CHALMERS
UNIVERSITY OF TECHNOLOGY

# JrCUP: Joint RIS Calibration and User Positioning for 6G Wireless Systems

Downloaded from: https://research.chalmers.se, 2024-05-02 01:39 UTC

Citation for the original published paper (version of record):
Zheng, P., Chen, H., Ballal, T. et al (2023). JrCUP: Joint RIS Calibration and User Positioning for 6G Wireless Systems. IEEE Transactions on Wireless Communications, In press.
http://dx.doi.org/10.1109/TWC.2023.3336177

N.B. When citing this work, cite the original published paper.

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

This document was downloaded from http://research.chalmers.se, where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec, 8.1.9. (http://www.ieee.org/documents/opsmanual.pdf).

(article starts on next page)

# JrCUP: Joint RIS Calibration and User Positioning for 6G Wireless Systems

Pinjun Zheng, *Student Member, IEEE*, Hui Chen, *Member, IEEE*, Tarig Ballal, *Member, IEEE*, Mikko Valkama, *Fellow, IEEE*, Henk Wymeersch, *Fellow, IEEE*, and Tareq Y. Al-Naffouri, *Senior Member, IEEE*

***Abstract*—Reconfigurable intelligent surface (RIS)-assisted localization has attracted extensive attention as it can enable and enhance localization services in extreme scenarios. However, most existing works treat RISs as anchors with known positions and orientations, which is not realistic in applications with mobile or uncalibrated RISs. This work considers the *joint RIS calibration and user positioning* (JrCUP) problem with an active RIS. We propose a novel two-stage method to solve the considered JrCUP problem. The first stage comprises a tensor-estimation of signal parameters via rotational invariance techniques (tensor-ESPRIT), followed by a channel parameters refinement using least-squares. In the second stage, a two-dimensional search algorithm is proposed to estimate the three-dimensional user and RIS positions, one-dimensional RIS orientation, and clock bias from the estimated channel parameters. The Cramér-Rao lower bounds of the channel parameters and localization parameters are derived to verify the effectiveness of the proposed tensor-ESPRIT-based algorithms. In addition, simulation results reveal that the active RIS can significantly improve the localization performance compared to the passive case under the same system power supply in practical regions. Moreover, we observe the presence of blind areas with limited JrCUP localization performance, which can be mitigated by either leveraging more prior information or deploying extra base stations.**

***Index Terms*—5G/6G, reconfigurable intelligent surface, tensor-ESPRIT, RIS calibration, positioning.**

## I. Introduction

Reconfigurable intelligent surface (RIS) is an emerging technology for 5G/6G and beyond, which consists of an array of reflecting elements and offers distinctive characteristics that make the propagation environment controllable [1], [2]. With the flexibility to reshape wireless channels and the low cost of deployment, RIS has become one of the key enablers for future millimeter wave (mmWave) and terahertz (THz) band communication systems [3]–[5]. Over the years, different types of RISs have been proposed and widely studied, including passive RISs, active RISs, hybrid RISs, and simultaneous transmitting and receiving (STAR) RISs [6]–[8].

Pinjun Zheng, Tarig Ballal, and Tareq Y. Al-Naffouri are with the Division of Computer, Electrical and Mathematical Science & Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, KSA. (Email: {pinjun.zheng; tarig.ahmed; tareq.alnaffouri}@kaust.edu.sa). Hui Chen and Henk Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden (Email: {hui.chen; henkw}@chalmers.se). Mikko Valkama is with the Department of Electrical Engineering, Tampere University, Finland (Email: mikko.valkama@tuni.fi).

This publication is based upon the work supported in part by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. ORA-CRG2021-4695, in part by the EU H2020 RISE-6G project under grant 101017011, and in part by the Finnish Funding Agency for Innovation under the 6G-ISAC project.

Apart from the benefits to communication, the inclusion of RISs also opens new opportunities for radio localization. Radio location via wireless networks has been regarded as an indispensable function in advanced 5G/6G systems, which plays an increasingly important role in various applications. For example, radio and network localization systems in scenarios where the global positioning system (GPS) is insufficient or not available are well-studied [9]–[11]. A significant advantage that RISs offer in radio localization is reducing the number of base stations (BSs) required to perform localization. A RIS can not only act as a new synchronized location reference but also provide additional geometric measurements thanks to its high angular resolution. With RISs being introduced properly, it is possible to perform localization using a single BS [12] or even without any BSs [13]–[15] at all. Recent studies have shown the potential of RIS-assisted localization systems in various scenarios, e.g., localization under user mobility [16], simultaneous indoor and outdoor localization [17], received-signal-strength based localization [18], etc. Besides, the position and orientation estimation error bounds for the RIS-assisted localization are derived in [19]. A few RIS beamforming design optimization works can be found in [20]–[22].

Although promising results on RIS-assisted localization are shown in the literature, most of the existing works regard RIS as an anchor with known position and orientation, which is not realistic in some application scenarios involving RISs with calibration errors or mobile RISs. As a matter of fact, calibration errors in the RIS placement and geometric layout are unavoidable in practice, making RIS calibration a necessity for performing a high-precision localization. The results in [23] reveal that a minor calibration error on RIS geometry can cause a non-negligible model mismatch and result in performance degradation, especially in a high signal-to-noise ratio (SNR) scenarios. In this context, Bayesian analysis for a RIS-aided localization problem under RIS position and orientation offsets is carried out in [24], which discussed the possibility of correcting the RIS position and orientation under the near-field and far-field models. Recently, some efforts have also been directed towards the integration of unmanned aerial vehicless (UAVs) and RISs [25], [26], which extends the application scenarios where RISs locations vary with time. In such scenarios, the localization of the RIS becomes a newly introduced issue that needs to be tackled. As a consequence, localizing the RIS itself while localizing the user has become an increasingly important problem to solve today.

The joint RIS calibration and user positioning (JrCUP) problem refers to localizing the RIS and user simultaneously (with

or without a priori information about the RIS position and orientation). The three-dimensional (3D) JrCUP localization problem was first formulated in [27], which explored the relationship between the channel parameters and localization unknowns, with the corresponding Fisher information matrix derived and analyzed. Nonetheless, the adopted passive RIS in [27] limits the localization performance, and the design of an efficient channel estimator for JrCUP is still missing. In [28], a multi-stage solution for the two-dimensional (2D) JrCUP problem in a hybrid RIS-assisted system is reported. However, the hybrid RIS setup requires an extra central processing unit (CPU) for the receiver and RIS to share observations, which increases the system complexity. Recently, active RIS, which can simultaneously reflect and amplify the incident signals without requiring extra CPU, has shown the potential to provide better localization performance compared to the passive RIS, as it enhances the reflected signals thus avoiding the overwhelming dominance of the TX-RX line-of-sight (LOS) channel [13], [29], [30].

In this work, we extend the 3D JrCUP problem in [27] by developing an efficient channel estimation algorithm and utilizing active RISs to improve localization performance. The contributions of this work can be summarized as follows:

- We formulate the JrCUP problem in an uplink single-input-multiple-output (SIMO) scenario with an active RIS. Motivated by the mobile RIS with vertical attitude adjustment capability, we define the unknown localization parameters as the 3D position of user equipment (UE), the 3D position and 1D orientation of the RIS, and the clock bias between the UE and BS. By performing a localizability analysis, we show that this problem is solvable and that more dimensions of RIS orientation can be estimated by introducing more UEs.
- We propose a two-stage solution to the JrCUP problem. In the first stage, we perform a coarse channel estimation via tensor-Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) followed by a channel parameter refinement through a least-squares (LS)-based algorithm. In the second stage, a 2D-search-based algorithm is proposed to estimate the localization parameters.
- The fundamental Cramér-Rao lower bounds (CRLBs) for the JrCUP problem are derived, which consist of the CRLBs for the estimations of the channel parameters and localization parameters. We show that the received noise in the active RIS-involved system is colored and unknown which causes the proposed algorithm to fail to reach the theoretical bounds. However, the very minor gaps between the tested root mean square errors (RMSEs) and the derived CRLBs still show the effectiveness of our algorithms.
- Based on the derived CRLBs, we compare the active RIS and the passive RIS setups. It is shown that the active RIS can outperform the passive RIS within practical power supply regions. The localization performance of the active RIS setup is improved with the power supply increasing up to a certain level where the performance saturates. Furthermore, we show that blind areas exist in the JrCUP problem, which can be restrained by leveraging prior geometric information and/or deploying more BSs.

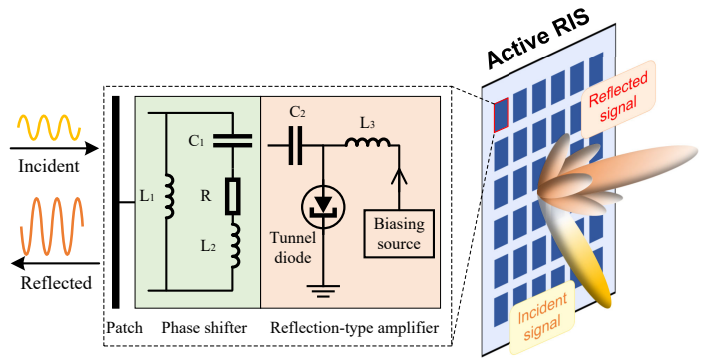



Fig. 1: Schematic diagram of the hardware architecture of a typical active RIS. Each RIS element integrates a phase shifter [31] and a reflection-type amplifier [32].

The remainder of this paper is organized as follows. Section II introduces the system model and formulates the JrCUP problem. Section III reviews the necessary mathematical preliminaries for the tensor decomposition and ESPRIT algorithms, based on which a two-stage solution for the JrCUP problem is proposed in Section IV. The CRLBs of the underlying estimation problems are derived in Section V. Simulation results are presented in Section VI followed by the conclusion of this work in Section VII.

Notations: Italic letters denote scalars (e.g., $a$), bold lower-case letters denote vectors (e.g., $\mathbf{a}$), bold upper-case letters denote matrices (e.g., $\mathbf{A}$), and bold calligraphic letters denote tensors (e.g., $\boldsymbol{\mathcal{A}}$). The notations $(\cdot)^{\mathsf{T}}$, $(\cdot)^{*}$, $(\cdot)^{\mathsf{H}}$, $(\cdot)^{-1}$, $(\cdot)^{\dagger}$, and $\mathrm{tr}(\cdot)$ are reserved for the transpose, conjugate, conjugate transpose, inverse, Moore-Penrose pseudo-inverse, and the matrix trace operations. The notation $\otimes$ denotes the Kronecker product, $\circ$ denotes the outer product, and $\odot$ denotes the Hadamard product. We use $[\mathbf{x}]_i$ to represent the $i$th entry of a vector $\mathbf{x}$, and $[\mathbf{X}]_{i,j}$ to represent the entry in the $i$th row and $j$th column of a matrix $\mathbf{X}$. The notations $\Re(\cdot)$ and $\Im(\cdot)$ denote the operations of taking the real and imaginary parts of a complex quantity, respectively.

## II. System Model

### A. Active RIS

Active RIS is an array of active elements that can scatter the incident signals with both amplification and tunable phases. Fig. 1 presents a schematic diagram of the hardware architecture of a typical active RIS, where the RIS element consists of a phase shifter and a reflection-type amplifier [6]. Many circuit implements have been proposed to achieve the function of phase shift and amplification. For example, the phase shift can be realized through a parallel resonant circuit [31], and the amplification can be realized by a tunnel diode circuit [32], as depicted in Fig. 1. Usually, the circuit network in an active RIS element can be modeled as a two-port component whose characteristics can be described by an S-parameter

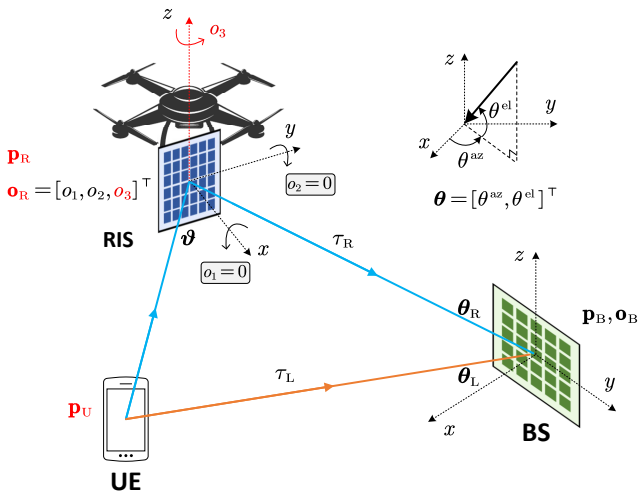



Fig. 2: Illustration of an uplink SIMO JrCUP scenario, where the states of the UE and RIS are unknown. The RIS is deployed on a drone with gravity sensors, where the RIS is always perpendicular to the ground; hence, only one orientation angle ($[\mathbf{o}_\mathrm{R}]_3$) needs to be estimated.

matrix [33].[1] By modulating the load impedance of each RIS element, the desired amplitude and phase change can be obtained (examples can be found in, e.g., [31], [32]).

Consider an active RIS of size $N_{\mathrm{R},1}\times N_{\mathrm{R},2}$. Assuming all the RIS elements keep the same amplification coefficient, the RIS profile for a transmission can be denoted as $\boldsymbol{\gamma}=p\tilde{\boldsymbol{\gamma}}$, where $p>1$ is the amplification coefficient and $\tilde{\boldsymbol{\gamma}}\in\mathbb{C}^{N_{\mathrm{R},1}N_{\mathrm{R},2}}$ is the phase-shift vector with unit-modulus entries. We define the incident signal power *per RIS element* (assumed identical across RIS elements) as $P_\mathrm{in}$. Then the required power supply of the active RIS can be calculated as [34]

$$P_\mathrm{R}=(p^2-1)N_{\mathrm{R},1}N_{\mathrm{R},2}(P_\mathrm{in}+\sigma_r^2), \tag{1}$$

where $\sigma_r^2$ denotes the power of the thermal noise introduced in each active RIS element. Note that (1) is an idealized model, and the actual energy consumption of an active RIS needs to further account for the RIS mutual coupling (MC) effect,[2] electronic circuit energy consumption, energy efficiency, etc.

### B. Geometry Model

This subsection describes the geometric relationship among the system devices. Here, we only focus on the LOS components of the UE-BS and the UE-RIS-BS paths which are used for localization, while the non-line-of-sight (NLOS) multipath will be discussed in Subsection II-D.

We consider an uplink SIMO wireless system consisting of a single-antenna UE, an $N_{\mathrm{R},1}\times N_{\mathrm{R},2}$-element active RIS, and an $N_{\mathrm{B},1}\times N_{\mathrm{B},2}$-element BS located at $\mathbf{p}_\mathrm{U}\in\mathbb{R}^3$, $\mathbf{p}_\mathrm{R}\in\mathbb{R}^3$, and $\mathbf{p}_\mathrm{B}\in\mathbb{R}^3$, respectively, as shown in Fig. 2. The orientations of the RIS and the BS are denoted as Euler angles $\mathbf{o}_\mathrm{R}\in\mathbb{R}^3$ and $\mathbf{o}_\mathrm{B}\in\mathbb{R}^3$. Note that the orientations can also be represented by the rotation matrices $\mathbf{R}_\mathrm{R}\in\mathrm{SO}(3)$ and $\mathbf{R}_\mathrm{B}\in\mathrm{SO}(3)$, which is constrained in the group of 3D rotations defined as $\mathrm{SO}(3)=\{\mathbf{R}|\mathbf{R}^\mathsf{T}\mathbf{R}=\mathbf{I}_3,\det(\mathbf{R})=1\}$. The mapping between a rotation matrix and its Euler angles can be found in [5]. In this work, the states of the UE and the RIS are the unknowns to be estimated, while the BS is used as the reference point in the considered coordinate system with known states. Specifically, we consider a scenario where the RIS is mobile and has vertical attitude adjustment capability, making a single degree of freedom in the RIS's orientation. For example, the RIS can be deployed on a drone with gravity sensors/accelerometers [26], [37]. Then we can assume the RIS plane to always be perpendicular to the ground (fixed pitch and yaw angles) and only a 1D orientation (roll angle) in the horizontal plane needs to be estimated, as depicted in Fig. 2. Hence, the unknown geometric parameters consist of the 3D position of UE, the 3D position, and the 1D orientation of RIS.

In the BS's local coordinate system (LCS), the angle-of-arrival (AOA) for the LOS path and the RIS reflection path are denoted as $\boldsymbol{\theta}_\mathrm{L}$ and $\boldsymbol{\theta}_\mathrm{R}$. Note that each AOA pair consists of an azimuth angle and an elevation angle, i.e., $\boldsymbol{\theta}_\mathrm{L}=[\theta_\mathrm{L}^\mathrm{az},\theta_\mathrm{L}^\mathrm{el}]^\mathsf{T}$ and $\boldsymbol{\theta}_\mathrm{R}=[\theta_\mathrm{R}^\mathrm{az},\theta_\mathrm{R}^\mathrm{el}]^\mathsf{T}$. Those angles are related to the geometric parameters as follows:

$$\theta_\mathrm{L}^\mathrm{az}=\arctan 2\left([\mathbf{R}_\mathrm{B}^\mathsf{T}(\mathbf{p}_\mathrm{U}-\mathbf{p}_\mathrm{B})]_2,[\mathbf{R}_\mathrm{B}^\mathsf{T}(\mathbf{p}_\mathrm{U}-\mathbf{p}_\mathrm{B})]_1\right), \tag{2}$$

$$\theta_\mathrm{L}^\mathrm{el}=\arcsin\left([\mathbf{R}_\mathrm{B}^\mathsf{T}(\mathbf{p}_\mathrm{U}-\mathbf{p}_\mathrm{B})]_3/\|\mathbf{p}_\mathrm{U}-\mathbf{p}_\mathrm{B}\|_2\right). \tag{3}$$

A similar relationship holds for $\theta_\mathrm{R}^\mathrm{az}$ and $\theta_\mathrm{R}^\mathrm{el}$. Analogously, in the RIS's LCS, the AOA from the UE can be denoted as $\boldsymbol{\phi}_\mathrm{A}=[\phi_\mathrm{A}^\mathrm{az},\phi_\mathrm{A}^\mathrm{el}]^\mathsf{T}$ and the angle-of-departure (AOD) towards the BS is $\boldsymbol{\phi}_\mathrm{D}=[\phi_\mathrm{D}^\mathrm{az},\phi_\mathrm{D}^\mathrm{el}]^\mathsf{T}$. Furthermore, we assume an unknown clock bias $\Delta\in\mathbb{R}$ exists between the UE and BS [5], [20]. Therefore, delays over the UE-BS and the UE-RIS-BS paths are given by

$$\tau_\mathrm{L}=\|\mathbf{p}_\mathrm{B}-\mathbf{p}_\mathrm{U}\|_2/c+\Delta, \tag{4}$$

$$\tau_\mathrm{R}=\|\mathbf{p}_\mathrm{R}-\mathbf{p}_\mathrm{U}\|_2/c+\|\mathbf{p}_\mathrm{B}-\mathbf{p}_\mathrm{R}\|_2/c+\Delta, \tag{5}$$

where $c$ is the speed of light.

### C. Signal Model

The BS observes and processes the uplink signals received through the UE-BS channel and the UE-RIS-BS channel simultaneously. We consider the transmission of $G$ orthogonal frequency-division multiplexing (OFDM) pilot symbols with $K$ subcarriers. The frequency of the $k$th subcarrier is denoted as $f_k=f_c+\frac{(2k-1-K)}{2}\Delta_f, k=1,\ldots,K$, where $f_c$ is the carrier frequency, $\Delta_f=B/K$ is the subcarrier spacing, and $B$ is the bandwidth. With reflection-type amplifiers supported by a power supply, the active RIS profile at the $g$th transmission is denoted by $\boldsymbol{\gamma}_g=[pe^{j\beta_{1,g}},\ldots,pe^{j\beta_{N_{\mathrm{R},1}N_{\mathrm{R},2},g}}]^\mathsf{T}\in\mathbb{C}^{N_{\mathrm{R},1}N_{\mathrm{R},2}}$, $g=1,\ldots,G$, where $p>1$ denotes the amplification factor and $\beta_{n,g}$ denotes the phase shift of the $n$th RIS element at the $g$th transmission. Again, here we assume that each active RIS element keeps the same amplification factor. The reflection matrix of the active RIS can be defined as $\boldsymbol{\Gamma}_g=\mathrm{diag}(\boldsymbol{\gamma}_g)$.

[1]The investigation into detailed circuit networks is beyond the scope of this work, so we simply adopt two cascading mathematical operations (i.e., phase shift and amplification) to model the function of the active RIS.

[2]The MC effect on RIS refers to coupling that arises between adjacent RIS elements, which causes a higher energy consumption and lower achievable data rates in wireless communication systems [35], [36]. The usage of active RIS can further accentuate this impact. The method development in this paper is based on a RIS MC-free scenario, while the impact of RIS MC will be examined in Subsection VI-B4.

Suppose that all the BS antennas are connected to an $N_1 \times N_2$ radio-frequency chain (RFC) array. We assume $N_1 > 1$ and $N_2 > 1$. The received baseband signal for the $g$th transmission and the $k$th subcarrier, $\mathbf{y}_{g,k} \in \mathbb{C}^{N_1 N_2 \times 1}$, can be expressed as

$$\mathbf{y}_{g,k} = \mathbf{W}^{\mathsf{H}} \left( \mathbf{h}_{\mathrm{L}}^k x_{g,k} + \mathbf{H}_{\mathrm{R},2}^k \boldsymbol{\Gamma}_g \mathbf{h}_{\mathrm{R},1}^k x_{g,k} + \mathbf{H}_{\mathrm{R},2}^k \boldsymbol{\Gamma}_g \mathbf{n}_r + \mathbf{n}_0 \right), \tag{6}$$

where $x_{g,k} \in \mathbb{C}$ is the transmitted signal with average transmission power $|x_{g,k}|^2 = P_{\mathrm{T}}$, $\mathbf{h}_{\mathrm{L}}^k$ is the UE-BS channel vector, $\mathbf{h}_{\mathrm{R},1}^k$ is the UE-RIS channel vector, $\mathbf{H}_{\mathrm{R},2}^k$ is the RIS-BS channel matrix, $\mathbf{n}_r \sim \mathcal{CN}(\mathbf{0}, \sigma_r^2 \mathbf{I})$ denotes the thermal noise introduced in the active RIS, $\mathbf{n}_0 \sim \mathcal{CN}(\mathbf{0}, \sigma_0^2 \mathbf{I})$ denotes the thermal noise at the receiver, and $\mathbf{W} \in \mathbb{C}^{N_{\mathrm{B},1} N_{\mathrm{B},2} \times N_1 N_2}$ is the combiner matrix which will be specified in Subsection IV-A. Note that model (6) is reduced to the passive RIS case when $p = 1$ and $\sigma_r = 0$.

According to (6), we have the total received noise for the $g$th transmission and the $k$th subcarrier as

$$\mathbf{n}_{g,k} = \mathbf{W}^{\mathsf{H}} (\mathbf{H}_{\mathrm{R},2}^k \boldsymbol{\Gamma}_g \mathbf{n}_r + \mathbf{n}_0). \tag{7}$$

Here we notice the received noise $\mathbf{n}_{g,k}$ is colored and related to the unknown RIS-BS channel $\mathbf{H}_{\mathrm{R},2}^k$. The statistics of $\mathbf{n}_{g,k}$ will be derived in Subsection V-A. The channel model (6) can be rewritten as

$$\mathbf{y}_{g,k} = \mathbf{W}^{\mathsf{H}} \mathbf{h}_{g,k} x_{g,k} + \mathbf{n}_{g,k}, \tag{8}$$

where $\mathbf{h}_{g,k} = \mathbf{h}_{\mathrm{L}}^k + \mathbf{H}_{\mathrm{R},2}^k \boldsymbol{\Gamma}_g \mathbf{h}_{\mathrm{R},1}^k$.

### D. Channel Model

By introducing the NLOS multipath, the UE-BS channel $\mathbf{h}_{\mathrm{L}}^k$, the UE-RIS channel $\mathbf{h}_{\mathrm{R},1}^k$, and the RIS-BS channel $\mathbf{H}_{\mathrm{R},2}^k$ are given by [5]

$$\mathbf{h}_{\mathrm{L}}^k = \sum_{i=0}^{I_{\mathrm{L}}} \alpha_{\mathrm{L}}^i e^{-j2\pi(k-1)\Delta_f \tau_{\mathrm{L}}^i} \mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{L}}^i), \tag{9}$$

$$\mathbf{h}_{\mathrm{R},1}^k = \sum_{i=0}^{I_{\mathrm{R},1}} \alpha_{\mathrm{R},1}^i \mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{A}}^i), \tag{10}$$

$$\mathbf{H}_{\mathrm{R},2}^k = \sum_{i=0}^{I_{\mathrm{R},2}} \alpha_{\mathrm{R},2}^i e^{-j2\pi(k-1)\Delta_f \tau_{\mathrm{R}}^i} \mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{R}}^i) \mathbf{a}_{\mathrm{R}}^{\mathsf{T}}(\boldsymbol{\phi}_{\mathrm{D}}^i), \tag{11}$$

where $\mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}) \in \mathbb{C}^{N_{\mathrm{B},1} N_{\mathrm{B},2}}$ and $\mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}) \in \mathbb{C}^{N_{\mathrm{R},1} N_{\mathrm{R},2}}$ denote the array response vectors of the BS and the RIS, and $\alpha_{\mathrm{L}}^i$, $\alpha_{\mathrm{R},1}^i$ and $\alpha_{\mathrm{R},2}^i$ are the complex channel gains for the UE-BS, the UE-RIS, and the RIS-BS channels, respectively. Here, $i = 0$ represents the LOS channel and the rest are the NLOS multipath channels reflected by scattering points (SPs). The numbers of the NLOS paths of the UE-BS, the UE-RIS and the RIS-BS channels are denoted as $I_{\mathrm{L}}$, $I_{\mathrm{R},1}$, and $I_{\mathrm{R},2}$.

In this work, the multipath components are not used to perform localization, thus, it acts as a negative effect that generates additional noise. Given that, the algorithm development in this paper is based on the multipath-free model; however, the impact of the multipath effect will be evaluated in Section VI. Ignoring the multipath, the channel models (9)–(11) are reduced to

$$\mathbf{h}_{\mathrm{L}}^k = \alpha_{\mathrm{L}} e^{-j2\pi(k-1)\Delta_f \tau_{\mathrm{L}}} \mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{L}}), \tag{12}$$

$$\mathbf{h}_{\mathrm{R},1}^k = \alpha_{\mathrm{R},1} \mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{A}}), \tag{13}$$

$$\mathbf{H}_{\mathrm{R},2}^k = \alpha_{\mathrm{R},2} e^{-j2\pi(k-1)\Delta_f \tau_{\mathrm{R}}} \mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{R}}) \mathbf{a}_{\mathrm{R}}^{\mathsf{T}}(\boldsymbol{\phi}_{\mathrm{D}}), \tag{14}$$

where $\alpha_{\mathrm{L}}$, $\alpha_{\mathrm{R},1}$ and $\alpha_{\mathrm{R},2}$ are the complex channel gains for the corresponding LOS channels. The array response vectors of the BS and the RIS are defined as

$$[\mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta})]_i = e^{j\frac{2\pi f_c}{c} \mathbf{t}(\boldsymbol{\theta})^{\mathsf{T}} \mathbf{p}_{\mathrm{B},i}}, \tag{15}$$

$$[\mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi})]_i = e^{j\frac{2\pi f_c}{c} \mathbf{t}(\boldsymbol{\phi})^{\mathsf{T}} \mathbf{p}_{\mathrm{R},i}}, \tag{16}$$

where $\mathbf{p}_{\mathrm{B},i}$ and $\mathbf{p}_{\mathrm{R},i}$ are respectively the positions of the $i$th element of the BS and the RIS given in their LCS, and $\mathbf{t}(\boldsymbol{\theta})$ is the direction vector defined as $\mathbf{t}(\boldsymbol{\theta}) \triangleq \left[\cos(\theta^{\mathrm{az}})\cos(\theta^{\mathrm{el}}), \sin(\theta^{\mathrm{az}})\cos(\theta^{\mathrm{el}}), \sin(\theta^{\mathrm{el}})\right]^{\mathsf{T}}$. For later derivation, we assume both the antenna arrays in the BS and the RIS to be uniform planar arrays (UPAs) and their element spacings are denoted as $d_{\mathrm{B}}$ and $d_{\mathrm{R}}$.

We assume that all the BS and RIS elements are deployed on the YOZ plane of their LCSs, i.e., the $x$-coordinates of these elements' positions are zeros. Since the first entry of the $\mathbf{p}_{\mathrm{R},i}$ is zero, we introduce the intermediate RIS-related angles $\boldsymbol{\vartheta} = [\vartheta_2, \vartheta_3]^{\mathsf{T}}$ as [27]

$$\vartheta_2 = \sin(\phi_{\mathrm{A}}^{\mathrm{az}})\cos(\phi_{\mathrm{A}}^{\mathrm{el}}) + \sin(\phi_{\mathrm{D}}^{\mathrm{az}})\cos(\phi_{\mathrm{D}}^{\mathrm{el}}), \tag{17}$$

$$\vartheta_3 = \sin(\phi_{\mathrm{A}}^{\mathrm{el}}) + \sin(\phi_{\mathrm{D}}^{\mathrm{el}}). \tag{18}$$

Then, the total RIS array response for both signal arrival and departure can be represented through $\vartheta_2$ and $\vartheta_3$ as $[\mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{A}}) \odot \mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{D}})]_i = e^{j\frac{2\pi f_c}{c}(\mathbf{t}(\boldsymbol{\phi}_{\mathrm{A}}) + \mathbf{t}(\boldsymbol{\phi}_{\mathrm{D}}))^{\mathsf{T}} \mathbf{p}_{\mathrm{R},i}} = e^{j\frac{2\pi f_c}{c}(\vartheta_2 [\mathbf{p}_{\mathrm{R},i}]_2 + \vartheta_3 [\mathbf{p}_{\mathrm{R},i}]_3)}$. Thus, a more compact formulation for the UE-RIS-BS channel can be represented as

$$\mathbf{H}_{\mathrm{R},2}^k \boldsymbol{\Gamma}_g \mathbf{h}_{\mathrm{R},1}^k = \alpha_{\mathrm{R}} \left[ (\mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{A}}) \odot \mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{D}}))^{\mathsf{T}} \boldsymbol{\gamma}_g \right] \times e^{-j2\pi(k-1)\Delta_f \tau_{\mathrm{R}}} \mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{R}}), \tag{19}$$

where $\alpha_{\mathrm{R}} = \alpha_{\mathrm{R},1} \alpha_{\mathrm{R},2}$.

### E. Localizability Analysis and Problem Formulation

This work aims to jointly estimate the RIS position and orientation, the UE position, and the clock bias based on the received signals. We adopt a two-stage estimation framework consisting of a channel parameters estimation followed by a RIS and UE states estimation from the obtained channel parameters. Based on (12)–(14), we define the vector $\boldsymbol{\eta}_{\mathrm{ch}}$ of *all* unknown channel parameters and the vector $\boldsymbol{\eta}$ that contains only localization-related channel parameters as

$$\boldsymbol{\eta}_{\mathrm{ch}} \triangleq [\boldsymbol{\eta}^{\mathsf{T}}, \Re(\alpha_{\mathrm{L}}), \Im(\alpha_{\mathrm{L}}), \Re(\alpha_{\mathrm{R}}), \Im(\alpha_{\mathrm{R}})]^{\mathsf{T}} \in \mathbb{R}^{12}, \tag{20}$$

$$\boldsymbol{\eta} \triangleq [\theta_{\mathrm{L}}^{\mathrm{az}}, \theta_{\mathrm{L}}^{\mathrm{el}}, \theta_{\mathrm{R}}^{\mathrm{az}}, \theta_{\mathrm{R}}^{\mathrm{el}}, \tau_{\mathrm{L}}, \tau_{\mathrm{R}}, \vartheta_2, \vartheta_3]^{\mathsf{T}} \in \mathbb{R}^8. \tag{21}$$

Note that the parameters $\alpha_{\mathrm{L}}$ and $\alpha_{\mathrm{R}}$ that contribute to $\boldsymbol{\eta}_{\mathrm{ch}}$ are nuisance parameters that will not be used for solving the JrCUP problem. As an objective of the general JrCUP

problem, the localization parameter vector that contains the RIS and UE states is defined as

$$\boldsymbol{\xi}_{\text{Loc}} \triangleq [\mathbf{p}_{\text{U}}^{\mathsf{T}}, \mathbf{p}_{\text{R}}^{\mathsf{T}}, \mathbf{o}_{\text{R}}^{\mathsf{T}}, \Delta]^{\mathsf{T}} \in \mathbb{R}^{10}. \tag{22}$$

The JrCUP problem refers to using estimates of the channel parameters $\boldsymbol{\eta}$ to determine the state vector $\boldsymbol{\xi}_{\text{Loc}}$. However, in an estimation problem, the number of unknowns cannot exceed the number of observations $\boldsymbol{\eta}$. In the general case, it is unlikely to find a unique value of $\boldsymbol{\xi}_{\text{Loc}}$ based solely on the observations $\boldsymbol{\eta}$. The restriction of the RIS orientation as highlighted in Subsection II-B reduces the unknowns pertaining to the RIS orientation to a single parameter, making the number of observations equal to the number of unknowns.[3] Consequently, the localization parameters to be estimated in this work can be redefined as

$$\boldsymbol{\xi} \triangleq [\mathbf{p}_{\text{U}}^{\mathsf{T}}, \mathbf{p}_{\text{R}}^{\mathsf{T}}, o_3, \Delta]^{\mathsf{T}} \in \mathbb{R}^{8}, \tag{23}$$

where $o_3 = [\mathbf{o}_{\text{R}}]_3$ is the Euler angles of the RIS orientation around the $Z$-axis; the rest of Euler angles (i.e., $o_1$ and $o_2$) are assumed to be known.

In this work, we focus on developing a solution to the JrCUP problem with a single UE and 1D RIS orientation. When the BS received the OFDM symbols from the UE through both LOS and RIS reflected channels, we first estimate $\boldsymbol{\eta}$ based on the received signals $\mathbf{y}_{g,k}, k = 1, \ldots, K, g = 1, \ldots, G$. Afterwards, $\boldsymbol{\xi}$ is estimated based on $\boldsymbol{\eta}$.

## III. Mathematical Preliminaries

To make the paper self-contained, we provide a brief review of the canonical polyadic (CP) decomposition of tensors and the fundamentals of the ESPRIT method. More details on these topics can be found in [38]–[41].

### A. Tensors & CP Decomposition

A tensor, also known as *multi-way array*, is a generalization of data arrays to three or higher dimensions. Let $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ denote an $N$th-order tensor. The *order* $N$ indicates the number of dimensions, and each dimension is called a *mode*.

The polyadic decomposition approximates a tensor with a sum of $R$ rank-one tensors. If the number of rank-one terms $R$ is minimum, the corresponding decomposition is called a CP decomposition and the minimum achievable $R$ is referred to as the *rank* of the tensor. Suppose $\boldsymbol{\mathcal{T}}$ is a rank-$R$ tensor, the CP decomposition decomposes $\boldsymbol{\mathcal{T}}$ as

$$\boldsymbol{\mathcal{T}} = \sum_{r=1}^{R} \lambda_r \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \cdots \circ \mathbf{u}_r^{(N)}. \tag{24}$$

Here, $[\mathbf{u}_1^{(n)}, \ldots, \mathbf{u}_R^{(n)}] = \mathbf{U}^{(n)} \in \mathbb{C}^{I_n \times R}$ is the factor matrix along the $n$th mode. Each matrix $\mathbf{U}^{(n)}$ has orthonormal columns. A visual representation of (24) in the third-order case ($N = 3$) is shown in Fig. 3.

[3]In cases where the full 3D orientation of the RIS needs to be estimated, multiple UEs can be utilized to obtain more channel parameters. For example, two UEs at different locations can provide 16 localization-related channel parameters, and the dimension of the unknowns $\boldsymbol{\xi}_{\text{Loc}}$ becomes 14 (with one more UE position and clock bias), making the problem solvable.

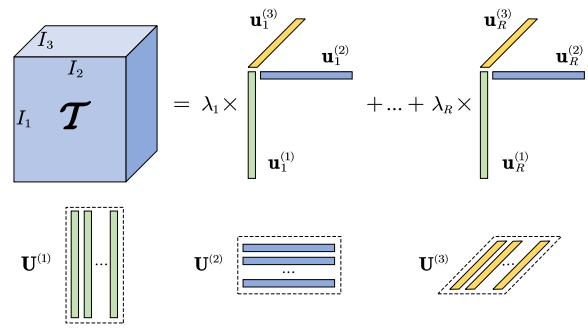



Fig. 3: A CP decomposition of a rank-$R$ third-order tensor.

### B. Tensor-ESPRIT

ESPRIT is a search-free signal subspace-based parameter estimation technique for multidimensional harmonic retrieval, which has been widely used in channel estimation [42], spectrum sensing [43], sub-Nyquist sampling [44], etc. This subsection recaps two ESPRIT variants, namely, element-space and beamspace tensor ESPRIT.

**Element-Space Tensor ESPRIT:** Consider a multidimensional harmonic retrieval problem with an observation tensor $\boldsymbol{\mathcal{T}}$ whose elements are given by

$$[\boldsymbol{\mathcal{T}}]_{i_1, i_2, \ldots, i_N} = \sum_{r=1}^{R} \lambda_r \prod_{n=1}^{N} e^{j(i_n - 1)\omega_{n,r}}. \tag{25}$$

To estimate the unknown angular frequencies $\omega_{n,r}$, we define

$$\mathbf{a}^{(I_n)}(\omega_{n,r}) = [1, e^{j\omega_{n,r}}, \ldots, e^{j(I_n-1)\omega_{n,r}}]^{\mathsf{T}} \in \mathbb{C}^{I_n}, \tag{26}$$

$$\mathbf{A}_n = \left[\mathbf{a}^{(I_n)}(\omega_{n,1}), \ldots, \mathbf{a}^{(I_n)}(\omega_{n,R})\right] \in \mathbb{C}^{I_n \times R}. \tag{27}$$

Then (25) can be rewritten as

$$\boldsymbol{\mathcal{T}} = \sum_{r=1}^{R} \lambda_r \mathbf{a}^{(I_1)}(\omega_{1,r}) \circ \mathbf{a}^{(I_2)}(\omega_{2,r}) \circ \cdots \circ \mathbf{a}^{(I_N)}(\omega_{N,r}). \tag{28}$$

We define two selection matrices $\mathbf{J}_{n,1} = \left[\mathbf{I}_{I_n-1}, \mathbf{0}_{(I_n-1)\times 1}\right] \in \mathbb{R}^{(I_n-1)\times I_n}$, $\mathbf{J}_{n,2} = \left[\mathbf{0}_{(I_n-1)\times 1}, \mathbf{I}_{I_n-1}\right] \in \mathbb{R}^{(I_n-1)\times I_n}$, and let $\boldsymbol{\Phi}_n = \text{diag}\left\{[e^{j\omega_{n,1}}, \ldots, e^{j\omega_{n,R}}]^{\mathsf{T}}\right\} \in \mathbb{C}^{R\times R}$. To estimate the angular frequency, ESPRIT relies on the shift-invariance property of $\mathbf{A}_n$, which is given by

$$\mathbf{J}_{n,1}\mathbf{A}_n\boldsymbol{\Phi}_n = \mathbf{J}_{n,2}\mathbf{A}_n. \tag{29}$$

Then the associated frequencies $\omega_{n,r}, r = 1, \ldots, R, n = 1, \ldots, N$, can be estimated as follows. By applying the CP decomposition to $\boldsymbol{\mathcal{T}}$, we obtain the factor matrices $\mathbf{U}^{(n)}, n = 1, \ldots, N$. Since $\mathbf{a}^{(I_n)}(\omega_{n,r})$ lies in the column space of $\mathbf{U}^{(n)}$, we replace $\mathbf{A}_n$ by $\mathbf{A}_n = \mathbf{U}^{(n)}\mathbf{D}_n$, where $\mathbf{D}_n \in \mathbb{C}^{R\times R}$ is a non-singular matrix. Then (29) becomes

$$\mathbf{J}_{n,1}\mathbf{U}^{(n)}\boldsymbol{\Theta}_n = \mathbf{J}_{n,2}\mathbf{U}^{(n)}, \tag{30}$$

where $\boldsymbol{\Theta}_n = \mathbf{D}_n\boldsymbol{\Phi}_n\mathbf{D}_n^{-1}$. Based on (30), the LS estimate of $\boldsymbol{\Theta}_n$ reads [40]

$$\hat{\boldsymbol{\Theta}}_n = \left(\mathbf{J}_{n,1}\mathbf{U}^{(n)}\right)^{\dagger}\mathbf{J}_{n,2}\mathbf{U}^{(n)}. \tag{31}$$

Since the $r$th eigenvalue of $\boldsymbol{\Theta}_n$ is given by $e^{j\omega_{n,r}}$, the angular frequencies $\hat{\omega}_{n,r}, r = 1, \ldots, R$, can be obtained from the phase components of the eigenvalues of $\hat{\boldsymbol{\Theta}}_n$.

**Beamspace Tensor ESPRIT:** For beamspace measurements, after the $n$th-mode product of $\boldsymbol{\mathcal{T}}$ with a linear transformation matrix [45], the model (28) is modified to

$$\boldsymbol{\mathcal{T}}_b = \sum_{r=1}^{R} \lambda_r \mathbf{b}^{(J_1)}(\omega_{1,r}) \circ \mathbf{b}^{(J_2)}(\omega_{2,r}) \circ \cdots \circ \mathbf{b}^{(J_N)}(\omega_{N,r}), \tag{32}$$

where $\mathbf{b}^{(J_n)}(\omega_{n,r}) = \mathbf{T}_n^{\mathsf{H}} \mathbf{a}^{(I_n)}(\omega_{n,r})$ with $\mathbf{T}_n \in \mathbb{C}^{I_n \times J_n}$. The beamspace array manifold becomes $\mathbf{B}_n = \left[\mathbf{b}^{(J_n)}(\omega_{n,1}), \ldots, \mathbf{b}^{(J_n)}(\omega_{n,R})\right] \in \mathbb{C}^{J_n \times R}$ and $\boldsymbol{\mathcal{T}}_b \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_n}$.

In the beamspace model with the transformation matrix $\mathbf{T}_n$, generally the shift-invariance property (29) does not hold, i.e., $\mathbf{J}_{n,1}\mathbf{B}_n\boldsymbol{\Phi}_n \neq \mathbf{J}_{n,2}\mathbf{B}_n$. However, if $\mathbf{T}_n$ has a proper structure, the following proposition [46], [47] shows that the lost shift-invariance property can be restored.

**Proposition 1.** *Assume that $\mathbf{T}_n$ has a shift-invariant structure that satisfies $\mathbf{J}_{n,1}\mathbf{T}_n = \mathbf{J}_{n,2}\mathbf{T}_n\mathbf{F}_n$ where $\mathbf{F}_n \in \mathbb{C}^{J_n \times J_n}$ is a non-singular matrix and $\mathbf{T}_n^{\mathsf{H}} = [\mathbf{t}_1, \ldots, \mathbf{t}_{I_n}]^{\mathsf{T}}$. If there exists a matrix $\mathbf{Q}_n \in \mathbb{C}^{J_n \times J_n}$ such that*

$$\begin{cases} \mathbf{Q}_n \mathbf{t}_{I_n} = \mathbf{0}_{J_n \times 1}, \\ \mathbf{Q}_n \mathbf{F}_n^{\mathsf{H}} \mathbf{t}_1 = \mathbf{0}_{J_n \times 1}, \end{cases} \tag{33}$$

*then we have*

$$\mathbf{Q}_n \mathbf{B}_n \boldsymbol{\Phi}_n^{\mathsf{H}} = \mathbf{Q}_n \mathbf{F}_n^{\mathsf{H}} \mathbf{B}_n, \tag{34}$$

*where the shift-invariance property is restored.*

*Proof.* See Appendix A in [48]. □

The matrices $\mathbf{Q}_n$ and $\mathbf{F}_n$ can be estimated as [46]

$$\hat{\mathbf{Q}}_n = \mathbf{I}_{J_n} - \mathbf{t}_{I_n}\mathbf{t}_{I_n}^{\mathsf{H}} - \left(\mathbf{F}_n^{\mathsf{H}}\mathbf{t}_1\right)\left(\mathbf{F}_n^{\mathsf{H}}\mathbf{t}_1\right)^{\mathsf{H}}, \tag{35}$$

$$\hat{\mathbf{F}}_n = \left(\mathbf{J}_{n,2}\mathbf{T}_n\right)^{\dagger} \mathbf{J}_{n,1}\mathbf{T}_n. \tag{36}$$

Then, similarly, after obtaining $\mathbf{U}^{(n)}$ by applying the CP decomposition, we can replace $\mathbf{B}_n$ by $\mathbf{B}_n = \mathbf{U}^{(n)}\mathbf{D}_n$, thus (34) becomes $\mathbf{Q}_n\mathbf{U}^{(n)}\boldsymbol{\Theta}_n = \mathbf{Q}_n\mathbf{F}_n^{\mathsf{H}}\mathbf{U}^{(n)}$. Therefore,

$$\hat{\boldsymbol{\Theta}}_n = \left(\mathbf{Q}_n\mathbf{U}^{(n)}\right)^{\dagger}\mathbf{Q}_n\mathbf{F}_n^{\mathsf{H}}\mathbf{U}^{(n)}. \tag{37}$$

Finally, the unknown angular frequencies are obtained from the phase of the eigenvalues of $\hat{\boldsymbol{\Theta}}_n$.

## IV. Proposed Localization Method

This section develops a two-stage localization method to solve the JrCUP problem formulated in Subsection II-E based on the preliminaries in Section III.

### A. Coarse Channel Estimation via Tensor-ESPRIT

Given $\mathbf{y}_{g,k}$ in (10), we can obtain the beamspace channel estimates by multiplying the received signals by the conjugate of the pilot symbols and dividing them by the average power as

$$\hat{\mathbf{h}}_{g,k}^{(\mathrm{b})} = \frac{\mathbf{y}_{g,k}x_{g,k}^*}{P_{\mathrm{T}}} = \mathbf{W}^{\mathsf{H}}\mathbf{h}_{g,k} + \Delta\mathbf{h}_{g,k}^{(\mathrm{b})}, \tag{38}$$

where $\Delta\mathbf{h}_{g,k}^{(\mathrm{b})}$ is the noise term. We organize the beamspace channels as

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{h}}_{1,1}^{(\mathrm{b})} & \cdots & \hat{\mathbf{h}}_{G,1}^{(\mathrm{b})} \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{h}}_{1,K}^{(\mathrm{b})} & \cdots & \hat{\mathbf{h}}_{G,K}^{(\mathrm{b})} \end{bmatrix} = [\hat{\mathbf{h}}_1^{(\mathrm{b})}, \ldots, \hat{\mathbf{h}}_G^{(\mathrm{b})}] \in \mathbb{C}^{KN_1N_2 \times G}. \tag{39}$$

The goal of this subsection is to estimate the localization-related channel parameters in $\boldsymbol{\eta}$ defined in (21) given $\hat{\mathbf{H}}$.

To apply the tensor-ESPRIT method, we design the combiner matrix $\mathbf{W}$ and the total RIS profile matrix $\boldsymbol{\Gamma}$ to follow the structure

$$\mathbf{W} = \mathbf{T}_1 \otimes \mathbf{T}_2 \in \mathbb{C}^{N_{\mathrm{B},1}N_{\mathrm{B},2} \times N_1N_2}, \tag{40}$$

$$\boldsymbol{\Upsilon} = [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_G] = \mathbf{T}_3 \otimes \mathbf{T}_4 \in \mathbb{C}^{N_{\mathrm{R},1}N_{\mathrm{R},2} \times G}, \tag{41}$$

where $\mathbf{T}_1 \in \mathbb{C}^{N_{\mathrm{B},1} \times N_1}$, $\mathbf{T}_2 \in \mathbb{C}^{N_{\mathrm{B},2} \times N_2}$, $\mathbf{T}_3 \in \mathbb{C}^{N_{\mathrm{R},1} \times \sqrt{G}}$, and $\mathbf{T}_4 \in \mathbb{C}^{N_{\mathrm{R},2} \times \sqrt{G}}$ with $G$ thus set to a square number. Note that we utilize a fixed combiner $\mathbf{W}$ over different transmissions while the RIS profile $\boldsymbol{\gamma}_g$ changes with $g$. Considering the structure of the array response vectors in (15), and according to (26), we can write

$$\mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{L}}) = \mathbf{a}^{(N_{\mathrm{B},1})}(\omega_{\mathrm{L},1}) \otimes \mathbf{a}^{(N_{\mathrm{B},2})}(\omega_{\mathrm{L},2}), \tag{42}$$

$$\mathbf{a}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{R}}) = \mathbf{a}^{(N_{\mathrm{B},1})}(\omega_{\mathrm{R},1}) \otimes \mathbf{a}^{(N_{\mathrm{B},2})}(\omega_{\mathrm{R},2}), \tag{43}$$

where

$$\omega_{\mathrm{L},1} = (2\pi f_c d_{\mathrm{B}}/c)\sin(\theta_{\mathrm{L}}^{\mathrm{az}})\cos(\theta_{\mathrm{L}}^{\mathrm{el}}), \tag{44}$$

$$\omega_{\mathrm{L},2} = (2\pi f_c d_{\mathrm{B}}/c)\sin(\theta_{\mathrm{L}}^{\mathrm{el}}), \tag{45}$$

$$\omega_{\mathrm{R},1} = (2\pi f_c d_{\mathrm{B}}/c)\sin(\theta_{\mathrm{R}}^{\mathrm{az}})\cos(\theta_{\mathrm{R}}^{\mathrm{el}}), \tag{46}$$

$$\omega_{\mathrm{R},2} = (2\pi f_c d_{\mathrm{B}}/c)\sin(\theta_{\mathrm{R}}^{\mathrm{el}}). \tag{47}$$

Furthermore, we define

$$\omega_{\tau_{\mathrm{L}}} = -2\pi\Delta_f\tau_{\mathrm{L}}, \quad \omega_{\tau_{\mathrm{R}}} = -2\pi\Delta_f\tau_{\mathrm{R}}. \tag{48}$$

*1) Estimating $\theta_L^{az}, \theta_L^{el}, \theta_R^{az}, \theta_R^{el}, \tau_L, \tau_R$:* We first use the sum of the channels over $g = 1, \ldots, G$, which is $\hat{\mathbf{h}}^{(\mathrm{b})} = \sum_{g=1}^{G}\hat{\mathbf{h}}_g^{(\mathrm{b})} \in \mathbb{C}^{KN_1N_2}$. Suppose $\hat{\mathbf{h}}^{(\mathrm{b})} = \mathbf{h}^{(\mathrm{b})} + \Delta\mathbf{h}^{(\mathrm{b})}$ with $\mathbf{h}^{(\mathrm{b})}$ represents the true beamspace channels. The true beamspace channel $\mathbf{h}^{(\mathrm{b})}$ can be naturally represented by a tensor $\boldsymbol{\mathcal{H}}^{(b)} \in \mathbb{C}^{K \times N_1 \times N_2}$ as

$$\begin{aligned} \boldsymbol{\mathcal{H}}^{(b)} &= G\alpha_{\mathrm{L}}\mathbf{a}^{(K)}(\omega_{\tau_{\mathrm{L}}}) \circ \mathbf{T}_1^{\mathsf{H}}\mathbf{a}^{(N_{\mathrm{B},1})}(\omega_{\mathrm{L},1}) \circ \mathbf{T}_2^{\mathsf{H}}\mathbf{a}^{(N_{\mathrm{B},2})}(\omega_{\mathrm{L},2}) \\ &+ (\sum_{g=1}^{G}\beta_{\mathrm{R},g})\mathbf{a}^{(K)}(\omega_{\tau_{\mathrm{R}}}) \circ \mathbf{T}_1^{\mathsf{H}}\mathbf{a}^{(N_{\mathrm{B},1})}(\omega_{\mathrm{R},1}) \circ \mathbf{T}_2^{\mathsf{H}}\mathbf{a}^{(N_{\mathrm{B},2})}(\omega_{\mathrm{R},2}), \end{aligned} \tag{49}$$

where $\beta_{\mathrm{R},g} = \alpha_{\mathrm{R}}\left[(\mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{A}}) \odot \mathbf{a}_{\mathrm{R}}(\boldsymbol{\phi}_{\mathrm{D}}))^{\mathsf{T}}\boldsymbol{\gamma}_g\right]$. We can see that $\boldsymbol{\mathcal{H}}^{(\mathrm{b})}$ is a rank-two third-order tensor based on the definition in Subsection III-A, as it is the sum of two rank-one tensors (i.e., the UE-BS channel and the UE-RIS-BS channel) and each of them is an outer product of three vectors. The first mode of $\boldsymbol{\mathcal{H}}^{(\mathrm{b})}$ lies in the element-space while the rest of modes lie in the beamspace generated by the transformation matrices $\mathbf{T}_1$ and $\mathbf{T}_2$. Further, we can define $\boldsymbol{\Phi}_1 = \mathrm{diag}\left\{[e^{j\omega_{\tau_{\mathrm{L}}}}, e^{j\omega_{\tau_{\mathrm{R}}}}]^{\mathsf{T}}\right\}$, $\boldsymbol{\Phi}_2 = \mathrm{diag}\left\{[e^{j\omega_{\mathrm{L},1}}, e^{j\omega_{\mathrm{R},1}}]^{\mathsf{T}}\right\}$, and $\boldsymbol{\Phi}_3 = \mathrm{diag}\left\{[e^{j\omega_{\mathrm{L},2}}, e^{j\omega_{\mathrm{R},2}}]^{\mathsf{T}}\right\}$.

**Algorithm 1** Coarse Channel Estimation via Tensor-ESPRIT

**Input:** $\mathbf{y}_{g,k}, g = 1, \ldots, G, k = 1, \ldots, K$.
**Output:** channel parameters estimate $\hat{\boldsymbol{\eta}}$.

1: Obtain $\mathbf{H} = [\mathbf{h}_1^{(\mathrm{b})}, \ldots, \mathbf{h}_G^{(\mathrm{b})}]$ based on (38) and (39).
2: Calculate $\mathbf{h}^{(\mathrm{b})} = \sum_{g=1}^{G} \mathbf{h}_g^{(\mathrm{b})}$ and form tensor $\boldsymbol{\mathcal{H}}^{(b)}$ in (49).
3: Apply CP decomposition to $\boldsymbol{\mathcal{H}}^{(b)}$ and obtain the factor matrices $\mathbf{U}^{(n)}, n = 1, \ldots, 3$.
4: Estimate $\hat{\boldsymbol{\Theta}}_1$ through (31) and obtain $\{\hat{\omega}_{\tau_\mathrm{L}}, \hat{\omega}_{\tau_\mathrm{R}}\}$ from the phases of the eigenvalues of $\hat{\boldsymbol{\Theta}}_1$.
5: Estimate $\hat{\boldsymbol{\Theta}}_2$ and $\hat{\boldsymbol{\Theta}}_3$ through (35)–(37), and obtain $\{\hat{\omega}_{\mathrm{L},1}, \hat{\omega}_{\mathrm{R},1}, \hat{\omega}_{\mathrm{L},2}, \hat{\omega}_{\mathrm{R},2}\}$ from the phases of the eigenvalues of $\hat{\boldsymbol{\Theta}}_2$ and $\hat{\boldsymbol{\Theta}}_3$.
6: Recover $\{\hat{\theta}_\mathrm{L}^{\mathrm{az}}, \hat{\theta}_\mathrm{L}^{\mathrm{el}}, \hat{\theta}_\mathrm{R}^{\mathrm{az}}, \hat{\theta}_\mathrm{R}^{\mathrm{el}}, \hat{\tau}_\mathrm{L}, \hat{\tau}_\mathrm{R}\}$ based on (44)–(48).
7: Obtain $\mathbf{s}_\mathrm{R}$ by (51) and form tensor $\boldsymbol{\mathcal{S}}^{(b)}$ in (54).
8: Obtain $\{\hat{\omega}_{\vartheta_2}, \hat{\omega}_{\vartheta_3}\}$ by applying CP decomposition and repeating Step 5 to $\boldsymbol{\mathcal{S}}^{(b)}$.
9: Recover $\{\hat{\vartheta}_2, \hat{\vartheta}_3\}$ based on (53).
10: Return $\hat{\boldsymbol{\eta}} = [\hat{\theta}_\mathrm{L}^{\mathrm{az}}, \hat{\theta}_\mathrm{L}^{\mathrm{el}}, \hat{\theta}_\mathrm{R}^{\mathrm{az}}, \hat{\theta}_\mathrm{R}^{\mathrm{el}}, \hat{\tau}_\mathrm{L}, \hat{\tau}_\mathrm{R}, \hat{\vartheta}_2, \hat{\vartheta}_3]^\mathsf{T}$.

To estimate the underlying angular frequencies, we first apply the CP decomposition to $\hat{\boldsymbol{\mathcal{H}}}$ (generated from $\hat{\mathbf{h}}^{(\mathrm{b})}$) and obtain the factor matrices $\mathbf{U}^{(n)}, n = 1, 2, 3$. Then $\boldsymbol{\Theta}_1$ can be estimated by (31) while $\boldsymbol{\Theta}_2$ and $\boldsymbol{\Theta}_3$ can be estimated through (35), (36), and (37). Next, the estimated angular frequencies $\hat{\omega}_{\tau_\mathrm{L}}$, $\hat{\omega}_{\tau_\mathrm{R}}$, $\hat{\omega}_{\mathrm{L},1}$, $\hat{\omega}_{\mathrm{R},1}$, $\hat{\omega}_{\mathrm{L},2}$, and $\hat{\omega}_{\mathrm{R},2}$ can be obtained from the phases of the eigenvalues of $\hat{\boldsymbol{\Theta}}_1$, $\hat{\boldsymbol{\Theta}}_2$, and $\hat{\boldsymbol{\Theta}}_3$, and the corresponding channel parameters can be recovered based on (44)–(48).

*2) Estimating $\vartheta_2$ and $\vartheta_3$:* Based on the estimated parameters, we construct a matrix $\mathbf{R}$ as

$$\mathbf{R} = \left[\mathbf{a}^{(K)}(\hat{\omega}_{\tau_\mathrm{L}}) \otimes \mathbf{W}^\mathsf{H}\mathbf{a}_\mathrm{B}(\hat{\boldsymbol{\theta}}_\mathrm{L}) \quad \mathbf{a}^{(K)}(\hat{\omega}_{\tau_\mathrm{R}}) \otimes \mathbf{W}^\mathsf{H}\mathbf{a}_\mathrm{B}(\hat{\boldsymbol{\theta}}_\mathrm{R})\right].$$

Then, we have the following system of equations:

$$\mathbf{R}[\alpha_\mathrm{L}, \beta_{\mathrm{R},g}]^\mathsf{T} = \mathbf{h}_g^{(\mathrm{b})}, \ g = 1, \ldots, G. \tag{50}$$

Thus, we have a set of LS estimates $\{\hat{\alpha}_\mathrm{L}, \hat{\beta}_{\mathrm{R},g}\}, g = 1, \ldots, G$ given by

$$[\hat{\alpha}_\mathrm{L}, \hat{\beta}_{\mathrm{R},g}]^\mathsf{T} = (\mathbf{R}^\mathsf{H}\mathbf{R})^{-1}\mathbf{R}^\mathsf{H}\mathbf{h}_g^{(\mathrm{b})}. \tag{51}$$

Now, let

$$\hat{\mathbf{s}}_\mathrm{R} = [\hat{\beta}_{\mathrm{R},1}, \hat{\beta}_{\mathrm{R},2}, \ldots, \hat{\beta}_{\mathrm{R},G}]^\mathsf{T} = \mathbf{s}_\mathrm{R} + \Delta\mathbf{s}_\mathrm{R} \in \mathbb{C}^G, \tag{52}$$

where $\mathbf{s}_\mathrm{R} = [\beta_{\mathrm{R},1}, \beta_{\mathrm{R},2}, \ldots, \beta_{\mathrm{R},G}]^\mathsf{T} \in \mathbb{C}^G$ and $\Delta\mathbf{s}_\mathrm{R}$ is the estimation error of $\hat{\mathbf{s}}_\mathrm{R}$. By defining

$$\omega_{\vartheta_2} = (2\pi f_c d_\mathrm{R}/c)\vartheta_2, \quad \omega_{\vartheta_3} = (2\pi f_c d_\mathrm{R}/c)\vartheta_3, \tag{53}$$

the noise-free vector $\mathbf{s}_\mathrm{R}$ can be represented as a second-order tensor (matrix) $\boldsymbol{\mathcal{S}}^{(b)} \in \mathbb{C}^{\sqrt{G}\times\sqrt{G}}$ as

$$\boldsymbol{\mathcal{S}}^{(b)} = \alpha_\mathrm{R}(\mathbf{T}_3 \otimes \mathbf{T}_4)\Big(\mathbf{a}^{(N_{\mathrm{R},1})}(\omega_{\vartheta_2}) \otimes \mathbf{a}^{(N_{\mathrm{R},2})}(\omega_{\vartheta_3})\Big) \tag{54}$$

$$= \alpha_\mathrm{R}\mathbf{T}_3\mathbf{a}^{(N_{\mathrm{R},1})}(\omega_{\vartheta_2}) \circ \mathbf{T}_4\mathbf{a}^{(N_{\mathrm{R},2})}(\omega_{\vartheta_3}). \tag{55}$$

Therefore, the parameters $\vartheta_2$ and $\vartheta_3$ can be estimated from $\hat{\mathbf{s}}_\mathrm{R}$ using the same routine given by (35), (36), and (37).

As a summary, the complete steps of the coarse channel estimation process are presented in Algorithm 1.

### *B. Channel Parameters Refinement via Least Squares*

Channel parameter estimates $\tilde{\boldsymbol{\eta}}$ are refined based on the LS criterion initialized using the coarse estimates from Subsection IV-A. By defining

$$\boldsymbol{\mu}_\mathrm{L}^{g,k} = \mathbf{W}^\mathsf{H} e^{-j2\pi(k-1)\Delta_f\tau_\mathrm{L}}\mathbf{a}_\mathrm{B}(\boldsymbol{\theta}_\mathrm{L})x_{g,k}, \tag{56}$$

$$\boldsymbol{\mu}_\mathrm{R}^{g,k} = \mathbf{W}^\mathsf{H}\left[(\mathbf{a}_\mathrm{R}(\boldsymbol{\phi}_\mathrm{A}) \odot \mathbf{a}_\mathrm{R}(\boldsymbol{\phi}_\mathrm{D}))^\mathsf{T}\boldsymbol{\gamma}_g\right] \tag{57}$$

$$\times\, e^{-j2\pi(k-1)\Delta_f\tau_\mathrm{R}}\mathbf{a}_\mathrm{B}(\boldsymbol{\theta}_\mathrm{R})x_{g,k}, \tag{58}$$

we have the noise-free signals given by

$$\boldsymbol{\mu}_{g,k} = \alpha_\mathrm{L}\boldsymbol{\mu}_\mathrm{L}^{g,k} + \alpha_\mathrm{R}\boldsymbol{\mu}_\mathrm{R}^{g,k}. \tag{59}$$

We further define vectors $\{\mathbf{y}, \boldsymbol{\mu}_\mathrm{L}, \boldsymbol{\mu}_\mathrm{R}\} \in \mathbb{C}^{GKN_1N_2}$ as the concatenation of $\mathbf{y}_{g,k}$, $\boldsymbol{\mu}_\mathrm{L}^{g,k}$ and $\boldsymbol{\mu}_\mathrm{R}^{g,k}$ over $g = 1, \ldots, G, k = 1, \ldots, K$. Since the noise in (8) is colored with unknown statistics, we perform a sub-optimal estimator based on the LS criterion as

$$\hat{\boldsymbol{\eta}}_\mathrm{LS} = \arg\min_{\boldsymbol{\eta}} \ \|\mathbf{y} - \alpha_\mathrm{L}\boldsymbol{\mu}_\mathrm{L}(\boldsymbol{\eta}) - \alpha_\mathrm{R}\boldsymbol{\mu}_\mathrm{R}(\boldsymbol{\eta})\|_2^2. \tag{60}$$

The value of the complex channel gains $\alpha_\mathrm{L}$ and $\alpha_\mathrm{R}$ can be obtained as a function of $\mathbf{y}$, $\boldsymbol{\mu}_\mathrm{L}$ and $\boldsymbol{\mu}_\mathrm{R}$ by solving

$$\begin{cases} \partial\|\mathbf{y} - \alpha_\mathrm{L}\boldsymbol{\mu}_\mathrm{L}(\boldsymbol{\eta}) - \alpha_\mathrm{R}\boldsymbol{\mu}_\mathrm{R}(\boldsymbol{\eta})\|_2^2/\partial\Re(\alpha_\mathrm{L}) = 0, \\ \partial\|\mathbf{y} - \alpha_\mathrm{L}\boldsymbol{\mu}_\mathrm{L}(\boldsymbol{\eta}) - \alpha_\mathrm{R}\boldsymbol{\mu}_\mathrm{R}(\boldsymbol{\eta})\|_2^2/\partial\Re(\alpha_\mathrm{R}) = 0, \end{cases} \tag{61}$$

which give

$$\hat{\alpha}_\mathrm{L} = \frac{\boldsymbol{\mu}_\mathrm{L}^\mathsf{H}\mathbf{y}\|\boldsymbol{\mu}_\mathrm{R}\|_2^2 - \boldsymbol{\mu}_\mathrm{R}^\mathsf{H}\mathbf{y}\boldsymbol{\mu}_\mathrm{L}^\mathsf{H}\boldsymbol{\mu}_\mathrm{R}}{\|\boldsymbol{\mu}_\mathrm{L}\|_2^2\|\boldsymbol{\mu}_\mathrm{R}\|_2^2 - |\boldsymbol{\mu}_\mathrm{L}^\mathsf{H}\boldsymbol{\mu}_\mathrm{R}|^2}, \tag{62}$$

$$\hat{\alpha}_\mathrm{R} = \frac{\boldsymbol{\mu}_\mathrm{R}^\mathsf{H}\mathbf{y}\|\boldsymbol{\mu}_\mathrm{L}\|_2^2 - \boldsymbol{\mu}_\mathrm{L}^\mathsf{H}\mathbf{y}\boldsymbol{\mu}_\mathrm{R}^\mathsf{H}\boldsymbol{\mu}_\mathrm{L}}{\|\boldsymbol{\mu}_\mathrm{L}\|_2^2\|\boldsymbol{\mu}_\mathrm{R}\|_2^2 - |\boldsymbol{\mu}_\mathrm{L}^\mathsf{H}\boldsymbol{\mu}_\mathrm{R}|^2}. \tag{63}$$

Then, (60) can be solved by using, e.g., gradient descent method given the initialization $\hat{\boldsymbol{\eta}}$. We denote the refined localization-related channel parameters as $\tilde{\boldsymbol{\eta}}$.

### *C. Conversion to the Localization Domain*

Based on the refined channel parameters estimates $\tilde{\boldsymbol{\eta}}$, the localization parameters $\boldsymbol{\xi}$ can be recovered by carrying out a 2D search over $o_3$ and $\Delta$; the rest of the localization parameters can be determined from each search point $[o_3, \Delta]^\mathsf{T}$ and a cost metric can be defined to compare the fitness of different search points, as will be explained imminently.

Given channel parameters $[\theta_\mathrm{L}^{\mathrm{az}}, \theta_\mathrm{L}^{\mathrm{el}}, \theta_\mathrm{R}^{\mathrm{az}}, \theta_\mathrm{R}^{\mathrm{el}}, \tau_\mathrm{L}, \tau_\mathrm{R}, \vartheta_2, \vartheta_3]^\mathsf{T}$ and any $\check{o}_3$ and $\check{\Delta}$, we can first determine the propagation distance of the LOS UE-BS path and RIS reflection path as

$$\check{d}_\mathrm{L} = c(\tau_\mathrm{L} - \check{\Delta}), \qquad \check{d}_\mathrm{R} = c(\tau_\mathrm{R} - \check{\Delta}). \tag{64}$$

which further determine the UE position as

$$\check{\mathbf{p}}_\mathrm{U} = \mathbf{p}_\mathrm{B} + \check{d}_\mathrm{L}\mathbf{R}_\mathrm{B}\mathbf{t}(\theta_\mathrm{L}^{\mathrm{az}}, \theta_\mathrm{L}^{\mathrm{el}}). \tag{65}$$

The RIS position can be obtained as the intersection of the ellipsoid $\|\check{\mathbf{p}}_\mathrm{R} - \mathbf{p}_\mathrm{B}\|_2 + \|\check{\mathbf{p}}_\mathrm{R} - \check{\mathbf{p}}_\mathrm{U}\|_2 = \check{d}_\mathrm{R}$ and the line $\check{\mathbf{p}}_\mathrm{R} = \mathbf{p}_\mathrm{B} + x\mathbf{R}_\mathrm{B}\mathbf{t}(\theta_\mathrm{R}^{\mathrm{az}}, \theta_\mathrm{R}^{\mathrm{el}})$, with $x > 0$ being the distance between BS and RIS. The intersection is determined by solving for $x$ to obtain

$$x = \frac{\check{d}_\mathrm{R}^2 - \|\mathbf{p}_\mathrm{B} - \check{\mathbf{p}}_\mathrm{U}\|_2^2}{2\left(\check{d}_\mathrm{R} + (\mathbf{R}_\mathrm{B}\mathbf{t}(\theta_\mathrm{R}^{\mathrm{az}}, \theta_\mathrm{R}^{\mathrm{el}}))^\mathsf{T}(\mathbf{p}_\mathrm{B} - \check{\mathbf{p}}_\mathrm{U})\right)}. \tag{66}$$

**Algorithm 2** 2D Search-Based Localization Algorithm

**Input:** refined channel parameters $\tilde{\boldsymbol{\eta}}$, number of refinement $Q$.
**Output:** localization parameters estimates $\hat{\boldsymbol{\xi}}$.

Initialize search space and resolution $\{\mathcal{P}_{o_3}^0, \mathcal{P}_{\Delta}^0, d_{o_3}^0, d_{\Delta}^0\}$.
**for** $i = 0, \ldots, Q$ **do**
  **for** every candidate $\check{o}_3 \in \mathcal{P}_{o_3}^i$ **do**
    **for** every candidate $\check{\Delta} \in \mathcal{P}_{\Delta}^i$ **do**
      Compute $\check{d}_{\mathrm{L}}$ and $\check{d}_{\mathrm{R}}$ through (64).
      Estimate UE position $\check{\mathbf{p}}_{\mathrm{U}}$ through (65) and estimate RIS position $\check{\mathbf{p}}_{\mathrm{R}}$ through (66).
      Predict the intermediate measurements $\{\check{\vartheta}_2, \check{\vartheta}_3\}$ based on $\check{o}_3$, $\check{\mathbf{p}}_{\mathrm{U}}$, $\check{\mathbf{p}}_{\mathrm{R}}$ and $\mathbf{p}_{\mathrm{B}}$.
      Compute the cost metric $f(\check{o}_3, \check{\Delta})$ based on (67).
    **end for**
  **end for**
  Select the candidate pair $\{\hat{o}_3^i, \hat{\Delta}^i\}$ that minimize $f(o_3, \Delta)$ and determine $\hat{\mathbf{p}}_{\mathrm{U}}^i$ and $\hat{\mathbf{p}}_{\mathrm{R}}^i$ based on $\{\hat{o}_3^i, \hat{\Delta}^i\}$.
  Shrink $\{\mathcal{P}_{o_3}^{i+1}, \mathcal{P}_{\Delta}^{i+1}, d_{o_3}^{i+1}, d_{\Delta}^{i+1}\}$ by (68)–(70).
**end for**
Return $\hat{\boldsymbol{\xi}} = [(\hat{\mathbf{p}}_{\mathrm{U}}^Q)^{\mathsf{T}}, (\hat{\mathbf{p}}_{\mathrm{R}}^Q)^{\mathsf{T}}, \hat{o}_3^Q, \hat{\Delta}^Q]^{\mathsf{T}}$.

Then, we can predict the intermediate measurements $\{\check{\vartheta}_2, \check{\vartheta}_3\}$ according to (17) and (18) based on $\check{o}_3$, $\check{\mathbf{p}}_{\mathrm{U}}$, $\check{\mathbf{p}}_{\mathrm{R}}$ and $\mathbf{p}_{\mathrm{B}}$. Thus, we can compute the cost metric as

$$f(\check{o}_3, \check{\Delta}) = \left\| [\check{\vartheta}_2, \check{\vartheta}_3]^{\mathsf{T}} - [\hat{\vartheta}_2, \hat{\vartheta}_3]^{\mathsf{T}} \right\|_2^2, \tag{67}$$

which allows us to perform a 2D search over all the $\{o_3, \Delta\}$ candidates in a pre-defined search space. The optimal values that minimize $f(\check{o}_3, \check{\Delta})$ in (67) are then returned as the estimated $\hat{o_3}$ and $\hat{\Delta}$, and the corresponding $\hat{\mathbf{p}}_{\mathrm{U}}$ and $\hat{\mathbf{p}}_{\mathrm{R}}$ can be determined accordingly.

Let $\{\mathcal{P}_{o_3}, \mathcal{P}_{\Delta}\}$ denotes the search spaces and $\{d_{o_3}, d_{\Delta}\}$ denotes the search resolutions. We can further perform multiple rounds of search grid refinement. As an example, for the $i$th round refined search over the search space $\{\mathcal{P}_{o_3}^i, \mathcal{P}_{\Delta}^i\}$ and resolution $\{d_{o_3}^i, d_{\Delta}^i\}$ that returns estimates $\{\hat{o}_3^i, \hat{\Delta}^i\}$, we can shrink the resolution and the search space in round $i+1$ such that

$$d_{o_3}^{i+1} = \kappa d_{o_3}^i, \qquad d_{\Delta}^{i+1} = \kappa d_{\Delta}^i, \tag{68}$$

$$\mathcal{P}_{o_3}^{i+1} = \{\ldots, \hat{o}_3^i - d_{o_3}^{i+1}, \hat{o}_3^i, \hat{o}_3^i + d_{o_3}^{i+1}, \ldots\}, \tag{69}$$

$$\mathcal{P}_{\Delta}^{i+1} = \{\ldots, \hat{\Delta}^i - d_{\Delta}^{i+1}, \hat{\Delta}^i, \hat{\Delta}^i + d_{\Delta}^{i+1}, \ldots\}, \tag{70}$$

where $\kappa \in (0, 1)$ and the cardinalities of $\mathcal{P}_{o_3}^i$ and $\mathcal{P}_{\Delta}^i$ are fixed as $\mathrm{card}(\mathcal{P}_{o_3}^i) = C_{o_3}$ and $\mathrm{card}(\mathcal{P}_{\Delta}^i) = C_{\Delta}$ for all $i$ values. The pseudo-code of the proposed search method is summarized in Algorithm 2.

### D. Complexity Analysis

This subsection evaluates the computational complexity of the proposed algorithms. Among the proposed method, Algorithm 1 involves the channel matrix recovery with a complexity $\mathcal{O}(KGN_1N_2)$, the CP decomposition of $\boldsymbol{\mathcal{H}}^{(b)}$ and the corresponding matrix multiplications with a complexity $\mathcal{O}(KN_1N_2) + \mathcal{O}(K^2) + \mathcal{O}(N_1^2) + \mathcal{O}(N_2^2)$, and the CP decomposition of $\boldsymbol{\mathcal{S}}^{(b)}$ and the corresponding matrix multiplications with a complexity $\mathcal{O}(G)$. The LS refinement of the channel parameters performs an iterative procedure, which gives a complexity $\mathcal{O}(TKGN_1N_2)$ where $T$ denotes the total number of iterations. Finally, a complexity of $\mathcal{O}(QC_{o_3}C_{\Delta})$ is introduced by Algorithm 2. In summary, the overall complexity of the proposed solution for the JrCUP problem is given by

$$\begin{aligned}\mathcal{O}_{\text{total}} = \mathcal{O}(K^2) + \mathcal{O}(N_1^2) + \mathcal{O}(N_2^2) \\ + \mathcal{O}(TKGN_1N_2) + \mathcal{O}(QC_{o_3}C_{\Delta}).\end{aligned} \tag{71}$$

## V. Localization Error Bounds Derivation

### A. CRLB for Channel Parameters Estimation

Considering the fact that the received noise (7) is zero-mean noncircular complex Gaussian random variables, the Fisher information matrix (FIM) of all the unknown channel parameters $\boldsymbol{\eta}_{\text{ch}}$ is given by the following Proposition 2.

**Proposition 2.** *Based on the channel model* (7), (8) *and* (59), *the FIM of channel parameters $\boldsymbol{\eta}_{ch}$ can be computed as*

$$\mathbf{J}(\boldsymbol{\eta}_{ch}) = \sum_{g=1}^{G} \sum_{k=1}^{K} \mathbf{D}_{g,k}^{\mathsf{T}} \left[\mathbf{C}_0 + \mathbf{C}_r^{g,k}\right]^{-1} \mathbf{D}_{g,k}, \tag{72}$$

*where*

$$\mathbf{D}_{g,k} = \left[\Re\left(\frac{\partial \boldsymbol{\mu}_{g,k}}{\partial \boldsymbol{\eta}_{ch}}\right)^{\mathsf{T}}, \ \Im\left(\frac{\partial \boldsymbol{\mu}_{g,k}}{\partial \boldsymbol{\eta}_{ch}}\right)^{\mathsf{T}}\right]^{\mathsf{T}}, \tag{73}$$

$$\mathbf{C}_0 = \frac{\sigma_0^2}{2} \begin{bmatrix} \Re\left(\mathbf{A}_0\mathbf{A}_0^{\mathsf{H}}\right) & \Im\left(\mathbf{A}_0\mathbf{A}_0^{\mathsf{H}}\right)^{\mathsf{T}} \\ \Im\left(\mathbf{A}_0\mathbf{A}_0^{\mathsf{H}}\right) & \Re\left(\mathbf{A}_0\mathbf{A}_0^{\mathsf{H}}\right) \end{bmatrix}, \tag{74}$$

$$\mathbf{C}_r^{g,k} = \frac{\sigma_r^2}{2} \begin{bmatrix} \Re\left(\mathbf{A}_r^{g,k}(\mathbf{A}_r^{g,k})^{\mathsf{H}}\right) & \Im\left(\mathbf{A}_r^{g,k}(\mathbf{A}_r^{g,k})^{\mathsf{H}}\right)^{\mathsf{T}} \\ \Im\left(\mathbf{A}_r^{g,k}(\mathbf{A}_r^{g,k})^{\mathsf{H}}\right) & \Re\left(\mathbf{A}_r^{g,k}(\mathbf{A}_r^{g,k})^{\mathsf{H}}\right) \end{bmatrix}, \tag{75}$$

*with $\mathbf{A}_0 = \mathbf{W}^{\mathsf{H}}$ and $\mathbf{A}_r^{g,k} = \mathbf{W}^{\mathsf{H}}\mathbf{H}_{R,2}^k\boldsymbol{\Gamma}_g$.*

*Proof.* See Appendix A □

**Remark 1.** *Note that in* (72), *the covariance matrix $\mathbf{C}_r^{g,k}$ is a function of $\boldsymbol{\eta}_{ch}$, which also contributes to the FIM of $\boldsymbol{\eta}_{ch}$ in principle [49, B.3.3]. Nonetheless, since the noise statistics information is not used in the channel estimation processes in Subsection IV-A and IV-B, we ignore this relationship in this work.*

Based on (72), we can compute the FIM of localization-related parameters $\boldsymbol{\eta}$ using Schur's complement: we partition $\mathbf{J}(\boldsymbol{\eta}_{\text{ch}}) = [\mathbf{X}, \mathbf{Y}; \mathbf{Y}^{\mathsf{T}}, \mathbf{Z}]$, where $\mathbf{X} \in \mathbb{R}^{8\times 8}$ so that $\mathbf{J}(\boldsymbol{\eta}) = \mathbf{X} - \mathbf{Y}\mathbf{Z}^{-1}\mathbf{Y}^{\mathsf{T}}$. Then the estimation error bounds for $\boldsymbol{\theta}_{\mathrm{L}}$, $\boldsymbol{\theta}_{\mathrm{R}}$, $\tau_{\mathrm{L}}$, $\tau_{\mathrm{R}}$ and $\boldsymbol{\vartheta}$ can be derived as

$$\mathrm{EB}(\boldsymbol{\theta}_{\mathrm{L}}) = \sqrt{\mathrm{tr}([\mathbf{J}(\boldsymbol{\eta})^{-1}]_{1:2,1:2})}, \tag{76}$$

$$\mathrm{EB}(\boldsymbol{\theta}_{\mathrm{R}}) = \sqrt{\mathrm{tr}([\mathbf{J}(\boldsymbol{\eta})^{-1}]_{3:4,3:4})}, \tag{77}$$

$$\mathrm{EB}(\tau_{\mathrm{L}}) = \sqrt{[\mathbf{J}(\boldsymbol{\eta})^{-1}]_{5,5}}, \tag{78}$$

$$\mathrm{EB}(\tau_{\mathrm{R}}) = \sqrt{[\mathbf{J}(\boldsymbol{\eta})^{-1}]_{6,6}}, \tag{79}$$

$$\mathrm{EB}(\boldsymbol{\vartheta}) = \sqrt{\mathrm{tr}([\mathbf{J}(\boldsymbol{\eta})^{-1}]_{7:8,7:8})}, \tag{80}$$

which lower bound the estimation RMSEs for the corresponding parameters.

TABLE I: Default Simulation Parameters

| Parameter | Value |
|---|---|
| Propagation Speed $c$ | $2.9979 \times 10^8$ m/s |
| Carrier Frequency $f_c$ | 28 GHz |
| Bandwidth $B$ | 100 MHz |
| # Subcarriers $K$ | 32 |
| # Transmissions $G$ | 9 |
| Clock Offset $\rho$ | 100 ns |
| Transmission Power $P_{\mathrm{T}}$ | 10 dBm |
| Active RIS Power $P_{\mathrm{R}}$ | 7 dBm |
| Noise PSD of Receiver & RIS | $-174$ dBm/Hz |
| Noise Figure of Receiver & RIS | 10 dB |
| Array Size of BS / RFC / RIS | $10 \times 10$ / $5 \times 5$ / $15 \times 15$ |
| Position & Orientation of BS | $\mathbf{p}_{\mathrm{B}} = [0, 5, 3]^{\mathsf{T}}$, $\mathbf{o}_{\mathrm{B}} = [0, 0, -\pi/2]^{\mathsf{T}}$ |
| Position & Orientation of RIS | $\mathbf{p}_{\mathrm{R}} = [-5, 0, 3]^{\mathsf{T}}$, $\mathbf{o}_{\mathrm{R}} = [0, 0, 0]^{\mathsf{T}}$ |
| Position of UE | $\mathbf{p}_{\mathrm{B}} = [3, 2, 1]^{\mathsf{T}}$ |

### B. CRLB for Localization Parameters Estimation

Based on the calculated $\mathbf{J}(\boldsymbol{\eta})$ and the geometric model in (2)–(5) and (17)–(18), we can further derive the FIM of the localization parameters $\boldsymbol{\xi}$ using the chain rule of the FIM transformation as [50]

$$\mathbf{J}(\boldsymbol{\xi}) = \mathbf{T}^{\mathsf{T}}\mathbf{J}(\boldsymbol{\eta})\mathbf{T}, \tag{81}$$

where $\mathbf{T} = \partial\boldsymbol{\eta}/\partial\boldsymbol{\xi} \in \mathbb{R}^{8\times 8}$ is the Jacobian matrix. Then the lower bounds for the estimation RMSE of $\mathbf{p}_{\mathrm{U}}$, $\mathbf{p}_{\mathrm{R}}$, $o_3$ are

$$\mathrm{EB}(\mathbf{p}_{\mathrm{U}}) = \sqrt{\mathrm{tr}([\mathbf{J}(\boldsymbol{\xi})^{-1}]_{1:3,1:3})}, \tag{82}$$

$$\mathrm{EB}(\mathbf{p}_{\mathrm{R}}) = \sqrt{\mathrm{tr}([\mathbf{J}(\boldsymbol{\xi})^{-1}]_{4:6,4:6})}, \tag{83}$$

$$\mathrm{EB}(o_3) = \sqrt{[\mathbf{J}(\boldsymbol{\xi})^{-1}]_{7,7}}. \tag{84}$$

## VI. Numerical Results

### A. Evaluation Setup

We consider an indoor localization scenario within a $10\,\mathrm{m} \times 10\,\mathrm{m} \times 3\,\mathrm{m}$ space. We use random signal symbols $x_{g,k}$ with the power constraint $|x_{g,k}| = \sqrt{P_{\mathrm{T}}}$, and random precoder $\mathbf{W}$ and RIS profiles $\boldsymbol{\Upsilon}$ satisfying the structure constraints (40) and (41). Based on (1), the active RIS amplification factor $p$ is calculated as

$$p = \sqrt{\frac{P_{\mathrm{R}}}{N_{\mathrm{R},1}N_{\mathrm{R},2}(P_{\mathrm{T}}|\alpha_{\mathrm{R},1}|^2 + \sigma_r^2)} + 1}. \tag{85}$$

The channel gains of the LOS UE-BS, UE-RIS and RIS-BS paths are generated by $\alpha_{\mathrm{L}} = \frac{\lambda_c}{4\pi\|\mathbf{p}_{\mathrm{U}}-\mathbf{p}_{\mathrm{B}}\|_2}e^{j\psi_{\mathrm{L}}}$, $\alpha_{\mathrm{R},1} = \frac{\lambda_c}{4\pi\|\mathbf{p}_{\mathrm{U}}-\mathbf{R}_{\mathrm{R}}\|_2}e^{j\psi_{\mathrm{R},1}}$, $\alpha_{\mathrm{R},2} = \frac{\lambda_c}{4\pi\|\mathbf{p}_{\mathrm{R}}-\mathbf{R}_{\mathrm{B}}\|_2}e^{j\psi_{\mathrm{R},2}}$, where $\lambda_c = c/f_c$ and $\psi_{\mathrm{L}}$, $\psi_{\mathrm{R},1}$ and $\psi_{\mathrm{R},2}$ are independently generated from a uniform distribution $\mathcal{U}(0, 2\pi)$. When the multipath effect is introduced, as an example, the channel gains of the NLOS paths between the UE and BS in (9) are set as $\alpha_{\mathrm{L}}^i = \frac{\sqrt{4\pi c_{\mathrm{L},i}}\lambda_c}{16\pi^2 d_{\mathrm{U},i}d_{\mathrm{B},i}}e^{j\psi_{\mathrm{L},i}}$, $i = 1, \ldots, I_{\mathrm{L}}$, where $c_{\mathrm{L},i}$ represents the radar cross section (RCS) coefficient and $\psi_{\mathrm{L},i}$ is the random phase. Here, $d_{\mathrm{U},i}$ and $d_{\mathrm{B},i}$ are the distances between the UE and the $i$th SP and the distance between the BS and the $i$th SP. The channel gains $\alpha_{\mathrm{R},1}^i$ and $\alpha_{\mathrm{R},2}^i$ in (10) and (11) are defined in a similar manner. In addition, we define the received SNR as

$$\mathrm{SNR} \triangleq \frac{\sum_{g=1}^{G}\sum_{k=1}^{K}\|\boldsymbol{\mu}_{g,k}\|_2^2}{\sum_{g=1}^{G}\sum_{k=1}^{K}\mathrm{tr}(\mathbf{C}_0 + \mathbf{C}_r^{g,k})}. \tag{86}$$

In this paper, the three Euler angles, i.e. $[\mathbf{o}]_1$, $[\mathbf{o}]_2$ and $[\mathbf{o}]_3$, represent the rotations around $X$-axis, $Y$-axis and $Z$-axis, respectively. The default orientation $\mathbf{o} = [0, 0, 0]^{\mathsf{T}}$ is set to face the positive $X$-axis. The element spacings of the BS and RIS are set as $0.5\,\lambda_{\mathrm{c}}$ and $0.2\,\lambda_{\mathrm{c}}$, respectively. Other default simulation parameters are listed in Table I. Throughout the simulation examples, all the involved RMSEs are computed over 500 Monte Carlo trials. The channel delays and the clock bias are presented in units of meters by multiplying them by the constant propagation speed $c$ for better intuition.

### B. Performance Evaluation of the Proposed Algorithms

*1) Channel Estimation Performance:* We first evaluate the performance of the proposed channel estimators. To provide a benchmark in addition to the CRLBs, the proposed channel estimator is compared with the existing simultaneous orthogonal matching pursuit (SOMP) algorithm which has been verified to offer better channel estimation performance than the original orthogonal matching pursuit (OMP) algorithm [51]. We use SOMP to estimate the two strongest paths as the UE-BS and the UE-RIS-BS channels. For implementation details of SOMP, the readers are referred to [51], [52]. To meet the *restricted isometry property* [53] that SOMP requires, we use random combiner $\mathbf{W}$ and random RIS profile $\boldsymbol{\Upsilon}$ when performing SOMP, while the proposed $\{\mathbf{W}, \boldsymbol{\Upsilon}\}$ in (40) and (41) are used when evaluating the proposed methods. We derive and present the CRLBs for both cases. The SOMP dictionary sizes for each parameter in $\{\boldsymbol{\theta}_{\mathrm{L}}, \boldsymbol{\theta}_{\mathrm{R}}, \tau_{\mathrm{L}}, \tau_{\mathrm{R}}, \boldsymbol{\vartheta}\}$ are set equally as $D$, where $D = \{2^{12}, 2^{14}, 2^{16}\}$. Besides, the LS refinement in (60) is solved using the trust-region method, which is implemented through Manopt toolbox [54] and the number of iterations is set as $T = 40$.

Fig. 4 shows the evaluation of the RMSEs of $\boldsymbol{\theta}_{\mathrm{L}}$, $\boldsymbol{\theta}_{\mathrm{R}}$, $\tau_{\mathrm{L}}$, $\tau_{\mathrm{R}}$ and $\boldsymbol{\vartheta}$ versus the received SNR for the SOMP algorithm, the proposed tensor-ESPRIT coarse estimation, and the proposed LS-based refinement. It is observed that the proposed algorithm performs better in high-SNR regions compared to the existing SOMP algorithm but is inferior in low-SNR regions. While the RMSEs of both coarse estimation methods exhibit large gaps from the CRLBs, the proposed LS refinement can significantly reduce the distance to the CRLBs in high-SNR regions for both methods.[4] Nevertheless, there are still non-negligible gaps between the results of LS refinement (especially for $\boldsymbol{\theta}_{\mathrm{L}}$, $\boldsymbol{\theta}_{\mathrm{R}}$ and $\tau_{\mathrm{L}}$) and the theoretical bounds, which result from the mismatch between the used LS criterion (60) and the actual statistics of the noise (7). By comparing the CRLBs of the random and the proposed

[4]In the region that $\mathrm{SNR} < 0\,\mathrm{dB}$, the LS refinement cannot improve the performance for both SOMP and tensor-ESPRIT methods. This can be referred to as the threshold of *no information region*, which is a well-documented phenomenon in maximum likelihood estimators [55]. The same phenomenon can be observed in Fig. 5.

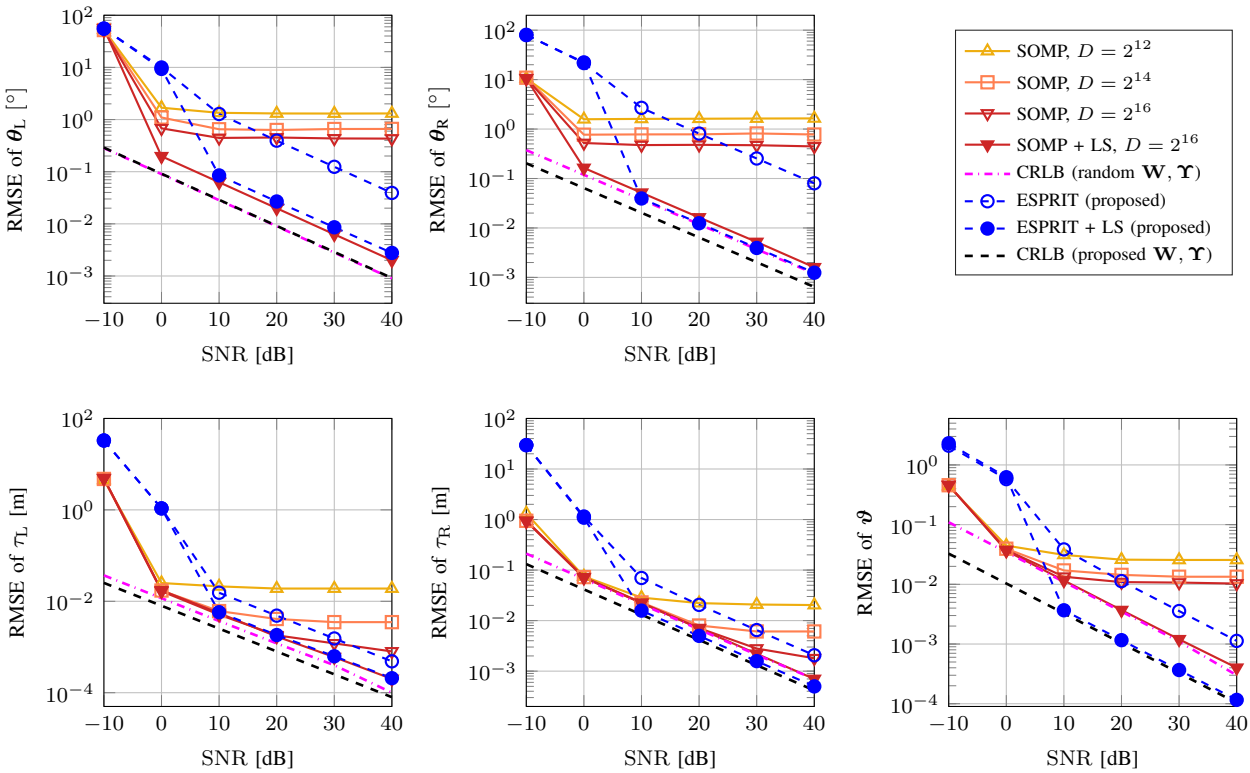



Fig. 4: The evaluation of RMSE of $\boldsymbol{\theta}_{\mathrm{L}}$, $\boldsymbol{\theta}_{\mathrm{R}}$, $\tau_{\mathrm{L}}$, $\tau_{\mathrm{R}}$ and $\vartheta$ versus received SNR for the existing SOMP algorithm, the proposed tensor-ESPRIT coarse estimation, and the proposed LS-based refinement.

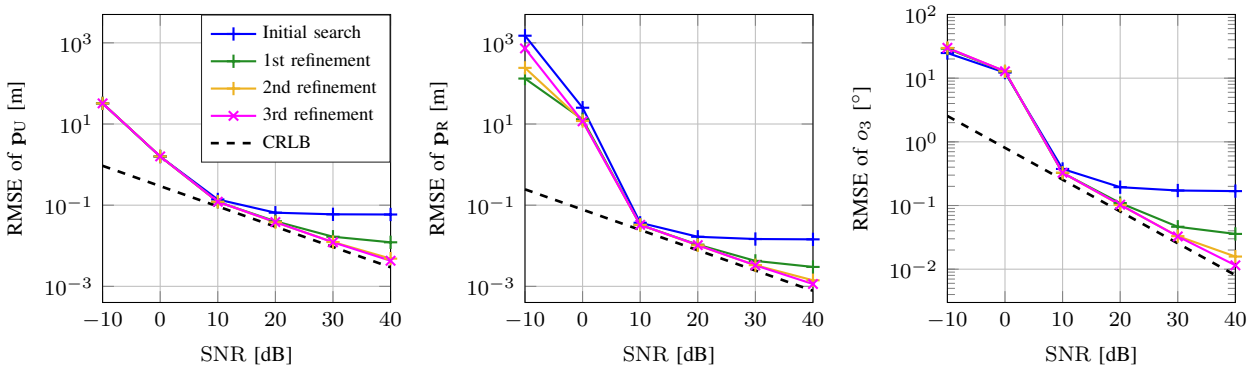
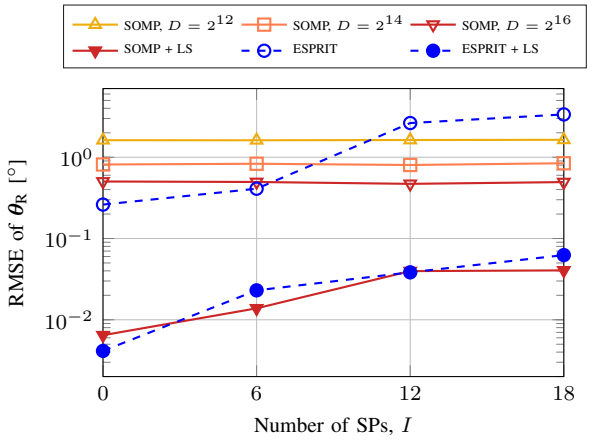



Fig. 5: The evaluation of RMSE of $\mathbf{p}_{\mathrm{U}}$, $\mathbf{p}_{\mathrm{R}}$ and $o_3$ versus received SNR for different numbers $(0, 1, 2, 3)$ of grid refinements.

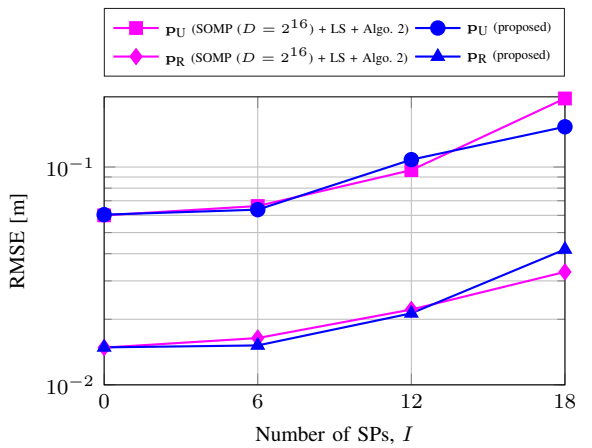
$\{\boldsymbol{\Upsilon}, \mathbf{W}\}$, we can observe that the proposed design offers lower bounds on channel parameter estimation, especially for $\boldsymbol{\theta}_{\mathrm{R}}$, $\tau_{\mathrm{R}}$, and $\vartheta$. Consequently, it is noticed that the proposed method (i.e., tensor-ESPRIT+LS refinement) provides more accurate estimation of most channel parameters ($\boldsymbol{\theta}_{\mathrm{R}}$, $\tau_{\mathrm{R}}$, and $\vartheta$) than the SOMP+LS solution in high-SNR regions.

For reference, the computational complexity of the SOMP algorithm used in this paper is provided as $\mathcal{O}(DN_1N_2N_{\mathrm{B},1}N_{\mathrm{B},2}) + \mathcal{O}(DKN_1N_2) + \mathcal{O}(DKG) + \mathcal{O}(DGN_{\mathrm{R},1}N_{\mathrm{R},2})$. According to (71), the computational complexity of the proposed tensor-ESPRIT solution is $\mathcal{O}(KGN_1N_2) + \mathcal{O}(K^2) + \mathcal{O}(N_1^2) + \mathcal{O}(N_2^2)$, which is not a function of $D$ (i.e., search-free). The performance of SOMP relies on the dictionary size $D$. A large dictionary that brings heavy computation is needed for the SOMP to offer satisfactory performance, making tensor-ESPRIT preferred in scenarios that require a fast response and low computational load (e.g., as an initialization).

*2) Localization Performance:* Then, we assess the performance of Algorithm 2 for the second stage of localization parameters estimation. As the second stage in JrCUP is a specialized problem, there exists no corresponding benchmark method, and only CRLB are compared. Fig. 5 presents the RMSEs of estimating $\mathbf{p}_{\mathrm{U}}$, $\mathbf{p}_{\mathrm{R}}$ and $o_3$ versus the received SNR for different numbers $(0, 1, 2, 3)$ of grid-search refinement iterations and $\kappa = 0.1$. Here, the input of Algorithm 2 is the result of the proposed tensor-ESPRIT+LS refinement in the first stage. It can be observed that in the low SNR regions




(a) RMSE evaluation of channel parameter $\boldsymbol{\theta}_{\mathrm{R}}$




(b) RMSE evaluation of localization parameters $\mathbf{p}_{\mathrm{U}}$ and $\mathbf{p}_{\mathrm{R}}$

Fig. 6: The evaluation of RMSEs of the estimated channel parameters and localization parameters under the multipath effect. The tested numbers of SPs are set as $I = \{0, 6, 12, 18\}$. (a) RMSEs of $\boldsymbol{\theta}_{\mathrm{R}}$ by SOMP ($D = \{2^{12}, 2^{14}, 2^{16}\}$), SOMP+LS refinement, the proposed tensor-ESPRIT, and the proposed tensor-ESPRIT+LS refinement; (b) RMSEs of $\mathbf{p}_{\mathrm{U}}$ and $\mathbf{p}_{\mathrm{R}}$ by SOMP ($D = 2^{16}$)+LS refinement+ Algorithm 2 and the proposed method (i.e., tensor-ESPRIT+LS refinement+ Algorithm 2).

(lower than 10 dB), the RMSEs stay far from the theoretical bound. In these regions, the input channel parameter estimates contain large errors that lead to localization failure. Thus, increasing the number of grid-search refinements does not improve performance. In the high SNR regions (10 dB or higher), however, we can see that the RMSEs decrease as more search refinements are carried out, which indicates localization success. The RMSEs follow the CRLB closely after two or more search iterations are performed. These results confirm that our proposed algorithms can achieve a nearly efficient localization performance at practical SNRs (higher than 10 dB). The refinement dependence of performance presents an unavoidable trade-off between localization accuracy and computational complexity in practice.

*3) Impact of Multipath:* The impact of the multipath effect is evaluated in Fig. 6. In this trial, we set the number of SPs in different channels as $I_{\mathrm{L}} = I_{\mathrm{R},1} = I_{\mathrm{R},2} = I$. For each of the UE-BS, the UE-RIS, and the RIS-BS channel, we randomly generate $I$ SPs within the space defined by $-5\,\mathrm{m} < x < 5\,\mathrm{m}, -5\,\mathrm{m} < y < 5\,\mathrm{m}, 0\,\mathrm{m} < z < 5\,\mathrm{m}$, to produce a total

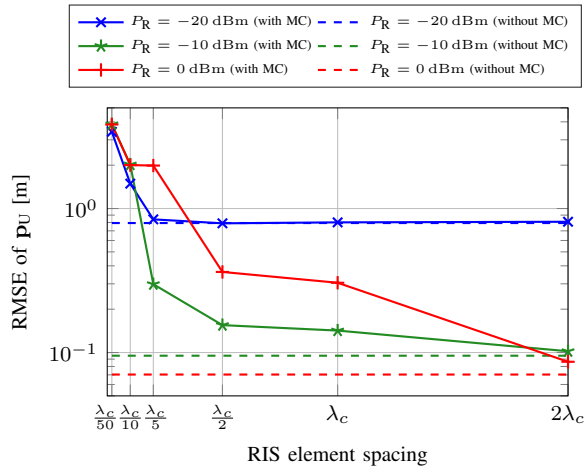



Fig. 7: Evaluation of RMSEs of $\mathbf{p}_{\mathrm{U}}$ versus RIS element spacing considering MC. Different active RIS powers $P_{\mathrm{R}} = \{-20, -10, 0\}$ dBm are tested.

of $3I$ SPs. The RCS coefficients of all NLOS paths are fixed as $c_{L,i} = c_{\mathrm{R},1,i} = c_{\mathrm{R},2,i} = 0.5\,\mathrm{m}^2,\ i = 1, \ldots, I$ [15]. The received SNR is set as 30 dB.

Fig. 6-(a) illustrates the multipath effect on the channel estimation performance of the proposed method together with benchmark methods. The parameter $\boldsymbol{\theta}_{\mathrm{R}}$ is considered as a representative. From Fig. 6-(a), we can see that the SOMP algorithm is more robust to the multipath effect, as its estimation error remains stable with the increase of SPs. On the other hand, the estimation error of the tensor-ESPRIT increases with the increase of SPs. This can be attributed to the SOMP's strategy that involves matching the atoms with the highest correlation in the dictionary, making the matching results less sensitive to weak multipath noise. However, since the proposed tensor-ESPRIT approach is based on tensor decomposition, the structured noise introduced by NLOS multipath can increase the rank of the channel tensor, which in turn directly affects the decomposition result. Nonetheless, both the performance of SOMP and tensor-ESPRIT can be effectively improved (to a similar level) by applying the LS refinement. The corresponding RMSEs of $\mathbf{p}_{\mathrm{U}}$ and $\mathbf{p}_{\mathrm{R}}$ are shown in Fig. 6-(b). It is clearly shown that after the proposed LS refinement and running the localization algorithm, the SOMP and tensor-ESPRIT reach a similar localization accuracy. The more severe the multipath effect, the higher the estimation errors for both methods. It is worth noting that in environments with sparse NLOS multipath (e.g., $I < 6$) that most mmWave/THz wireless systems can satisfy [3], [4], the final positioning accuracy remains very close to that of the multipath-free case (i.e., $I = 0$), which demonstrates the robustness of the proposed LS refinement and localization algorithm in both cases of initialization using SOMP and tensor-ESPRIT.

*4) Impact of RIS Mutual Coupling:* As previously mentioned in Subsection II-A, the utilization of active RIS amplifies the impact of the MC among RIS elements, rendering it unignorable. When MC is taken into account, the reflection matrix of RIS (denoted as $\tilde{\boldsymbol{\Gamma}}_g$) is given by [56], [57]

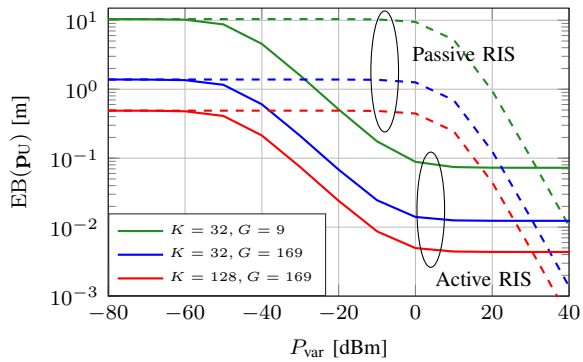
$$\tilde{\boldsymbol{\Gamma}}_g = (\boldsymbol{\Gamma}_g^{-1} - \mathbf{S})^{-1},\ g = 1, \ldots, G, \tag{87}$$

where $\mathbf{S}$ denotes the scattering matrix of RIS elements. According to microwave network theory [56], [58], the scattering matrix $\mathbf{S}$ is given by $\mathbf{S} = (\mathbf{Z} + Z_0\mathbf{I})^{-1}(\mathbf{Z} - Z_0\mathbf{I})$, where $\mathbf{Z}$ denotes the impedance matrix of RIS elements and $Z_0$ is the reference impedance (typically $Z_0 = 50\,\Omega$). In general, the matrices $\mathbf{S}$ and $\mathbf{Z}$ can be acquired through standard electromagnetic solvers such as CST Microwave Studio [33]. For the sake of simulation convenience, we adopt the analytical model in [59]. By assuming all the RIS antennas are cylindrical thin wires of perfectly conducting material, the mutual impedances between every pair of scattering elements of RIS can be explicitly calculated using [59, Eq. (2)] or [60, Eq. (3)]. The results in [60], [61] reveal that a denser integration of RIS elements generally generates a greater impact on, e.g., received signal power and channel estimation performance.

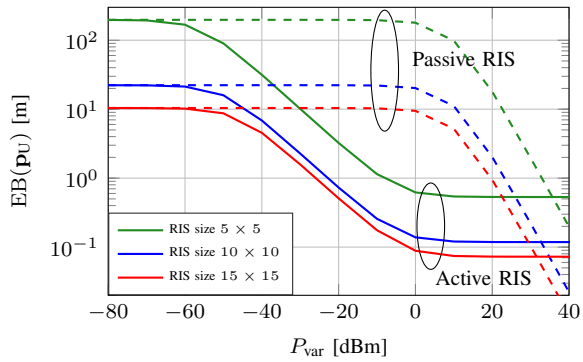
Fig. 7 presents the RMSEs of $\mathbf{p}_{\mathrm{U}}$ versus RIS element spacings for different active RIS powers $P_{\mathrm{R}} = \{-20, -10, 0\}$ dBm. The cases with MC use the reflection matrix $\tilde{\mathbf{\Gamma}}_g$ in (87), while the cases without MC use $\mathbf{\Gamma}_g$. To obtain the best performance, we perform both tensor-ESPRIT and SOMP at the coarse channel estimation stage and choose the result with lower residual error in (60) to initialize the LS refinement; then the localization parameters are obtained through Algorithm 2. We fix the received SNR as 30 dB, and the other parameters are set according to Table I. It can be observed that the shorter the RIS element spacing, the higher the estimation RMSE, which coincides with the results in [60], [61]. Furthermore, the gap between the cases with and without MC increases as we enlarge $P_{\mathrm{R}}$ at fixed RIS element spacing, revealing that a higher active RIS power accentuates the impact of MC. A noteworthy phenomenon is that in the absence of MC, the higher the active RIS power, the lower the estimation error. But this rule no longer holds when MC is considered. Higher RIS power helps to increase the signal strength but also amplifies the impact of MC, which implies that increasing the RIS power is not always beneficial. For instance, the case with $P_{\mathrm{R}} = -10\,\mathrm{dBm}$ can provide better localization accuracy than $P_{\mathrm{R}} = 0\,\mathrm{dBm}$ when the RIS element spacing is less than $\lambda_c$. This result reveals that an optimal active RIS power exists when MC is taken into account.

### C. Active RIS versus passive RIS

Based on the channel model (6), the active RIS provides power gain while also introducing additional noise. Therefore, the combined effect needs to be evaluated and compared with the passive case. As a representation of the localization performance, we evaluate the value of $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ over different power supplies in a multipath-free and MC-free scenario. In this trial, we define $P_{\mathrm{var}}$ denoting the additional system power. For active RIS cases, we fix the transmission power as $P_{\mathrm{T}}$ in Table I, and set $P_{\mathrm{R}} = P_{\mathrm{var}}$. The passive RIS cases are simulated by setting RIS power $P_{\mathrm{R}} = 0$ (thus $p = 1$ according to (85)) and $\sigma_r = 0$. To guarantee a fair comparison, we set the transmission power in the passive RIS cases as $P_{\mathrm{T}} + P_{\mathrm{var}}$, thus the total power supply of the system stays the same as in active cases. We evaluate $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ over different values of $P_{\mathrm{var}}$ from -80 dBm to 40 dBm.



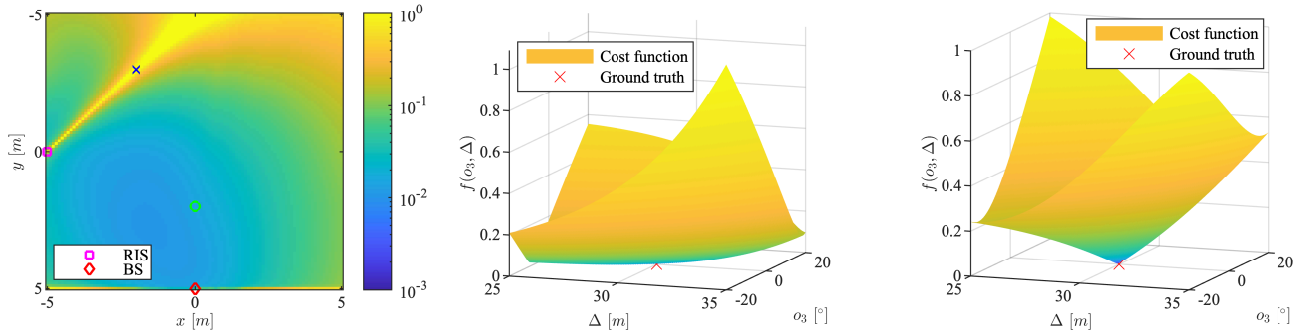

(a) $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ evaluations for different $K, G$




(b) $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ evaluations for different RIS size

Fig. 8: The evaluation of $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ versus active RIS power supply level for the active and passive RIS setups. (a) Evaluation for the cases $\{K = 32, G = 9\}$, $\{K = 32, G = 169\}$, and $\{K = 128, G = 169\}$ while other parameters are fixed as in Table I; (b) Evaluation for the cases $N_{\mathrm{R},1} \times N_{\mathrm{R},2} = 5 \times 5, 10 \times 10, 15 \times 15$ while other parameters are fixed as in Table I.

Fig. 8 (a) demonstrates the results for different numbers of subcarriers ($K = \{32, 128\}$) and transmissions ($G = \{9, 169\}$). It is clearly shown that $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ of the active RIS decreases with the increase of RIS power supply until saturating at around $P_{\mathrm{var}} = 0\,\mathrm{dBm}$. This shows that, even with the introduction of more noise, the power gain from active RIS can still provide a positive improvement in localization performance. The performance saturation for the active RIS can be explained by analyzing the noise pattern. When the RIS power is large, the noise introduced by active RIS $\mathbf{n}_r$ dominates $\mathbf{n}_0$. As the noise and signal powers at the RIS channel are boosted equally, the estimation performance saturates. In contrast, allocating the additional power to the transmitter (i.e., the passive RIS case) can continuously improve localization performance when $P_{\mathrm{var}} > 0\,\mathrm{dBm}$. However, it is also noted that in the passive RIS case, increasing $P_{\mathrm{var}}$ offers little improvement to localization performance when $P_{\mathrm{var}} < 0\,\mathrm{dBm}$. The comparison concludes that in a practical region (i.e., $-60\,\mathrm{dBm} < P_{\mathrm{var}} < 30\,\mathrm{dBm}$), allocating an extra power budget to the RIS can provide a higher localization accuracy than allocating the same power to the transmitter. Additionally, Fig. 8 (a) shows that increasing the number of subcarriers $K$ and transmissions $G$ helps to improve the localization performance for both the active and passive RIS cases. Furthermore, larger $\{K, G\}$ enables the passive RIS to outperform the active RIS starting from a lower $P_{\mathrm{var}}$ value.




(a) (b) (c)

Fig. 9: The visualization of the blind areas. (a) $\mathrm{EB}(\mathbf{p}_\mathrm{U})$ versus $\mathbf{p}_\mathrm{U}$, where the geometries of the BS and the RIS are fixed as in Table I. In addition, two sample locations of the blind area (blue cross) and non-blind area (green circle) are shown; (b) $f(o_3, \Delta)$ of the blind location in (a); (c) $f(o_3, \Delta)$ of the non-blind location in (a).

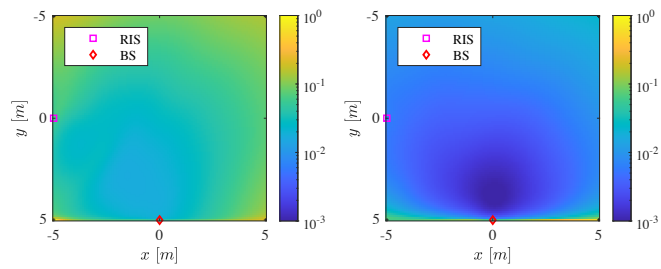
Fig. 8 (b) demonstrates the results for different RIS sizes, which reveal that increasing the RIS size can also improve the localization performance for both RIS types.

### D. The Blind Areas Analysis

*1) Blind Areas Visualization and Interpretation:* So far we have presented results for scenarios with fixed UE and RIS positions and orientations. In this subsection, we examine the spatial variability of performance. To this end, the localization CRLBs (take $\mathrm{EB}(\mathbf{p}_\mathrm{U})$ as a representative) are computed over different UE positions while the BS and RIS positions and orientations are kept fixed. We assume the UE to be placed across a $10\,\mathrm{m} \times 10\,\mathrm{m}$ space at a fixed height of $1\,\mathrm{m}$, and the BS and RIS are deployed as default parameters in Table I. The corresponding results are shown in Fig. 9 (a). From Fig. 9 (a), we can observe the presence of areas with extremely high CRLB (yellow areas). Since these areas with high CRLB yield a poor localization performance or are even unable to perform localization because of the existence of the ambiguity [27], we name those areas as the *blind areas*. In Fig. 9 (a), we select two sample locations of the blind area (blue cross) and non-blind area (green circle), respectively, for further investigation.

Fig. 9 (b) and (c) visualize the cost function $f(o_3, \Delta)$ in (67) with the noise-free observations $\boldsymbol{\eta}$ for the selected blind and non-blind locations of the UE as marked in Fig. 9 (a). The ground-truth values of $o_3$ and $\Delta$ are marked with a red cross. In both cases, we can see that the ground truth coincides with a local minimum point, which indicates that Algorithm 2 can converge to the ground truth given a proper initialization search of the interval. However, the cost function of the blind location is flat around the ground truth, while the non-blind location shows a sharper descending structure. This implies that the uncertainty in the blind location is much higher than that in the non-blind location, and so the blind locations would produce a larger estimation error than the non-blind locations in the noisy case, which results in the extremely high CRLBs in Fig. 9 (a). To avoid localization in blind areas, we propose two strategies, namely, leveraging more prior information and adding extra BSs.

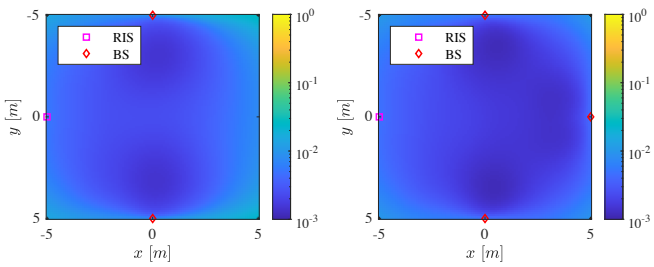
*2) Evaluation of the Impact of Extra Prior Information:* We first assess the impact of using extra prior information on unknown localization parameters. We test two types of prior information, i.e., the RIS's orientation $o_3$ and the clock bias $\Delta$. Assume that we know the values of these two parameters in advance. Then, we remove the corresponding columns in the Jacobian matrix $\mathbf{J}$ in (81) and calculate a new CRLB accordingly. The results are shown in Fig. 10. It can be seen that with prior knowledge of $o_3$ or $\Delta$, the blind area is greatly reduced. Furthermore, prior knowledge of $\Delta$ seems to provide a better performance in eliminating the blind area effect compared to using $o_3$. This can be explained by the fact that $\Delta$ contains more information than $o_3$ regarding the geometry of the UE and the RIS. For example, once we know the clock bias $\Delta$, we can determine the positions of the UE and the RIS immediately through (64)–(66), while the RIS orientation $o_3$ cannot provide further information other than itself.




(a) (b)

Fig. 10: $\mathrm{EB}(\mathbf{p}_\mathrm{U})$ versus $\mathbf{p}_\mathrm{U}$. (a) The RIS's orientation $o_3$ is known; (b) The clock bias $\Delta$ is known.

*3) Evaluation of the Impact of Additional BSs:* Finally, we evaluate the impact of adding more BSs. Fig. 11 demonstrates the value of $\mathrm{EB}(\mathbf{p}_\mathrm{U})$ for UE at different positions under different amount of BSs. In Fig. 11 (a), an additional BS is introduced at $[0, -5, 3]^\mathsf{T}$ with orientation $[0, 0, \pi/2]^\mathsf{T}$, while one more BS is added at $[5, 0, 3]^\mathsf{T}$ with orientation $[0, 0, \pi]^\mathsf{T}$ in Fig. 11 (b). For each case, we collect parallel observations $\mathbf{y}_{g,k}$ from multiple BSs, which is dependent on the same geometric parameters of the UE and the RIS. Then the CRLBs are rederived and evaluated. We can observe that adding more BSs can significantly reduce the blind areas. The more BSs we deploy, the lower the overall localization bounds.




(a) (b)

Fig. 11: $\mathrm{EB}(\mathbf{p}_{\mathrm{U}})$ versus $\mathbf{p}_{\mathrm{U}}$. (a) Add one BS at $[0,-5,3]^{\mathsf{T}}$ with orientation $[0,0,\pi/2]^{\mathsf{T}}$; (b) Add two BSs at $\{[0,-5,3]^{\mathsf{T}},[5,0,3]^{\mathsf{T}}\}$ with orientation $\{[0,0,\pi/2]^{\mathsf{T}},[0,0,\pi]^{\mathsf{T}}\}$.

## VII. Conclusion

In this paper, we formulated and solved a joint RIS calibration and user positioning problem for an active RIS-assisted uplink SIMO system, where the 3D user position, 3D RIS position, 1D RIS orientation, and clock bias were estimated. A two-stage localization method has been proposed, which consists of a coarse channel parameter estimation using tensor-ESPRIT, a channel parameters refinement via LS estimation, and a 2D search-based localization algorithm. The fundamental CRLBs for the channel parameter and localization parameter estimation were derived. Through simulation studies, we demonstrated the effectiveness of the proposed algorithms by comparing their estimation RMSEs with those of the existing SOMP algorithm and the derived CRLBs. In addition, a comparison between the performance of the active and passive RISs was carried out, which showed that active RISs outperform the passive RISs in terms of localization accuracy within practical power supply regions. Furthermore, we show that blind areas exist in the JrCUP problem, which can be interpreted by the uncertainty in the cost function. Two strategies are proposed to combat blind areas, namely, using additional prior information (e.g., from extra sensors) or deploying more BSs. These strategies have been verified by numerical simulations. Future research on the JrCUP problem includes extending the formulation to multi-user/multi-RIS scenarios, addressing 3D RIS orientation estimation, as well as developing the RIS mutual coupling-aware estimators.

## Appendix A

Consider estimating the deterministic unknowns $\mathbf{x}\in\mathbb{R}^l$ from $\tilde{\mathbf{y}}\in\mathbb{C}^m$ based on the noisy observation model

$$\tilde{\mathbf{y}}=\tilde{\mathbf{f}}(\mathbf{x})+\tilde{\mathbf{A}}\tilde{\mathbf{n}}, \tag{88}$$

where $\tilde{\mathbf{A}}\in\mathbf{C}^{m\times n}$, $\tilde{\mathbf{n}}\in\mathbb{C}^n$ and $\tilde{\mathbf{n}}\sim\mathcal{CN}(\mathbf{0},\sigma^2\mathbf{I}_n)$. Suppose $\tilde{\mathbf{y}}=\mathbf{y}_{\mathrm{R}}+j\mathbf{y}_{\mathrm{I}}$, $\tilde{\mathbf{f}}(\mathbf{x})=\mathbf{f}_{\mathrm{R}}(\mathbf{x})+j\mathbf{f}_{\mathrm{I}}(\mathbf{x})$, $\tilde{\mathbf{A}}=\mathbf{A}_{\mathrm{R}}+j\mathbf{A}_{\mathrm{I}}$ and $\tilde{\mathbf{n}}=\mathbf{n}_{\mathrm{R}}+j\mathbf{n}_{\mathrm{I}}$. By separating the real and imaginary parts of the observations $\tilde{\mathbf{y}}$, we can rewrite the model (88) as

$$\begin{bmatrix}\mathbf{y}_{\mathrm{R}}\\ \mathbf{y}_{\mathrm{I}}\end{bmatrix}=\underbrace{\begin{bmatrix}\mathbf{f}_{\mathrm{R}}(\mathbf{x})\\ \mathbf{f}_{\mathrm{I}}(\mathbf{x})\end{bmatrix}}_{\mathbf{f}(\mathbf{x})}+\underbrace{\begin{bmatrix}\mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{R}}-\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{I}}\\ \mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{I}}+\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{R}}\end{bmatrix}}_{\mathbf{n}}. \tag{89}$$

Since

$$\mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{R}}-\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{I}}\sim\mathcal{N}\Big(\mathbf{0},\frac{\sigma^2}{2}(\mathbf{A}_{\mathrm{R}}\mathbf{A}_{\mathrm{R}}^{\mathsf{T}}+\mathbf{A}_{\mathrm{I}}\mathbf{A}_{\mathrm{I}}^{\mathsf{T}})\Big), \tag{90}$$

$$\mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{I}}+\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{R}}\sim\mathcal{N}\Big(\mathbf{0},\frac{\sigma^2}{2}\underbrace{(\mathbf{A}_{\mathrm{R}}\mathbf{A}_{\mathrm{R}}^{\mathsf{T}}+\mathbf{A}_{\mathrm{I}}\mathbf{A}_{\mathrm{I}}^{\mathsf{T}})}_{=\Re(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^{\mathsf{H}})}\Big), \tag{91}$$

and

$$\begin{aligned}&\mathbb{E}\left[(\mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{R}}-\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{I}})(\mathbf{A}_{\mathrm{R}}\mathbf{n}_{\mathrm{I}}+\mathbf{A}_{\mathrm{I}}\mathbf{n}_{\mathrm{R}})^{\mathsf{T}}\right]\\ =&\frac{\sigma^2}{2}(\mathbf{A}_{\mathrm{R}}\mathbf{A}_{\mathrm{I}}^{\mathsf{T}}-\mathbf{A}_{\mathrm{I}}\mathbf{A}_{\mathrm{R}}^{\mathsf{T}})=\frac{\sigma^2}{2}\Im(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^{\mathsf{H}})^{\mathsf{T}},\end{aligned} \tag{92}$$

we have

$$\mathbf{n}\sim\mathcal{N}\left(\mathbf{0},\underbrace{\frac{\sigma^2}{2}\begin{bmatrix}\Re\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right) & \Im\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right)^{\mathsf{T}}\\ \Im\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right) & \Re\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right)\end{bmatrix}}_{\mathbf{C}_n}\right). \tag{93}$$

Then the FIM of $\mathbf{x}$ can be derived as [49, B.3.3]

$$\mathbf{J}(\mathbf{x})=\begin{bmatrix}\Re(\frac{\partial\tilde{\mathbf{f}}(\mathbf{x})}{\partial\mathbf{x}})\\ \Im(\frac{\partial\mathbf{f}(\mathbf{x})}{\partial\mathbf{x}})\end{bmatrix}^{\mathsf{T}}\mathbf{C}_n^{-1}\begin{bmatrix}\Re(\frac{\partial\tilde{\mathbf{f}}(\mathbf{x})}{\partial\mathbf{x}})\\ \Im(\frac{\partial\mathbf{f}(\mathbf{x})}{\partial\mathbf{x}})\end{bmatrix}. \tag{94}$$

Letting $\mathbf{A}_0=\mathbf{W}^{\mathsf{H}}$, $\mathbf{A}_r^{g,k}=\mathbf{W}^{\mathsf{H}}\mathbf{H}_{\mathrm{R},2}^k\mathbf{\Gamma}_g$ and adding up the FIMs over all the transmissions $g=1,\ldots,G$ and subcarriers $k=1,\ldots,K$ yield Proposition 2.

## References


[1] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

[2] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. D. Renzo, A. Lee Swindlehurst, R. Zhang, and A. Y. Zhang, "An overview of signal processing techniques for RIS/IRS-Aided wireless systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 883–917, 2022.

[3] J. He, F. Jiang, K. Keykhosravi, J. Kokkoniemi, H. Wymeersch, and M. Juntti, "Beyond 5G RIS mmWave systems: Where communication and localization meet," *IEEE Access*, vol. 10, pp. 68 075–68 084, 2022.

[4] H. Sarieddeen, M.-S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for terahertz communications," *Proceedings of the IEEE*, vol. 109, no. 10, pp. 1628–1665, 2021.

[5] H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M.-S. Alouini, and T. Y. Al-Naffouri, "A tutorial on terahertz-band localization for 6G communication systems," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1780–1815, 2022.

[6] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. Vincent Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1707–1725, 2023.

[7] R. Schroeder, J. He, G. Brante, and M. Juntti, "Two-stage channel estimation for hybrid RIS assisted MIMO systems," *IEEE Transactions on Communications*, vol. 70, no. 7, pp. 4793–4806, 2022.

[8] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3083–3098, 2022.

[9] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018.

[10] P. Zheng, X. Liu, T. Ballal, and T. Y. Al-Naffouri, "5G-aided RTK positioning in GNSS-deprived environments," in *IEEE European Signal Processing Conference (EUSIPCO)*, 2023.

[11] X. Fang, X. Li, and L. Xie, "3-D distributed localization with mixed local relative measurements," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5869–5881, 2020.

[12] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, 2022.

[13] H. Chen, H. Kim, M. Ammous, G. Seco-Granados, G. C. Alexandropoulos, S. Valaee, and H. Wymeersch, "RISs and sidelink communications in smart cities: The key to seamless localization and sensing," *accepted by IEEE Communications Magazine*, 2023.

[14] H. Kim, H. Chen, M. F. Keskin, Y. Ge, K. Keykhosravi, G. C. Alexandropoulos, S. Kim, and H. Wymeersch, "RIS-enabled and access-point-free simultaneous radio localization and mapping," *preprint arXiv:2212.07141*, 2022.

[15] H. Chen, P. Zheng, M. F. Keskin, T. Al-Naffouri, and H. Wymeersch, "Multi-RIS-enabled 3D sidelink positioning," *preprint arXiv:2302.12459*, 2023.

[16] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, P. Popovski, and H. Wymeersch, "RIS-Enabled SISO localization under user mobility and spatial-wideband effects," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1125–1140, 2022.

[17] J. He, A. Fakhreddine, and G. C. Alexandropoulos, "Simultaneous indoor and outdoor 3D localization with STAR-RIS-assisted millimeter wave systems," in *IEEE Vehicular Technology Conference (VTC)*, 2022.

[18] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Metalocalization: Reconfigurable intelligent surface aided multi-user wireless indoor localization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7743–7757, 2021.

[19] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5386–5402, 2021.

[20] A. Fascista, M. F. Keskin, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "RIS-aided joint localization and synchronization with a single-antenna receiver: Beamforming design and low-complexity estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1141–1156, 2022.

[21] P. Gao, L. Lian, and J. Yu, "Wireless area positioning in RIS-Assisted mmWave systems: Joint passive and active beamforming design," *IEEE Signal Processing Letters*, vol. 29, pp. 1372–1376, 2022.

[22] M. F. Keskin, F. Jiang, F. Munier, G. Seco-Granados, and H. Wymeersch, "Optimal spatial signal design for mmWave positioning under imperfect synchronization," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5558–5563, 2022.

[23] P. Zheng, H. Chen, T. Ballal, H. Wymeersch, and T. Y. Al-Naffouri, "Misspecified Cramér-Rao bound of RIS-aided localization under geometry mismatch," in *IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*, 2023.

[24] D.-R. Emenonye, H. S. Dhillon, and R. M. Buehrer, "RIS-Aided localization under position and orientation offsets in the near and far field," *preprint arXiv:2210.03599*, 2022.

[25] M. Samir, M. Elhattab, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3978–3983, 2021.

[26] L. Ge, H. Zhang, J.-B. Wang, and G. Y. Li, "Reconfigurable wireless relaying with Multi-UAV-carried intelligent reflecting surfaces," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 4932–4947, 2023.

[27] Y. Lu, H. Chen, J. Talvitie, H. Wymeersch, and M. Valkama, "Joint RIS calibration and multi-user positioning," in *IEEE Vehicular Technology Conference (VTC)*, 2022.

[28] R. Ghazalian, H. Chen, G. C. Alexandropoulos, G. Seco-Granados, H. Wymeersch, and R. Jäntti, "Joint user localization and location calibration of a hybrid reconfigurable intelligent surface," *preprint arXiv:2210.10150*, 2022.

[29] G. Mylonopoulos, C. D'Andrea, and S. Buzzi, "Active reconfigurable intelligent surfaces for user localization in mmWave MIMO systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2022.

[30] G. Mylonopoulos, L. Venturino, S. Buzzi, and C. D'Andrea, "Maximum-likelihood user localization in active-RIS empowered mmWave wireless networks," in *17th European Conference on Antennas and Propagation*, 2023.

[31] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849–5863, 2020.

[32] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4962–4975, 2021.

[33] J. Rao, Y. Zhang, S. Tang, Z. Li, C.-Y. Chiu, and R. Murch, "An active reconfigurable intelligent surface utilizing phase-reconfigurable reflection amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 71, no. 7, pp. 3189–3202, 2023.

[34] Z. Peng, X. Liu, C. Pan, L. Li, and J. Wang, "Multi-pair D2D communications aided by an active RIS over spatially correlated channels with phase noise," *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2090–2094, 2022.

[35] S. Saab, A. Mezghani, and R. W. Heath, "Optimizing the mutual information of frequency-selective multi-port antenna arrays in the presence of mutual coupling," *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 2072–2084, 2022.

[36] M. Di Renzo, F. H. Danufane, and S. Tretyakov, "Communication models for reconfigurable intelligent surfaces: From surface electromagnetics to wireless networks optimization," *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1164–1209, 2022.

[37] Y. Mitikiri and K. Mohseni, "Acceleration compensation for gravity sense using an accelerometer in an aerodynamically stable UAV," in *IEEE Conference on Decision and Control*, 2019, pp. 1177–1182.

[38] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. PHAN, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[39] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.

[40] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-Based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, 2008.

[41] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.

[42] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmWave MIMO systems via Tensor-ESPRIT algorithms in DFT beamspace," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 816–831, 2021.

[43] S. Jiang, N. Fu, Z. Wei, X. Li, L. Qiao, and X. Peng, "Joint spectrum, carrier, and DOA estimation with beamforming MWC sampling system," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.

[44] N. Fu, Z. Wei, L. Qiao, and Z. Yan, "Short-observation measurement of multiple sinusoids with multichannel sub-Nyquist sampling," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6853–6869, 2020.

[45] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[46] F. Wen, H. C. So, and H. Wymeersch, "Tensor decomposition-based beamspace ESPRIT algorithm for multidimensional harmonic retrieval," in *IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*, 2020, pp. 4572–4576.

[47] F. Jiang, F. Wen, Y. Ge, M. Zhu, H. Wymeersch, and F. Tufvesson, "Beamspace multidimensional ESPRIT approaches for simultaneous localization and communications," *preprint arXiv:2111.07450*, 2021.

[48] F. Wen, N. Garcia, J. Kulmer, K. Witrisal, and H. Wymeersch, "Tensor decomposition based beamspace ESPRIT for millimeter wave MIMO channel estimation," in *IEEE Global Communications Conference (GLOBECOM)*, 2018.

[49] P. Stoica, R. L. Moses *et al.*, *Spectral analysis of signals*. Pearson Prentice Hall Upper Saddle River, NJ, 2005, vol. 452.

[50] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[51] S. Tarboush, A. Ali, and T. Y. Al-Naffouri, "Compressive estimation of near field channels for ultra massive-MIMO wideband THz systems," in *IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*, 2023.

[52] ——, "Cross-field channel estimation for ultra massive-MIMO THz systems," *preprint arXiv:2305.13757*, 2023.

[53] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[54] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, no. 42, pp. 1455–1459, 2014. [Online]. Available: https://www.manopt.org

[55] F. Athley, "Threshold region performance of maximum likelihood direction of arrival estimators," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1359–1373, 2005.

[56] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 1229–1243, 2022.

[57] D. Wijekoon, A. Mezghani, and E. Hossain, "Beamforming optimization in RIS-aided MIMO systems under multiple-reflection effects," in *IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*, 2023.

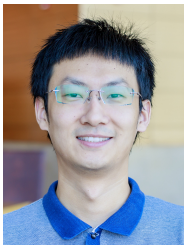
[58] D. M. Pozar, *Microwave engineering*. John wiley & sons, 2011.

[59] M. Di Renzo, V. Galdi, and G. Castaldi, "Modeling the mutual coupling of reconfigurable metasurfaces," in *17th European Conference on Antennas and Propagation*, 2023.

[60] P. Zheng, X. Ma, and T. Y. Al-Naffouri, "On the impact of mutual coupling on RIS-assisted channel estimation," *preprint arXiv:2309.04990*, 2023.

[61] X. Qian and M. D. Renzo, "Mutual coupling and unit cell aware optimization for reconfigurable intelligent surfaces," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1183–1187, 2021.

**Pinjun Zheng** received the B.S. and M.S. degree from the Harbin Institute of Technology, Harbin, China. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. His research interests include 5G/6G localization and communication.



**Hui Chen** received the B.S. degree in electrical engineering from Beijing Forestry University, Beijing, China, in 2013, the M.S. Degree in computer engineering from University of Chinese Academy of Sciences (UCAS), Beijing, China, in 2016, and the Ph.D. Degree in Electrical and Computer Engineering from King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2021. He is currently a senior researcher at Technology Innovation Institute (TII), Abu Dhabi, UAE. From 2021 to 2023, he worked as a Postdoctoral Researcher at Chalmers University of Technology, Gothenburg, Sweden. His current research interests include 5G/6G (mmWave/THz) and RIS-aided localization.

**Tarig Ballal** received a B.Sc. degree with honors in electrical engineering from the University of Khartoum, Sudan, in 2001, and an M.Sc. degree in telecommunications from Blekinge Institute of Technology, Sweden, in 2005. He completed his Ph.D. degree from the School of Computer Science and Informatics, University College Dublin, Ireland, in 2011. In the period from April 2011 to July 2012, Dr. Ballal worked as a research engineer with BiancaMed Ltd. and University College Dublin. Since September 2012, he has been with the Electrical Engineering Department, King Abdullah University of Science and Technology (KAUST), Saudi Arabia, as a postdoctoral fellow and currently as a research scientist. His current research focuses on signal processing, localization, and GNSS positioning and attitude determination.

**Mikko Valkama** received his M.Sc. (Tech.) and D.Sc. (Tech.) degrees (both with honors) from Tampere University of Technology, Finland, in 2000 and 2001, respectively. In 2003, he was with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA, as a visiting research fellow. Currently, he is a Full Professor and the Head of the Unit of Electrical Engineering at the newly formed Tampere University, Finland. His general research interests include radio communications, radio localization, and radio-based sensing, with particular emphasis on 5G and 6G mobile radio networks.

**Henk Wymeersch** obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. He is also a Distinguished Research Associate with Eindhoven University of Technology. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication Letters (2009-2013), IEEE Transactions on Wireless Communications (since 2013), and IEEE Transactions on Communications (2016-2018) and is currently Senior Member of the IEEE Signal Processing Magazine Editorial Board. During 2019-2021, he was an IEEE Distinguished Lecturer with the Vehicular Technology Society. His current research interests include the convergence of communication and sensing, in a 5G and Beyond 5G context.

**Tareq Al-Naffouri** received the B.S. degrees in mathematics and electrical engineering (with first honors) from King Fahd University of Petroleum and Minerals, Saudi Arabia, the M.S. degree in electrical engineering from the Georgia Institute of Technology, and the Ph.D. degree in electrical engineering from Stanford University, Stanford in 2004. He was a visiting scholar at California Institute of Technology, Pasadena, CA in 2005 and summer 2006. He was a Fulbright scholar at the University of Southern California in 2008. He is currently a Professor at the Electrical Engineering Department, King Abdullah University of Science and Technology (KAUST). His research interests lie in the areas of sparse, adaptive, and statistical inference/learning and their applications to wireless communications, localization, smart cities, and smart health. He has over 370 publications in journal and conference proceedings and 24 issued/pending patents. He has won the IEEE Education Society Chapter Achievement Award (2008), Almarie Award for Innovative research in communication (2009), and Abdul Hameed Shoman Prize for innovative research in IoT (2022).