_h} = e^{j \frac{2\pi d_h}{\lambda_c} (n_h - 1) \sin \theta \sin(\phi - \delta_h)}, \quad (6a)$$

$$[\mathbf{a}_v(\phi, \theta; \delta_h, \delta_v)]_{n_v} = e^{j \frac{2\pi d_v}{\lambda_c} (n_v - 1) \left( \sin \theta \sin \delta_v \cos(\phi - \delta_h) + \cos \theta \cos \delta_v \right)}, \quad (6b)$$

where  $n_{h(v)} \in \mathcal{I}(N_{h(v)})$ . Now, the uplink channel matrices via the  $r$ th secondary IRS can be represented by (2) with the radiation patterns and steering vectors of (3), (4) being replaced by those of (5), (6), respectively. One can also generate (3), (4) as  $F(\phi, \theta; 0, 0)$  and  $\mathbf{a}_1(\phi, \theta; 0, 0) = \mathbf{a}_h(\phi, \theta; 0, 0) \otimes \mathbf{a}_v(\phi, \theta; 0, 0)$ , respectively. Note that (5), (6) can be derived by multiplying  $[\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^T$  with the yaw and pitch rotation matrices defined by  $\delta_h$  and  $\delta_v$ , respectively [28].

### III. TENSOR-BASED CHANNEL ESTIMATION

In this section, we develop channel estimation schemes to recover the channel matrices, i.e.,  $\mathbf{H}_{\text{BI},r,k}$ ,  $\mathbf{H}_{\text{IU},r,k}$ . The IRS plane adopts dynamic reflection coefficients  $\{\psi_{r,k,p}\}_{p=1}^{P_{\text{tr}}}$  across the training frames. We concatenate the received training signals of (1) across  $P_{\text{tr}} T_{\text{tr}}$  training time slots in total as  $\mathbf{Y}_{r,k} \triangleq [\mathbf{y}_{r,k,1,1}, \mathbf{y}_{r,k,1,2}, \dots, \mathbf{y}_{r,k,P_{\text{tr}},T_{\text{tr}}}] \in \mathbb{C}^{M_{\text{tr}} \times P_{\text{tr}} T_{\text{tr}}}$ . It can be equivalently regarded as a matricization of a third-order tensor as  $\mathbf{Y}_{r,k} = \text{Matr}(\mathcal{Y}_{r,k}; 1, [3, 2])$ , where  $\mathcal{Y}_{r,k} \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}} \times P_{\text{tr}}}$  fits the CP tensor model as [21], [22]

$$\begin{aligned} \mathbf{Y}_{r,k} &= \sum_{n=1}^{N_{\text{I}}} \mathbf{W}_{r,k}^H [\mathbf{H}_{\text{BI},r,k}]_{:,n} \circ \mathbf{F}_{r,k}^T [\mathbf{H}_{\text{IU},r,k}]_{n,:}^T \circ [\Psi_{r,k}]_{:,n} \\ &\quad + \mathcal{N}_{r,k} \\ &= \mathcal{I}_{3,N_{\text{I}}} \times_1 \tilde{\mathbf{H}}_{\text{BI},r,k} \times_2 \tilde{\mathbf{H}}_{\text{IU},r,k} \times_3 \Psi_{r,k} + \mathcal{N}_{r,k}, \quad (7) \end{aligned}$$

where  $\Psi_{r,k} \triangleq [\psi_{r,k,1}, \dots, \psi_{r,k,P_{\text{tr}}}]^T \in \mathbb{C}^{P_{\text{tr}} \times N_{\text{I}}}$ ;  $\mathbf{F}_{r,k} \triangleq [\mathbf{F}_{r,k,1} \mathbf{x}_{r,k,1}, \dots, \mathbf{F}_{r,k,T_{\text{tr}}} \mathbf{x}_{r,k,T_{\text{tr}}}] \in \mathbb{C}^{N_{\text{U}} \times T_{\text{tr}}}$ ;  $\tilde{\mathbf{H}}_{\text{BI},r,k} \triangleq \mathbf{W}_{r,k}^H \mathbf{H}_{\text{BI},r,k} \in \mathbb{C}^{M_{\text{tr}} \times N_{\text{I}}}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k} \triangleq \mathbf{F}_{r,k}^T \mathbf{H}_{\text{IU},r,k} \in \mathbb{C}^{T_{\text{tr}} \times N_{\text{I}}}$  denote the combined and precoded equivalent channels, respectively;  $\mathcal{N}_{r,k} \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}} \times P_{\text{tr}}}$  is the equivalent noise tensor;  $\mathcal{I}_{3,N_{\text{I}}} \in \{0, 1\}^{N_{\text{I}} \times N_{\text{I}} \times N_{\text{I}}}$  denotes a third-order identity tensor. Performing the channel estimation is now equivalent to solving the factorization problem of (7) with a maximal tensor rank  $\min(\text{r}(\tilde{\mathbf{H}}_{\text{BI},r,k}), \text{r}(\tilde{\mathbf{H}}_{\text{IU},r,k}), \text{r}(\Psi_{r,k}), N_{\text{I}})$ . Since the wideband training signal  $\{\mathcal{Y}_{r,k}\}_{k=1}^{K_{\text{tr}}}$  cannot be modeled as a fourth-order CP tensor with frequency-sensitive

$\{\Psi_{r,k}\}_{k=1}^{K_{\text{tr}}}$ , we seek to subcarrier-wisely perform the factorization of  $\mathcal{Y}_{r,k}$ . One can average the estimation results of frequency-flat channel parameters along multiple subcarriers to enhance the recovery accuracy.

Obviously, the training IRS coefficients significantly affect the estimation strategy and performance. First of all, we consider a training pattern  $\Psi_{r,k}$  with full column rank, which requires the number of measurements to exceed the number of signal components, i.e.,  $P_{\text{tr}} \geq N_{\text{I}}$ . A cascaded beamformed channel can be naturally derived by a least squares (LS) solution as follows:

$$\begin{aligned} \tilde{\mathbf{H}}_{\text{cas},r,k} &\triangleq \tilde{\mathbf{H}}_{\text{BI},r,k} \odot \tilde{\mathbf{H}}_{\text{IU},r,k} \\ &= \text{Matr}(\mathcal{Y}_{r,k}; [2, 1], 3) (\Psi_{r,k}^T)^{\dagger}. \quad (8) \end{aligned}$$

The  $n$ th column of (8) can be reshaped into a rank-1 product, i.e.,  $\tilde{\mathbf{h}}_{\text{BI},r,k,n} \tilde{\mathbf{h}}_{\text{IU},r,k,n}^T \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}}}$ , where  $\tilde{\mathbf{h}}_{\text{BI},r,k,n} \triangleq [\tilde{\mathbf{H}}_{\text{BI},r,k}]_{:,n}$ ,  $\tilde{\mathbf{h}}_{\text{IU},r,k,n} \triangleq [\tilde{\mathbf{H}}_{\text{IU},r,k}]_{:,n}$ . By performing rank-1 truncated singular value decomposition (tSVD)  $\lambda_1 \mathbf{u}_1 \mathbf{v}_1^H = \tilde{\mathbf{h}}_{\text{BI},r,k,n} \tilde{\mathbf{h}}_{\text{IU},r,k,n}^T$  with the dominant singular value  $\lambda_1$ , we can take  $\sqrt{\lambda_1} \mathbf{u}_1 \in \mathbb{C}^{M_{\text{tr}}}$  and  $\sqrt{\lambda_1} \mathbf{v}_1^* \in \mathbb{C}^{T_{\text{tr}}}$  as the estimates of  $\tilde{\mathbf{h}}_{\text{BI},r,k,n}$  and  $\tilde{\mathbf{h}}_{\text{IU},r,k,n}$ , respectively.

For the case of  $P_{\text{tr}} < N_{\text{I}}$  or  $\Psi_{r,k}^{\dagger} \Psi_{r,k} \neq \mathbf{I}_{N_{\text{I}}}$ , however, the LS solution in (8) is invalid, which undermines the integration of large-scale IRSs with a massive number of unit cells. To address this critical problem, we present four different IRS reflection pattern designs, as well as, the corresponding uniqueness condition analysis.

#### A. Random Pattern

One of the simplest designs is to assign random reflection coefficients to the IRS elements [30]. For instance, we employ a training pattern with unit moduli and random phases, i.e.,  $|\Psi_{r,k}[p,n]| = 1, \forall p, n$ . A uniqueness condition of factorizing (7) is presented as follows:

*Lemma 1* ([35, Theorem 1], [36, Lemma 4.3]): Consider a third-order CP tensor  $\mathcal{Y}_{r,k} \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}} \times P_{\text{tr}}}$  with  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$ ,  $\Psi_{r,k}$  being the factor matrices in (7). If

$$k(\tilde{\mathbf{H}}_{\text{BI},r,k}) + k(\tilde{\mathbf{H}}_{\text{IU},r,k}) + k(\Psi_{r,k}) \geq 2 N_{\text{I}} + 2, \quad (9)$$

then,  $\text{r}(\mathcal{Y}_{r,k}) = N_{\text{I}}$ , and the factorization of  $\mathcal{Y}_{r,k}$  is unique up to column scaling and permutation. Furthermore, if  $\Psi_{r,k}$  is known with  $k(\Psi_{r,k}) = P_{\text{tr}} \geq 2$ , and

$$\min \left( k(\tilde{\mathbf{H}}_{\text{BI},r,k}), k(\tilde{\mathbf{H}}_{\text{IU},r,k}) \right) + k(\Psi_{r,k}) \geq N_{\text{I}} + 2, \quad (10)$$

then,  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$  can be found algebraically up to column scaling ambiguities.

If  $\mathbf{F}_{r,k}$ ,  $\mathbf{W}_{r,k}$  have full column ranks with  $M_{\text{tr}} \geq L_{\text{BI},r}$ ,  $T_{\text{tr}} \geq L_{\text{IU},r}$ , the beamformed channels inherit the sparse nature, i.e.,  $\text{r}(\tilde{\mathbf{H}}_{\text{BI},r,k}) \leq \text{r}(\mathbf{H}_{\text{BI},r,k}) \leq L_{\text{BI},r}$ ,  $\text{r}(\tilde{\mathbf{H}}_{\text{IU},r,k}) \leq \text{r}(\mathbf{H}_{\text{IU},r,k}) \leq L_{\text{IU},r}$ . According to the property of  $k$ -rank:  $k(\mathbf{A}) \leq \text{r}(\mathbf{A}) \leq \min(I, J), \forall \mathbf{A} \in \mathbb{C}^{I \times J}$  [37], in the generic case, (10) can be relaxed to  $\min(L_{\text{BI},r}, L_{\text{IU},r}) + P_{\text{tr}} \geq N_{\text{I}} + 2$ . Hence, to estimate low-rank mmWave channels with usually 3–5 paths [38], the quantity of training frames should be of

the same order of magnitude as the number of IRS elements. According to [36, Lemma 4.3], the factorization of (7) can be algebraically implemented by generalized eigenvalue decomposition. In practice, when  $P_{\text{tr}} \min(M_{\text{tr}}, T_{\text{tr}}) \geq N_{\text{I}}$ , we can adopt the bilinear alternating least squares (BALS) method [14] to iteratively update  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$ .

### B. Structured Pattern

If additional structural information is embedded into the IRS coefficients, one can develop a more flexible solution with a more relaxed uniqueness condition. For instance, we can design the training coefficients as  $[\Psi_{r,k}]_{p,n} = \lambda_n^{p-1}$  with distinct  $\{\lambda_n\}_{n=1}^{N_{\text{I}}}$ , which has a column-wise phase *rotational-invariant* characteristic. We note that this codebook may be valid only for a particular subcarrier due to the frequency-sensitive characteristics of IRS reflectors [30]. As such, this strategy is more suitable for single-carrier training. In practice, the IRS is usually implemented by finite-resolution phase shifters to reduce the hardware complexity. Hence, one can employ a discrete Fourier transform (DFT) codebook, i.e.,  $\lambda_n \triangleq e^{-j2\pi(n-1)/N_{\text{I}}}$ . In this way, the training codewords multiplex  $N_{\text{I}}$  distinct phases uniformly spaced within  $[0, 2\pi]$ , which can be realized by  $b = \lceil \log_2 N_{\text{I}} \rceil$ -bit phase shifters. A corresponding uniqueness condition is presented as follows:

*Lemma 2 ([39, Theorem III.3]):* Consider a third-order CP tensor  $\mathcal{Y}_{r,k} \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}} \times P_{\text{tr}}}$  with  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$ ,  $\Psi_{r,k}$  being the factor matrices in (7). Define  $\{P_1, P_2\} : P_1 + P_2 = P_{\text{tr}} + 1$ . If  $[\Psi_{r,k}]_{p,n} = \lambda_n^{p-1}$ , and

$$r\left(\Psi_{r,k}^{[P_1-1]} \odot \tilde{\mathbf{H}}_{\text{BI},r,k}\right) = r\left(\Psi_{r,k}^{[P_2]} \odot \tilde{\mathbf{H}}_{\text{IU},r,k}\right) = N_{\text{I}}, \quad (11)$$

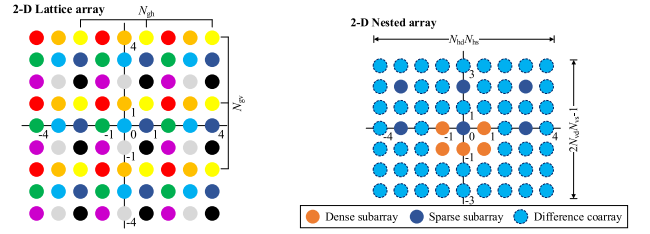
then,  $r(\mathcal{Y}_{r,k}) = N_{\text{I}}$ , and  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$  can be found algebraically up to column scaling ambiguities.

According to [35, Lemma 1],  $\mathbf{A} \odot \mathbf{B}$  has full column rank if  $k(\mathbf{A}) + k(\mathbf{B}) \geq D + 1, \forall \mathbf{A} \in \mathbb{C}^{I \times D}, \mathbf{B} \in \mathbb{C}^{J \times D}$ , which implies  $I + J \geq D + 1 \Rightarrow IJ \geq D$  with  $k(\mathbf{A}) \leq I, k(\mathbf{B}) \leq J$ . In the generic case, (11) can be relaxed to  $\min((P_1 - 1)L_{\text{BI},r}, P_2 L_{\text{IU},r}) \geq N_{\text{I}}$ . Compared with the random pattern case, the necessary number of training frames for estimating mmWave channels with 3–5 paths is now roughly halved to about  $N_{\text{I}}/2$ .

The decomposition of (7) can now be algebraically solved by the structured CP decomposition (SCPD) method [40], where the original eigenvalue decomposition for the generator recovery is replaced by the eigenvector derivation with *a priori* eigenvalues. One can adopt a set of valid candidates  $\{P_1, P_2\}$  and average the obtained results to enhance the estimation performance.

### C. Grouping Pattern

We can cluster the IRS reflectors into  $G \ll N_{\text{I}}$  groups, where the elements of the  $g$ th group are indexed by  $\mathcal{N}_g : |\mathcal{N}_g| = N_g$ . The reflectors belonging to the same group adopt identical training coefficients, i.e.,  $[\Psi_{r,k}]_{:,n} \triangleq \psi_{r,k}^g, \forall n \in \mathcal{N}_g$ .



(a) Grouping pattern: 2-D lattice array. (b) Sparse pattern: 2-D nested array.

Fig. 2. IRS training reflection patterns. (a) Grouping pattern with a 2-D lattice array composed of  $G (N_{\text{gh}} \times N_{\text{gv}})$ -subarrays, where identical training reflection coefficients are allocated to the elements in the same color. (b) Sparse pattern with a 2-D nested array composed of an  $(N_{\text{hd}} \times N_{\text{vd}})$ -dense subarray and an  $(N_{\text{hs}} \times N_{\text{vs}})$ -sparse subarray.

The training signal (7) can be modified as

$$\begin{aligned} \mathcal{Y}_{r,k} &= \sum_{g=1}^G \sum_{n \in \mathcal{N}_g} [\tilde{\mathbf{H}}_{\text{BI},r,k}]_{:,n} \circ [\tilde{\mathbf{H}}_{\text{IU},r,k}]_{:,n} \circ \psi_{r,k}^g + \mathcal{N}_{r,k} \\ &= \mathcal{C}_r \times_1 \tilde{\mathbf{H}}_{\text{BI},r,k} \times_2 \tilde{\mathbf{H}}_{\text{IU},r,k} \times_3 \bar{\Psi}_{r,k} + \mathcal{N}_{r,k}, \quad (12) \end{aligned}$$

where  $\bar{\Psi}_{r,k} \triangleq [\psi_{r,k}^1, \dots, \psi_{r,k}^G] \in \mathbb{C}^{P_{\text{tr}} \times G}$  concatenates the grouping training coefficients;  $\mathcal{C}_r \triangleq \mathcal{I}_{3, N_{\text{I}}} \times_3 \Gamma_r \in \{0, 1\}^{N_{\text{I}} \times N_{\text{I}} \times G}$  is a core tensor constrained by the grouping indicator  $\Gamma_r \in \{0, 1\}^{G \times N_{\text{I}}}$  with  $[\Gamma_r]_{g,n} = 1, \forall n \in \mathcal{N}_g$ . Since  $\Psi_{r,k} = \bar{\Psi}_{r,k} \Gamma_r$  is rank-deficient, (12) fits a *parallel profiles with linear dependencies (PARALIND)* model [41, (41)–(43)].

We propose to divide the IRS reflector array into latticed  $(N_{\text{gh}} \times N_{\text{gv}})$ -subarrays with  $G_{\text{h}} \triangleq N_{\text{h}}/N_{\text{gh}}, G_{\text{v}} \triangleq N_{\text{v}}/N_{\text{gv}}$  and  $G = G_{\text{h}} G_{\text{v}}$ , as illustrated in Fig. 2(a). The horizontal and vertical indices of the IRS entries belonging to the  $g$ th group are defined as

$$\begin{aligned} \mathcal{N}_{\text{h}}^{[g]} &= \{G_{\text{h}}(n_{\text{gh}} - \lceil g/G_{\text{h}} \rceil) + g\}_{n_{\text{gh}}=1}^{N_{\text{gh}}}, \\ \mathcal{N}_{\text{v}}^{[g]} &= \{G_{\text{v}}(n_{\text{gv}} - 1) + \lceil g/G_{\text{h}} \rceil\}_{n_{\text{gv}}=1}^{N_{\text{gv}}}, \quad (13a) \end{aligned}$$

$$\begin{aligned} \mathcal{N}_g &= \{N_{\text{v}}(n_{\text{h}}^{[g]} - 1) + n_{\text{v}}^{[g]}\}, \\ \forall n_{\text{h}}^{[g]} \in \mathcal{N}_{\text{h}}^{[g]}, \quad n_{\text{v}}^{[g]} \in \mathcal{N}_{\text{v}}^{[g]}. \quad (13b) \end{aligned}$$

We call a factor matrix *partially unique* if its columns can be partitioned into disjoint subsets and each subset is identified up to its linear span [42, Definition 3.2]. Then, a partial uniqueness condition is presented as follows:

*Lemma 3:* Consider a third-order tensor  $\mathcal{Y}_{r,k} \in \mathbb{C}^{M_{\text{tr}} \times T_{\text{tr}} \times P_{\text{tr}}}$  with  $\tilde{\mathbf{H}}_{\text{BI},r,k}$ ,  $\tilde{\mathbf{H}}_{\text{IU},r,k}$ ,  $\Psi_{r,k}$  being the factor matrices in (7). If  $\Psi_{r,k} = \bar{\Psi}_{r,k} \Gamma_r$  with  $r(\bar{\Psi}_{r,k}^\dagger) = G$  and  $[\Gamma_r]_{g,n} = 1, \forall n \in \mathcal{N}_g$ , then,  $\tilde{\mathbf{H}}_{\text{BI},r,k}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}$  are partially unique.

*Proof:* By removing the grouping codebook  $\bar{\Psi}_{r,k}$  from (12) via the LS method, one can obtain  $G$  sets of rank- $N_g$  products as

$$\mathcal{C}_r \times_1 \tilde{\mathbf{H}}_{\text{BI},r,k} \times_2 \tilde{\mathbf{H}}_{\text{IU},r,k} = \mathcal{Y}_{r,k} \times_3 \bar{\Psi}_{r,k}^\dagger, \quad (14a)$$

$$\left[ \mathcal{C}_r \times_1 \tilde{\mathbf{H}}_{\text{BI},r,k} \times_2 \tilde{\mathbf{H}}_{\text{IU},r,k} \right]_{:, :, g} = \tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_g)} \left( \tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_g)} \right)^T, \quad (14b)$$

where  $\tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_g)} \in \mathbb{C}^{M_{\text{tr}} \times N_g}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_g)} \in \mathbb{C}^{T_{\text{tr}} \times N_g}$  contain the columns of  $\mathbf{H}_{\text{BI},r,k}$  and  $\mathbf{H}_{\text{IU},r,k}$  allocated to the  $g$ th group, respectively. By performing rank- $N_g$  tSVD on the  $g$ th slice of (14),  $\tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_g)}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_g)}$  are factorized uniquely up to a nonsingular transformation. ■

The partial uniqueness condition is generally satisfied with  $P_{\text{tr}} \geq G$  training frames, which is generally independent of the channel ranks. The grouping pattern (13) does not directly lead to a complete estimate of  $\mathbf{H}_{\text{BI(IU)},r,k}$ . By leveraging the phase shift feature of grouped subarrays, one can formulate an equivalent  $(G_{\text{h}} \times G_{\text{v}})$ -array. As long as  $\min(G_{\text{h}}, G_{\text{v}}) \geq 2$ , we are able to acquire the cascaded channel parameters, which in turn support a unique channel separation. The detailed derivations will be provided in the next section.

#### D. Sparse Pattern

It can be observed from the earlier designs that the required amount of training frames is generally proportional to the number of reflectors. By leveraging the concept of sparse arrays [43], we are able to obtain the parameterized channels with fewer IRS elements, leading to reduced training overhead. Specifically,  $N_{\text{A}} \ll N_{\text{I}}$  reflectors are installed regularly in the  $(N_{\text{h}} \times N_{\text{v}})$ -array aperture, leaving  $(N_{\text{I}} - N_{\text{A}})$  virtual unit cells with no circuits. We denote the index set of selected unit cells by  $\mathcal{N}_{\text{A}} : |\mathcal{N}_{\text{A}}| = N_{\text{A}}$ , and modify (7) as

$$\begin{aligned} \mathcal{Y}_{r,k} &= \sum_{n \in \mathcal{N}_{\text{A}}} [\tilde{\mathbf{H}}_{\text{BI},r,k}]_{:,n} \circ [\tilde{\mathbf{H}}_{\text{IU},r,k}]_{:,n} \circ [\Psi_{r,k}]_{:,n} + \mathcal{N}_{r,k} \\ &= \mathcal{I}_{3,N_{\text{A}}} \times \mathbf{1} \tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_{\text{A}})} \times \mathbf{2} \tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_{\text{A}})} \times \mathbf{3} \Psi_{r,k}^{(N_{\text{A}})} + \mathcal{N}_{r,k}, \end{aligned} \quad (15)$$

where  $\Psi_{r,k}^{(N_{\text{A}})} \in \mathbb{C}^{P_{\text{tr}} \times N_{\text{A}}}$  contains columns of  $\Psi_{r,k}$  indexed by  $\mathcal{N}_{\text{A}}$ ;  $\tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_{\text{A}})} \in \mathbb{C}^{M_{\text{tr}} \times N_{\text{A}}}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_{\text{A}})} \in \mathbb{C}^{T_{\text{tr}} \times N_{\text{A}}}$  contain the columns indexed by  $\mathcal{N}_{\text{A}}$  of  $\mathbf{H}_{\text{BI},r,k}$  and  $\mathbf{H}_{\text{IU},r,k}$ , respectively.

One can leverage sparse array topologies for the training pattern. Since the *nested array* geometry admits a closed-form expression for the locations of the antenna elements and offers a consecutive difference coarray with large degrees of freedom, we consider a 2-D symmetric nested array [43], as illustrated in Fig. 2(b). The training pattern consists of an  $(N_{\text{hd}} \times N_{\text{vd}})$ -dense subarray and an  $(N_{\text{hs}} \times N_{\text{vs}})$ -sparse subarray with  $N_{\text{hd}}, N_{\text{hs}}$  being odd values [23]. The 2-D subarray coordinates are defined as<sup>4</sup>

$$\begin{aligned} \mathcal{C}_{\text{hd}} &= \{n_{\text{hd}} - (N_{\text{hd}} + 1)/2\}_{n_{\text{hd}}=1}^{N_{\text{hd}}}, \\ \mathcal{C}_{\text{vd}} &= \{n_{\text{vd}} - N_{\text{vd}}\}_{n_{\text{vd}}=1}^{N_{\text{vd}}}, \end{aligned} \quad (16a)$$

$$\begin{aligned} \mathcal{C}_{\text{hs}} &= \{N_{\text{hd}}(n_{\text{hs}} - (N_{\text{hs}} + 1)/2)\}_{n_{\text{hs}}=1}^{N_{\text{hs}}}, \\ \mathcal{C}_{\text{vs}} &= \{N_{\text{vd}}(n_{\text{vs}} - 1)\}_{n_{\text{vs}}=1}^{N_{\text{vs}}}, \end{aligned} \quad (16b)$$

where  $N_{\text{hd}}, N_{\text{vd}}, N_{\text{hs}}, N_{\text{vs}} \in \mathbb{Z}^+$  satisfy the geometric constraint, i.e.,  $N_{\text{hd}}(N_{\text{hs}} - 1) + 1 \leq N_{\text{h}}$ ,  $N_{\text{vd}}N_{\text{vs}} \leq N_{\text{v}}$ . The total number of activated IRS elements is  $N_{\text{A}} = N_{\text{hd}}N_{\text{vd}} +$

<sup>4</sup>The 2-D coordinates  $(n_{\text{hd}(\text{hs})}, n_{\text{vd}(\text{vs})})$  of the nested array defined by (14) are uniquely mapped to the 2-D coordinates  $(n_{\text{h}}, n_{\text{v}})$  of (5) as  $n_{\text{h}} = \lceil N_{\text{h}}/2 \rceil - n_{\text{hd}(\text{hs})}$ ,  $n_{\text{v}} = \lceil N_{\text{v}}/2 \rceil - n_{\text{vd}(\text{vs})}$ , which correspond to the indices  $(n_{\text{h}} - 1)N_{\text{v}} + n_{\text{v}} \in \mathcal{N}_{\text{A}}$ .

TABLE II  
RECOMMENDED TRAINING SETTINGS FOR IRS PATTERN DESIGNS

IRS Pattern Design	Recommended Training Settings
Random	$\min(L_{\text{BI},r}, L_{\text{IU},r}) + P_{\text{tr}} \geq N_{\text{I}} + 2.$
Structured	$P_{\text{tr}} = P_1 + P_2 - 1,$ $\min((P_1 - 1)L_{\text{BI},r}, P_2 L_{\text{IU},r}) \geq N_{\text{I}}.$
Grouping	$\min(G_{\text{h}}, G_{\text{v}}) \geq 2, P_{\text{tr}} \geq G = G_{\text{h}}G_{\text{v}}.$
Sparse	$N_{\text{hd}}(N_{\text{hs}} - 1) + 1 \leq N_{\text{h}}, N_{\text{vd}}N_{\text{vs}} \leq N_{\text{v}},$ $P_{\text{tr}} \geq N_{\text{A}} = N_{\text{hd}}N_{\text{vd}} + N_{\text{hs}}N_{\text{vs}} - 1.$

$N_{\text{hs}}N_{\text{vs}} - 1$ , and the difference coarray size is  $N_{\text{hd}}N_{\text{hs}} \times (2N_{\text{vd}}N_{\text{vs}} - 1)$ . If  $r(\Psi_{r,k}^{(N_{\text{A}})}) = N_{\text{A}}$ , one can follow the LS and tSVD solution as in (8) to factorize  $\tilde{\mathbf{H}}_{\text{BI},r,k}^{(N_{\text{A}})}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}^{(N_{\text{A}})}$  uniquely up to column scaling. The necessary number of training frames is hence  $P_{\text{tr}} \geq N_{\text{A}}$ , which is independent of the channel ranks. Generally, the sparsity of IRS arrays enhances the channel estimation accuracy, but may yield reduced beamforming gains or system capacity. Similar to the grouping pattern (13), the sparse pattern (16) can only return a portion of the subchannel columns. Fortunately, we are able to recover the cascaded channel parameters and reconstruct the channel matrices. The detailed derivation will be presented in the next section.

*Remark 1:* Note that all the proposed training patterns do not place any constraint on the specific values of reflection coefficients. They can be designed as directional beams to search across the cascaded angular parameters on the IRS plane, while the BS can select the received signals with the highest power to implement the channel estimation or dynamically adjust the searching directions of the subsequent frames through the controller [19]. Moreover, apart from the structured pattern, the others can be applied to multi-subcarriers without being interfered by the frequency-sensitive nature of reflectors. Based on the uniqueness conditions and other physical constraints, we summarize the recommended training settings for IRS pattern designs in Table II.

#### IV. MULTIPATH PARAMETER RECOVERY

We now recall that many wireless applications, e.g., environment mapping, user positioning, mobility tracking, etc., require precise information of the propagation paths, i.e., AoAs, AoDs, and time delays. Moreover, the channel estimates with missing or partially unique columns, derived from the sparse or grouping training pattern, can be well calibrated based on the propagation parameters. Specifically, by leveraging the characteristics of  $\mathbf{H}_{\text{BI(IU)},r,k}$  in (2), we can equivalently represent the cascaded beamformed channel  $\tilde{\mathbf{H}}_{\text{cas},r,k}$  in (8) as

$$\begin{aligned} \tilde{\mathbf{H}}_{\text{cas},r,k} &= (\mathbf{W}_{r,k}^H \mathbf{A}_{\text{B},r} \text{Diag}(\boldsymbol{\beta}_{\text{BI},r,k}) \\ &\quad \otimes \mathbf{F}_{r,k}^T \mathbf{A}_{\text{U},r}^* \text{Diag}(\boldsymbol{\beta}_{\text{IU},r,k})) (\mathbf{A}_{\text{D},r}^H \odot \mathbf{A}_{\text{A},r}^T) \\ &= (\mathbf{W}_{r,k}^H \otimes \mathbf{F}_{r,k}^T) \sum_{\ell_1=1}^{L_{\text{BI},r}} \sum_{\ell_2=1}^{L_{\text{IU},r}} \rho_{r,k,\ell_1\ell_2} \\ &\quad \times (\mathbf{a}_{\text{B}}(\phi_{\text{B},r,\ell_1}, \theta_{\text{B},r,\ell_1}) \otimes \mathbf{a}_{\text{U}}^*(\phi_{\text{U},r,\ell_2}, \theta_{\text{U},r,\ell_2})) \\ &\quad \times \mathbf{a}_{\text{I}}^T(\pi_{\text{h},r,\ell_1\ell_2}, \pi_{\text{v},r,\ell_1\ell_2}), \end{aligned} \quad (17)$$

where  $\tilde{\mathbf{H}}_{\text{cas},r,k} = \left( \mathbf{W}_{r,k}^H \otimes \mathbf{F}_{r,k}^T \right) \mathbf{H}_{\text{cas},r,k}$  with  $\mathbf{H}_{\text{cas},r,k} \triangleq \mathbf{H}_{\text{BI},r,k} \odot \mathbf{H}_{\text{IU},r,k}^T$ . With a little abuse of notation,  $\mathbf{a}_{\text{I}}(\pi_{\text{h}}, \pi_{\text{v}}) \triangleq \mathbf{a}_{\text{h}}(\pi_{\text{h}}) \otimes \mathbf{a}_{\text{v}}(\pi_{\text{v}})$  denotes the IRS response vector with  $[\mathbf{a}_{\text{h}}(\pi_{\text{h}})]_n = e^{j \frac{2\pi}{\lambda_c} d_{\text{h}}(n-1)\pi_{\text{h}}}$ . The cascaded channel parameters via the primary IRS are defined as

$$\begin{cases} \pi_{\text{h},r,\ell_1\ell_2} \triangleq -\sin\theta_{\text{D},r,\ell_1} \sin\phi_{\text{D},r,\ell_1} \\ \quad + \sin\theta_{\text{A},r,\ell_2} \sin\phi_{\text{A},r,\ell_2}, \\ \pi_{\text{v},r,\ell_1\ell_2} \triangleq -\cos\theta_{\text{D},r,\ell_1} + \cos\theta_{\text{A},r,\ell_2}, \\ \rho_{r,k,\ell_1\ell_2} \triangleq \beta_{\text{BI},r,k,\ell_1} \beta_{\text{IU},r,k,\ell_2}. \end{cases} \quad (18)$$

Based on (17), (18),  $\mathbf{H}_{\text{cas},r,k}$  can be reconstructed by the channel parameters  $\{\phi(\theta)_{\text{B},\ell_1}, \phi(\theta)_{\text{U},\ell_2}, \pi_{\text{h}}(\text{v}),r,\ell_1\ell_2, \rho_{r,k,\ell_1\ell_2}\}$ , and be further separated into  $\tilde{\mathbf{H}}_{\text{BI},r,k}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k}$  by the tSVD. Note that in the noiseless case, the separated estimates of beamformed channels can be represented as  $\tilde{\mathbf{H}}_{\text{BI},r,k} \mathbf{\Lambda}_{r,k}$  and  $\tilde{\mathbf{H}}_{\text{IU},r,k} \mathbf{\Lambda}_{r,k}^{-1}$  with a diagonal matrix  $\mathbf{\Lambda}_{r,k} \in \mathbb{C}^{N_{\text{I}} \times N_{\text{I}}}$  containing the scaling ambiguities. Clearly, an arbitrary  $\mathbf{\Lambda}_{r,k}$  does not change the values of  $\mathcal{Y}_{r,k}$  or  $\tilde{\mathbf{H}}_{\text{cas},r,k}$ , which cannot be exactly determined. Also, the row-subspace information of  $\tilde{\mathbf{H}}_{\text{BI}(\text{IU}),r,k}$  has been compromised by those uncertainties. Generally, without *a priori* assumptions or conditions, one can only determine the cascaded parameters in (18) but not the exact path parameters via a single IRS plane, i.e.,  $\{\phi_{\text{D}(\text{A}),r,\ell}, \theta_{\text{D}(\text{A}),r,\ell}, \tau_{\text{BI}(\text{IU}),r,\ell}\}$ . It can be physically explained by that the passive IRS cannot sense the environment, while the receiver can only process signals propagated through the entire two-hop channel. To address these issues, we propose the channel parameter recovery schemes, as well as, the cascaded parameter decoupling strategy in this section.

### A. Beamforming Design

We seek to adopt special beamforming to preserve the structural information of BS/UE antenna arrays. By leveraging the concept of beamspace ESPRIT method [44], [45], we adopt a Kronecker product-formed DFT combining strategy for the UPA-sized BS antennas as follows

$$\mathbf{W}_{r,k} \triangleq \mathbf{W}_{\text{h},r,k} \otimes \mathbf{W}_{\text{v},r,k}, \quad (19a)$$

$$[\mathbf{W}_{\text{h}(\text{v}),r,k}]_{n,m} = N_{\text{B},\text{h}(\text{v})}^{-\frac{1}{2}} e^{-j2\pi(n-1)\frac{\omega_{\text{h}(\text{v}),m}}{N_{\text{B},\text{h}(\text{v})}}}, \quad (19b)$$

where  $\mathbf{W}_{\text{h}(\text{v}),r,k} \in \mathbb{C}^{N_{\text{B},\text{h}(\text{v})} \times M_{\text{h}(\text{v})}}$  denotes the combiner along the horizontal (vertical) direction with  $N_{\text{B}} = N_{\text{B},\text{h}} N_{\text{B},\text{v}}$ ,  $M_{\text{tr}} = M_{\text{h}} M_{\text{v}}$ ;  $\{\omega_{\text{h}(\text{v}),m}\}$  are beamspace codewords selected from the candidate codebook  $\mathcal{I}(N_{\text{B},\text{h}(\text{v})})$ . It can be verified that the DFT beamforming preserves the rotational-invariant feature of IRS array phases since

$$\bar{\mathbf{J}}_{\text{h1}(\text{v1})} \mathbf{W}_{\text{h}(\text{v}),r,k} = \bar{\mathbf{J}}_{\text{h2}(\text{v2})} \mathbf{W}_{\text{h}(\text{v}),r,k} \mathbf{T}_{\text{h}(\text{v})}, \quad (20a)$$

$$\begin{aligned} \mathbf{Q}_{\text{h}(\text{v})} [\mathbf{W}_{\text{h}(\text{v}),r,k}]_{N_{\text{B},\text{h}(\text{v})}}^H &= \mathbf{Q}_{\text{h}(\text{v})} \mathbf{T}_{\text{h}(\text{v})} [\mathbf{W}_{\text{h}(\text{v}),r,k}]_{1,1}^H \\ &= \mathbf{0}_{N_{\text{B},\text{h}(\text{v})} \times 1}, \end{aligned} \quad (20b)$$

$$\begin{aligned} \mathbf{Q}_{\text{T}} \mathbf{W}_{r,k}^H \mathbf{a}_{\text{B}}(\phi, \theta) &= \mathbf{Q} \mathbf{W}_{r,k}^H \mathbf{a}_{\text{B}}(\phi, \theta) \\ &\times e^{j \frac{2\pi}{\lambda_c} (d_{\text{B},\text{h}} \sin \phi + d_{\text{B},\text{v}} \cos \theta)}, \end{aligned} \quad (20c)$$

where  $\bar{\mathbf{J}}_{\text{h1}(\text{v1})}, \bar{\mathbf{J}}_{\text{h2}(\text{v2})} \in \{0, 1\}^{(N_{\text{B},\text{h}(\text{v})}-1) \times N_{\text{B},\text{h}(\text{v})}}$  select the first and last  $(N_{\text{B},\text{h}(\text{v})} - 1)$  rows of matrices, respectively;

$\mathbf{T}_{\text{h}(\text{v})} \in \mathbb{C}^{M_{\text{h}(\text{v})} \times M_{\text{h}(\text{v})}}$  is calculated as the LS solution of (20a);  $\mathbf{Q}_{\text{h}(\text{v})} \in \mathbb{C}^{N_{\text{B},\text{h}(\text{v})} \times M_{\text{h}(\text{v})}}$  is derived by forming a projection corresponding to the orthogonal subspace of  $\{[\mathbf{W}_{\text{h}(\text{v}),r,k}]_{N_{\text{B},\text{h}(\text{v})}}^H, \dots, \mathbf{T}_{\text{h}(\text{v})} [\mathbf{W}_{\text{h}(\text{v}),r,k}]_{1,1}^H\}$ ;  $\mathbf{Q} \triangleq \mathbf{Q}_{\text{h}} \otimes \mathbf{Q}_{\text{v}}$ ,  $\mathbf{Q}_{\text{T}} \triangleq \mathbf{Q} (\mathbf{T}_{\text{h}} \otimes \mathbf{T}_{\text{v}})^H$ . The precoded pilot design at the UE side can be similarly derived as in (20). As long as the search region formed by the beamspace codewords covers the channel path direction formed by  $\{\phi_{\text{B}(\text{U}),\ell}, \theta_{\text{B}(\text{U}),\ell}\}$ , we are able to receive the training signal with a desired power level [44]. Furthermore, before transmitting  $\mathcal{Y}_{r,k}$  in (7), one can perform a coarse sensing to find appropriate  $\{\omega_{\text{h}(\text{v}),m}\}$  as in [45].

*Remark 2:* The beamforming method enables us to recover  $\{\phi(\theta)_{\text{B},\ell}\}$  and  $\{\phi(\theta)_{\text{U},\ell}\}$  directly by exploiting the column-subspaces of  $\mathbf{Y}_{r,k}^{(1)} \triangleq \text{Matr}(\mathcal{Y}_{r,k}; 1, [3, 2])$  and  $\mathbf{Y}_{r,k}^{(2)} \triangleq \text{Matr}(\mathcal{Y}_{r,k}; 2, [3, 1])$ , respectively. After obtaining the angular parameters at the BS/UE ends, we can compute a smaller-sized tensor signal (in the noiseless case) as

$$\begin{aligned} \mathcal{Z}_{r,k} &\triangleq \mathcal{Y}_{r,k} \times_1 (\mathbf{W}_{r,k}^H \mathbf{A}_{\text{B},r})^\dagger \times_2 (\mathbf{F}_{r,k}^T \mathbf{A}_{\text{U},r}^*)^\dagger \\ &= \mathcal{I}_{3, N_{\text{I}}} \times_1 \text{Diag}(\beta_{\text{BI},r,k}) \mathbf{A}_{\text{D},r}^H \times_2 \text{Diag}(\beta_{\text{IU},r,k}) \mathbf{A}_{\text{A},r}^T \\ &\quad \times_3 \Psi_{r,k}. \end{aligned} \quad (21)$$

The rank-deficient factor matrices of (7) are transformed into full-row-rank ones. Performing channel factorization or parameter recovery of  $\mathcal{Z}_{r,k} \in \mathbb{C}^{L_{\text{BI},r} \times L_{\text{IU},r} \times P_{\text{tr}}}$  can effectively reduce the computational complexity.

### B. Cascaded Parameter Recovery

For the training design of random or structured IRS reflection pattern, one can directly compute

$$\begin{aligned} \mathbf{G}_{\text{cas},r,k} &\triangleq (\mathbf{W}_{r,k}^H \mathbf{A}_{\text{B},r} \otimes \mathbf{F}_{r,k}^T \mathbf{A}_{\text{U},r}^*)^\dagger (\tilde{\mathbf{H}}_{\text{BI},r,k} \odot \tilde{\mathbf{H}}_{\text{IU},r,k}) \\ &= (\text{Diag}(\beta_{\text{BI},r,k}) \otimes \text{Diag}(\beta_{\text{IU},r,k})) (\mathbf{A}_{\text{D},r}^H \odot \mathbf{A}_{\text{A},r}^T), \end{aligned} \quad (22)$$

which concatenates  $L_{\text{BI},r} L_{\text{IU},r}$  parallel response vectors of  $\{\rho_{r,k,\ell_1\ell_2} \mathbf{a}_{\text{I}}(\pi_{\text{h},r,\ell_1\ell_2}, \pi_{\text{v},r,\ell_1\ell_2})\}$  in (17). By leveraging the 2-D uniform array geometry of  $\mathbf{a}_{\text{I}}(\pi_{\text{h}}, \pi_{\text{v}})$ , we can easily recover the cascaded channel parameters indexed by  $(\ell_1, \ell_2)$  in (18) from the  $((\ell_1 - 1)L_{\text{IU},r} + \ell_2)$  row of  $\mathbf{G}_{\text{cas},r,k}$ . We note that  $\pi_{\text{h}}(\text{v}),r,\ell_1\ell_2$  ranges across  $[-2, 2]$ , which may cause a *spatial aliasing* problem, i.e., the extracted phase of  $e^{j \frac{2\pi}{\lambda_c} d_{\text{h}}(\text{v})\pi_{\text{h}}(\text{v}),r,\ell_1\ell_2}$  may exhibit an uncertain bias of  $2i\pi, i \in \mathbb{Z}$ . In order to achieve a unique value of  $\pi_{\text{h}}(\text{v}),r,\ell_1\ell_2$ , we propose to configure the inter-element spacing of IRS arrays as  $d_{\text{h}} = d_{\text{v}} = \lambda_c/4$ , yielding a maximal phase range of  $[-\pi, \pi]$ . Furthermore, by exploiting the rotational-invariant phase feature of  $\{\rho_{r,k,\ell_1\ell_2}\}_{k=1}^{K_{\text{tr}}}$  along multiple training subcarriers as

$$\rho_{r,k_1,\ell_1\ell_2}^{-1} \rho_{r,k_2,\ell_1\ell_2} = e^{j \frac{2\pi f_s}{K} (k_1 - k_2) (\tau_{\text{BI},r,\ell_1} + \tau_{\text{IU},r,\ell_2})}, \quad (23)$$

with  $k_1, k_2 \in \mathcal{I}(K_{\text{tr}})$ , we can recover the cascaded time delays, i.e.,  $\tau_{\text{BI},r,\ell_1} + \tau_{\text{IU},r,\ell_2}$ .

Next, we develop two dedicated channel parameter recovery solutions for the grouping and sparse training patterns respectively, as well as, a generalized solution for an arbitrary reflection pattern.

1) *Grouping Scheme*: Recall that the grouping pattern in Fig. 2(a) employs  $G$  distinct training codebooks  $\{\psi_{r,k}^g\}$ .

Following (14), we define  $\bar{\mathcal{G}}_{\text{cas},r,k} \triangleq \mathcal{Z}_{r,k} \times_3 \bar{\Psi}_{r,k}^\dagger$ , whose mode-3 slice matrices can be expressed as

$$[\bar{\mathcal{G}}_{\text{cas},r,k}]_{:,:,g} = \text{Diag}(\beta_{\text{BI},r,k})[\mathbf{A}_{\text{D},r}]_{\mathcal{N}_g}^H \times [\mathbf{A}_{\text{A},r}]_{\mathcal{N}_g} \text{Diag}(\beta_{\text{IU},r,k}). \quad (24)$$

With the regularly grouped lattice layout, the mode-3 (tube) fibers of  $\bar{\mathcal{G}}_{\text{cas},r,k} \in \mathbb{C}^{L_{\text{BI},r} \times L_{\text{IU},r} \times G}$  have a phase rotational-invariant characteristic as

$$(\mathbf{J}_{\text{gh}2} \otimes \mathbf{J}_{\text{gv}2})[\bar{\mathcal{G}}_{\text{cas},r,k}]_{\ell_1, \ell_2, :} = (\mathbf{J}_{\text{gh}1} \otimes \mathbf{J}_{\text{gv}1})[\bar{\mathcal{G}}_{\text{cas},r,k}]_{\ell_1, \ell_2, :} \times e^{j\frac{2\pi}{\lambda_c}(d_h \pi_{h,r,\ell_1 \ell_2} + d_v \pi_{v,r,\ell_1 \ell_2})}, \quad (25)$$

where  $\ell_1 \in \mathcal{I}(L_{\text{BI},r})$ ,  $\ell_2 \in \mathcal{I}(L_{\text{IU},r})$ ;  $\mathbf{J}_{\text{gh}1(\text{gv}1)}$ ,  $\mathbf{J}_{\text{gh}2(\text{gv}2)} \in \{0, 1\}^{(G_{\text{h}(v)}-1) \times G_{\text{h}(v)}}$  are the selection matrices that select the first and last  $(G_{\text{h}(v)} - 1)$  entries of the fibers, respectively. Therefore,  $[\bar{\mathcal{G}}_{\text{cas},r,k}]_{\ell_1, \ell_2, :} \in \mathbb{C}^G$  can be equivalently regarded as the steering vector of a  $(G_{\text{h}} \times G_{\text{v}})$ -UPA topology, from which we are able to derive the cascaded parameters in (18).

2) *Sparse Scheme*: Recall that the sparse pattern in Fig. 2(b) activates  $N_{\text{A}}$  out of  $N_{\text{I}}$  reflectors. One can derive a submatrix of (22) as  $\mathbf{G}_{\text{cas},r,k}^{(N_{\text{A}})} = \text{Matr}(\mathcal{Z}_{r,k}; [2, 1], 3)(\Psi_{r,k}^{(N_{\text{A}})})^\dagger$ . Then, the following steps are performed: (i) Copy and insert data at the overlapped origin of the sparse/dense subarrays, yielding  $\mathbf{G}_{\text{cas},r,k}^{(N_{\text{A}})} \in \mathbb{C}^{L_{\text{BI},r} L_{\text{IU},r} \times (N_{\text{A}}+1)}$  with  $\bar{N}_{\text{A}} \triangleq \{N_{\text{A}}, N_{\text{v}}(\lceil N_{\text{h}}/2 \rceil - 1) + \lceil N_{\text{v}}/2 \rceil\}$ ; (ii) Compute  $(\mathbf{G}_{\text{cas},r,k}^{(N_{\text{A}})})^H \odot (\mathbf{G}_{\text{cas},r,k}^{(N_{\text{A}})})^T$  and sort its rows to fit the 2-D nested array geometry; (iii) Remove the repetitive or discontinuous rows, yielding  $L_{\text{BI},r} L_{\text{IU},r}$  response vectors of an  $(N_{\text{hd}} N_{\text{hs}} \times (2N_{\text{vd}} N_{\text{vs}} - 1))$ -difference coarray; (iv) Leverage rank-1 tSVD to recover the cascaded parameters in (18). Note that these steps leverage the principle of nested array processing, but do not require any statistical information of the signal correlation.

3) *ANM-SDP/ADMM Scheme*: We leverage the atomic norm denoising technique to develop a general solution of (18) for arbitrary designs of IRS training patterns [25], [26]. The  $(\ell_1, \ell_2)$ th mode-3 fiber of  $\mathcal{Z}_{r,k}$ , i.e.,  $[\mathcal{Z}_{r,k}]_{\ell_1, \ell_2, :}$ , is expressed as  $\mathbf{z}_{r,k,\ell_1 \ell_2} = \Psi_{r,k} \mathbf{a}_{\text{I}}(\pi_{\text{h},r,\ell_1 \ell_2}, \pi_{\text{v},r,\ell_1 \ell_2}) \rho_{r,k,\ell_1 \ell_2} \cdot \mathbf{a}_{\text{I}}(\pi_{\text{h}}, \pi_{\text{v}})$  can be regarded as a 2-D atom, whilst the 2-D atomic norm of arbitrary vector  $\mathbf{x} \in \mathbb{C}^{N_{\text{I}}}$  with respect to the atomic set  $\mathcal{A} \triangleq \{\mathbf{a}_{\text{I}}(\pi_{\text{h}}, \pi_{\text{v}}) | \pi_{\text{h}}, \pi_{\text{v}} \in [-2, 2]\}$  is then defined as

$$\begin{aligned} \|\mathbf{x}\|_{\mathcal{A}} &= \inf \{t : \mathbf{x} \in t \text{conv}(\mathcal{A})\} \\ &= \inf_{\pi_{\text{h}(v)}, \ell, \rho \ell} \left\{ \sum_{\ell} |\rho \ell| \|\mathbf{x} = \sum_{\ell} \rho \ell \mathbf{a}_{\text{I}}(\pi_{\text{h},\ell}, \pi_{\text{v},\ell})\right\} \\ &= \inf_{\mathbf{u}, t} \left\{ \frac{t}{2} + \frac{1}{2N_{\text{I}}} \text{Tr}(\mathbb{T}_2(\mathbf{u})) \right\}, \\ &\text{s.t. } \begin{bmatrix} \mathbb{T}_2(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & t \end{bmatrix} \succeq \mathbf{0}, \end{aligned} \quad (26)$$

where  $\text{conv}(\mathcal{A})$  is the convex hull of  $\mathcal{A}$ ;  $\mathbb{T}_2(\mathbf{u}) \in \mathbb{C}^{N_{\text{I}} \times N_{\text{I}}}$  denotes a two-level block Toeplitz matrix defined by  $\mathbf{u} \in \mathbb{C}^{(2N_{\text{h}}-1)(2N_{\text{v}}-1)}$  [25], [26]. We can now individually solve

$L_{\text{BI},r} L_{\text{IU},r}$  atomic norm minimization (ANM) subproblems of (26) as

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{x}, t} \quad & \frac{\mu t}{2} + \frac{\mu}{2N_{\text{I}}} \text{Tr}(\mathbb{T}_2(\mathbf{u})) + \frac{1}{2} \|\mathbf{z}_{r,k,\ell_1 \ell_2} - \Psi_{r,k} \mathbf{x}\|^2 \\ \text{s.t.} \quad & \begin{bmatrix} \mathbb{T}_2(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & t \end{bmatrix} \succeq \mathbf{0}, \ell_1 \in \mathcal{I}(L_{\text{BI},r}), \ell_2 \in \mathcal{I}(L_{\text{IU},r}), \end{aligned} \quad (27)$$

where  $\mu$  denotes the regularization parameter;  $\mathbf{x}$  approximates  $\rho_{r,k,\ell_1 \ell_2} \mathbf{a}_{\text{I}}(\pi_{\text{h},r,\ell_1 \ell_2}, \pi_{\text{v},r,\ell_1 \ell_2})$ . These optimization subproblems can be solved by the semidefinite programming method (SDP). In order to improve the computational efficiency of parameter recovery for large-scale IRSs, one can adopt the alternating direction method of multipliers (ADMM). The detailed derivations of the Lagrangian function and variable update are omitted due to space limitations, and the interested readers are referred to [25], [26] and references therein. Finally, one can exploit the rotational-invariant feature of  $\mathbb{T}_2(\mathbf{u})$  to recover the cascaded parameters.

### C. Twin-IRS-Assisted Parameter Decoupling

Given arbitrary pairs of  $(\phi_{\text{D},r,\ell_1}, \phi_{\text{A},r,\ell_2})$ ,  $(\theta_{\text{D},r,\ell_1}, \theta_{\text{A},r,\ell_2})$ ,  $(\beta_{\text{BI},r,k,\ell_1}, \beta_{\text{IU},r,k,\ell_1})$  that satisfy the constraints of (18), one can recover the cascaded channel  $\mathbf{H}_{\text{cas},r,k}$  as in (17). However, the information of (18) is not sufficient to support an exact decoupling of the cascaded parameters. Concretely, there are  $2(L_{\text{BI},r} + L_{\text{IU},r})$  AoA/AoDs at the IRS to be estimated but only  $2(L_{\text{BI},r} + L_{\text{IU},r} - 1)$  out of  $2L_{\text{BI},r} L_{\text{IU},r}$  effective phase equations, whilst the amplitude equations cannot be directly used due to the unknown path gains. Therefore, with a single IRS plane, one can only formulate an underdetermined nonlinear equation system of  $\{\phi_{\text{A}(\text{D}),r,\ell}, \theta_{\text{A}(\text{D}),r,\ell}\}$ , which does not have a unique solution. In order to provide sufficient information for the parameter decoupling, we propose the twin-IRS structure, enabling the nonlinear solver to converge to the precise solution with a much higher probability. Based on (5), (6), we can represent the cascaded parameters via the secondary IRS as

$$\begin{cases} \varpi_{\text{h},r,\ell_1 \ell_2} \triangleq -\sin \theta_{\text{D},r,\ell_1} \sin(\phi_{\text{D},r,\ell_1} - \delta_{\text{h}}) \\ \quad + \sin \theta_{\text{A},r,\ell_2} \sin(\phi_{\text{A},r,\ell_2} - \delta_{\text{h}}), \\ \varpi_{\text{v},r,\ell_1 \ell_2} \triangleq -\sin \delta_{\text{v}} \sin \theta_{\text{D},r,\ell_1} \cos(\phi_{\text{D},r,\ell_1} - \delta_{\text{h}}) \\ \quad + \sin \delta_{\text{v}} \sin \theta_{\text{A},r,\ell_2} \cos(\phi_{\text{A},r,\ell_2} - \delta_{\text{h}}) \\ \quad + (-\cos \theta_{\text{D},r,\ell_1} + \cos \theta_{\text{A},r,\ell_2}) \cos \delta_{\text{v}}, \\ \varrho_{r,k,\ell_1 \ell_2} \triangleq \beta_{\text{BI},r,k,\ell_1} \beta_{\text{IU},r,k,\ell_2} \frac{F(\phi_{\text{D},r,\ell_1}, \theta_{\text{D},r,\ell_1}; \delta_{\text{h}}, \delta_{\text{v}})}{F(\phi_{\text{D},r,\ell_1}, \theta_{\text{D},r,\ell_1}; 0, 0)} \\ \quad \times \frac{F(\phi_{\text{A},r,\ell_2}, \theta_{\text{A},r,\ell_2}; \delta_{\text{h}}, \delta_{\text{v}})}{F(\phi_{\text{A},r,\ell_2}, \theta_{\text{A},r,\ell_2}; 0, 0)}. \end{cases} \quad (28)$$

By combining the *cascaded* parameters in (18), (28) and leveraging the trigonometric function theory, extra phase equations of  $-\sin \theta_{\text{D},r,\ell_1} \cos \phi_{\text{D},r,\ell_1} + \sin \theta_{\text{A},r,\ell_2} \cos \phi_{\text{A},r,\ell_2}$ , and magnitude equations of  $\frac{F(\phi_{\text{D},r,\ell_1}, \theta_{\text{D},r,\ell_1}; \delta_{\text{h}}, \delta_{\text{v}}) F(\phi_{\text{A},r,\ell_2}, \theta_{\text{A},r,\ell_2}; \delta_{\text{h}}, \delta_{\text{v}})}{F(\phi_{\text{D},r,\ell_1}, \theta_{\text{D},r,\ell_1}; 0, 0) F(\phi_{\text{A},r,\ell_2}, \theta_{\text{A},r,\ell_2}; 0, 0)}$  can be introduced by the secondary IRS. These additional constraints make it possible to realize the parameter decoupling. Here, we present two decoupling modes of channel parameters against different cases of system settings and channel conditions.



1) *“All” Decoupling Mode*: If the IRS reflectors are omnidirectional, i.e.,  $F(\phi, \theta; \delta_h, \delta_v) \equiv 1$ , the magnitude equations of (28) become invalid. Hence, the nonlinear solver is able to converge to the exact azimuth/elevation AoA/AoDs only when the number of effective phase equations is no less than the number of angles to be estimated, i.e.,

$$\begin{aligned} 3(L_{\text{BI},r} + L_{\text{IU},r} - 1) &\geq 2(L_{\text{BI},r} + L_{\text{IU},r}) \\ &\Rightarrow L_{\text{BI},r} + L_{\text{IU},r} \geq 3. \end{aligned} \quad (29)$$

We jointly solve a nonlinear system of  $3(L_{\text{BI},r} + L_{\text{IU},r} - 1)$  phase equations in (18), (28) involving geometric information of  $\sin \theta \cos \phi$ ,  $\sin \theta \sin \phi$ ,  $\cos \theta$  along the  $x, y, z$ -axes, respectively, to acquire precise estimates of **all** the  $2(L_{\text{BI},r} + L_{\text{IU},r})$  angular parameters  $\{\phi(\theta)_{\text{D},r,\ell_1}, \phi(\theta)_{\text{A},r,\ell_2}\}$  at once. One can directly employ existing nonlinear algorithms, e.g., Gauss-Newton method, Levenberg-Marquardt method, trust-region (dogleg) method, etc., to efficiently solve this problem.

2) *“Pair” Decoupling Mode*: If there exists only one propagation path in both the BS-IRS and IRS-MS channels, i.e.,  $L_{\text{BI},r} + L_{\text{IU},r} = 2$ , one cannot just rely on three phase equations to obtain unique estimates of four angles.<sup>5</sup> We need to recover four angular parameters  $\{\phi_{\text{D}(\text{A}),1}, \theta_{\text{D}(\text{A}),1}\}$  by jointly solving four nonlinear magnitude/phase equations of (18), (28). Note that this approach can also be applied to the case (29), returning the angles of a **pair** of paths at a time. One needs to average the recovery results of each angle to obtain  $2(L_{\text{BI},r} + L_{\text{IU},r})$  distinct parameters from in total  $4L_{\text{BI},r}L_{\text{IU},r}$  estimates.

After obtaining the angles and the cascaded time delays, we can remove their contributions from the equivalent path gains in (18) or (28) to derive  $\alpha_{\text{BI},r,k} \alpha_{\text{IU},r,k}$ . Due to the randomly distributed complex gains including both the large-scale fading and the small-scale fading, it is usually difficult to decouple the path gains, as well as, the time delays solely based on the training signals. Fortunately, we can leverage the 3-D geometric relationship of communication devices to realize the environment mapping, determining the precise propagation distances. The detailed derivations are provided in the next section.

## V. CHANNEL PARAMETER-BASED USER LOCALIZATION

After acquiring precise estimates of the channel parameters via the decoupling mode described in Section IV. C. 1) or 2), we can combine the information of propagation channels and device physical configurations to realize an application of user localization. The BS and UE are located at  $\mathbf{p}_\text{B}, \mathbf{p}_\text{U} \in \mathbb{R}^3$ , respectively, whose orientations, i.e., the normals to the antenna array planes, are defined as  $\mathbf{n}_\text{B}, \mathbf{n}_\text{U} \in \mathbb{R}^3$ , respectively. The adjacent primary/secondary IRSs of the  $r$ th twin-IRS structure are assumed to share identical coordinates  $\mathbf{p}_r \in \mathbb{R}^3$ , while the normal direction to the primary IRS is  $\mathbf{n}_r \in \mathbb{R}^3$ . The scatterers of the  $\ell_1$ th NLoS path of  $\mathbf{H}_{\text{BI},r,k}$  and

<sup>5</sup>If there exists a LoS path with a large Rician  $K$ -factor, the other NLoS components are much weaker and may be regarded as environmental noise. Then, the channel matrix  $\mathbf{H}_{\text{BI}(\text{IU}),r,k}$  has approximately rank-1 with  $L_{\text{BI}(\text{IU}),r} \approx 1$ .

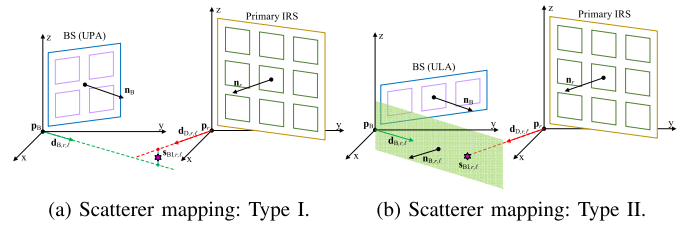


Fig. 3. Environment mapping of the BS-IRS channels. (a) The scatterer is approximated as the median point on the common perpendicular of two spatial lines. (b) The scatterer is approximated as the intersection point of a spatial line and a spatial plane.

the  $\ell_2$ th NLoS path of  $\mathbf{H}_{\text{IU},r,k}$  are located at  $\mathbf{s}_{\text{BI},r,\ell_1} \in \mathbb{R}^3$  and  $\mathbf{s}_{\text{IU},r,\ell_2} \in \mathbb{R}^3$ , respectively.

### A. Preliminary Environment Mapping

The positioning performance depends on the environment mapping that determines the distribution of scatterers and propagation paths. Note that the AoA/AoDs at the IRS are defined relative to the primary IRS on the  $yz$ -plane with a default normal  $\mathbf{n}_0 \triangleq [1, 0, 0]^T$ , as illustrated in Fig. 1(b). We parameterize the rotation of an IRS plane with a normal  $\mathbf{n}_r \in \mathbb{R}^3$  from the default orientation by the *rotation axis* and *rotation angle*, which are defined as<sup>6</sup>

$$\mathbf{c}_r = \frac{\mathbf{n}_0 \times \mathbf{n}_r}{\|\mathbf{n}_0 \times \mathbf{n}_r\|}, \quad \xi_r = \cos^{-1}(\mathbf{n}_0 \bullet \mathbf{n}_r). \quad (30)$$

The actual direction of a signal wave parameterized by azimuth/elevation angles  $\{\phi, \theta\}$ , or equivalently a direction  $\mathbf{d}_0 \triangleq [\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^T$ , relative to the primary IRS can be derived by the *Rodrigues' rotation formula* as [46]

$$\mathbf{d}_r = \mathbf{d}_0 \cos \xi_r + (\mathbf{c}_r \times \mathbf{d}_0) \sin \xi_r + \mathbf{c}_r (\mathbf{c}_r \bullet \mathbf{d}_0) (1 - \cos \xi_r). \quad (31)$$

Given the angular estimates and known device orientations, we can derive the actual path directions at the BS and IRSs, denoted by  $\mathbf{d}_{\text{B},r,\ell_1}, \mathbf{d}_{\text{D},r,\ell_1}, \mathbf{d}_{\text{A},r,\ell_2} \in \mathbb{R}^3, \forall \ell_1 \in \mathcal{I}(L_{\text{BI},r}), \ell_2 \in \mathcal{I}(L_{\text{IU},r})$ , respectively. For a LoS path,  $\mathbf{d}_{\text{B},r,\ell_1}, \mathbf{d}_{\text{D},r,\ell_1}$  should be (approximately) collinear.

Recall that only the NLoS paths scattered or reflected once are considered. The scatterer located at  $\mathbf{s}_{\text{BI},r,\ell_1} \in \mathbb{R}^3$  is the intersection point of two spatial lines passing through  $\mathbf{p}_\text{B}$  and  $\mathbf{p}_r$  with directions  $\mathbf{d}_{\text{B},r,\ell_1}$  and  $\mathbf{d}_{\text{D},r,\ell_1}$ , respectively. In the noisy case, these estimated spatial lines may not exactly intersect. As illustrated in Fig. 3(a), the scatterer can be approximated by the median point on the common perpendicular of the two spatial lines as

$$\begin{aligned} &\min_{\mathbf{x} \in \mathbb{R}^2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2, \\ &\text{s.t. } \mathbf{s}_{\text{BI},r,\ell_1} = \frac{1}{2}(\mathbf{p}_\text{B} + \mathbf{d}_{\text{B},r,\ell_1}[\mathbf{x}]_1 + \mathbf{p}_r - \mathbf{d}_{\text{D},r,\ell_1}[\mathbf{x}]_2), \\ &\mathbf{A} = [\mathbf{d}_{\text{B},r,\ell_1}, \mathbf{d}_{\text{D},r,\ell_1}], \quad \mathbf{b} = \mathbf{p}_r - \mathbf{p}_\text{B}. \end{aligned} \quad (32)$$

Note that if the BS antennas are ULA-shaped, which are commonly assumed to be parallel with the  $xy$ -plane, only the

<sup>6</sup>If  $\mathbf{n}_r, \mathbf{n}_0$  are collinear,  $\mathbf{c}_r \in \mathbb{R}^3$  can be arbitrary vector orthogonal to  $\mathbf{n}_r(\mathbf{n}_0)$ , and  $\xi_r = 0$  or  $\pi$ .

azimuth angle can be estimated (elevation angle is  $\pi/2$  by default). In this case, as illustrated in Fig. 3(b), the scatterer can be derived as the intersection point of a spatial line passing through  $\mathbf{p}_r$  with a direction  $\mathbf{d}_{D,r,\ell_1}$  and a spatial plane passing through  $\mathbf{p}_B$  with a normal  $\mathbf{n}_{B,r,\ell_1} \perp \mathbf{d}_{B,r,\ell_1}$  as [47]

$$\begin{aligned} [\mathbf{s}_{B,r,\ell_1}]_i &= \frac{1}{\mathbf{n}_{B,r,\ell_1} \bullet \mathbf{d}_{D,r,\ell_1}} \\ &\times \left( [\mathbf{d}_{D,r,\ell_1}]_i \sum_{j \neq i} [\mathbf{n}_{B,r,\ell_1}]_j ([\mathbf{p}_B]_j - [\mathbf{p}_r]_j) \right. \\ &+ [\mathbf{d}_{D,r,\ell_1}]_i [\mathbf{n}_{B,r,\ell_1}]_i [\mathbf{p}_B]_i \\ &\left. + [\mathbf{p}_r]_i \sum_{j \neq i} [\mathbf{n}_{B,r,\ell_1}]_j [\mathbf{d}_{D,r,\ell_1}]_j \right), \end{aligned} \quad (33)$$

where  $i \in \{1, 2, 3\}$ . Once the scatterer position is determined, the propagation distance, as well as, the path time delay from the IRS to the BS can be directly computed as

$$\tau_{B,r,\ell_1} = \frac{\|\mathbf{p}_B - \mathbf{s}_{B,r,\ell_1}\| + \|\mathbf{p}_r - \mathbf{s}_{B,r,\ell_1}\|}{v_c}, \quad (34)$$

where  $v_c = 3 \times 10^8$  m/s is the light velocity. One can now easily derive the UE-IRS path distance by combining (34) and the estimated cascaded time delay, i.e.,  $\tau_{B,r,\ell_1} + \tau_{U,r,\ell_2}$ , from (18), (28). Without the information of the user coordinates, it is difficult to realize a unique environment mapping between the IRS and UE. Fortunately, after acquiring the UE position and orientation based on the LoS component of  $\mathbf{H}_{U,r,k}$ , we can finally determine the scattering details.

### B. User Localization Implementation

We present two dedicated user localization schemes for the single-carrier and multi-carrier training strategies, respectively. When the UE-IRS channel contains a LoS path with a relatively large Rician  $K$ -factor, it can be detected and distinguished by its dominant complex gains.<sup>7</sup>

1) *Single-Carrier Strategy*: In the single-carrier strategy, the system activates a single subcarrier to perform channel estimation. With only one estimate of  $\beta_{B,r,k,\ell_1} \beta_{U,r,k,\ell_2}$ , we cannot recover the cascaded time delays  $\tau_{B,r,\ell_1} + \tau_{U,r,\ell_2}$ , as well as, the LoS path distance of  $\mathbf{H}_{U,r,k}$ . We propose to localize the UE by an AoA-based positioning scheme with  $R = 2$  distributed twin-IRS structures. The user position can be derived as the intersection point of two spatial lines passing through  $\mathbf{p}_1$  and  $\mathbf{p}_2$  with directions  $\mathbf{d}_{A,1,1}$  and  $\mathbf{d}_{A,2,1}$ , respectively, which can be approximated by following (32) as

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^2} & \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2, \\ \text{s.t.} & \mathbf{A} = [\mathbf{d}_{A,1,1}, \mathbf{d}_{A,2,1}], \mathbf{b} = \mathbf{p}_2 - \mathbf{p}_1, \\ & \mathbf{p}_U = \frac{1}{2}(\mathbf{p}_1 + \mathbf{d}_{A,1,1}[\mathbf{x}]_1 + \mathbf{p}_2 - \mathbf{d}_{A,2,1}[\mathbf{x}]_2). \end{aligned} \quad (35)$$

2) *Multi-Carrier Strategy*: In the multi-carrier strategy, we are able to precisely determine the LoS path distance of  $\mathbf{H}_{U,r,k}$  by realizing the scatterer mapping. We propose to localize the UE by a hybrid AoA-delay positioning scheme with  $R = 1$  twin-IRS structure. The user position can be

<sup>7</sup>Note that the proposed twin-IRS structure also provides precise information required by NLoS-based positioning schemes, e.g., [48], [49]. Due to space limitations, we consider only the case of LoS-based positioning here.

### Algorithm 1 Channel Estimation and User Localization

**Require:** observation  $\mathcal{Y}_{r,k}$ , IRS pattern  $\Psi_{r,k}$ , beamformers  $\mathbf{F}_{r,k}, \mathbf{W}_{r,k}$ , channel ranks  $L_{B,r}, L_{U,r}$ .

- 1: **procedure** CH-EST
- 2: Derive  $\mathbf{Y}_{r,k}^{(1)}, \mathbf{Y}_{r,k}^{(2)}$  by unfolding  $\mathcal{Y}_{r,k}$  as in Remark 2.
- 3: Derive AoA/AoDs  $\{\phi(\theta)_{B,\ell_1}\}$  and  $\{\phi(\theta)_{U,\ell_2}\}$  from the subspaces of  $\mathbf{Y}_{r,k}^{(1)}$  and  $\mathbf{Y}_{r,k}^{(2)}$ , respectively, by (20).
- 4: Compute the compressed tensor  $\mathcal{Z}_{r,k}$  by (21).
- 5: Derive  $\{\pi_{h(v),r,\ell_1\ell_2}, \rho_{r,k,\ell_1\ell_2}\}$  from  $\mathcal{Z}_{r,k}$  with specific training pattern as in Section IV. B.
- 6: Reconstruct  $\mathbf{H}_{cas,r,k}$  with the recovered (cascaded) channel parameters by (17).
- 7: **end procedure**
- 8: **procedure** CAS-DEC
- 9: Derive  $\{\pi_{h(v),r,\ell_1\ell_2}, \rho_{r,k,\ell_1\ell_2}\}, \{\varpi_{h(v),r,\ell_1\ell_2}, \varrho_{r,k,\ell_1\ell_2}\}$  by running CH-EST with the primary and secondary IRS planes, respectively.
- 10: Decouple  $\{\phi(\theta)_{D,r,\ell_1}, \phi(\theta)_{A,r,\ell_2}\}$  by solving the non-linear system of equations formulated by (18), (28) as in Section IV. C.
- 11: **end procedure**
- 12: **procedure** UE-LOC
- 13: Determine the actual signal directions by (30), (31).
- 14: Derive the scatterer positions by (32), (33).
- 15: Derive the UE position and orientation by (34)–(37).
- 16: **end procedure**

easily derived as the point,  $v_c \tau_{U,r,1}$  meters away from  $\mathbf{p}_r$ , on a spatial line passing through  $\mathbf{p}_1$  with the direction  $\mathbf{d}_{A,1,1}$  as

$$\mathbf{p}_U = \mathbf{p}_r + \mathbf{d}_{A,r,1} v_c \tau_{U,r,1}. \quad (36)$$

3) *Orientation Determination*: We can further derive the device orientation. The default direction of a LoS path parameterized by  $\{\phi_{U,r,1}, \theta_{U,r,1}\}$  at the UE is defined as  $\mathbf{d}_{U,r,1} \triangleq [\sin \theta_{U,r,1} \cos \phi_{U,r,1}, \sin \theta_{U,r,1} \sin \phi_{U,r,1}, \cos \theta_{U,r,1}]^T$ , and the actual direction of this path is  $-\mathbf{d}_{A,r,1} \approx \mathbf{p}_r - \mathbf{p}_U$ . The relative rotation is identical to that of the UE orientation from  $\mathbf{n}_0 = [1, 0, 0]^T$  to  $\mathbf{n}_U \in \mathbb{R}^3$ , which can be derived by the Rodrigues' rotation theory as

$$\begin{aligned} \mathbf{c}_{U,r,1} &= \frac{\mathbf{d}_{U,r,1} \times (-\mathbf{d}_{A,r,1})}{\|\mathbf{d}_{U,r,1} \times (-\mathbf{d}_{A,r,1})\|}, \\ \xi_{U,r,1} &= \cos^{-1}(\mathbf{d}_{U,r,1} \bullet (-\mathbf{d}_{A,r,1})), \end{aligned} \quad (37a)$$

$$\begin{aligned} \mathbf{n}_U &= \mathbf{n}_0 \cos \xi_{U,r,1} + (\mathbf{c}_{U,r,1} \times \mathbf{n}_0) \sin \xi_{U,r,1} \\ &+ \mathbf{c}_{U,r,1} (\mathbf{c}_{U,r,1} \bullet \mathbf{n}_0) (1 - \cos \xi_{U,r,1}). \end{aligned} \quad (37b)$$

Finally, we summarize the channel estimation and user localization schemes as Algorithm 1, where the terms ‘‘CH-EST’’, ‘‘CAS-DEC’’ and ‘‘UE-LOC’’ are abbreviations for channel estimation, cascaded parameter decoupling and UE localization, respectively.

## VI. NUMERICAL RESULTS

Typical simulation settings are listed here:  $N_B = N_U = 24$  (ULA),  $M_{tr} = T_{tr} = 12$ ,  $K = 128$ ,  $K_{tr} = 6$ ,  $f_s = 0.32$  GHz,

TABLE III  
PERFORMANCE COMPARISON OF DIFFERENT BEAMFORMING SCHEMES

Approaches		OMP [13]		Coarse [45]		Baseline		Uniform DFT		Optimized DFT	
$P_{\text{tr}}$	SNR (dB)	acc	pow	acc	pow	acc	pow	acc	pow	acc	pow
8	15	1.862	6.082	8.229	5.931	0.210	6.173	0.281	6.261	0.201	11.533
	30	1.732	6.077	0.904	5.929	0.034	6.182	0.043	6.264	0.024	11.588
14	15	1.797	6.203	8.153	6.015	0.148	6.280	0.201	6.373	0.139	11.918
	30	1.702	6.221	0.879	5.999	0.025	6.267	0.032	6.360	0.017	12.103

$f_c = 28$  GHz;  $L_{\text{BI},r} = L_{\text{IU},r} = 3$ ,  $\phi_{\text{B}(\text{U}),r,\ell}$  follows a uniform distribution  $\mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$ ,  $\phi_{\text{D}(\text{A}),r,\ell}$  follows  $\mathcal{U}(-\frac{\pi}{3}, \frac{\pi}{3})$ ,  $\theta_{\text{D}(\text{A}),r,\ell}$  follows  $\mathcal{U}(\frac{\pi}{6}, \frac{5\pi}{6})$ ,  $\alpha_{\text{BI}(\text{IU}),r,\ell}$  follows a normal distribution  $\mathcal{CN}(0, 1)$ .<sup>8</sup> The received signal-to-noise ratio (SNR) is defined as  $\|\mathcal{Y}_{r,k} - \mathcal{N}_{r,k}\|_{\text{F}}^2 / \|\mathcal{N}_{r,k}\|_{\text{F}}^2$ . The simulation of channel estimation and parameter recovery considers a full-NLoS propagation; the simulation of environment mapping and user localization assumes that  $\mathbf{H}_{\text{IU},r,k}$  contains a LoS path, while  $\mathbf{H}_{\text{BI},r,k}$  consists of only NLoS components. Apart from the narrowband positioning that requires two twin-IRS structures, most simulations are realized with a single twin-IRS structure. The two twin-IRS pairs are located at  $\mathbf{p}_1 = [0, 0, 0]^T$ ,  $\mathbf{p}_2 = [8, 0, 0]^T$  with orientations  $\mathbf{n}_1 = [1, 0, 0]^T$ ,  $\mathbf{n}_2 = [-1, 0, 0]^T$ , respectively. We leverage the first one to perform the channel estimation and parameter recovery by default. The proposed channel estimation schemes, i.e., the sparse, grouping and ANM-SDP/ADMM schemes described in Section IV. B. 1), 2) and 3), respectively, are compared with the OMP [13], BALS [14] and SCPD [40] schemes.<sup>9</sup>

Table III tabulates the estimation performance of BS/UE angular parameters and normalized signal power with random training pattern for different beamforming schemes, where  $N_{\text{h}} = N_{\text{v}} = 4$  and  $q = 1.5$ .<sup>10</sup> It shows that the beamspace beamforming (19) achieves the best recovery accuracy and highest received signal strength when the search regions generally cover the AoA/AoDs at the BS/UE. The baseline scheme achieves slightly better performance than the uniform DFT scheme but yields much lower effective signal power than the optimized DFT scheme, leading to higher transmitted power for a desired received SNR. Hereinafter, we introduce a coarse estimation scheme with one extra frame, and use the obtained estimates to construct the beamformers  $\mathbf{F}_{r,k}$ ,  $\mathbf{W}_{r,k}$ , as well as, the training signal tensor  $\mathcal{Y}_{r,k}$ .

<sup>8</sup>One can leverage the principal component analysis or minimum length description to estimate the number of paths [40]. For simplifying the evaluation of the parameter recovery performance, we assume that  $L_{\text{BI}(\text{IU}),r}$  is *a priori* known or perfectly estimated.

<sup>9</sup>The maximum iterations of BALS is 20; the angular resolution of OMP is  $\frac{2\pi}{4N_{\text{h}}} \times \frac{2\pi}{4N_{\text{v}}}$ ; the maximum iterations and penalty parameter of ANM-ADMM are 20 and 0.1, respectively; the ANM-SDP is handled by the CVX toolbox [50].

<sup>10</sup>“acc” is defined as the rooted mean square error (RMSE) (degrees):  $(\sum_{\ell=1}^{L_{\text{BI}(\text{IU}),r}} (\hat{\phi}_{\text{B}(\text{U}),r,\ell} - \phi_{\text{B}(\text{U}),r,\ell})^2 / L_{\text{BI}(\text{IU}),r})^{1/2}$ ; “pow” is defined as the normalized signal power:  $\|\mathcal{Y}_{r,k} - \mathcal{N}_{r,k}\|_{\text{F}} / \sqrt{M_{\text{tr}} T_{\text{tr}} P_{\text{tr}}}$ ; the resolution of OMP is  $2\pi / 4N_{\text{B}(\text{U})}$ ; “Coarse” applies the coarse estimation scheme [45] within a single frame; “Baseline” applies the coarse estimation scheme [45] across  $(P_{\text{tr}} + 1)$  frames; “Uniform DFT” applies the DFT beamforming (19) across  $(P_{\text{tr}} + 1)$  frames with uniformly-spaced  $\{\omega_{\text{h}(\text{v}),m}\}$ ; “Optimized DFT” applies the DFT beamforming (19) across  $P_{\text{tr}}$  frames with  $\{\omega_{\text{h}(\text{v}),m}\}$  closest to the coarse estimates.

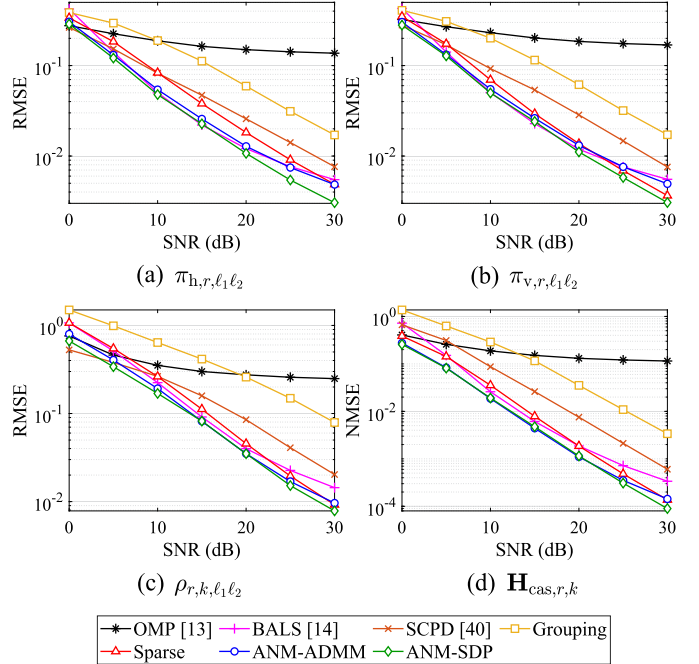


Fig. 4. RMSE/NMSE of the cascaded parameters/channel vs. SNR,  $N_{\text{h}} = N_{\text{v}} = 4$ ,  $P_{\text{tr}} = 14$ ,  $q = 1.5$ .

Table IV tabulates the RMSEs of cascaded phase/amplitude parameters (18) and normalized mean square error (NMSE) of the cascaded channel with different configurations of the sparse and grouping patterns.<sup>11</sup> It shows that when the quantity of training frames is sufficient, the estimation accuracy is mainly determined by the degrees of freedom of the difference coarray and equivalent array of the sparse and grouping patterns, respectively. Hereinafter, we adopt the best pattern designs marked in bold in Table IV to evaluate the channel estimation performance.

Figs. 4 and 5 plot the RMSEs of cascaded phase/amplitude parameters (18) and normalized mean square error (NMSE) of the cascaded channel versus the received SNR with  $N_1 = 16$  and 64, respectively. It shows that the OMP returns SNR-insensitive results due to its fixed angular resolution. The grouping scheme has relatively worse performance with the lowest complexity. The BALS and SCPD yield deteriorated performance when the quantity of training frames fails to keep pace with the increasing number of IRS unit cells. Moreover, the ANM-SPD achieves the best estimation

<sup>11</sup>The RMSE of a parameter  $x_{\ell_1, \ell_2}$  is defined as  $(\sum_{\ell_1, \ell_2} (\hat{x}_{\ell_1, \ell_2} - x_{\ell_1, \ell_2})^2 / L_{\text{BI},r} L_{\text{IU},r})^{1/2}$ ; the NMSE of the cascaded channel is defined as  $\|\hat{\mathbf{H}}_{\text{cas},r,k} - \mathbf{H}_{\text{cas},r,k}\|_{\text{F}}^2 / \|\mathbf{H}_{\text{cas},r,k}\|_{\text{F}}^2$ .

TABLE IV  
PERFORMANCE COMPARISON OF DIFFERENT TRAINING PATTERN CONFIGURATIONS

System Settings	$N_h = N_v = 4, P_{tr} = 14, \text{SNR} = 30 \text{ dB}$				$N_h = N_v = 8, P_{tr} = 48, \text{SNR} = 30 \text{ dB}$				
	Sparse			Grouping	Sparse			Grouping	
	$N_{hs}, N_{vs}, N_{hd}, N_{vd}; N_A$			$N_{gh}, N_{gv}; G$	$N_{hs}, N_{vs}, N_{hd}, N_{vd}; N_A$			$N_{gh}, N_{gv}; G$	
$\pi_{h,r,k,\ell_1\ell_2}$	3,1,1,2;4	1,1,3,2;6	<b>1,2,3,2;7</b>	<b>2,2;4</b>	1,2,3,3;10	3,2,3,2;11	<b>3,2,3,3;14</b>	4,4;4	<b>2,2;16</b>
$\pi_{v,r,k,\ell_1\ell_2}$	5.72e-3	5.42e-3	4.85e-3	1.65e-2	2.10e-3	9.00e-4	9.24e-4	1.47e-2	4.37e-3
$\rho_{r,k,\ell_1\ell_2}$	8.32e-3	8.90e-3	3.64e-3	1.73e-3	1.14e-3	1.37e-3	9.17e-4	1.57e-2	4.27e-3
$\mathbf{H}_{cas,r,k}$	1.57e-2	1.30e-2	9.29e-3	8.56e-2	3.91e-3	2.97e-3	2.85e-3	1.58e-1	3.18e-2
	3.28e-4	3.12e-4	1.37e-4	3.39e-3	4.33e-5	2.42e-5	2.02e-5	1.03e-2	6.17e-4

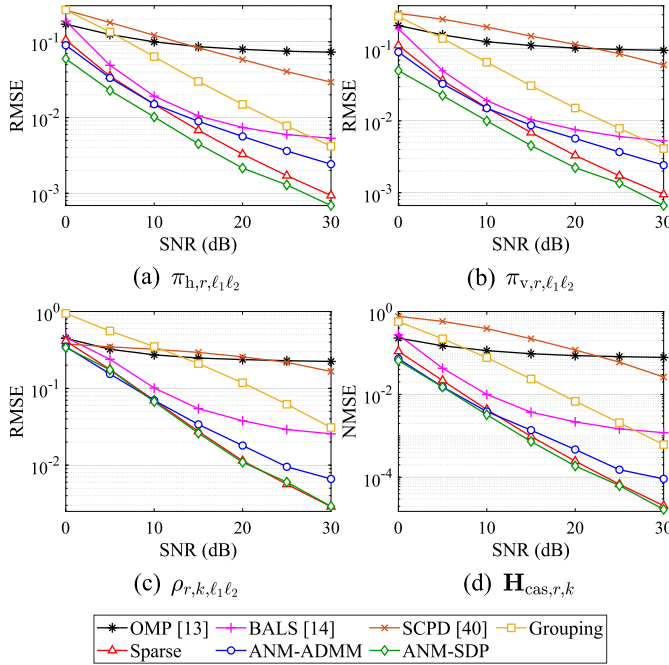


Fig. 5. RMSE/NMSE of the cascaded parameters/channel vs. SNR,  $N_h = N_v = 8, P_{tr} = 48, q = 1.5$ .

accuracy with the highest complexity. The sparse scheme and ANM-ADMM can effectively acquire RMSE/NMSEs close to those of ANM-SDP with reduced computational burden, especially in the case of small-scale IRSs with sufficient training measurements.

Figs. 6 and 7 plot the estimation performance curves of the cascaded parameters and channels versus the number of training frames with  $N_I = 16$  and 64, respectively. It shows that as  $P_{tr}$  increases, the recovery accuracy of the BALS rapidly improves, while the SCPD is working when the uniqueness condition (10) holds. Moreover, the performance of the ANM-ADMM gradually approximates to that of ANM-SDP. By contrast, the results of the sparse, grouping and ANM-SDP schemes are relatively less sensitive to the quantity of measurements, which originates from the robustness of LS and SDP operations. One can observe from Figs. 4–7 that the sparse scheme and ANM-SDP can achieve better performance than other counterparts. This can be attributed to the fact that the former effectively reduces the number of subchannels to be decomposed and reconstructed, while the latter leverages the optimization approaches with dominant computational complexity. Hereinafter, we combine these two

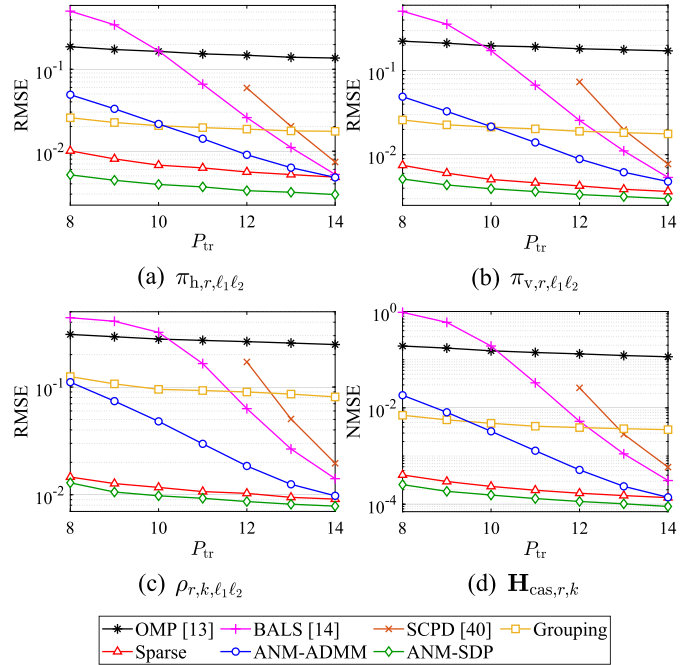


Fig. 6. RMSE/NMSE of the cascaded parameters/channel vs. training frames,  $N_h = N_v = 4, \text{SNR} = 30 \text{ dB}, q = 1.5$ .

methods with the proposed twin-IRS structure to implement the decoupling of channel multipath parameters.

Figs. 8 and 9 plot the RMSEs of the decoupled angular parameters versus the relative rotation angle and the power radiation coefficient of twin-IRS structures, respectively. The ANM-SDP and sparse schemes work with  $N_I = 16, P_{tr} = 14$  and  $N_I = 64, P_{tr} = 48$ , respectively. The figures indicate that the proposed twin-IRS structure enables us to precisely recover the angular parameters at the IRS end with a resolution less than  $0.1^\circ$ . Concretely, it can be observed that as  $\delta_h, \delta_v$  increase from  $5^\circ$  to  $30^\circ$ , the RMSEs firstly decrease and then increase. This can be explained as follows: too small rotation angles cannot yield significant amplitude/phase variations along the IRS array aperture, whilst too large ones will lead to a disappearance of some weak paths with near-zero AoA/AoDs relative to the IRS planes due to the non-ideal power radiation pattern. Furthermore, as  $q$  increases from 0 to 3, the performance of the ‘‘All’’ decoupling mode gradually deteriorates, while that of the ‘‘Pair’’ counterpart first improves and then worsens. The former does not rely on amplitude equations derived by the power radiation pattern, while the latter depends on effective amplitude equations with

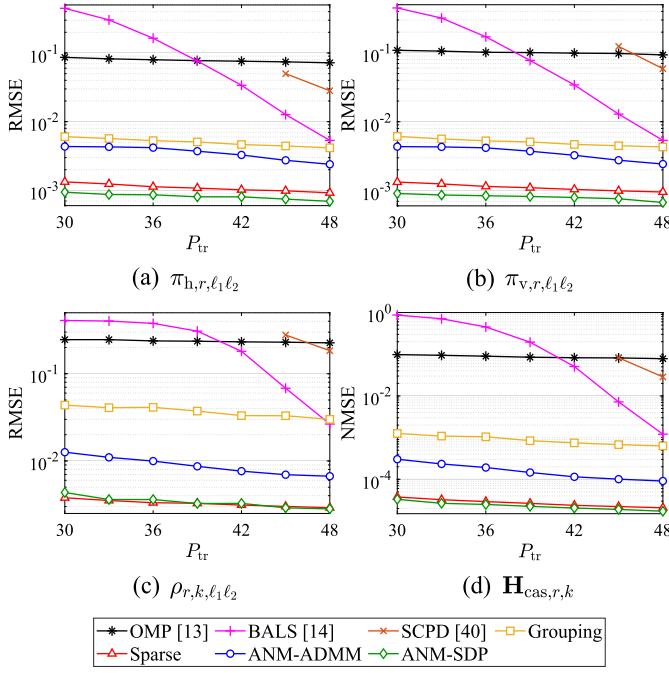


Fig. 7. RMSE/NMSE of the cascaded parameters/channel vs. training frames,  $N_h = N_v = 8$ , SNR = 30 dB,  $q = 1.5$ .

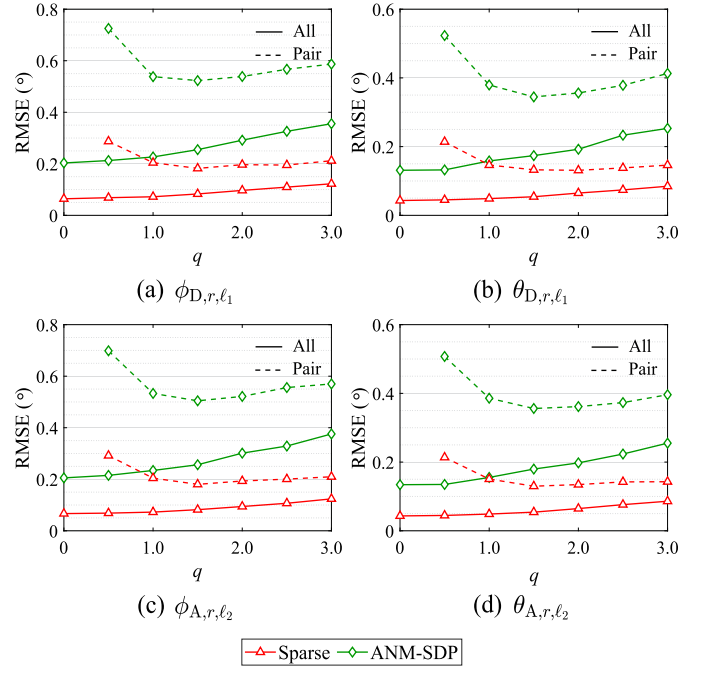


Fig. 9. RMSE of the decoupled angular parameters vs. IRS power radiation coefficient, SNR = 30 dB,  $\delta_h = \delta_v = 15^\circ$ .

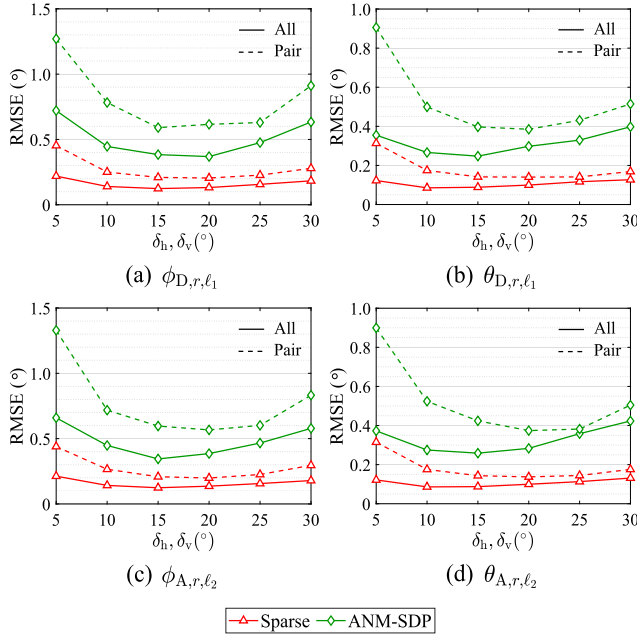


Fig. 8. RMSE of the decoupled angular parameters vs. IRS relative rotation angles, SNR = 30 dB,  $q = 3$ .

an appropriate value of  $q$ . In conclusion, the optimal values of  $\delta_h$ ,  $\delta_v$  and  $q$  are shown to be about  $15^\circ$  and  $1.5$ , respectively.

Figs. 10 and 11 plot the RMSEs of the environment mapping and user localization with the sparse scheme, where the proposed multi-carrier (MC) and single-carrier (SC) strategies are implemented with the aid of one and two twin-IRS structures, respectively.<sup>12</sup> The simulation results show that even with no

<sup>12</sup>The BS and UE are randomly located at  $\mathbf{p}_B = [2 \sim 6, -6 \sim -2, -2 \sim 2]^T$ ,  $\mathbf{p}_U = [2 \sim 6, 2 \sim 6, -2 \sim 2]^T$  with orientations  $\mathbf{n}_B = [\cos \eta_B, \sin \eta_B, 0]^T$ ,  $\mathbf{n}_U = [\cos \eta_U, \sin \eta_U, 0]^T$  with  $\eta_B = 30^\circ \sim 150^\circ$ ,  $\eta_U = -150^\circ \sim -30^\circ$ , respectively.

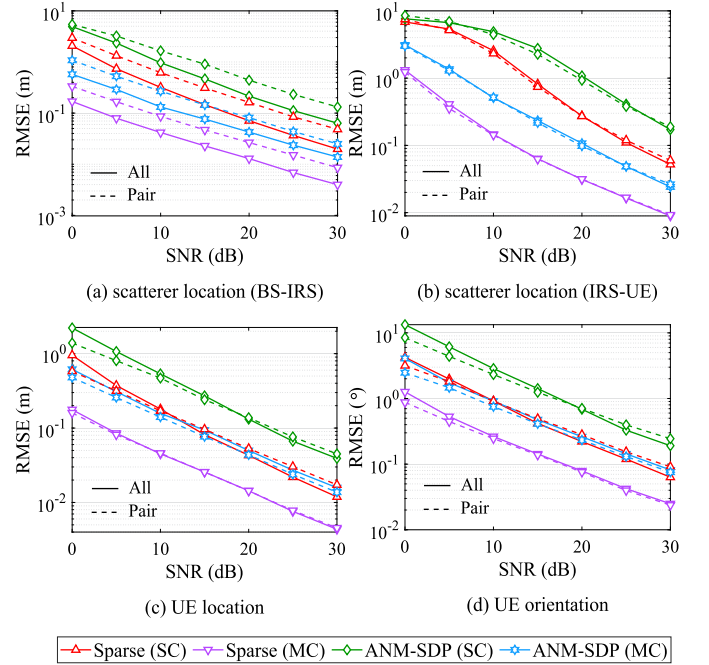


Fig. 10. RMSE of the environment mapping and user localization vs. SNR,  $K_{tr} = 6$ ,  $\delta_h = \delta_v = 15^\circ$ ,  $q = 1.5$ ,  $K_{I,U,r} = 20$  dB.

LoS path between the BS and IRSs, the proposed twin-IRS structure and the corresponding channel estimation schemes can achieve centimeter-level and degree-level resolution of the scatterer/user locations and orientations, respectively. The MC strategy consistently outperforms the SC counterpart, which benefits from more training subcarriers and less error accumulation of twin-IRS structures. More specifically, the “All” decoupling mode returns more accurate results of scatterer mapping, whilst the “Pair” counterpart generally performs

TABLE V  
COMPUTATIONAL COMPLEXITY OF CHANNEL ESTIMATION APPROACHES

Training IRS Pattern	Channel Estimation Method	Computational Complexity
Random	BALS [14]	$\mathcal{O}(M_{\text{tr}}T_{\text{tr}}P_{\text{tr}}N_{\text{I}})$ per iteration
Structured	SCPD [40]	$\mathcal{O}(M_{\text{tr}}T_{\text{tr}}P_{\text{I}}P_{\text{2}}N_{\text{I}})$
Grouping	LS + ESPRIT	$\mathcal{O}(M_{\text{tr}}T_{\text{tr}}P_{\text{tr}}G)$
Sparse	LS + ESPRIT	$\mathcal{O}(M_{\text{tr}}T_{\text{tr}}P_{\text{tr}}N_{\text{A}})$
Arbitrary	ANM-SDP/ADMM	$\mathcal{O}(L_{\text{BI},r}L_{\text{IU},r}N_{\text{I}}^{3.5})/\mathcal{O}(L_{\text{BI},r}L_{\text{IU},r}N_{\text{I}}^2)$ per iteration

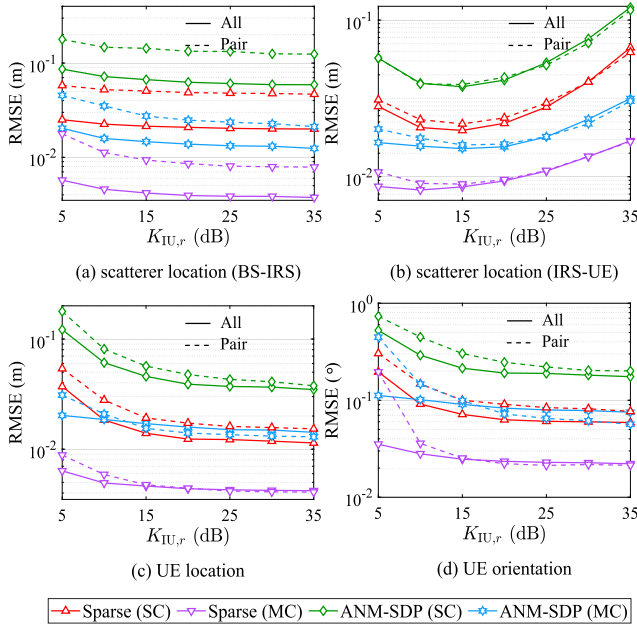


Fig. 11. RMSE of the environment mapping and user localization vs. Rician  $K$ -factor, SNR = 30 dB,  $K_{\text{tr}} = 6$ ,  $\delta_{\text{h}} = \delta_{\text{v}} = 15^\circ$ ,  $q = 1.5$ .

better at user positioning. The former can more robustly recover the parameters of multiple NLoS paths, while the latter is more suitable for solving the AoA/AoDs of a specific pair of paths involving the LoS component. Furthermore, one can observe that an increasing Rician  $K$ -factor leads to enhanced performance of environment mapping and user localization except that of IRS-UE scatterer positioning. This is due to the fact that higher power of a LoS path can improve its probability of being correctly detected, but will dominate the NLoS components and make them more likely to be treated as environmental noise.

Finally, we analyze the computational complexity of the involved channel estimation algorithms, as tabulated in Table V. Our analysis indicates that the complexity of the BALS and SCPD is proportional to the number of reflectors  $N_{\text{I}}$ . The algebraic grouping and sparse schemes efficiently reduce the complexity by activating fewer reflection unit cells and formulating smaller-scale equivalent arrays, respectively. Moreover, the optimization-based ANM-SDP/ADMM can handle arbitrary designs of the IRS training coefficients at the expense of relatively higher complexity.

## VII. CONCLUSION

We considered the channel estimation, as well as, the user localization problems of an IRS-assisted mmWave

MIMO-OFDM system. We proposed a novel twin-IRS structure consisting of two IRS planes with a relative spatial rotation to extract the 3-D propagation channel. By leveraging the techniques of tensor factorization, sparse array processing and atomic norm denoising, we presented four IRS training pattern designs and the corresponding parameter recovery schemes. By concatenating the cascaded phase/amplitude parameters recovered from the twin-IRS structure and leveraging the geometric relationship of devices, we achieved the decoupling of the exact channel angular and temporal parameters. We also proposed wideband and narrowband training strategies with single and multiple twin-IRS structures for the scatterer mapping and user positioning. Numerical results showed that the proposed strategy can precisely extract the channel parameters, and, therefore, can support a centimeter-level positioning resolution. In our future work, we are going to exploit the potential of conformal array topologies for the IRS designs alongside the corresponding channel estimation strategies and parameter recovery schemes.

## REFERENCES

- [1] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges," *Wireless Netw.*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015.
- [2] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, Jan. 2021.
- [3] M. Xiao *et al.*, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.
- [4] J. Zhang *et al.*, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.
- [5] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.
- [6] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Towards large intelligent surface (LIS)-based communications," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6568–6582, Oct. 2020.
- [7] W. Tang *et al.*, "Wireless communications with programmable meta-surface: Transceiver design and experimental results," *China Commun.*, vol. 16, no. 5, pp. 46–61, May 2019.
- [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] X. Yang, C.-K. Wen, and S. Jin, "MIMO detection for reconfigurable intelligent surface-assisted millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1777–1792, Aug. 2020.
- [10] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.
- [11] 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.

- [12] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8315–8329, Dec. 2020.
- [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, May 2020.
- [14] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.
- [15] J. Mirza and B. Ali, "Channel estimation method and phase shift design for reconfigurable intelligent surface assisted MIMO networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 2, pp. 441–451, Jun. 2021.
- [16] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, early access, Apr. 12, 2021, doi: [10.1109/TCOMM.2021.3072729](https://doi.org/10.1109/TCOMM.2021.3072729).
- [17] H. Liu, X. Yuan, and Y.-J.-A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2621–2636, Nov. 2020.
- [18] C. Huang *et al.*, "Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1663–1677, Jun. 2021.
- [19] S. Eddine Zegrar, L. Afeef, and H. Arslan, "A general framework for RIS-aided mmWave communication networks: Channel estimation and mobile user tracking," 2020, *arXiv:2009.01180*. [Online]. Available: <https://arxiv.org/abs/2009.01180>
- [20] W. Wang and W. Zhang, "Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Wireless Commun.*, early access, Apr. 21, 2021, doi: [10.1109/TWC.2021.3073140](https://doi.org/10.1109/TWC.2021.3073140).
- [21] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors—Part I: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, Jul. 2013.
- [22] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors—Part II: Uniqueness of the overall decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, Jul. 2013.
- [23] P. Pal and P. P. Vaidyanathan, "Nested arrays in two dimensions, Part I: Geometrical considerations," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4694–4705, Sep. 2012.
- [24] T. G. Kolda and B. W. Bader, "Tensor decomposition and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
- [25] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.
- [26] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.
- [27] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer. (Mar. 2016). *Tensorlab 3.0*. [Online]. Available: <https://www.tensorlab.net/>
- [28] J. Diebel, "Representing attitude: Euler angles, unit quaternions, and rotation vectors," *Matrix*, vol. 58, nos. 15–16, pp. 1–35, Oct. 2006.
- [29] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [30] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, Sep. 2020.
- [31] W. Tang *et al.*, "MIMO transmission through reconfigurable intelligent surface: System design, analysis, and implementation," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2683–2699, Nov. 2020.
- [32] M.-T. Martínez-Inglés, D. P. Gaillot, J. Pascual-García, J.-M. Molina-García-Pardo, M. Lienard, and J.-V. Rodríguez, "Deterministic and experimental indoor mmW channel modeling," *IEEE Antennas Wireless Propag. Lett.*, vol. 13, pp. 1047–1050, May 2014.
- [33] J. Yang *et al.*, "A geometry-based stochastic channel model for the millimeter-wave band in a 3GPP high-speed train scenario," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3853–3865, May 2018.
- [34] Z. Lin, T. Lv, and P. T. Mathiopoulos, "3-D indoor positioning for millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2472–2486, Jun. 2018.
- [35] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of  $N$ -way arrays," *J. Chemometrics*, vol. 14, no. 3, pp. 229–239, Jun. 2000.
- [36] I. Domanov and L. D. Lathauwer, "Canonical polyadic decomposition of third-order tensors: Reduction to generalized eigenvalue decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 2, pp. 636–660, Jan. 2014.
- [37] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
- [38] S. Geng, J. Kivinen, X. Zhao, and P. Vainikainen, "Millimeter-wave propagation channel characterization for short-range wireless communications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 3–13, Jan. 2009.
- [39] M. Sørensen and L. De Lathauwer, "Blind signal separation via tensor decomposition with Vandermonde factor: Canonical polyadic decomposition," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5507–5519, Aug. 2013.
- [40] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based channel estimation for millimeter wave MIMO-OFDM with dual-wideband effects," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4218–4232, Jul. 2020.
- [41] G. Favier and A. de Almeida, "Overview of constrained PARAFAC models," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–25, 2014.
- [42] A. Stegeman and A. L. F. de Almeida, "Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 3, pp. 1469–1490, Dec. 2009.
- [43] Z. Lin, T. Lv, W. Ni, J. A. Zhang, and R. P. Liu, "Nested hybrid cylindrical array design and DoA estimation for massive IoT networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 919–933, Apr. 2021.
- [44] F. Wen, H. C. So, and H. Wymeersch, "Tensor decomposition-based beamspace esprit algorithm for multidimensional harmonic retrieval," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 4572–4576.
- [45] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmWave MIMO systems via tensor-ESPRIT algorithms in DFT beamspace," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 816–831, Apr. 2021.
- [46] J. S. Dai, "Euler–Rodrigues formula variations, quaternion conjugation and intrinsic connections," *Mechanism Mach. Theory*, vol. 92, pp. 144–152, Oct. 2015.
- [47] D. M. Y. Sommerville, *Analytical Geometry of Three Dimensions*. Cambridge, U.K.: Cambridge Univ. Press, 1934.
- [48] A. Shahmansoori, G. E. Garcia, G. Destino, G. S.-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.
- [49] R. Mendrzik, H. Wymeersch, and G. Bauch, "Joint localization and mapping through millimeter wave MIMO in 5G systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
- [50] M. Grant and S. Boyd. (Jan. 2020). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.2*. [Online]. Available: <http://cvxr.com/cvx>



**Yuxing Lin** (Student Member, IEEE) received the B.S. degree in information engineering from Southeast University, Nanjing, China, in 2016, where he is currently pursuing the Ph.D. degree with the School of Information Science and Engineering. His research interests include massive MIMO, millimeter wave wireless communications, hybrid beamforming, channel estimation, beam training/tracking, tensor signal processing, and intelligent reconfigurable surface-assisted communications.



**Shi Jin** (Senior Member, IEEE) received the B.S. degree in communications engineering from Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in information and communications engineering from Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow with University College London, Adastral Park Research Campus, London, U.K. He is currently with the Faculty of the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He and his coauthors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society. He serves as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, and *IET Communications*.



**Michail Matthaiou** (Senior Member, IEEE) was born in Thessaloniki, Greece, in 1981. He received the Diploma degree (five years) in electrical and computer engineering from Aristotle University of Thessaloniki, Greece, in 2004, the M.Sc. degree (Hons.) in communication systems and signal processing from the University of Bristol, U.K., in 2005, and the Ph.D. degree from The University of Edinburgh, U.K., in 2008. From September 2008 to May 2010, he was with the Institute for Circuit Theory and Signal Processing, Technical University of Munich (TUM), Germany, as a Post-Doctoral Research Associate. He is currently a Professor of communications engineering and signal processing and the Deputy Director of the Centre for Wireless Innovation (CWI), Queen's University Belfast, U.K., after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, beyond massive MIMO, intelligent reflecting surfaces, mm-wave/THz systems, and deep learning for communications.

Dr. Matthaiou and his coauthors received the IEEE Communications Society (ComSoc) Leonard G. Abraham Prize in 2017. He currently holds the ERC Consolidator Grant BEATRICE (2021–2026) focused on the interface between information and electromagnetic theories. He was awarded the prestigious 2018/2019 Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship and also received the 2019 EURASIP Early Career Award. His team was also the Grand Winner of the 2019 Mobile World Congress Challenge. He was a recipient of the 2011 IEEE ComSoc Best Young Researcher Award for Europe, Middle East, and Africa Region, and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the Best M.Sc. Dissertation in the area of communications. He has coauthored papers that received best paper awards at the 2018 IEEE WCSP and 2014 IEEE ICC. He was an Exemplary Reviewer of IEEE COMMUNICATIONS LETTERS in 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He is currently the Editor-in-Chief of *Physical Communication* (Elsevier), a Senior Editor of IEEE WIRELESS COMMUNICATIONS LETTERS, and an Associate Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Series on Machine Learning for Communications and Networks.



**Xiaohu You** (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Nanjing Institute of Technology, Nanjing, China, in 1982, 1985, and 1989, respectively. From 1987 to 1989, he was a Lecturer with Nanjing Institute of Technology. Since 1990, he has been with Southeast University, first as an Associate Professor and then as a Professor. He contributed over 40 IEEE journal articles and two books in the areas of adaptive signal processing and neural networks and their applications to communication systems. He was the Premier Foundation Investigator of the China National Science Foundation. From 1999 to 2002, he was the Principal Expert of the C3G Project, responsible for organizing China's 3G mobile communications research and development activities. From 2001 to 2006, he was the Principal Expert of the National 863 FuTURE Project. His research interests include mobile communications, adaptive signal processing, and artificial neural networks with applications to communications and biomedical engineering. He received the Excellent Paper Award from China Institute of Communications in 1987 and the Elite Outstanding Young Teacher Award from Southeast University in 1990, 1991, and 1993. He is currently the Chairman of the IEEE Nanjing Section. In 2012, he was selected as an IEEE Fellow for his contributions to the development of mobile communications in China.