Channel Estimation for RIS via Tensor-ESPRIT in DFT Beamspace (CERISE)

Paul Armand Ugo Lysandre Göttlich, *Damir Rakhimov*, and *Martin Haardt*
Communications Research Laboratory, Ilmenau University of Technology
P. O. Box 100565, D-98684 Ilmenau, Germany
Email: {paul-armand-ugo-lysandre.goettlich, damir.rakhimov, martin.haardt}@tu-ilmenau.de

***Abstract*—The integration of Reconfigurable Intelligent Surfaces (RISs) into wireless systems presents a transformative opportunity to redefine communication strategies by actively shaping the propagation environment. Despite the significant potential of RISs to enhance the spectral efficiency and the spatial coverage, their implementation faces several challenges, such as channel estimation. In this paper, we propose a novel framework for Channel Estimation for RISs via Tensor-ESPRIT in DFT Beamspace (CERISE). By exploiting the sparsity of the beamformed channels and introducing a two-stage estimation scheme, CERISE can achieve high resolution estimates of the target parameters. The performance of CERISE is verified via numerical simulations, proving its ability to provide reliable Channel State Information (CSI).**

I. Introduction

Reconfigurable Intelligent Surfaces (RISs) have recently emerged as a transformative technology in wireless communications. At the same time, they have a long history, originating from early studies on smart reflect-arrays and passive beamforming structures [1]. The integration of RISs into modern wireless networks brings several notable advantages, such as enhancing signal coverage, improving spectral efficiency, and enabling fine-grained control over the radio environment. This can be implemented via dynamical adjustments of the phases of the impinging signals at the RIS [2]. Multiple works have highlighted the advantages of including RISs in the environment, such as the "square law" [3], which results in a scaled received power of signals by the square of the number of elements present at the RIS. These features render RISs an attractive enabler for realizing smart radio environments in 6G and beyond wireless networks, especially in massive machine-type communications (mMTC) and enhanced mobile broadband (EMB) scenarios.

However, there are several significant challenges on the way to a practical deployment. Among them, we can distinguish the problem of acquisition of CSI. It is crucial for optimizing the phase shift coefficients of the RIS to ensure reliable communications. At the same time, conventional techniques for channel estimation are inapplicable due to the high dimensionality of the channels and the passive nature of RISs [4].

There exists a variety of channel estimation methods based on different approaches such as Expectation-Maximization (EM) or Maximum-Likelihood-Estimation (MLE) [5], Compressive Sensing (CS) [6] or heuristic approaches, e.g., Monopulse [7]. However, there is only a limited number of studies in this field covering subspace-based techniques for channel estimation.

In this paper, we propose a gridless channel estimation algorithm based on ESPRIT in DFT beamspace for a Millimeter Wave (mmWave) multiple-input-multiple-output (MIMO) system operating in the far field. To this end, a generic Tensor-ESPRIT in DFT beamspace algorithm is developed. The proposed algorithm focuses on a single-user scenario, aiming to exploit the structure of the RIS-assisted channels for an efficient estimation of the target parameters. It comprises two-stages allowing to substantially reduce training overhead and computational complexity.

This paper has the following structure. First, we introduce a data model for the tensor representation of the measured signals. Then we review subspace-based DoA/DoD estimation techniques and explain the proposed algorithm. Lastly, numerical results are presented to demonstrate the system's performance.

II. Data model

We consider a single-user (SU) multiple-input-multiple-output (MIMO) communication system assisted by an RIS. It operates in the far field and is equipped with multiple transmit and receive antennas, M_{T} and M_{R}, respectively, and an RIS with M_{RIS} elements. The notation in this paper follows the form of [8]. The array elements at the transmitter, receiver, and RIS are arranged in the form of uniform linear arrays (ULAs). It is further assumed that a Line-of-Sight (LOS) connection between the transmitter and the receiver is unavailable, but they are communicating via the RIS visible by both sides. In this case, the channel tensor $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}} \times M_{\mathrm{RIS}}}$ can be expressed as

$$\boldsymbol{\mathcal{H}} = \boldsymbol{a}_{\mathrm{R}} \circ \boldsymbol{a}_{\mathrm{T}}^{*} \circ \boldsymbol{a}_{\mathrm{RIS}}^{*} \circ \rho = \boldsymbol{\mathcal{I}}_{4,1} \times_1 \boldsymbol{a}_{\mathrm{R}} \times_2 \boldsymbol{a}_{\mathrm{T}}^{*} \times_3 \boldsymbol{a}_{\mathrm{RIS}}^{*} \times_4 \rho, \quad (1)$$

where $\rho = \alpha\beta \in \mathbb{C}$ denotes the overall path gain coefficient, α is the path gain of the channel from the transmitter to the RIS, and β is the path gain of the channel from the RIS to the receiver. The resulting rank-one structure is assumed to be known a priori. We assume a quasi-stationary environment such that ρ stays constant during training. The vector $\boldsymbol{a}_r(\mu_r) = \begin{bmatrix} 1 & e^{j\mu_r} & \cdots & e^{j(M_r-1)\mu_r} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{M_r \times 1}$, $r \in \{\mathrm{R}, \mathrm{T}, \mathrm{RIS}\}$ is the array steering vector with a Vandermonde structure. The spatial frequency $\mu_r = \frac{2\pi}{\lambda_{\mathrm{CR}}} \Delta \sin\varphi_r$ is defined by the angle of arrival (AoA) or departure (AoD) φ_r, and $\Delta = \lambda_{\mathrm{CR}}/2$ denotes

The work of P.A.U.L. Göttlich was supported in part by the Friedrich Naumann Foundation For Freedom.
The work of D. Rakhimov and M. Haardt was supported in part by the German Research Foundation (DFG) under Grant 402834619 (AdAMMM-II, HA 2239/14-3).

the spacing between two array elements which is defined by the wavelength λ_{CR}.

A visual interpretation of the system model is given in Figure 1a, where $\{\texttt{imp}, \texttt{ex}\}$ denotes the impinging and excitation directions.

The effective channel including the RIS can be expressed as

$$\boldsymbol{H}_e = \boldsymbol{a}_{\text{R}} \underbrace{\boldsymbol{a}_{\text{RIS}}^{\text{ex}\,\text{H}} \boldsymbol{\Phi} \boldsymbol{a}_{\text{RIS}}^{\text{imp}}}_{h_{\text{RIS}}} \boldsymbol{a}_{\text{T}}^{\text{H}} \in \mathbb{C}^{M_{\text{R}} \times M_{\text{T}}}, \tag{2}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{M_{\text{RIS}} \times M_{\text{RIS}}}$ defines the RIS configuration. It can be written as

$$\boldsymbol{\Phi} = \text{diag}(\boldsymbol{\phi}), \ \boldsymbol{\phi} = \begin{bmatrix} \beta_1 e^{j\alpha_1} & \cdots & \beta_{M_{\text{RIS}}} e^{j\alpha_{M_{\text{RIS}}}} \end{bmatrix}^{\text{T}} \in \mathbb{C}^{M_{\text{RIS}}}, \tag{3}$$

where we assume that there is no coupling between the individual RIS elements. Additionally, due to the previous far field and LOS assumptions, the $\boldsymbol{\phi}$ in (3) can be simplified to

$$\boldsymbol{\phi} = \begin{bmatrix} 1 & e^{j\alpha} & \cdots & e^{j(M_{\text{RIS}}-1)\alpha} \end{bmatrix}^{\text{T}}, \tag{4}$$

where we have assumed $\beta_i = 1$, and a progressive structure of the phases, i.e., $\alpha_i = (i-1)\alpha$, $\forall i = \{1, \ldots, M_{\text{RIS}}\}$. This definition for the RIS coefficients allows us to reformulate h_{RIS} in (2) using the Hadamard product as

$$\boldsymbol{a}_{\text{RIS}}^{\text{ex}\,\text{H}} \boldsymbol{\Phi} \boldsymbol{a}_{\text{RIS}}^{\text{imp}} = \boldsymbol{\phi}^{\text{T}} \left(\boldsymbol{a}_{\text{RIS}}^{\text{ex}\,*} \odot \boldsymbol{a}_{\text{RIS}}^{\text{imp}} \right) = \boldsymbol{\phi}^{\text{T}} \boldsymbol{a}_{\text{RIS}}^{*}, \tag{5}$$

where $\boldsymbol{a}_{\text{RIS}} = \boldsymbol{a}(\mu_{\text{RIS}})$ is the combined RIS steering vector, with a Vandermonde structure. The obtained RIS spatial frequency $\mu_{\text{RIS}} = \mu_{\text{RIS}}^{\text{ex}} - \mu_{\text{RIS}}^{\text{imp}}$ denotes the difference between the spatial frequencies of the impinging and excitation signal. The direction of the excitation signal after the RIS can be controlled by choosing an appropriate configuration in (5). In this work, we use sets of DFT beamformers for precoding, decoding, and RIS coefficients.

The q^{th} DFT beamformer in the r^{th} mode is represented by the scaled q^{th} column of a DFT matrix $\boldsymbol{W}_r \in \mathbb{C}^{M_r \times B_r}$, $r \in \{\text{R, T, RIS}\}$ such that

$$\boldsymbol{w}_{r,q} = e^{j\left(\frac{M_r-1}{2}\right)\gamma_{r,q}} \begin{bmatrix} 1 & e^{-j\gamma_{r,q}} & e^{-j2\gamma_{r,q}} & \cdots & e^{-j(M_r-1)\gamma_{r,q}} \end{bmatrix}^{\text{T}}, \tag{6}$$

where $\gamma_{r,q} = \frac{2\pi}{M_r}(q-1)$, $\forall q = \{1, \ldots, M_r\}$ is the center of the q^{th} beam. The scalar B_r denotes the number of used beams. For the case of $B_r = M_r$, the matrix $\boldsymbol{W}_r$ is a full size scaled DFT matrix.

The mirrored reflection of the impinging signal at the RIS, which is defined by Snell's law, corresponds to the RIS configuration of a DFT beamformer with $q = 1$. It is illustrated in Figure 1b. All other beams deflect the excitation direction with a beam-specific offset $\gamma_{r,q}$ from the mirrored response.

For the proposed training scheme, we send pilots to acquire CSI. The received signal $y_{i,j,k}[l] \in \mathbb{C}$ for a pilot symbol $s_{i,j,k}[l] \in \mathbb{C}$ at the i^{th} beam at the receiver, j^{th} beam at the transmitter, k^{th} beam at the RIS, and l^{th} training frame can be written as

$$\begin{aligned} y_{i,j,k}[l] &= \boldsymbol{w}_{\text{R},i}^{\text{H}} \left(\rho \boldsymbol{a}_{\text{R}} \boldsymbol{w}_{\text{RIS},k}^{\text{T}} \boldsymbol{a}_{\text{RIS}}^{*} \boldsymbol{a}_{\text{T}}^{\text{H}} \boldsymbol{w}_{\text{T},j} s_{i,j,k}[l] + \boldsymbol{n}_{i,j,k}[l] \right) \\ &= \boldsymbol{\mathcal{H}} \times_1 \boldsymbol{w}_{\text{R},i}^{\text{H}} \times_2 \boldsymbol{w}_{\text{T},j}^{\text{T}} \times_3 \boldsymbol{w}_{\text{RIS},k}^{\text{T}} \times_4 s_{i,j,k}[l] + \boldsymbol{w}_{\text{R},i}^{\text{H}} \boldsymbol{n}_{i,j,k}[l], \end{aligned}$$

where $\{\boldsymbol{w}_{\text{R},i}, \boldsymbol{w}_{\text{T},j}, \boldsymbol{w}_{\text{RIS},k}\} \in \mathbb{C}^{M_r \times 1}$, $r \in \{\text{R, T, RIS}\}$ are the beamforming vectors at the receiver, transmitter, and RIS drawn from the DFT beamforming matrices $\boldsymbol{W}_{\text{R}}, \boldsymbol{W}_{\text{T}}, \boldsymbol{W}_{\text{RIS}} \in \mathbb{C}^{M_r \times B_r}$, respectively. The scalars M_r, B_r denote the number of sensor elements and beams in each mode, respectively, and $i = \{1, \ldots, B_{\text{R}}\}, j = \{1, \ldots, B_{\text{T}}\}, k = \{1, \ldots, B_{\text{RIS}}\}, l = \{1, \ldots, N_{\text{TF}}\}$. The scalar N_{TF} is the number of training frames, where each training frame comprises a full iteration cycle through all beam triplets, and $\boldsymbol{n}_{i,j,k}[l] \in \mathbb{C}^{M_{\text{R}} \times 1}$ is the zero mean circularly symmetric complex Gaussian (ZMCSCG) noise vector with mean $\mathbb{E}\{\boldsymbol{n}\} = \boldsymbol{0}_{M_{\text{R}}}$ and covariance matrix $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^{\text{H}}\} = \sigma_n^2 \boldsymbol{I}_{M_{\text{R}}}$.

We assume a complex-valued, constant-modulus sequence of mutually known pilots similar to [8] as

$$|s_n|^2 = \frac{P_s}{M_{\text{T}}}, \tag{7}$$

such that $\|\boldsymbol{w}_{\text{T},j} s_n\|_2^2 = P_s$, where $P_s = \frac{P}{N}$ is the power per pilot symbol, and the total training power is P. The scalar $N = N_{\text{TF}} B_{\text{R}} B_{\text{T}} B_{\text{RIS}}$ is the total number of pilot symbols, and $n = \{1, \ldots, N\}$ denotes the index of the current pilot symbol.

The demodulated signal $\tilde{y}_{i,j,k}[l] = y_{i,j,k}[l] \cdot s^* \frac{M_{\text{T}}}{P_s}$ can be written as

$$\tilde{y}_{i,j,k}[l] = \boldsymbol{\mathcal{H}} \times_1 \boldsymbol{w}_{\text{R},i}^{\text{H}} \times_2 \boldsymbol{w}_{\text{T},j}^{\text{T}} \times_3 \boldsymbol{w}_{\text{RIS},k}^{\text{T}} + \tilde{n}_{i,j,k}[l], \tag{8}$$

where $\tilde{n}_{i,j,k}[l] = \boldsymbol{w}_{\text{R},i}^{\text{H}} \boldsymbol{n}_{i,j,k}[l] s_{i,j,k}^*[l] \frac{M_{\text{T}}}{P_s}$ is the effective noise after demodulation. The collection of demodulated signals for all pilots during the l^{th} training frame can be represented as $\tilde{\boldsymbol{\mathcal{Y}}}_l \in \mathbb{C}^{B_{\text{R}} \times B_{\text{T}} \times B_{\text{RIS}}}$ which can be written as

$$\begin{aligned} \tilde{\boldsymbol{\mathcal{Y}}}_l &= \boldsymbol{\mathcal{H}} \times_1 \boldsymbol{W}_{\text{R}}^{\text{H}} \times_2 \boldsymbol{W}_{\text{T}}^{\text{T}} \times_3 \boldsymbol{W}_{\text{RIS}}^{\text{T}} + \tilde{\boldsymbol{\mathcal{N}}}_l \\ &= \underbrace{\boldsymbol{\mathcal{I}}_{4,1} \times_1 \boldsymbol{W}_{\text{R}}^{\text{H}} \boldsymbol{a}_{\text{R}} \times_2 \boldsymbol{W}_{\text{T}}^{\text{T}} \boldsymbol{a}_{\text{T}}^{*} \times_3 \boldsymbol{W}_{\text{RIS}}^{\text{T}} \boldsymbol{a}_{\text{RIS}}^{*}}_{\boldsymbol{\mathcal{B}}} \times_4 \rho + \tilde{\boldsymbol{\mathcal{N}}}_l, \end{aligned} \tag{9}$$

where $\tilde{\boldsymbol{\mathcal{N}}}_l \in \mathbb{C}^{B_{\text{R}} \times B_{\text{T}} \times B_{\text{RIS}}}$ is the effective noise tensor after demodulation. The beamspace steering tensor $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{B_{\text{R}} \times B_{\text{T}} \times B_{\text{RIS}}}$ consists of the beamspace steering vectors $\boldsymbol{b}_r \in \mathbb{C}^{B_r \times 1}$, $\forall r \in \{\text{R, T, RIS}\}$ such that

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{I}}_{3,1} \times_1 \boldsymbol{b}_{\text{R}} \times_2 \boldsymbol{b}_{\text{T}}^{*} \times_3 \boldsymbol{b}_{\text{RIS}}^{*}, \tag{10}$$

where

$$\boldsymbol{b}_r = \boldsymbol{W}_r^{\text{H}} \boldsymbol{a}_r, \ \forall r \in \{\text{R, T, RIS}\}. \tag{11}$$

III. Review of Algorithms

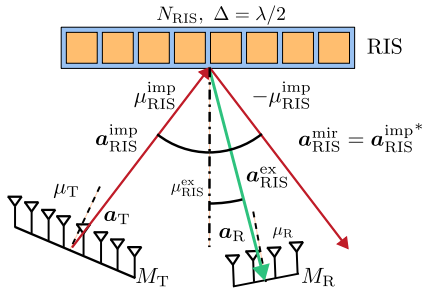
A. 1D Standard Beamspace ESPRIT

Standard ESPRIT in DFT Beamspace ESPRIT (SBE) exploits the shift invariance equation for parameter estimation as shown in [9]

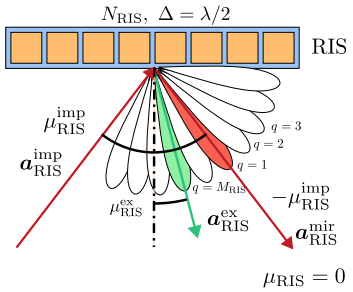
$$\boldsymbol{G}_{1,B} \boldsymbol{b} \omega = \boldsymbol{G}_{2,B} \boldsymbol{b}, \tag{12}$$

where $\omega = e^{j\mu}$ contains the phase information of the source impinging on the antenna array. For this work, we use the simplified notation for a single source. The vector $\boldsymbol{b} \in \mathbb{C}^B$ denotes the beamspace steering vector, and the $\boldsymbol{G}_{i,B} \in \mathbb{C}^{(B-1) \times B}$, $\forall i \in \{1, 2\}$ denote the beamspace selection matrices [10]. We assume that only B beams which comprise the Sector of Interest (SoI) are used for parameter estimation. As a result, SBE operates on a reduced dimensional signal space, which allows to lower the computational complexity and the number of training symbols. However, it requires prior knowledge about the SoI.

Furthermore, the beamspace steering vector $\boldsymbol{b} \in \mathbb{C}^B$ is not available during estimation, but it can be substituted by a



(a) Relationships between impinging, mirrored and desired excitation signal at the RIS.



(b) Interpretation of beamforming coefficients deflecting from the mirrored response at the RIS.

Fig. 1: System Model of a RIS assisted Channel.

vector $\boldsymbol{u}_S \in \mathbb{C}^B$, sharing the same signal subspace with $\boldsymbol{b}$, such that $\boldsymbol{b} = \boldsymbol{u}_S t$, where $t \in \mathbb{C}$. The invariance equation can then be expressed as $\boldsymbol{G}_{1,B}\boldsymbol{u}_S\omega = \boldsymbol{G}_{2,B}\boldsymbol{u}_S$, where B denotes the number of selected beams. The scalar ω can be computed by, e.g., Least Squares, $\omega = (\boldsymbol{G}_{1,B}\boldsymbol{u}_S)^{+}\boldsymbol{G}_{2,B}\boldsymbol{u}_S \in \mathbb{C}$. The target parameters μ can then be estimated from ω where the target parameters are obtained as $\hat{\mu} = \arg\{\omega\}$.

B. Maximum Power Beam Algorithm

The Maximum Power Beam Algorithm (MPB), also referred to as Beam Sweeping [11], returns the spatial frequency corresponding to the center of the beam with the largest power. In the case where multiple snapshots are available, additional pre-processing is included, which accumulates the received power across different snapshots.

C. Tensor Power Method

The Tensor Power Method (TPM) is an effective algorithm to compute the rank-one Canonical Polyadic Decomposition (CPD) [12], [13]. One iteration of the TPM algorithm can be written as

$$\boldsymbol{f}_v = \boldsymbol{\mathcal{Y}} \bigtimes_{\substack{r=1 \\ r\neq v}}^{V} \frac{\boldsymbol{f}_r^{\text{H}}}{\|\boldsymbol{f}_r\|_2^2},\ \forall v = \{1, \ldots, V\}, \tag{13}$$

where V denotes the number of modes of $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{M_1 \times \cdots \times M_V}$, and $\boldsymbol{f}_r \in \mathbb{C}^{M_r \times 1}$ is the factor vector of the r^{th} mode. An initial set of factor vectors can either be initialized with prior knowledge or generated randomly. Since the resulting factor vectors span the signal subspaces of the tensor, the TPM can be seen as a low complexity replacement for a rank-one HOSVD.

The TPM algorithm includes two stopping criteria, a maximum number of iterations Iter_{Max} and a threshold for the change of the relative error between two iterations $\delta_{\text{rel}} = \frac{|\delta_t - \delta_{t-1}|}{\delta_0}$, where $\delta_t = \left\|\boldsymbol{\mathcal{Y}} - \hat{\boldsymbol{\mathcal{Y}}}_t\right\|_{\text{F}}$ is the decomposition error at the t^{th} iteration, $\hat{\boldsymbol{\mathcal{Y}}}_t$ is the reconstructed tensor, and δ_0 is the decomposition error given the initialization vectors.

IV. Proposed Solution

In this section, we describe the main stages of the proposed CERISE algorithm. It comprises a two stage procedure consisting of a coarse estimation stage and a fine estimation stage. For the coarse estimation, we use aperture reduction to reduce the total number of beams covering the full beamspace. For RIS elements, this task is non-trivial, however, recently proposed methods such as using an RIS with absorption properties [14] allow to realize this part of the algorithm. As a result, the width of the beams is broader, allowing a full coverage of the spatial spectrum with fewer beams, i.e., $M_r^{(\text{crs})} < M_r,\ \forall r \in \{\text{R, T, RIS}\}$. We assume that there is a total power P available for the entire training procedure, which is split among the coarse and fine stages, such that $P = P^{(\text{crs})} + P^{(\text{fine})}$. The pilot symbols are then normalized as shown in (7) with the stage specific power and number of training symbols of the coarse stage. After acquiring the demodulated measurement tensor $\tilde{\boldsymbol{\mathcal{Y}}}^{(\text{crs})} \in \mathbb{C}^{M_{\text{R}}^{(\text{crs})} \times M_{\text{T}}^{(\text{crs})} \times M_{\text{RIS}}^{(\text{crs})} \times N_{\text{TF}}^{(\text{crs})}}$, an initial coarse estimate of target parameters $\hat{\mu}_{r,\text{I}}^{(\text{crs})}$ is calculated for each mode $r \in \{\text{R, T, RIS}\}$ via MPB.

These estimates are then further refined via TPM with the following preprocessing. We use the initial coarse estimates to define a coarse SoI, by selecting the $B_r^{(\text{crs})}$ closest beams to $\hat{\mu}_{r,\text{I}}^{(\text{crs})}$ in each mode. Given this set of beam indices of the coarse SoI in each mode, we select a subset of demodulated measurements $\tilde{\boldsymbol{\mathcal{Y}}}_{\text{sub}}^{(\text{crs})} \in \mathbb{C}^{B_{\text{R}}^{(\text{crs})} \times B_{\text{T}}^{(\text{crs})} \times B_{\text{RIS}}^{(\text{crs})} \times N_{\text{TF}}^{(\text{crs})}}$ for further processing via TPM.

The factor vectors are initialized in each mode with the corresponding beamspace steering vectors $\hat{\boldsymbol{b}}_{r,\text{I}}^{(\text{crs})}(\hat{\mu}_{r,\text{I}}^{(\text{crs})})$ (11). The refined coarse estimates of the target parameters $\hat{\mu}_{r,\text{II}}^{(\text{crs})}$ are then obtained by applying SBE on each factor vector.

The resulting estimates after the coarse stage $\hat{\mu}_{r,\text{II}}^{(\text{crs})}$ are used to specify a SoI for training in the fine estimation stage for each mode. In the fine estimation stage, the full size aperture is used in each mode to regain the fine grid of beams. We select $B_r^{(\text{fine})}$ beams for the SoI, where the SoI in the r^{th} mode is centered around the closest beam to $\hat{\mu}_{r,\text{II}}^{(\text{crs})}$ such that

$$\hat{\gamma}_{r,q} \leftarrow \hat{q} = \arg\min_{q} |\gamma_{r,q} - \hat{\mu}_{r,\text{II}}^{(\text{crs})}|. \tag{14}$$

After acquiring the measurement tensor in the fine estimation stage $\tilde{\boldsymbol{\mathcal{Y}}}^{(\text{fine})} \in \mathbb{C}^{B_{\text{R}}^{(\text{fine})} \times B_{\text{T}}^{(\text{fine})} \times B_{\text{RIS}}^{(\text{fine})} \times N_{\text{TF}}^{(\text{fine})}}$, beamspace steering vectors (11) based on the coarse estimate are computed and used as initialization for the TPM similarly to the coarse stage. Fine estimates of the target parameters $\hat{\mu}_r^{(\text{fine})}$ are then obtained after applying SBE on the resulting approximates of the beamspace steering vectors.

A summary of CERISE is given in Algorithm 1.

Algorithm 1 Two Stage Channel Estimation "CERISE"

Coarse Est. Stage: Reduce aperture to $M_r^{(\mathrm{crs})}$ elements $\forall r \in \{\mathrm{R, T, RIS}\}$
Pilot transmission: Accumulate demodulated symbols $\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{crs})}$
Obtain initial estimates: $\hat{\mu}_{r,\mathrm{I}}^{(\mathrm{crs})} \leftarrow \mathrm{MPB}(\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{crs})})$
Sectorization: $\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{sub}}^{(\mathrm{crs})} \leftarrow$ Select $B_r^{(\mathrm{crs})}$ beams around $\hat{\mu}_{r,\mathrm{I}}^{(\mathrm{crs})}$
Tensor Decomposition:
    $\hat{\boldsymbol{b}}_{r,\mathrm{I}}^{(\mathrm{crs})} \leftarrow$ Calculate TPM initializations based on $\hat{\mu}_{r,\mathrm{I}}^{(\mathrm{crs})}$ (11)
    $[\hat{\boldsymbol{b}}_{\mathrm{R,II}}^{(\mathrm{crs})}\ \hat{\boldsymbol{b}}_{\mathrm{T,II}}^{(\mathrm{crs})}\ \hat{\boldsymbol{b}}_{\mathrm{RIS,II}}^{(\mathrm{crs})}] \leftarrow \mathrm{TPM}(\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{sub}}^{(\mathrm{crs})}, \{\hat{\boldsymbol{b}}_{r,\mathrm{I}}^{(\mathrm{crs})}\})$
Refine estimates: $\hat{\mu}_{r,\mathrm{II}}^{(\mathrm{crs})} \leftarrow \mathrm{SBE}(\{\hat{\boldsymbol{b}}_{r,\mathrm{II}}^{(\mathrm{crs})}\})$, $\forall r \in \{\mathrm{R, T, RIS}\}$
Fine Est. Stage: Use full aperture in each mode $\forall r \in \{\mathrm{R, T, RIS}\}$
Sectorization: Select $B_r^{(\mathrm{fine})}$ Beams around $\hat{\mu}_{r,\mathrm{II}}^{(\mathrm{crs})}$ by (14)
Pilot transmission: Accumulate demodulated symbols $\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{fine})}$
Tensor Decomposition:
    $\hat{\boldsymbol{b}}_{r,\mathrm{I}}^{(\mathrm{fine})} \leftarrow$ Calculate TPM initializations based on $\hat{\mu}_{r,\mathrm{II}}^{(\mathrm{crs})}$ (11)
    $[\hat{\boldsymbol{b}}_{\mathrm{R,II}}^{(\mathrm{fine})}\ \hat{\boldsymbol{b}}_{\mathrm{T,II}}^{(\mathrm{fine})}\ \hat{\boldsymbol{b}}_{\mathrm{RIS,II}}^{(\mathrm{fine})}] \leftarrow \mathrm{TPM}(\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{fine})}, \{\hat{\boldsymbol{b}}_{r,\mathrm{I}}^{(\mathrm{fine})}\})$
Obtain final estimates:$\hat{\mu}_r^{(\mathrm{fine})} \leftarrow \mathrm{SBE}(\{\hat{\boldsymbol{b}}_{r,\mathrm{II}}^{(\mathrm{fine})}\})\ \forall r \in \{\mathrm{R, T, RIS}\}$

V. Simulation Results

In this section, we present selected simulation results. These results reflect the performance of the estimators, which are measured in terms of the Root Mean Squared Error (RMSE). It is defined for each mode $r \in \{\mathrm{R, T, RIS}\}$ as

$$\mathrm{RMSE}_r = \sqrt{\mathbb{E}\Big\{(\mu_r - \hat{\mu}_{r,i})^2\Big\}}, \tag{15}$$

and the total RMSE is given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{3}\sum_{r=\{\mathrm{R, T, RIS}\}} \mathrm{RMSE}_r^2}. \tag{16}$$

They are compared with the Cramer Rao Lower Bound (CRLB) [15]. While recognizing that metrics of the spectral efficiency are more suitable to indicate performance, they exceed the scope of this paper.

The covered setup in this work contains $M_{\mathrm{R}} = 32$, $M_{\mathrm{T}} = 64$, $N_{\mathrm{RIS}} = 128$ elements at the receiver, transmitter, and RIS, respectively, and $N_{\mathrm{TF}} = 1$ training frame, with the chosen path gain $\rho = 1$. We define $B_r^{(\mathrm{crs})} = 3$, $\forall r \in \{\mathrm{R, T, RIS}\}$ for the coarse SoI. The convergence criteria for the TPM are set to $\mathrm{Iter}_{\mathrm{Max}} = 5$ and $\delta_{\mathrm{rel}} = 10^{-6}$. The available total power $P = 1$ is split evenly on the coarse and the fine estimation stage, i.e., $P^{(\mathrm{crs})} = P^{(\mathrm{fine})} = 0.5$, if not stated otherwise, and the SNR is given as $\mathrm{SNR} = \frac{P}{\sigma_n^2}$, where σ_n^2 denotes the noise power at the receiver before decoding.

First, the results for the coarse estimation stage are shown in Figure 2. The presented accuracy includes estimates for all modes: R, T, and RIS. In this setup, the entire power is allocated to the coarse estimation step, and different estimation approaches are compared. The coarse estimator named MPB stops the coarse estimation after the initial MPB estimate. In the low-SNR regime, both MPB and TPM+SBE with MPB initialization exhibit a similar accuracy. However, TPM+SBE surpasses the error floor of MPB due to its gridless nature, whereas the accuracy of MPB depends on the grid resolution. The initialization of the factor vectors plays a crucial role. Random initialization not only increases the required SNR for a given accuracy but also introduces inconsistencies, as the TPM may fail to converge reliably. This leads to non-uniform distortions in the RMSE curve, even at higher SNRs.

TABLE I: Approximate Complexity of CERISE.

MPB	$\approx (7N_{\mathrm{TF}} + 1) B_{\mathrm{R}} B_{\mathrm{T}} B_{\mathrm{RIS}}$
TPM	$\approx 48 \mathrm{Iter} B_{\mathrm{R}} B_{\mathrm{T}} B_{\mathrm{RIS}} N_{\mathrm{TF}}$
SBE, per mode	$\approx 2(B-1)^3$

In Figure 3, the accuracy of the proposed algorithm for estimating the RIS spatial frequency is shown after the fine estimation stage. Additionally, the accuracies of the fine estimation with the MPB coarse estimator and the estimation with perfect SoI information are provided for comparison. In low-to-medium SNR regimes, the CERISE estimator achieves a similar accuracy regardless of the SoI estimation method. Enhanced processing at the coarse estimation stage accelerates the approach of the estimator's curve toward the accuracy asymptote defined by perfect knowledge of the SoI in terms of the SNR, but introduces a moderate additional complexity as shown in Table I. Even though the power per pilot in the fine estimation stage is lower, it has a comparable performance to a purely coarse estimation. Furthermore, introducing an SoI significantly reduces the dimensionality of the signal space, leading to a much lower training overhead.

In Figure 4, the impact of varying the number of fine beams on the accuracy is analyzed. The figure compares the RMSE of RIS spatial frequency estimates, assuming perfect knowledge of the SoI. A smaller number of beams results in an increased power per pilot signal (7), thereby achieving a lower error asymptote. However, in practical scenarios, where a priori knowledge might not be available, choosing too few beams might result in higher errors, as the small number of beams might not cover the actual spatial position of the target.

Finally, Figure 5 presents the accuracy profiles for all modes under different power distributions between the coarse and fine estimation stages. The comparisons in Figure 3 have shown that the performance of the coarse estimator is critical for fine estimation, which can also be observed here. Assigning more power to the coarse estimation stage enables earlier divergence from the error ceiling at smaller SNRs. Allocating more resources ensures a higher accuracy in estimating the Sector of Interest, which normally limits the performance of CERISE in low SNR regimes. However, since the error asymptote in high-SNR regimes is primarily determined by the power allocated to pilots in the fine estimation stage, excessive power allocation to the coarse estimation stage can degrade the overall accuracy.

Based on the presented simulation results, we can confirm the ability of CERISE to obtain accurate estimates of target parameters in different configurations.

VI. Conclusions

In this paper, a novel two stage channel estimation algorithm for reconfigurable intelligent surface (RIS)-assisted channels CERISE is presented. It is based on the estimation of signal parameters via rotational invariance techniques (ESPRIT) in DFT beamspace. Numerical results have proven the ability of CERISE to provide high resolution parameter estimates. This two stage estimation approach offers strong performance

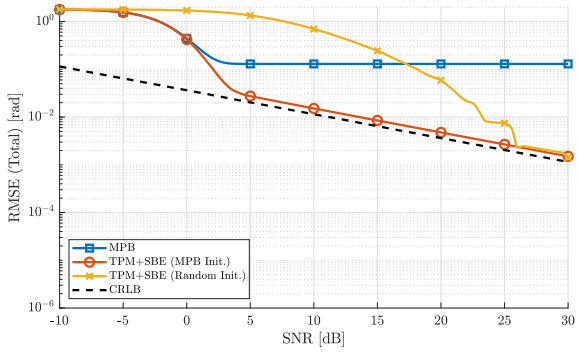


Fig. 2: Coarse estimation RMSE in all modes. $N_{\text{Trials}} = 10000$, $\rho = 1$, $[\mu_{\text{R}}, \mu_{\text{T}}, \mu_{\text{RIS}}] = [2.11, 1.05, 2.19]$, $[M_{\text{R}}, M_{\text{T}}, N_{\text{RIS}}] = [8, 16, 32]$, $N_{\text{TF}} = 1$, $P = 1$, $\text{Iter}_{\text{Max}} = 5$, $\delta_{\text{rel}} = 10^{-6}$.

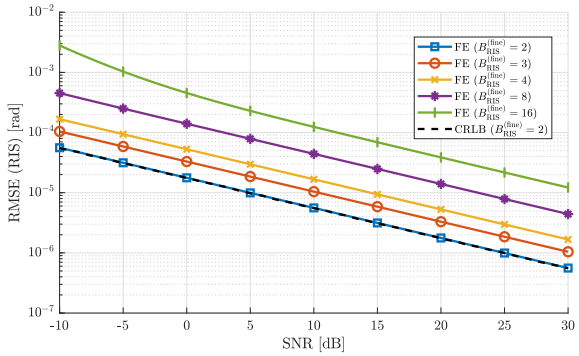


Fig. 4: Fine estimation (FE) accuracy for RIS with varying number of fine beams given perfect SoI knowledge. $N_{\text{Trials}} = 10000$, $\rho = 1$, $[\mu_{\text{R}}, \mu_{\text{T}}, \mu_{\text{RIS}}] = [2.11, 1.05, 2.19]$, $[M_{\text{R}}, M_{\text{T}}, N_{\text{RIS}}] = [32, 64, 128]$, $N_{\text{TF}}^{(\text{crs})} = N_{\text{TF}}^{(\text{fine})} = 1$, $P^{(\text{crs})} = P^{(\text{fine})} = 0.5$, $\text{Iter}_{\text{Max}} = 5$, $[M_{\text{R}}^{(\text{crs})}, M_{\text{T}}^{(\text{crs})}, M_{\text{RIS}}^{(\text{crs})}] = [8, 16, 32]$, $B_r^{(\text{fine})} = 4$, $\forall r \in \{\text{R, T}\}$

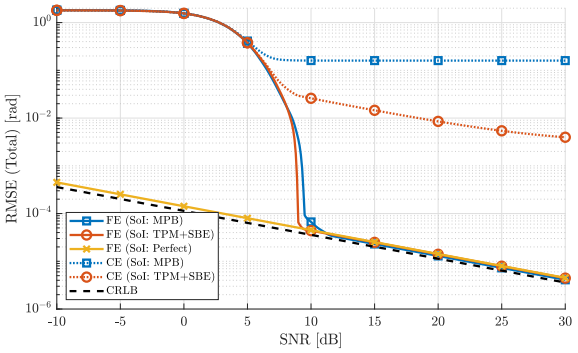


Fig. 3: RMSE for fine estimation (FE) and coarse estimation (CE). $N_{\text{Trials}} = 10000$, $\rho = 1$, $[\mu_{\text{R}}, \mu_{\text{T}}, \mu_{\text{RIS}}] = [2.11, 1.05, 2.19]$, $[M_{\text{R}}, M_{\text{T}}, N_{\text{RIS}}] = [32, 64, 128]$, $N_{\text{TF}}^{(\text{crs})} = N_{\text{TF}}^{(\text{fine})} = 1$, $P^{(\text{crs})} = P^{(\text{fine})} = 0.5$, $\text{Iter}_{\text{Max}} = 5$, $\delta_{\text{rel}} = 10^{-6}$, $[M_{\text{R}}^{(\text{crs})}, M_{\text{T}}^{(\text{crs})}, M_{\text{RIS}}^{(\text{crs})}] = [8, 16, 32]$, $B_r^{(\text{fine})} = 4$, $\forall r \in \{\text{R, T, RIS}\}$

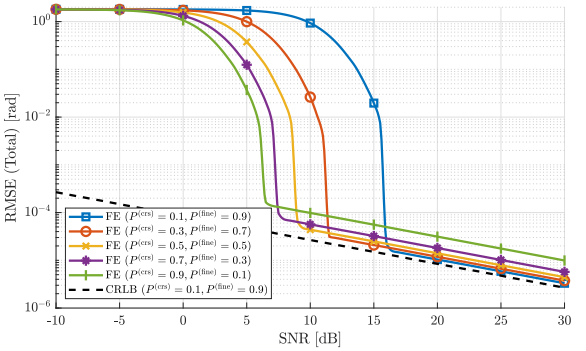


Fig. 5: Fine estimation (FE) accuracy for total setup with varying power distribution on coarse and fine estimation stage. $N_{\text{Trials}} = 10000$, $\rho = 1$, $[\mu_{\text{R}}, \mu_{\text{T}}, \mu_{\text{RIS}}] = [2.11, 1.05, 2.19]$, $[M_{\text{R}}, M_{\text{T}}, N_{\text{RIS}}] = [32, 64, 128]$, $N_{\text{TF}}^{(\text{crs})} = N_{\text{TF}}^{(\text{fine})} = 1$, $\text{Iter}_{\text{Max}} = 5$, $[M_{\text{R}}^{(\text{crs})}, M_{\text{T}}^{(\text{crs})}, M_{\text{RIS}}^{(\text{crs})}] = [8, 16, 32]$, $B_r^{(\text{fine})} = 4$, $\forall r \in \{\text{R, T, RIS}\}$

improvements in medium to high SNR regimes closely approaching the CRLB, while maintaining a low computational complexity.

References

[1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[2] M. Jian, G. C. Alexandropoulos, E. Basar, C. Huang, R. Liu, Y. Liu, and C. Yuen, "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 1–32, Mar. 2022.

[3] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[4] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel Estimation With Reconfigurable Intelligent Surfaces—A General Framework," *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1312–1338, Sep. 2022.

[5] C. Huang, J. Xu, W. Zhang, W. Xu, and D. W. K. Ng, "Semi-Blind Channel Estimation for RIS-Assisted MISO Systems Using Expectation Maximization," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 9, pp. 10 173–10 178, Sep. 2022.

[6] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel Estimation for Reconfigurable Intelligent Surface Aided Multi-User mmWave MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 10, pp. 6853–6869, Oct. 2023.

[7] E. Saltykova, "A Channel Estimation Between a Base Station, Intelligent Reflecting Surface, and a User Equipment," Master's thesis, Chalmers University of Technology, Gothenburg, Sweden, 2023.

[8] 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, Apr. 2021.

[9] M. Zoltowski, M. Haardt, and C. Mathews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 316–328, Feb. 1996.

[10] D. Rakhimov and M. Haardt, "Analytical Performance Assessment of 1-D ESPRIT in DFT Beamspace in Terms of Physical Parameters," in *Proc. 57th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 2023, pp. 1264–1270.

[11] N. Michelusi and M. Hussain, "Optimal Beam-Sweeping and Communication in Mobile Millimeter-Wave Networks," in *Proc. IEEE International Conference on Communications (ICC 2018)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[12] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[13] E. Kofidis and P. A. Regalia, "On the Best Rank-1 Approximation of Higher-Order Supersymmetric Tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 3, pp. 863–884, Jan. 2002.

[14] A. Albanese, F. Devoti, V. Sciancalepore, M. Di Renzo, and X. Costa-Pérez, "MARISA: A Self-configuring Metasurfaces Absorption and Reflection Solution Towards 6G," in *Proc. IEEE Conference on Computer Communications (INFOCOM 2022)*, London, United Kingdom, May 2022, pp. 250–259.

[15] D. Rakhimov, A. Rakhimov, A. Nadeev, and M. Haardt, "Tensor Formulation of the Cramer-Rao Lower Bound for Beamspace Channel Estimation in mmWave MIMO-OFDM," in *Proc. 25th International ITG Workshop on Smart Antennas (WSA 2021)*, Nice, France, Nov. 2021.