

Gridless Channel Estimation for Hybrid mmWave MIMO Systems via Tensor-ESPRIT Algorithms in DFT Beamspace

Jianshu Zhang ¹, Member, IEEE, Damir Rakhimov, Student Member, IEEE, and Martin Haardt ², Fellow, IEEE

Abstract—In this paper, we present a gridless channel estimation algorithm for a hybrid millimeter wave (mmWave) MIMO-OFDM system assuming a frequency-selective channel. The proposed algorithm is based on the Tensor-ESPRIT in DFT beamspace algorithm framework. First, we derive the R -dimensional (R -D) Standard/Unitary Tensor-ESPRIT in DFT beamspace framework and its analytic performance. We show that ESPRIT-type algorithms in a reduced-dimensional DFT beamspace can provide a significant performance gain over ESPRIT-type algorithms in full DFT beamspace and in element space under mild conditions. Afterwards, we develop a gridless channel estimation algorithm that is based on 3-D Tensor-ESPRIT in DFT beamspace algorithms. Numerical simulation results show that the proposed channel estimation algorithm can provide accurate channel estimates using only a few training resources.

Index Terms—5G mmWave communications, channel estimation, ESPRIT in DFT beamspace, hybrid MIMO, tensor signal processing.

I. INTRODUCTION

HYBRID analog-digital multi-antenna systems allow the exploitation of the MIMO gain in a cost- and energy-efficient way especially for mmWave massive MIMO systems [1]–[3]. Since only phase shifters are used in the analog processing, new non-convex constraints are introduced, which lead to significant challenges for the design of MIMO processing algorithms as well as the corresponding channel estimation procedure. In order to estimate the channel, many efforts have been put on sparse recovery methods that exploit the sparsity in the angular and/or the delay domain of the mmWave MIMO channel. Compressed sensing (CS) can significantly reduce the training overhead. Therefore, CS based channel estimation algorithms have been studied extensively. CS based channel estimation approaches for a frequency-flat channel have been proposed in [4]–[9]. For frequency-selective mmWave MIMO

channels, CS based channel estimation methods have been studied in [10]–[13]. The major limitation of the aforementioned CS based methods is the on-grid assumption of the channel parameters, which requires a grid-offset estimation especially in the high SNR regime to avoid an error floor.

In order to achieve a gridless and high-resolution channel (parameter) estimation, several non-CS based channel (parameter) estimation algorithms have been proposed. In [14] it is shown that a multi-band antenna together with a fully digital sub-6 GHz hardware can be used to obtain a similar angular profile as in the mmWave band. An auxiliary beam pair enabled joint estimation of the angular profile, i.e., directions of arrival (DoAs) and the directions of departures (DoDs), has been proposed in [15]. In [16] a CANDECOMP/PARAFAC (CP) decomposition based method has been proposed to first estimate the array steering matrix and then followed by a correlation-based search to estimate the channel parameters including the delays, the DoAs and the DoDs. To ensure the uniqueness of the solution, the hybrid precoder and decoder have to be designed such that Kruskal's condition is fulfilled. A similar approach has been extended to a time-varying channel in [17]. In [18] a beamspace ESPRIT based channel estimation algorithm is proposed, where the shift invariance structure is not exact. That is, it needs to be derived for every choice of hybrid precoder and decoder. It is worth stressing that the entries of the analog precoder and decoder in [16]–[18] are drawn randomly from the unit circle. In [19] we have also proposed a gridless channel estimation algorithm based on three-dimensional (3-D) Standard ESPRIT in discrete Fourier transform (DFT) beamspace. Given a priori knowledge of the sectors of interest (SoIs), in which the spatial frequencies of time delays, DoAs, and DoDs are, the proposed algorithm provides high resolution estimates. In our design the analog precoder and decoder are selected columns (rows) of a DFT matrix. It is much easier to implement a DFT based analog precoder and decoder than the randomly drawn entries, since we can focus on a pre-defined SoI. Moreover, the DFT based codebook is easy to operate and maintain at both the transmitter and the receiver. In terms of feedback only the indices of the DFT matrix are required. Hence, compared to [16]–[18], the approach in [19] is more efficient in terms of hardware implementation and feedback requirements. Moreover, the shift invariance structure in our work is exact and does not depend on the realization of hybrid precoder and decoder. It is therefore more memory

Manuscript received June 30, 2020; revised December 11, 2020 and March 1, 2021; accepted March 2, 2021. Date of publication March 4, 2021; date of current version March 29, 2021. This work was supported by German Research Foundation (DFG) under Grants HA 2239/14-1 (AdAMMM) and ZH 640/2-1 (AdAMMM). The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Sergiy Vorobyov. (Corresponding author: Jianshu Zhang.)

The authors are with the Communications Research Laboratory, Ilmenau University of Technology, D-98684 Ilmenau, Germany (e-mail: Jianshu.zhang@tu-ilmenau.de; damir.rakhimov@tu-ilmenau.de; martin.haardt@tu-ilmenau.de).

Digital Object Identifier 10.1109/JSTSP.2021.3063908

and computationally efficient compared to [18]. Our systematic design also makes it easy to derive the analytic performance. To further reduce the computational complexity, we have proposed the Unitary ESPRIT in DFT beamspace based channel estimation algorithm in [20] and its tensor version in [21].

In the context of array signal processing, one-dimensional (1-D) Standard ESPRIT in DFT beamspace has been developed in [22] and two-dimensional (2-D) Unitary ESPRIT in DFT beamspace has been developed in [23]. It is shown in [23] that if there is a priori information on the DoAs, one may apply only a subset of rows of a DFT matrix to extract the angular locations. Thereby, the overall computational complexity is further reduced. Based on a large-scale analysis, i.e., under the assumption of an infinite number of snapshots, the analytic performance of 1-D Standard ESPRIT in DFT beamspace and that of 2-D Unitary ESPRIT in DFT beamspace has been derived in [22] and [24], respectively. It is observed from numerical simulations that the performance of ESPRIT-type algorithms in DFT beamspace is close to that of ESPRIT-type algorithms in element space.

In this paper, we propose a gridless channel estimation algorithm based on Tensor-ESPRIT in DFT beamspace for a hybrid mmWave MIMO-OFDM system under a frequency selective channel. To this end, a generic Tensor-ESPRIT in DFT beamspace algorithm framework is developed. In particular, our contributions are summarized as:

- A generic Tensor-ESPRIT in DFT beamspace algorithm framework is derived. The matrix version of ESPRIT in DFT beamspace [19], [20] can be interpreted as a special case of its tensor counterpart. In contrast to [23], where only adjacent DFT beams are used, i.e., a single SoI, we prove that the shift invariance property still holds even if non-adjacent DFT beams, i.e., multiple SoIs, are used. We show that two DFT beams are required to identify one source when the ESPRIT in DFT beamspace algorithm is used. The maximum number of resolvable sources of the ESPRIT in DFT beamspace algorithm is derived. The developed framework can also be used in other high resolution parameter estimation applications.
- The achieved mean squared error (MSE) is derived analytically by using a first-order perturbation analysis. We show that for a single scenario source ESPRIT in a reduced-dimensional DFT beamspace can achieve a smaller MSE than ESPRIT in the full DFT beamspace and in element space especially for a large array. The gain is due to the array gain of the DFT beams in the SoIs.
- Inspired by the Tensor-ESPRIT in DFT beamspace framework, a gridless channel estimation algorithm is proposed. To fully exploit the gain of the DFT-beamspace ESPRIT algorithm, a sliding-window based sectorization approach is proposed to find the required SoIs. Based on our design criteria, explicit cost functions are formulated and a convex analysis is used to derive the solution. The heuristic sectorization method in [21] is a special case of the proposed sectorization approach.

Our simulation results show that the empirical MSE of ESPRIT in DFT beamspace matches the derived analytical MSE

in the medium to high SNR regime. Tensor-ESPRIT in DFT beamspace achieves a smaller MSE when compared to its matrix counterpart. For channel estimation, the proposed ESPRIT in DFT beamspace based approach provides high resolution channel estimates by using only a few training resources.

The rest of this paper is organized as follows. In Section II the Tensor-ESPRIT in DFT beamspace algorithm framework and its analytic performance are derived. Section III specifies the system and channel model as well as the training procedure of the channel estimation problem. Section IV describes the proposed gridless channel estimation algorithm. Section V collects the simulation results. Our conclusions are drawn in VI.

Notation: Upper-case calligraphic, upper-case and lower-case bold-faced letters denote tensors, matrices and vectors, respectively. The expectation, transpose, conjugate, and Hermitian transpose are denoted by $\mathbb{E}\{\cdot\}$, $\{\cdot\}^T$, $\{\cdot\}^*$, and $\{\cdot\}^H$, respectively. The Euclidean norm of a vector and the absolute value are denoted by $\|\cdot\|$ and $|\cdot|$, respectively. The Kronecker product and the Khatri-Rao product are denoted as \otimes and \diamond , respectively. The $m \times m$ identity matrix, exchange matrix, and DFT matrix with unit modulus entries are \mathbf{I}_m , $\mathbf{\Pi}_m$, and \mathbf{D}_m , respectively. The $m \times n$ matrix with all zero elements and all ones are $\mathbf{0}_{m \times n}$ and $\mathbf{1}_{m \times n}$, respectively. The r -mode product of a tensor with a matrix is denoted by \times_r [25]. The concatenation of tensors or matrices along the r -th dimension is denoted as \sqcup_r [26]. The $\text{vec}\{\cdot\}$ operator stacks the columns of a matrix into a vector. The operator $\text{diag}\{\mathbf{v}\}$ creates a $m \times m$ diagonal matrix by aligning the elements of the vector $\mathbf{v} \in \mathbb{C}^m$ onto its main diagonal. The $\text{Card}\{\cdot\}$ operator denote the cardinality of a set. Rounding a number to the nearest integer is denoted by $\text{round}\{\cdot\}$. The floor function is denoted by $\lfloor \cdot \rfloor$. Useful tensor algebra that is used in the derivations of this paper can be found in [26].

II. R -DIMENSIONAL TENSOR-ESPRIT IN REDUCED-DIMENSION DFT BEAMSPACE

Our proposed channel estimation algorithm is based on R -dimensional (R -D) Tensor-ESPRIT algorithms in reduced-dimension DFT beamspace, which uses a subset of rows of a DFT matrix to encompass the SoI. Although in the channel estimation problem we have only $R = 3$, we develop the general R -D Standard and Unitary Tensor-ESPRIT in DFT beamspace algorithm framework and derive the analytic performance.

Let us start with a tensor representation of a R -D harmonic retrieval problem with N snapshots, c.f., [26]

$$\mathcal{X} = \mathcal{A} \times_{R+1} \mathbf{S}^T + \mathcal{N} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times N}$$

where $\mathcal{A} = \mathcal{I}_{R+1, d \times 1} \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_R \mathbf{A}^{(R)} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times d}$ is the array steering tensor with the steering matrix $\mathbf{A}^{(r)} = [\mathbf{a}(\mu_1^{(r)}) \dots \mathbf{a}(\mu_d^{(r)})] \in \mathbb{C}^{M_r \times d}$. We have the spatial frequencies $\mu_i^{(r)} \in (-\pi, \pi]$ for $i \in \{1, \dots, d\}$ and $r \in \{1, \dots, R\}$. The tensor $\mathcal{N} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times N}$ denotes the additive zero mean circularly symmetric complex Gaussian (ZMCSG) noise with $\mathbb{E}\{\text{vec}\{[\mathcal{N}]_{(R+1)}^T\} \text{vec}\{[\mathcal{N}]_{(R+1)}^T\}^H\} = \sigma_n^2 \mathbf{I}_{MN}$, where $M = \prod_{r=1}^R M_r$. The unknown source signal is assumed to have unit power, i.e., $\mathbb{E}\{|\mathcal{S}_{i,t}|^2\} = 1$, where $t \in \{1, \dots, N\}$.

A. R-D Standard Tensor-ESPRIT (STE) in DFT Beamspace

If R-D Standard Tensor-ESPRIT (STE) in DFT beamspace is used to estimate the spatial frequencies, we project the r -th mode of \mathcal{X} onto the subspace of the DFT beamspace by multiplying $\mathbf{F}^{(r)} \in \mathbb{C}^{B_r \times M_r}$ on the r -th mode, i.e.,

$$\begin{aligned} \bar{\mathcal{X}} &= \mathcal{X} \times_1 \mathbf{F}^{(1)} \times_2 \mathbf{F}^{(2)} \cdots \times_R \mathbf{F}^{(R)} \\ &= \mathcal{B} \times_{R+1} \mathbf{S}^T + \bar{\mathcal{N}} \in \mathbb{C}^{B \times N}, \end{aligned} \quad (1)$$

where $\mathcal{B} = \mathcal{I}_{R+1, d \times 1} \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \cdots \times_R \mathbf{B}^{(R)} \in \mathbb{C}^{B_1 \times B_2 \cdots \times B_R \times d}$ is the array steering tensor in DFT beamspace and $\mathbf{B}^{(r)} = \mathbf{F}^{(r)} \mathbf{A}^{(r)} \in \mathbb{C}^{B_r \times d}$. We have $B = \prod_{r=1}^R B_r$ and $B_r \leq M_r$ for all r . The matrices $\mathbf{F}^{(r)}$ comprise B_r different rows of a $M_r \times M_r$ DFT matrix \mathbf{D}_{M_r} and $B_r \leq M_r$. Lastly, the effective noise $\bar{\mathcal{N}}$ is given by

$$\bar{\mathcal{N}} = \mathcal{N} \times_1 \mathbf{F}^{(1)} \times_2 \mathbf{F}^{(2)} \cdots \times_R \mathbf{F}^{(R)}, \quad (2)$$

such that $\mathbb{E}\{\bar{\mathcal{N}}\bar{\mathcal{N}}^H\} = M\sigma_n^2 \mathbf{I}_{BN}$, where $\bar{\mathbf{n}} = \text{vec}\{[\bar{\mathcal{N}}]_{(R+1)}^T\}$.

Similar as R-D STE in element space, in order to apply R-D STE in DFT beamspace the R-D effective array \mathcal{B} should have shift invariance properties in each of its modes, i.e.,

$$\mathcal{B} \times_r \mathbf{J}_1^{(r)} \times_{R+1} \Phi^{(r)} = \mathcal{B} \times_r \mathbf{J}_2^{(r)}, \forall r \quad (3)$$

where $\Phi^{(r)} = \text{diag}\{[e^{j\mu_1^{(r)}} \cdots e^{j\mu_d^{(r)}}]^T\} \in \mathbb{C}^{d \times d}$. One possibility of selection matrices $\mathbf{J}_1^{(r)} \in \mathbb{C}^{M_{\text{sel}} \times B_r}$ and $\mathbf{J}_2^{(r)} \in \mathbb{C}^{M_{\text{sel}} \times B_r}$ is given in the following lemma.

Lemma 1 (Complex-valued selection matrices in DFT beamspace): Let us define $\mathbf{G}_1^{(r)} \in \mathbb{C}^{M_r \times M_r}$ and $\mathbf{G}_2^{(r)} \in \mathbb{C}^{M_r \times M_r}$ as

$$\mathbf{G}_1^{(r)} = \begin{bmatrix} 1 & -e^{-\frac{j2\pi}{M_r}} & 0 & \cdots & 0 & 0 \\ 0 & 1 & -e^{-\frac{j2\pi}{M_r}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & -e^{-\frac{j2\pi}{M_r}} \\ 1 & 0 & 0 & 0 & 0 & -e^{\frac{j2\pi}{M_r}} \end{bmatrix} \quad (4)$$

and

$$\mathbf{G}_2^{(r)} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & e^{\frac{j2\pi}{M_r}} & -e^{\frac{j2\pi}{M_r}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & e^{\frac{j2\pi(M_r-2)}{M_r}} & -e^{\frac{j2\pi(M_r-2)}{M_r}} \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \quad (5)$$

Let $\mathcal{S}_r = \{\mathcal{S}_r^{(1)}, \cdots, \mathcal{S}_r^{(g_r)}\} \subseteq \{1, \cdots, M_r\}$ represent the set of indices of the selected rows of the M_r -by- M_r DFT matrix. The g -th subset $\mathcal{S}_r^{(g)} = \{i_1^{(r,g)}, i_2^{(r,g)}, \cdots, i_{n_{r,g}}^{(r,g)}\}$ contains only consecutive numbers, where $g \in \{1, \cdots, g_r\}$, $n_{r,g} \geq 2$ and $i_1^{(r,1)} < i_2^{(r,1)} < \cdots < i_{n_{r,g}}^{(r,g)}$.¹ We have $\text{card}(\mathcal{S}_r^{(g)}) = n_{r,g}$, $\forall r, g$ and $\sum_g \text{card}(\mathcal{S}_r^{(g)}) = B_r \leq M_r$. Moreover, different subsets are disjoint, i.e., $\mathcal{S}_r^{(g)} \cap \mathcal{S}_r^{(\bar{g})} = \emptyset$ for $\bar{g} \in \{1, \cdots, g_r\}$ and $\bar{g} \neq g$. Let us define the row indices and the column indices in the r -th

¹Note that 1 and M_r are treated as consecutive numbers.

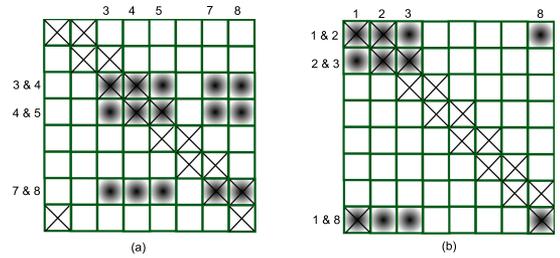


Fig. 1. Structure of the selection matrices $\mathbf{G}_1^{(1)}$ and $\mathbf{G}_2^{(1)}$ for the 1-D case with $M_1 = 8$. The symbol \times denotes the nonzero entries in both selection matrices. The shaded areas shows how to choose the subblocks of the selection matrices for reduced-dimensional processing, i.e., how to form $\mathbf{J}_1^{(1)}$ and $\mathbf{J}_2^{(1)}$. The following examples are given: (a) $B_1 = 5$ beams are applied. The corresponding rows are 3, 4, 5, 7, and 8. They are non-consecutive. (b) $B_1 = 4$ beams are applied. The rows 8, 1, 2, and 3 are consecutive.

dimension as

$$\begin{aligned} \text{row}^{(r)} &= i_1^{(r,1)} : i_{n_{r,1}}^{(r,1)} - 1, \cdots, i_1^{(r,g_r)} : i_{n_{r,g_r}}^{(r,g_r)} - 1 \\ \text{col}^{(r)} &= i_1^{(r,1)} : i_{n_{r,1}}^{(r,1)}, \cdots, i_1^{(r,g_r)} : i_{n_{r,g_r}}^{(r,g_r)} \end{aligned} \quad (6)$$

Then selection matrices $\mathbf{J}_1^{(r)} \in \mathbb{C}^{M_{\text{sel}} \times B_r}$ and $\mathbf{J}_2^{(r)} \in \mathbb{C}^{M_{\text{sel}} \times B_r}$ satisfying (3) are expressed as

$$\mathbf{J}_1^{(r)} = \mathbf{G}_1^{(r)}(\text{row}^{(r)}, \text{col}^{(r)}), \mathbf{J}_2^{(r)} = \mathbf{G}_2^{(r)}(\text{row}^{(r)}, \text{col}^{(r)}),$$

and $M_{\text{sel}} = B_r - g_r$.

Proof: The matrices $\mathbf{G}_1^{(r)}$ and $\mathbf{G}_2^{(r)}$ are derived by extending the results in [19] to the R-D case. When $B_r < M_r$, the construction of the selection matrices is based on the fact that each row of $\mathbf{G}_1^{(r)}$ and $\mathbf{G}_2^{(r)}$ has two non-zero elements, which correspond to two consecutive rows of the DFT matrix.

Lemma 1 implies the following design rule for $\mathbf{F}^{(r)}$.

Corollary 1: For ESPRIT-type algorithms in DFT beamspace $\mathbf{F}^{(r)}$ should consist of at least two consecutive rows of the DFT matrix, i.e., $B_r \geq 2$. But it is not necessary to form beams using only consecutive rows.

Based on Lemma 1, we also define the concept of sectors of interest.

Definition 1: Sectors of interest (SoIs) refers to the area that are embraced by the main lobes of the DFT beams $\mathbf{F}^{(r)}$. Mathematically, if $\mu_r \in [i_1^{(r,g)} - 1, i_{n_{r,g}}^{(r,g)} - 1] \frac{2\pi}{M_r}$, we say that μ_r lies in the g -th spatial SoI in the r -th dimension, where $i_{n_{r,g}}^{(r,g)}$ is defined as in Lemma 1.

Let us consider a 1-D example with $M_1 = 8$ for Lemma 1. The structure of the corresponding full dimensional selection matrices $\mathbf{G}_1^{(1)}$ and $\mathbf{G}_2^{(1)}$ is sketched in Fig. 1. The symbol \times denotes entries of both selection matrices that are nonzero. If $\mathbf{F}^{(1)}$ uses rows 3, 4, 5, 7, 8 of the DFT matrix to form 5 beams for estimating the spatial frequencies, the corresponding 3×5 subblock of the selection matrices $\mathbf{G}_1^{(1)}$ and $\mathbf{G}_2^{(1)}$, i.e., $\mathbf{J}_1^{(1)}$ and $\mathbf{J}_2^{(1)}$, is shaded in Fig. 1 (a). In such a case, two SoIs are formed. The first one is formed by beams 3, 4, and 5. The second one is formed by beams 7 and 8. Note that the first and last row of the DFT matrix steer beams that are physically adjacent to each other. If $\mathbf{F}^{(1)}$ uses rows 8, 1, 2, and 3 of the DFT matrix to form 4 beams for estimation, the corresponding 3×4 subblock of the

selection matrices $\mathbf{G}_1^{(1)}$ and $\mathbf{G}_2^{(1)}$, i.e., $\mathbf{J}_1^{(1)}$ and $\mathbf{J}_2^{(1)}$, is shaded in Fig. 1 (b). In this case only one SoI is formed.

To solve the shift invariance equation (3), let us first consider the noiseless case and $N \gg d$. Then we can compute the rank- d economy-size HOSVD of $\tilde{\mathcal{X}}$ as

$$\tilde{\mathcal{X}} = \underbrace{\mathbf{S}_s \times_1 \mathbf{U}_s^{(1)} \times_2 \mathbf{U}_s^{(2)} \cdots \times_R \mathbf{U}_s^{(R)}}_{\mathbf{u}_s} \times_{R+1} \mathbf{U}_s^{(R+1)} \quad (7)$$

where $\mathbf{U}_s^{(r)} \in \mathbb{C}^{B_r \times d}$ for $r \in \{1, \dots, R\}$ and $\mathbf{U}_s^{(R+1)} \in \mathbb{C}^{N \times d}$. Note that in the noisy case the rank- d truncated HOSVD is used. Clearly, $\mathbf{U}_s^{(r)}$ spans the column subspace of the r -mode unfolding of $\tilde{\mathcal{X}}$. Similarly as in [26], it can be proven that there should exist a nonsingular $d \times d$ matrix $\mathbf{Q} \in \mathbb{C}^{d \times d}$ such that

$$\mathbf{B} = \mathbf{U}_s \times_{R+1} \mathbf{Q}, \quad (8)$$

where $\mathbf{U}_s \in \mathbb{C}^{B_1 \times B_2 \times \cdots \times d}$ represents the tensor signal subspace. Thereby, after replacing the unknown \mathbf{B} in (3) by applying (8) we obtain

$$\mathbf{U}_s \times_r \mathbf{J}_1^{(r)} \times_{R+1} \Psi^{(r)} = \mathbf{U}_s \times_r \mathbf{J}_2^{(r)}, r \in \{1, \dots, R\}, \quad (9)$$

where an eigenvalue decomposition of $\Psi^{(r)}$ satisfies

$$\Psi^{(r)} = \mathbf{Q} \cdot \Phi^{(r)} \cdot \mathbf{Q}^{-1}. \quad (10)$$

To obtain the estimates of spatial frequencies from the invariance equations in (9), the same solutions as in [26] can be applied. For example, the solution of (9) in the least squares (LS) sense is calculated by

$$\hat{\Psi}^{(r)} = (\tilde{\mathbf{J}}_1^{(r)} [\mathbf{u}_s]_{(R+1)}^T)^+ \tilde{\mathbf{J}}_2^{(r)} [\mathbf{u}_s]_{(R+1)}^T, \forall r, \quad (11)$$

where $\tilde{\mathbf{J}}_j^{(r)} \in \mathbb{C}^{(B_r - g_r) \prod_{\bar{r}=1, \bar{r} \neq r}^R B_{\bar{r}} \times B}$ and

$$\tilde{\mathbf{J}}_j^{(r)} = \mathbf{I}_{\prod_{\bar{r}=r+1}^R B_{\bar{r}}} \otimes \mathbf{J}_j^{(r)} \otimes \mathbf{I}_{\prod_{\bar{r}=1}^{r-1} B_{\bar{r}}}, j = 1, 2. \quad (12)$$

Afterwards, the estimates of the spatial frequencies are the corresponding phases of the eigenvalues obtained by calculating the simultaneous Schur decomposition or the joint eigenvalue decomposition (EVD) of $\hat{\Psi}^{(r)}$, $\forall r$ [27]. The use of the simultaneous Schur decomposition or the joint EVD is to ensure automatic pairing of estimated frequencies in different dimensions. Regarding the number of resolvable resources, the following statement holds.

Lemma 2: The maximum number of identifiable sources d_{\max} of R -D STE in DFT beamspace is

$$d_{\max} = \min \left(\min_r \left((B_r - g_r) \prod_{\bar{r}=1, \bar{r} \neq r}^R B_{\bar{r}} \right), N \right). \quad (13)$$

Proof: Note that the existence of the LS solution in (11) requires that $\tilde{\mathbf{J}}_1^{(r)} [\mathbf{u}_s]_{(R+1)}^T$ has the full column rank, i.e., $d \leq (B_r - g_r) \prod_{\bar{r}=1, \bar{r} \neq r}^R B_{\bar{r}}, \forall r$. Furthermore, we require that $d \leq N$, c.f., (7). Thereby, (13) holds.

In the noisy case, the estimation of the signal subspace tensor \mathbf{U}_s depends not only on the noise level and the number of snapshots but also on $\mathbf{F}^{(r)}$. This can be explained by the fact that the DFT beamformer $\mathbf{F}^{(r)}$ has focused the power into one or several SoIs. The signal received from the spatial frequency $\mu_i^{(r)}$ within the SoIs will be enhanced while the signal received from the spatial frequencies outside of the SoIs is suppressed or even

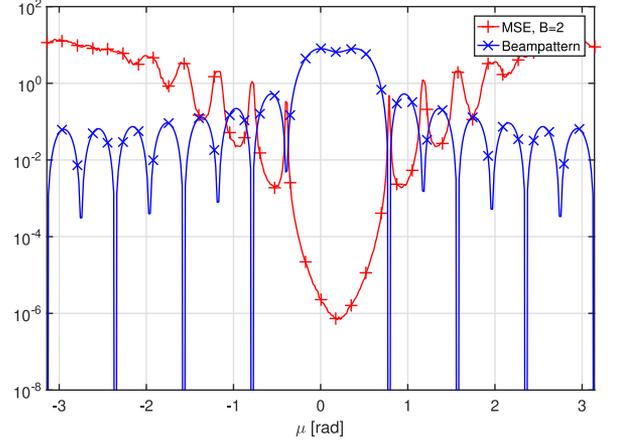


Fig. 2. Correspondence between the formed beampattern and the achieved MSE of 1-D Standard ESPRIT in DFT beamspace. $M = 16$, $B = 2$ and the first 2 rows of the DFT matrix are used. Red curve: achieved MSE $\mathbb{E}\{|\mu - \hat{\mu}|^2\}$ for $N = 20$ snapshots, $P_T = 1$ W, SNR = 20 dB, and 1000 noise realizations. Blue curve: The corresponding beampattern $\|\mathbf{B}^{(1)}\|_2^2$.

nulled. Therefore, if $\mathbf{F}^{(r)}$ is constructed such that its formed SoI is away from the spatial frequencies to be estimated, the corresponding columns of $\mathbf{B}^{(r)}$ will have much lower power, i.e., a much lower effective SNR. Consequently, there will be a worse estimation of the signal subspace. Alternatively, we can also interpret $\mathbf{F}^{(r)}$ as a band-pass filter which only extracts the information of spatial frequencies inside the pass-band. To better illustrate this, let us consider, for example, a 1-D case with $M = 16$ and $d = 1$. The achieved MSE $\mathbb{E}\{|\mu - \hat{\mu}|^2\}$ by using the 1-D Standard ESPRIT in DFT beamspace [22] and the corresponding beampattern $\|\mathbf{B}^{(1)}\|_2^2$ are demonstrated in Fig. 2. It can be seen that the achieved MSE is approximately inversely proportional to the beampattern. It is worth mentioning that a similar trend is obtained for $d > 1$ and/or $R > 1$ dimensional problems. Based on the facts above, we propose the next de-signing rule for $\mathbf{F}^{(r)}$.

Corollary 2: For ESPRIT-type algorithms in DFT beamspace the SoIs formed by $\mathbf{F}^{(r)}$ should include the spatial frequencies to be estimated. It is better that the spatial frequencies are near to the center of the SoIs. To this end, some prior information of the spatial frequencies $\mu_i^{(r)}$ is needed.

B. R -D Unitary Tensor-ESPRIT (UTE) in DFT Beamspace

Similarly as in the element space, R -D Unitary Tensor-ESPRIT (UTE) in DFT beamspace features real-valued signal processing for the signal subspace estimation and the subsequent spatial frequencies estimation step. Therefore, it has a lower computational complexity compared to R -D STE in DFT beamspace. In order to apply the real-valued subspace estimation, we multiply the r -th mode of the tensor \mathcal{X} by $\tilde{\mathbf{F}}^{(r)}$. The only difference between $\tilde{\mathbf{F}}^{(r)} \in \mathbb{C}^{B_r \times M_r}$ and $\mathbf{F}^{(r)}$ in Section II-A is that $\tilde{\mathbf{F}}^{(r)}$ comprises selected rows of the phase shifted DFT matrix $\tilde{\mathbf{D}}_{M_r} = \text{diag}\{[1, e^{j\frac{M_r-1}{2} \cdot \frac{2\pi}{M_r}}, \dots, e^{j\frac{M_r-1}{2} \cdot \frac{2\pi(M_r-1)}{M_r}}]\}^T \cdot \mathbf{D}_{M_r}$. The considered DFT matrix has unit modulus entries.

Note that $\mathbf{F}^{(r)H}$ is a left- Π real matrix, i.e., it satisfies $\Pi \tilde{\mathbf{F}}^{(r)T} = \tilde{\mathbf{F}}^{(r)H}$. The effective signal tensor is then expressed as

$$\begin{aligned} \tilde{\mathcal{X}} &= \mathcal{X} \times_1 \tilde{\mathbf{F}}^{(1)} \times_2 \tilde{\mathbf{F}}^{(2)} \cdots \times_R \tilde{\mathbf{F}}^{(R)} \\ &= \tilde{\mathcal{B}} \times_{R+1} \mathbf{S}^T + \tilde{\mathcal{N}} \in \mathbb{C}^{B \times N}, \end{aligned} \quad (14)$$

Next, we find a real-valued transformation of $\tilde{\mathcal{X}}$. One possible transformation is given in the following lemma.

Lemma 3 (Real-valued subspace in DFT beamspace): Let us define

$$\mathcal{Z} = \sqrt{2} \left[\text{Re}\{\tilde{\mathcal{X}}\} \sqcup_{R+1} \text{Im}\{\tilde{\mathcal{X}}\} \right], \quad (15)$$

which is a tensor extension of the concept in [23]. Then \mathcal{Z} corresponds to the real-valued transformation of $\tilde{\mathcal{X}}$. The rank- d economy-size (or truncated in the noisy case) HOSVD of \mathcal{Z} is real-valued and is computed by

$$\mathcal{Z} = \mathcal{E}_s \times_{R+1} \mathbf{E}_s^{(R+1)} \quad (16)$$

where $\mathcal{E}_s = \mathcal{Z}_s \times_1 \mathbf{E}_s^{(1)} \times_2 \mathbf{E}_s^{(2)} \cdots \times_R \mathbf{E}_s^{(R)}$.

Proof: See Appendix A.

Lemma 3 enables us to establish the real-valued shift invariance equations

$$\mathcal{E}_s \times_r \mathbf{K}_1^{(r)} \times_{R+1} \Upsilon^{(r)} = \mathcal{E}_s \times_r \mathbf{K}_2^{(r)}, \forall r, \quad (17)$$

and the real-valued EVD

$$\Upsilon^{(r)} = \mathbf{T} \cdot \Omega^{(r)} \cdot \mathbf{T}^{-1}, \quad (18)$$

where $\Omega^{(r)} = \text{diag}\left\{\left[\tan\left(\frac{\mu_1^{(r)}}{2}\right) \cdots \tan\left(\frac{\mu_d^{(r)}}{2}\right)\right]^T\right\} \in \mathbb{R}^{d \times d}$. The real-valued selection matrices $\mathbf{K}_1^{(r)}$ and $\mathbf{K}_2^{(r)}$ satisfying (17) can be constructed by following Lemma 4.

$$\begin{aligned} \bar{\mathbf{G}}_1^{(r)} &= \\ &\begin{bmatrix} 1 & \cos\left(\frac{\pi}{M_r}\right) & 0 & \cdots & 0 & 0 \\ 0 & \cos\left(\frac{\pi}{M_r}\right) \cos\left(\frac{2\pi}{M_r}\right) & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cos\left(\frac{\pi(M_r-2)}{M_r}\right) & \cos\left(\frac{\pi(M_r-1)}{M_r}\right) \\ (-1)^{M_r} & 0 & 0 & 0 & 0 & \cos\left(\frac{\pi(M_r-1)}{M_r}\right) \end{bmatrix} \end{aligned} \quad (19)$$

$$\begin{aligned} \bar{\mathbf{G}}_2^{(r)} &= \\ &\begin{bmatrix} 0 \sin\left(\frac{\pi}{M_r}\right) & 0 & \cdots & 0 & 0 \\ 0 \sin\left(\frac{\pi}{M_r}\right) \sin\left(\frac{2\pi}{M_r}\right) & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \sin\left(\frac{\pi(M_r-2)}{M_r}\right) & \sin\left(\frac{\pi(M_r-1)}{M_r}\right) \\ 0 & 0 & 0 & 0 & 0 & \sin\left(\frac{\pi(M_r-1)}{M_r}\right) \end{bmatrix} \end{aligned} \quad (20)$$

Lemma 4 (Real-valued selection matrices in DFT beamspace): Let us define $\bar{\mathbf{G}}_1^{(r)}$ in (19) and $\bar{\mathbf{G}}_2^{(r)}$ in (20). The matrices $\mathbf{K}_1^{(r)} \in \mathbb{R}^{(B_r-g_r) \times B_r}$ and $\mathbf{K}_2^{(r)} \in \mathbb{R}^{(B_r-g_r) \times B_r}$ fulfill (17) if they are constructed in the same way as $\mathbf{J}_1^{(r)}$ and $\mathbf{J}_2^{(r)}$, except that $\mathbf{G}_1^{(r)}$ and $\mathbf{G}_2^{(r)}$ are replaced by $\bar{\mathbf{G}}_1^{(r)}$ and $\bar{\mathbf{G}}_2^{(r)}$, respectively.

Proof: The proof is achieved by extending the derivations in [23] (or [20]) to the R -D case by following [27]. The details are omitted due to the space limitations.

As in the complex-valued case, the LS solution is calculated as

$$\hat{\Upsilon}^{(r)} = (\tilde{\mathbf{K}}_1^{(r)} [\mathcal{E}_s]_{(R+1)}^T)^+ \tilde{\mathbf{K}}_2^{(r)} [\mathcal{E}_s]_{(R+1)}^T, \forall r, \quad (21)$$

where $\tilde{\mathbf{K}}_1^{(r)} \in \mathbb{R}^{(B_r-g_r) \times \prod_{\bar{r}=1, \bar{r} \neq r}^R B_{\bar{r}} \times B}$

$$\tilde{\mathbf{K}}_1^{(r)} = \mathbf{I}_{\prod_{\bar{r}=r+1}^R B_{\bar{r}}} \otimes \mathbf{K}_j^{(r)} \otimes \mathbf{I}_{\prod_{\bar{r}=1}^{r-1} B_{\bar{r}}}, j = 1, 2. \quad (22)$$

Thanks to the forward-backward averaging (FBA) process (c.f. Appendix A), the number of effective snapshots used for the subspace estimation is doubled. Thereby, compared to R -D STE in DFT beamspace, R -D UTE in DFT beamspace algorithm can estimate more spatial frequencies or decorrelate correlated sources. The maximum number of resolvable sources of R -D UTE in DFT beamspace is given by

$$d_{\max} = \min \left(\min_r \left((B_r - g_r) \prod_{\bar{r}=1, \bar{r} \neq r}^R B_{\bar{r}} \right), 2N \right). \quad (23)$$

The derived R -D STE/UTE in DFT beamspace algorithm is summarized in Algorithm 1.

Algorithm 1: R -D Standard Tensor-ESPRIT (STE)/Unitary Tensor-ESPRIT (UTE) in DFT Beamspace.

1: **Tensor signal subspace estimation:**

- (STE) Compute \mathcal{U}_s from the rank- d truncated HOSVD of $\tilde{\mathcal{X}}$ in (1).
- (UTE) Compute \mathcal{E}_s from the rank- d truncated HOSVD of \mathcal{Z} in (15).

2: **Solution of the invariance equation:**

- (STE) Obtain $\hat{\Psi}^{(r)}$ as in (11) after solving (9) by LS.
- (UTE) Obtain $\hat{\Upsilon}^{(r)}$ as in (21) after solving (17) by LS.

3: **Spatial frequency estimation:**

- (STE) Obtain the estimates of $\hat{\Phi}^{(r)}$ by computing the joint EVD of $\hat{\Psi}^{(r)}$, $\forall r$. Compute $\hat{\mu}_i^{(r)} = \arg(\hat{\Psi}_{i,i}^{(r)})$.
 - (UTE) Obtain the estimates of $\hat{\Omega}^{(r)}$ by computing the joint EVD of $\hat{\Upsilon}^{(r)}$, $\forall r$. Compute $\hat{\mu}_i^{(r)} = 2 \arctan(\hat{\Omega}_{i,i}^{(r)})$.
-

Computational complexity: The computational complexity of ESPRIT-type algorithms in DFT beamspace is dominated by the computation of the tensor signal subspace \mathcal{U}_s (or \mathcal{E}_s). By following a similar analysis as in [26], the required number of complex-valued multiplications for R -D STE in DFT beamspace is $c_0 B \cdot N \cdot d \cdot (R+1) + B \cdot N \cdot d \cdot R + B \cdot d^2 \cdot R$, where c_0 is a constant design parameter for the SVD computation. The matrix version requires $c_0 B \cdot N \cdot d$ complex-valued multiplications. If one complex-valued multiplication is considered as four real-valued multiplications, the required (real-valued) multiplications of R -D UTE in DFT beamspace are only half of

that required by R -D STE in DFT beamspace. This calculation takes into account the fact that the number of snapshots in R -D UTE is $2N$ instead of N .

C. Hybrid R -D Tensor-ESPRIT

For applications such as the channel estimation problem in Section III, we can use the ESPRIT-type algorithms in element space in the frequency dimension because we do not have a reduced dimensional SoI in this dimension. Since the selection matrices applied in different dimensions are independent from each other, it is not difficult to see that the use of element-space or DFT-beamspace selection matrices will not violate the shift invariance property in (9) (or correspondingly in (17)). Therefore, the solution framework in Algorithm 1 can still be used with some modifications. More precisely, let us assume that element-space ESPRIT is used in the \bar{r} -th dimension. When STE is used, we have $\mathbf{F}^{(\bar{r})} = \mathbf{I}_{M_{\bar{r}}}$. The matrices $\mathbf{J}_j^{(\bar{r})}$ in (12) are replaced by selection matrices of STE in element space in [28]. When UTE is used, we use $\mathbf{F}^{(\bar{r})} = \mathbf{Q}_{M_{\bar{r}}}$ for constructing \mathcal{Z} of (15), where $\mathbf{Q}_{M_{\bar{r}}}$ can be equivalent to the unitary sparse left- Π -real matrices $\mathbf{Q}_{M_{\bar{r}}}^{(s)}$ defined accordingly in [29]. Similarly, the matrices $\mathbf{K}_j^{(\bar{r})}$ in (22) are replaced by the selection matrices of UTE in element space in [28]. In order to distinguish this type of ESPRIT algorithms from ESPRIT-type algorithms in pure element space or in pure DFT beamspace, we refer to it as hybrid R -D Tensor-ESPRIT.

D. Analytic Performance

To derive the analytic performance of R -D STE in DFT beamspace we use tensor-based first-order perturbation analysis in [28]. Let us define the SVD of the $(R+1)$ -th mode unfolding of \mathcal{X}_0 as

$$[\mathcal{X}_0]_{(R+1)}^T = \mathbf{U}_s \Sigma_s \mathbf{V}_s^H + \mathbf{U}_n \Sigma_n \mathbf{V}_n^H, \quad (24)$$

where \mathcal{X}_0 is the noiseless version of $\bar{\mathcal{X}}$ and $\Sigma_n = \mathbf{0}_{(B-d) \times (B-d)}$. The main result of the first-order perturbation analysis of the HOSVD from [28] can be directly applied, i.e., the tensor signal subspace estimate $\hat{\mathcal{U}}_s$ of $\bar{\mathcal{X}}$ satisfies

$$[\hat{\mathcal{U}}_s]_{(R+1)}^T = \mathbf{U}_s + [\Delta \mathcal{U}_s]_{(R+1)}^T, \quad (25)$$

where $\Delta \mathcal{U}_s$ is first-order estimation error defined in [28, Th. 2].

Let \mathbf{p}_i^T and \mathbf{q}_i denote the i -th row and the i -th column ($i \in \{1, \dots, d\}$) of \mathbf{Q}^{-1} and \mathbf{Q} as defined in (10), respectively. Let us write the perturbed spatial frequencies as $\hat{\mu}_i^{(r)} = \mu_i^{(r)} + \Delta \mu_i^{(r)}$. Assuming that the perturbation caused by the noise \mathcal{N} is small, we have

$$\Delta \mu_i^{(r)} \approx \text{Im} \left\{ \mathbf{p}_i^T \cdot \left(\tilde{\mathbf{J}}_1^{(r)} \cdot \mathbf{U}_s \right)^+ \cdot \left[\tilde{\mathbf{J}}_2^{(r)} / \lambda_i^{(r)} - \tilde{\mathbf{J}}_1^{(r)} \right] \cdot [\Delta \mathcal{U}_s]_{(R+1)}^T \cdot \mathbf{q}_i \right\}, \quad (26)$$

where $\lambda_i^{(r)} = e^{j\mu_i^{(r)}}$. Note that equation (26) is a direct extension of the result in [28].

Theorem 1: When circularly symmetric noise \mathcal{N} is present and R -D STE in DFT beamspace is used, the first-order approximation of the MSE for the i -th spatial frequency in the r -th

dimension is given by

$$\mathbb{E}\{(\Delta \mu_i^{(r)})^2\} = \frac{1}{2} M \sigma_n^2 \|\mathbf{W}_{\text{ten}}^T \mathbf{r}_i^{(r)}\|_2^2, \quad (27)$$

where $\mathbf{r}_i^{(r)} = \mathbf{q}_i \otimes \left(\left(\tilde{\mathbf{J}}_1^{(r)} \cdot \mathbf{U}_s \right)^+ \cdot \left(\tilde{\mathbf{J}}_2^{(r)} / \lambda_i^{(r)} - \tilde{\mathbf{J}}_1^{(r)} \right) \right)^T \cdot \mathbf{p}_i$. The term \mathbf{W}_{ten} is defined in a similar way as in (75) of [28] except that the corresponding subspace estimates are replaced by subspace estimates defined above.

Proof: The derivation is a direct extension of [28].

The derivation of the analytic performance of R -D UTE in DFT beamspace is achieved by following the same approach as for R -D STE in DFT beamspace. Let \mathcal{Z}_0 represent the noiseless version of \mathcal{Z} in (15). Then we have the following statement.

Theorem 2: Assume that circularly symmetric noise \mathcal{N} is present and R -D UTE in DFT beamspace is used. The first-order expansion of the estimation error for the i -th spatial frequency in the r -th dimension is expressed as

$$\Delta \mu_i^{(r)} \approx \left\{ \tilde{\mathbf{p}}_i^T \cdot \left(\tilde{\mathbf{K}}_1^{(r)} \cdot \mathbf{E}_s \right)^+ \cdot \left[\tilde{\mathbf{K}}_2^{(r)} - \tilde{\lambda}_i^{(r)} \tilde{\mathbf{K}}_1^{(r)} \right] \cdot [\Delta \mathcal{E}_s]_{(R+1)}^T \cdot \tilde{\mathbf{q}}_i \right\} \cdot 2 / (1 + \tilde{\lambda}_i^{(r)^2}), \quad (28)$$

where $\tilde{\lambda}_i^{(r)} = \tan(\mu_i^{(r)}/2)$, $\tilde{\mathbf{p}}_i^T$ and $\tilde{\mathbf{q}}_i$ denote the i -th row and the i -th column ($i \in \{1, \dots, d\}$) of \mathbf{T}^{-1} and \mathbf{T} as defined in (18), respectively. The corresponding MSE of the i -th spatial frequency in the r -th dimension is given by

$$\mathbb{E}\{(\Delta \mu_i^{(r)})^2\} = \sqrt{2} M \sigma_n^2 \|\mathbf{W}_{\text{ten}}^{(\text{UTE})T} \tilde{\mathbf{r}}_i^{(r)}\|_2^2, \quad (29)$$

where $\tilde{\mathbf{r}}_i^{(r)} = \tilde{\mathbf{q}}_i \otimes \left(\left(\tilde{\mathbf{K}}_1^{(r)} \cdot \mathbf{E}_s \right)^+ \cdot \left(\tilde{\mathbf{K}}_2^{(r)} - \tilde{\lambda}_i^{(r)} \tilde{\mathbf{K}}_1^{(r)} \right) \right)^T \cdot \tilde{\mathbf{p}}_i \cdot 2 / (1 + \tilde{\lambda}_i^{(r)^2})$. The matrix $\mathbf{W}_{\text{ten}}^{(\text{UTE})}$ is similar as \mathbf{W}_{ten} except that all the corresponding terms are replaced by their equivalents, which are computed from \mathcal{Z}_0 .

Proof: See Appendix B. ■

More insights regarding the difference between R -D ESPRIT-type algorithms in DFT beamspace and in element space can be obtained by studying scenarios with a single source. Since the use of Tensor-ESPRIT or Unitary ESPRIT does not bring a performance gain for a single source scenario, we consider R -D Standard ESPRIT in DFT beamspace. Moreover, we study the case with the full DFT beamspace, i.e., $B_r = M_r$, and the reduced-dimensional case with $B_r = 2, \forall r$.

Lemma 5: Assume that R -D ESPRIT in DFT beamspace is used. When $d = 1$ and $B_r = M_r, \forall r$, the MSE for the spatial frequency in the r -th mode is given by

$$\mathbb{E}\{(\Delta \mu^{(r)})^2\} = \frac{\sigma_n^2}{P_T M N} \cdot \sin^2 \left(\frac{\pi}{M_r} \right). \quad (30)$$

When $d = 1, B_r < M_r$, and $B_r = 2, \forall r$, the MSE for the spatial frequency in the r -th mode is given by

$$\mathbb{E}\{(\Delta \mu^{(r)})^2\} = \frac{M \sigma_n^2}{P_T N} \cdot \frac{2 - \text{Re} \left\{ \lambda^{(r)*} \mathbf{g}_{2,m^{(r)}}^{(r)} \mathbf{g}_{1,m^{(r)}}^{(r)H} \right\}}{\prod_{p=1, p \neq r}^R \|\mathbf{b}^{(p)}\|^2 |\mathbf{g}_{1,m^{(r)}}^{(r)} \mathbf{b}^{(r)}|^2}, \quad (31)$$

where $m^{(r)}$ denotes the index of the first row of the two selected consecutive rows of the DFT matrix. The vectors $\mathbf{g}_{1,m^{(r)}}^{(r)}$ and $\mathbf{g}_{1,m^{(r)}}^{(r)}$ denote the $m^{(r)}$ -th row of the selection matrices $\mathbf{G}_1^{(r)}$ and

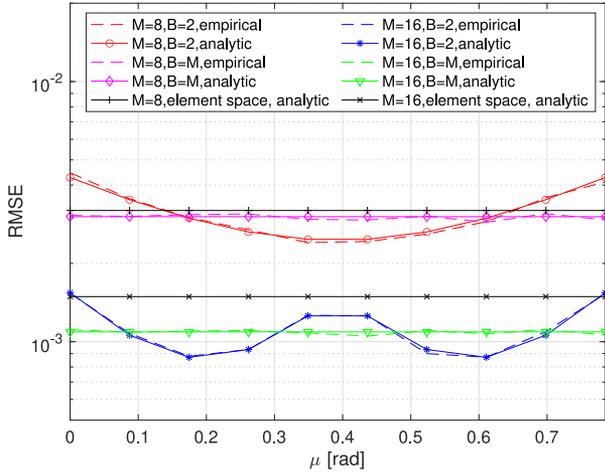


Fig. 3. Achievable empirical and analytic RMSE of 1-D ESPRIT in DFT beamspace or in element space. $\mu \in [0, 2\pi/8]$, $P_T = 1$ W, SNR = 20 dB, and $N = 20$ snapshots. When $M = 8$, the SoI is in $[0, 2\pi/8]$. When $M = 16$, the SoI is $[0, \pi/8]$ for $\mu \leq \pi/8$ and $(\pi/8, 2\pi/8]$ for $\mu > \pi/8$.

$\mathbf{G}_2^{(r)}$, respectively. We have defined that $\mathbf{b}^{(r)} = \mathbf{D}_{M_r}(m^{(r)} : m^{(r)} + 1, :) \cdot \mathbf{a}(\mu^{(r)})$.

Proof: See Appendix C.

Lemma 5 coincides with the observations in numerical simulations. That is, when $B_r < M_r$, the achievable MSE is also a function of the spatial frequencies. Lastly, if we compare the achievable MSE in (30) and that of the R -D ESPRIT in element space, we obtain the following result.

Corollary 3: When $d = 1$ and $B_r = M_r = 2$ or $B_r = M_r \geq 7$, R -D ESPRIT in DFT beamspace provides a smaller MSE for the spatial frequency in the r -th mode than R -D ESPRIT in element space.

Proof: According to [30], the MSE of $\mu^{(r)}$ by using the element space R -D Standard ESPRIT is given by

$$\mathbb{E}\{(\Delta\mu^{(r)})^2\} = \frac{\sigma_n^2}{P_T M N} \cdot \frac{M_r}{(M_r - 1)^2}, \quad (32)$$

when a single source is considered. Both functions in (30) and in (32) monotonically decrease when M_r increases. Numerically it is found that the value of (30) is smaller than the value of (32) when $M_r = 2$ or $M_r \geq 7$. Note that both functions are equal when $M_r = 3$.

In general, it is difficult to derive the exact gain of using ESPRIT-type algorithms in a reduced-dimensional ($B_r < M_r$) DFT beamspace. In Fig. 3 we verify the derived analytic performance and use it to numerically illustrate the gain of using ESPRIT-type algorithms in a reduced-dimensional DFT beamspace. The selected range of μ is the SoI covered by the first two DFT beams for $M = 8$ while it is the SoI covered by either the first two DFT beams ($\mu \leq \pi/8$) or by the second and the third DFT beams ($\mu > \pi/8$) for $M = 16$. As B decreases, the achievable root MSE (RMSE) $\sqrt{\mathbb{E}\{|\mu - \hat{\mu}|^2\}}$ decreases when μ is close to the center of SoI. The achievable RMSE also decreases when M increases. Hence, compared to ESPRIT-type algorithms in element space, a large performance gain is achievable when ESPRIT-type algorithms in reduced-dimensional DFT

beamspace are used, the spatial frequencies lie in a certain range, the array size is large, and we are in a high-SNR regime. This conclusion should be also valid for a R -D case.

Remark 1: Smoothing can significantly increase the number of effective snapshots and subsequently increase the number of identifiable sources [26]. In general, smoothing cannot be used on the received signal of ESPRIT-type algorithms in DFT beamspace. But it can be used together with hybrid R -D Tensor-ESPRIT, i.e., smoothing as described in [26] is used on the dimensions where we use the element space implementation of ESPRIT.

Remark 2: Fig. 2 shows that the performance of ESPRIT-type algorithms in DFT beamspace degrades significantly if the spatial frequencies are out of the SoI. A large SoI can reduce the possibility of out of SoI spatial frequencies. However, Fig. 3 implies that the enlarged SoIs result in a worse performance. Therefore, the optimization of the SoIs is critical for the performance of ESPRIT-type algorithms in DFT beamspace.

III. CHANNEL ESTIMATION PROBLEM

A. System Model and Training Procedure

We study a point-to-point hybrid mmWave massive MIMO-OFDM system. The received training signal on the n -th pilot subcarrier in the m -th OFDM symbol is given by [19]

$$\mathbf{y}_n[m] = \mathbf{W}^H[m](\mathbf{H}_n[m]\mathbf{F}[m]\mathbf{s}_n[m] + \mathbf{z}_n[m]), \quad (33)$$

where $\mathbf{z}_n[m]$ represents ZMCSCG noise with covariance matrix $\mathbb{E}\{\mathbf{z}_n[m]\mathbf{z}_n^H[m]\} = \sigma_n^2 \mathbf{I}_{M_R}$ for $n \in \{1, \dots, N_f\}$. Hybrid decoding and precoding matrices are denoted by $\mathbf{W}^H[m] \in \mathbb{C}^{N_R \times M_R}$ and $\mathbf{F}[m] \in \mathbb{C}^{M_T \times N_T}$, respectively. We have M_T antennas and N_T RF chains at the transmitter ($M_T > N_T$) and M_R antennas and N_R RF chains at the receiver ($M_R > N_R$). If it is not further specified, the digital precoder and decoder on each pilot subcarrier are identity matrices. Thereby, $\mathbf{W}^H[m]$ and $\mathbf{F}[m]$ also represent the effective analog decoder and precoder, which have only unit modulus entries [19]. Let k_n denote the index that satisfies $\{k_1, \dots, k_{N_f}\} \subset \{0, \dots, N_{\text{fft}} - 1\}$, where N_{fft} is the FFT size. The pilots have an equal distance $\Delta_{\text{pilot}} \in \mathbb{Z}$ and $k_n = \Delta_{\text{pilot}}(n - 1)$. The pilot symbols $\mathbf{s}_n[m] \in \mathbb{C}^{N_T}$ have to fulfill the transmit power constraint $\sum_{n=1}^{N_f} \|\mathbf{F}[m]\mathbf{s}_n[m]\|^2 = P_{\text{pilot}}, \forall m$.

Our proposed channel estimation algorithm is carried out in the training phase, which consists of κ_t training frames. For notational simplicity, the κ -th ($\kappa \in \{1, \dots, \kappa_t\}$) training frame contains N_T OFDM symbols. Hence, the training phase lasts $N_t = \kappa_t N_T$ OFDM symbols and $m \in \{1, \dots, N_t\}$. An identical time-frequency pilot pattern is used in all training frames. Let $\mathbf{S}_{n,\kappa} \in \mathbb{C}^{N_T \times N_T}$ denote the accumulated pilot matrix within the κ -th frame on the n -th pilot subcarrier. We design $\mathbf{S}_{n,\kappa}$ such that its Gramian is a scaled identity matrix, i.e., $\mathbf{S}_{n,\kappa} \mathbf{S}_{n,\kappa}^H = \frac{P_{\text{pilot}}}{M_T N_f} \mathbf{I}_{N_T}, \forall n, \kappa$. The hybrid precoding and decoding matrices remain constant in all OFDM symbols in each training frame. Moreover, \mathbf{W}_κ^H and \mathbf{F}_κ denote the unique hybrid decoder and precoder in the κ -th frame. Therefore, the frame index κ will replace the OFDM symbol index m in the rest of the paper. For simplicity, we assume that the channel is approximately time

invariant and thus $\mathbf{H}_n[m]$ is not a function of m . Finally, by stacking $\mathbf{y}_n[m]$ within the κ -th frame next to each other, the effective received pilot data matrix on the n -th subcarrier in the κ -th frame is expressed as,

$$\mathbf{Y}_{n,\kappa}^{(\text{orig})} = \mathbf{W}_{\kappa}^{\text{H}}(\mathbf{H}_n \mathbf{F}_{\kappa} \mathbf{S}_{n,\kappa} + \mathbf{Z}_{n,\kappa}) \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}}}, \quad (34)$$

where $\mathbf{Z}_{n,\kappa} \in \mathbb{C}^{M_{\text{R}} \times N_{\text{T}}}$ denotes the stacked noise vectors of the n -th subcarrier in the κ -th training frame. Note that the same training procedure is used in [12], [20]. The sampled frequency-domain channel transfer function (CTF) \mathbf{H}_n is given by [20], [31]

$$\mathbf{H}_n = \sum_{\ell=1}^L d_{\ell} \mathbf{a}_{\text{R}}(\mu_{\text{R},\ell}) \mathbf{a}_{\text{T}}^{\text{H}}(\mu_{\text{T},\ell}) e^{j(n-1)\mu_{\text{f},\ell}}, \quad (35)$$

where d_{ℓ} denotes the complex gain of the ℓ -th path, $\mu_{\text{R},\ell} \in [-\pi, \pi)$ denotes the DoA, and $\mu_{\text{T},\ell} \in [-\pi, \pi)$ denotes the DoD. The vectors $\mathbf{a}_{\text{T}}(\mu_{\text{T},\ell})$ and $\mathbf{a}_{\text{R}}(\mu_{\text{R},\ell})$ denote the array steering vectors of the transmitter and the receiver, respectively. We consider uniform linear arrays (ULAs), where the array steering vector is $\mathbf{a}(\mu) = [1 \quad e^{j\mu} \quad \dots \quad e^{j(M-1)\mu}]^{\text{T}} \in \mathbb{C}^M$. Here we consider a sampling period of $T_{\text{s}} = 1/(N_{\text{fft}} \cdot \Delta_{\text{f}})$, where Δ_{f} is the subcarrier spacing. For the frequency dimension we define $\mu_{\text{f},\ell} = -2\pi \Delta_{\text{pilot}} \tau_{\ell} \Delta_{\text{f}}$ and $0 \leq \tau_{\ell} \Delta_{\text{f}} \leq N_{\text{cp}}/N_{\text{fft}} < 1$. It is worth stressing that (35) can also be extended to incorporate a wideband model and Doppler shifts, e.g., [32]. The extension to uniform rectangular array is straightforward.

B. Pilot Demodulation and Tensor Model

By multiplying the result by $\mathbf{S}_{n,\kappa}^{\text{H}} M_{\text{T}} N_{\text{f}} / P_{\text{pilot}}$ from the right-hand side of (34), the demodulated data on the n -th pilot subcarrier is given by $\mathbf{Y}_{n,\kappa} = \mathbf{W}_{\kappa}^{\text{H}} \mathbf{H}_n \mathbf{F}_{\kappa} + \bar{\mathbf{Z}}_{n,\kappa} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}}}$, where $\bar{\mathbf{Z}}_{n,\kappa} = \mathbf{W}_{\kappa}^{\text{H}} \mathbf{Z}_{n,\kappa} \mathbf{S}_{n,\kappa}^{\text{H}} M_{\text{T}} N_{\text{f}} / P_{\text{pilot}}$. By stacking $\mathbf{y}_{n,\kappa} = \text{vec}\{\mathbf{Y}_{n,\kappa}\}$ on top of each other for $n = 1, 2, \dots, N_{\text{f}}$, we combine the demodulated signal from different pilot subcarriers, i.e.,

$$\begin{aligned} \mathbf{y}_{\kappa} &= (\mathbf{I}_{N_{\text{f}}} \otimes \mathbf{F}_{\kappa}^{\text{T}} \otimes \mathbf{W}_{\kappa}^{\text{H}}) \underbrace{(\mathbf{A}_{\text{f}} \diamond \mathbf{A}_{\text{T}} \diamond \mathbf{A}_{\text{R}})}_{\mathbf{h}_{\text{vec}}} \cdot \mathbf{d} + \bar{\mathbf{z}}_{\kappa} \\ &= (\mathbf{A}_{\text{f}} \diamond (\mathbf{F}_{\kappa}^{\text{T}} \mathbf{A}_{\text{T}}) \diamond (\mathbf{W}_{\kappa}^{\text{H}} \mathbf{A}_{\text{R}})) \cdot \mathbf{d} + \bar{\mathbf{z}}_{\kappa} \in \mathbb{C}^{N_{\text{R}} \bar{N}_{\text{t}}}, \end{aligned} \quad (36)$$

where $\mathbf{d} = [d_1 \dots d_L]^{\text{T}}$ and $\bar{\mathbf{z}}_{\kappa} = \text{vec}([\bar{\mathbf{Z}}_{1,\kappa} \bar{\mathbf{Z}}_{2,\kappa} \dots \bar{\mathbf{Z}}_{N_{\text{f},\kappa}}])$. The Vandermonde matrices \mathbf{A}_{f} , \mathbf{A}_{R} , and \mathbf{A}_{T} are defined by $\mathbf{A}_x = [\mathbf{a}_x(\mu_{x,1}) \dots \mathbf{a}_x(\mu_{x,L})]$, where $x \in \{\text{f}, \text{R}, \text{T}\}$.

The vector \mathbf{y}_{κ} can be interpreted as the transpose of the 4-mode unfolding of a tensor $\mathcal{Y}_{\kappa} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}} \times N_{\text{f}} \times 1}$, which is defined as

$$\begin{aligned} \mathcal{Y}_{\kappa} &= \mathcal{I}_{4,L} \times_1 (\mathbf{W}_{\kappa}^{\text{H}} \mathbf{A}_{\text{R}}) \times_2 (\mathbf{F}_{\kappa}^{\text{T}} \mathbf{A}_{\text{T}}) \times_3 \mathbf{A}_{\text{f}} \times_4 \mathbf{d}^{\text{T}} \\ &\quad + \mathcal{Z}_{\kappa} \times_1 \mathbf{W}_{\kappa}^{\text{H}} \times_2 \mathbf{S}_{n,\kappa}^{\text{H}} \times_3 \mathbf{P}, \end{aligned} \quad (37)$$

where $\mathcal{Z}_{\kappa} \in \mathbb{C}^{M_{\text{R}} \times N_{\text{T}} \times N_{\text{f}}}$ is the noise tensor with $\mathbf{Z}_{n,\kappa}$ as its n -th 3-mode slice, and $\mathbf{P} = M_{\text{T}} N_{\text{f}} / P_{\text{pilot}} \cdot \mathbf{I}_{N_{\text{f}}}$.

Given the above system assumptions, our goal is to design \mathbf{W}_{κ} and \mathbf{F}_{κ} , $\forall \kappa$ and the corresponding channel estimation algorithm, such that the channel is accurately estimated.

IV. TWO-STAGE GRIDLESS CHANNEL ESTIMATION

After comparing the system model (37) to the general R -D model in (1) or (14), we notice that (37) is a 3-D special case of (1) or (14) with a single snapshot ($N = 1$ because \mathbf{d} is a vector). Since the hybrid precoder at the transmitter does not affect the noise at the receiver, the noise model is not the same as in (2). Nevertheless, it can be easily proven that the n -mode unfolding of the noise tensor is still ZMCSCG distributed. The hybrid 3-D Tensor-ESPRIT algorithms in DFT beamspace (element space in the frequency dimension and DFT beamspace in the two spatial dimensions) can still be used to estimate the (spatial) frequencies in (37). Inspired by this fact, we propose a two-stage channel estimation algorithm. In the first stage, the spatial frequencies are estimated by using hybrid 3-D Tensor-ESPRIT as described in Section II-C. In the second stage, the complex path gain vector \mathbf{d} is estimated by using a LS method. Function-wise, the proposed channel estimation is achieved in three steps.

- 1) **Coarse estimation**, where rough estimates of the (dominant) spatial frequencies $\mu_{\text{R},\ell}$ and $\mu_{\text{T},\ell}$ are obtained. The estimated spatial frequencies are used to define SoIs in DFT beamspace. To obtain the rough estimates, we propose to form a wide-beam by switching off antenna elements. Then the spatial frequencies are estimated by using 3-D UTE in element space. The details are in Section IV-A.
- 2) **Sectorization**, where the SoIs to be scanned are determined according to the rough estimates of $\mu_{\text{R},\ell}$ and $\mu_{\text{T},\ell}$. The sectorization step determines the number of required training frames in the fine estimation stage. The details are in Section IV-B.
- 3) **Fine estimation**, where high-resolution estimates of the spatial frequencies and the path gain are computed by using the hybrid 3-D STE/UTE algorithm. In particular, the DFT beamspace ESPRIT algorithm is applied in the transmit and receive dimensions, and the element space ESPRIT algorithm is used in the frequency dimension. The hybrid decoder $\mathbf{W}_{\kappa}^{\text{H}}$ and the transpose of the hybrid precoder $\mathbf{F}_{\kappa}^{\text{T}}$ are constructed as $\mathbf{F}^{(r)}$ in Section II-A or $\tilde{\mathbf{F}}^{(r)}$ in Section II-B. The details are in Section IV-C.

A. Coarse Estimation

As already discussed in [20], there are several ways to obtain the prior knowledge of spatial frequencies, including antenna switch-off, determining the SoIs from previous CSI measurements or from Radar measurements [33], and so on. Note that our proposed two step procedure is ideally suited for time varying scenarios, since we can skip the coarse estimation step if previous estimates of parameters are available. In this case, we choose the sectors of interest such that the previous estimates are approximately in their center. The channel parameters must stay approximately constant during at least the $T_{\text{t}} = N_{\text{t}} \cdot T_{\text{s}} \cdot N_{\text{fft}}$ seconds, where T_{s} is the sampling interval.

If no prior knowledge is available we propose the following procedure. For simplicity we consider switching off the antennas as in [20]. More precisely, we switch off the last $(M_{\text{T}} - N_{\text{T}})$

transmit antennas and the last $(M_R - N_R)$ receive antennas for one training frame during the coarse estimation. In this way we obtain an antenna system where the number of RF chains is equal to the number of effective transmit/receive antennas. We assume that the loss of array gain due to the antenna switch-off will not cause any link failures. This is a valid assumption especially for the downlink communication. We use only one training frame. Hence, we drop the index κ . Let us define $\mathbf{A}_T^{(\text{eff})} = \mathbf{A}_T(1 : N_T, :) \in \mathbb{C}^{N_T \times L}$ and $\mathbf{A}_R^{(\text{eff})} = \mathbf{A}_R(1 : N_R, :) \in \mathbb{C}^{N_R \times L}$. Furthermore, we design the effective hybrid precoding matrix as $\mathbf{F} = \mathbf{M}_T \cdot \mathbf{F}_A \cdot \mathbf{F}_D = [\mathbf{I}_{N_T} \mathbf{0}_{(M_T - N_T) \times N_T}^T]^T$ for all subcarriers, where $\mathbf{M}_T = \text{diag}([\mathbf{I}_{N_T}^T \mathbf{0}_{M_T - N_T}^T]) \in \mathbb{R}^{M_T \times M_T}$ is a mask matrix due to the antenna switch-off, $\mathbf{F}_A = [\mathbf{D}_T^T \mathbf{R}_T^T]^T \in \mathbb{C}^{M_T \times N_T}$ is the analog precoding matrix where $\mathbf{D}_T \in \mathbb{C}^{N_T \times N_T}$ is a DFT matrix and $\mathbf{R}_T \in \mathbb{C}^{(M_T - N_T) \times N_T}$ contains random phases that are irrelevant because they are multiplied by zeros, and $\mathbf{F}_D = \mathbf{D}_T^H$ is the identical digital precoder on all pilot subcarriers. Similarly, we design the effective hybrid decoding matrix as $\mathbf{W} = [\mathbf{I}_{N_R} \mathbf{0}_{(M_R - N_R) \times N_R}^T]$ on all pilot subcarriers. Finally, in the coarse estimation phase, the system model of (37) simplifies to

$$\mathbf{Y} = \mathcal{I}_{4,L} \times_1 \mathbf{A}_R^{(\text{eff})} \times_2 \mathbf{A}_T^{(\text{eff})} \times_3 \mathbf{A}_f \times_4 \mathbf{d}^T + \mathcal{Z}^{(\text{eff})} \quad (38)$$

where $\mathcal{Z}^{(\text{eff})} = \mathcal{Z}(1 : N_R, :, :) \in \mathbb{C}^{N_R \times N_T \times N_f}$ denotes the effective noise tensor and the facts $\mathbf{A}_R^{(\text{eff})} = \mathbf{W}^H \mathbf{A}_R$ and $\mathbf{A}_T^{(\text{eff})} = \mathbf{F}^T \mathbf{A}_T$ are used. The system model of (38) is a wide beam model compared to (37). It corresponds to the 3-D harmonic-retrieval problem in [26] with one snapshot. Hence, 3-D Tensor-ESPRIT in element space can be used to estimate the spatial frequencies $(\mu_{T,\ell}, \mu_{R,\ell}), \forall \ell$. To increase the number of identifiable scatters, we apply the smoothing concept in [26] on the frequency dimension of \mathbf{Y} to increase the number of effective snapshots. Let $\mathbf{Y}_{\text{ss}} \in \mathbb{C}^{N_R \times N_T \times M_{\text{sub},f} \times L_f}$ denote the smoothed tensor that is constructed as in [26], i.e.,

$$\mathbf{Y}_{\text{ss}} = [\mathcal{Y}_{1,1,1} \quad \sqcup_4 \mathcal{Y}_{1,1,2} \quad \cdots \quad \sqcup_4 \mathcal{Y}_{1,1,L_f}], \quad (39)$$

where we have $M_{\text{sub},f} = N_f - L_f + 1$, L_f is the number of frequency subblocks for smoothing, and $M_{\text{sub},f}$ is the number of subcarriers per subblock. We have $\mathcal{Y}_{1,1,\ell_f} = \mathbf{Y} \times_3 \mathbf{J}_{\text{ss},\ell_f} \in \mathbb{C}^{N_R \times N_T \times M_{\text{sub},f} \times 1}$ for $\ell_f = 1, \dots, L_f$. The selection matrix $\mathbf{J}_{\text{ss},\ell_f}$ is defined as

$$\mathbf{J}_{\text{ss},\ell_f} = \begin{bmatrix} \mathbf{0}_{M_{\text{sub},f} \times (\ell_f - 1)} & \mathbf{I}_{M_{\text{sub},f}} & \mathbf{0}_{M_{\text{sub},f} \times (L_f - \ell_f)} \end{bmatrix}.$$

Then 3-D UTE is used on \mathbf{Y}_{ss} to provide estimates of the spatial frequencies. We use UTE because it provides the best performance in the low SNR regime [26], which can compensate the reduced array gain due to the antenna switch-off. Finally, the maximum number of resolvable scatterers in the coarse estimation stage is

$$L_{\text{max}} = \min((M_{\text{sub},f} - 1)N_T N_R, M_{\text{sub},f} \cdot (N_T - 1)N_R, M_{\text{sub},f} N_T (N_R - 1), 2L_f). \quad (40)$$

B. Sectorization

As discussed in Section II-A, the beamspace of the SoIs dominants the performance of ESPRIT-type algorithms in DFT

beamspace. A large SoI leverages the mis-detection in the coarse estimation but it also reduces the beamforming gain. In contrast, a small SoI fully exploits the gain of ESPRIT-type algorithms in DFT beamspace. It is non-trivial to find the optimal SoIs that achieve the minimum estimation error. Hence, we resort to a sub-optimal sliding-window based sectorization method, where the (minimum) size of a SoI is determined by the window length.

Taking the transmit dimension as an example, we sort L coarse estimates of the spatial frequencies $\mu_{T,\ell}$ in ascending order, i.e., $0 \leq \mu_{T,1} < \mu_{T,2} < \dots < \mu_{T,L} < 2\pi$. Let \mathbb{T} denote the set of indices, which correspond to the unclassified spatial frequencies. Let $W_T \frac{2\pi}{M_T}$ ($W_T \in \mathbb{Z}_+$) denote the window length in the transmit dimension. Starting from 0, we move the window at a step size of $\frac{2\pi}{M_T}$ until the following two objectives are fulfilled. First, there should be as many coarse estimates (including the first element $\mathbb{T}\{1\}$) as possible to be within the SoI. The goal is to reduce the total number of SoIs so that the required training frames can be reduced. Second, inspired by Corollary 2, the coarse estimates that lie in the SoI should be close to the geometric center of the SoI. To this end, we restrict the maximum distance between the center of the SoI and the coarse estimates within the SoI to be no larger than $\zeta_T \frac{\pi}{M_T}$ ($\zeta_T \in \mathbb{Z}_+$, $\zeta_T \leq W_T$). Once the two conditions are fulfilled, the beam indices within the window, i.e., $\{k_{\text{opt}}^{(\hat{g}^{(T)})}, \dots, k_{\text{opt}}^{(\hat{g}^{(T)})} + W_T\}$ (if $k_{\text{opt}}^{(\hat{g}^{(T)})} + W_T > M_T$, then $\{k_{\text{opt}}^{(\hat{g}^{(T)})}, \dots, M_T, 1, \dots, k_{\text{opt}}^{(\hat{g}^{(T)})} + W_T - M_T\}$), are the candidate beam indices for the $\hat{g}^{(T)}$ -th SoI, where $\hat{g}^{(T)} \in \{1, \dots, \hat{G}_T\}$. The fact that the first and last DFT beams are adjacent should be taken into account. To avoid an exhaustive search, we propose to formulate the two objectives as two sequential optimization problems. Let the set $\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}$ contains the indices of the coarse estimates that lie in the $\hat{g}^{(T)}$ -th window. The first optimization problem is,

$$\begin{aligned} & \max_{\bar{j}, k^{(\hat{g}^{(T)})}} \text{Cardi}\{\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}\} \\ & \text{s.t. } \Delta_{T,\bar{j}}(k^{(\hat{g}^{(T)})}) \leq \zeta_T \frac{\pi}{M_T}, \forall \bar{j} \\ & \bar{j} \in \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})} \subseteq \mathbb{T}, \mathbb{T}\{1\} \in \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}, \bar{j} \in \mathbb{Z}_+, \end{aligned} \quad (41a)$$

where $\Delta_{T,\bar{j}}(k^{(\hat{g}^{(T)})}) = \min(|\mu_{T,\bar{j}} - \frac{(2k^{(\hat{g}^{(T)})} + W_T - 2)\pi}{M_T}|, |\mu_{T,\bar{j}} + 2\pi - \frac{(2k^{(\hat{g}^{(T)})} + W_T - 2)\pi}{M_T}|)$ denotes the distance to the window center and takes into account the wrap-around property of the beam indices, and $k^{(\hat{g}^{(T)})}$ denotes the starting beam index. Problem (41) is a non-convex mixed-integer programming problem. The cardinality is maximized if the center of the SoI is larger than $\mu_{T,\mathbb{T}\{1\}}$. From (41a) we derive that $k_{\text{max}}^{(\hat{g}^{(T)})} = \lfloor \frac{\mu_{T,\mathbb{T}\{1\}} M_T + (\zeta_T + 2 - W_T)\pi}{2\pi} \rfloor$. Problem (41) is simplified to a feasibility problem, i.e.,

$$\begin{aligned} & \text{find } \bar{j} \\ & \text{s.t. } \Delta_{T,\bar{j}}(k_{\text{max}}^{(\hat{g}^{(T)})}) \leq \zeta_T \frac{\pi}{M_T} \\ & \bar{j} \in \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})} \subseteq \mathbb{T}, \mathbb{T}\{1\} \in \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}, \bar{j} \in \mathbb{Z}_+. \end{aligned} \quad (42)$$

Next, we find the optimal value $k_{\text{opt}}^{(\hat{g}^{(T)})}$ that minimizes the sum of the Euclidean distances between the coarse estimates and the center of the SoI, i.e.,

$$\min_{k^{(\hat{g}^{(T)})}} \sum_{\bar{j}} \Delta_{T,\bar{j}}(k^{(\hat{g}^{(T)})})$$

$$\text{s.t. } \bar{j} \in \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}, 1 \leq k^{(\hat{g}^{(T)})} \leq \left\lfloor \frac{\mu_{T,\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}\{1\}} M_T}{2\pi} \right\rfloor. \quad (43)$$

The optimal solution of the non-convex problem (43) is given by

$$k_{\text{opt}}^{(\hat{g}^{(T)})} = \min \left(\bar{k}^{(\hat{g}^{(T)})}, \lfloor \mu_{T,\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}\{1\}} M_T / (2\pi) \rfloor \right), \quad (44)$$

where $\bar{k}^{(\hat{g}^{(T)})} = \text{round}\{\frac{\bar{\mu}_T^{(\hat{g}^{(T)})} M_T}{2\pi}\} + 1$, when W_T is even, and $\bar{k}^{(\hat{g}^{(T)})} = \frac{1}{2}(\text{round}\{\frac{\bar{\mu}_T^{(\hat{g}^{(T)})} M_T - \pi}{2\pi}\} + 1 - W_T) + 1$, when W_T is odd, where $\bar{\mu}_T^{(\hat{g}^{(T)})} = \sum_{\bar{j}} \mu_{T,\bar{j}} / \text{Cardi}\{\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}\}$. Afterwards, we set $\mathbb{T} = \{\mathbb{T} \setminus \mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)})}\}$ and solve problems (41) and (43) to find the optimal solutions $k_{\text{opt}}^{(\hat{g}^{(T)+1)}$ and $\mathbb{T}_{\text{sel}}^{(\hat{g}^{(T)+1)}$ for the $(\hat{g}^{(T)} + 1)$ -th SoI. This procedure iterates until the set \mathbb{T} is empty. Finally, only unique indices in $\{\mathbb{T}_{\text{sel}}^{(1)}, \dots, \mathbb{T}_{\text{sel}}^{(\hat{G}^{(T)})}\}$ are selected and $G_T \leq \hat{G}^{(T)}$ independent SoIs are obtained, where the $g^{(T)}$ -th sector ($g^{(T)} \in \{1, \dots, G_T\}$) is characterized by the first index $k_F^{(g^{(T)})}$ and the last index $k_L^{(g^{(T)})}$. The coarse estimates in the receive dimension are classified in the same way to obtain G_R independent SoIs, where $k_F^{(g^{(R)})}$ and $k_L^{(g^{(R)})}$ denote the first index and the last index of the $g^{(R)}$ -th sector ($g^{(R)} \in \{1, \dots, G_R\}$), respectively. The corresponding window length and the maximum distance to the window center are denoted by $W_R \frac{2\pi}{M_R}$ ($W_R \in \mathbb{Z}_+$) and $\zeta_R \frac{\pi}{M_R}$ ($\zeta_R \in \mathbb{Z}_+, \zeta_R \leq W_R$), respectively. Note that the heuristic approach in [21], which selects the D left and right nearest beams of the coarse estimates, is a special case of the proposed sectorization approach, i.e., when $\zeta_T = 1$ and $W_T = 2 \cdot D - 1$. An example of the transmit side sectorization result with $L = 3$ paths is given in Fig. 4, where $\hat{\mu}_1, \hat{\mu}_2$, and $\hat{\mu}_3$ are coarse estimates at the transmitter. After sectorization, the windowing approach has identified $\hat{G}_T = 3$ SoI candidates but $G_T = 2$ SoIs have been formed.

C. Fine Estimation

After the sectorization is completed, we apply the frame-wise training procedure in Section III-A to sweep through the determined SoIs. Let us define $\bar{N}^{(R)} = \sum_{g^{(T)}} (k_L^{(g^{(R)})} - k_F^{(g^{(R)})} + 1)$ and $\bar{N}^{(T)} = \sum_{g^{(T)}} (k_L^{(g^{(T)})} - k_F^{(g^{(T)})} + 1)$ as the total number of required DFT beams in the receive dimension and the transmit dimension, respectively. When UTE is used, we have $\bar{\mathbf{W}} = \bar{\mathbf{D}}_{M_R}(:, k_F^{(1)} \dots k_L^{(1)} \dots k_F^{(G_R)} \dots k_L^{(G_R)}) \in \mathbb{C}^{M_R \times \bar{N}^{(R)}}$ and $\bar{\mathbf{F}} = \bar{\mathbf{D}}_{M_T}(:, k_F^{(1)} \dots k_L^{(1)} \dots k_F^{(G_T)} \dots k_L^{(G_T)}) \in \mathbb{C}^{M_T \times \bar{N}^{(T)}}$, where $\bar{\mathbf{D}}_{M_R}$ and $\bar{\mathbf{D}}_{M_T}$ are defined in Section II-B. When STE is used, we replace $\bar{\mathbf{D}}_{M_R}$ and $\bar{\mathbf{D}}_{M_T}$ by \mathbf{D}_{M_R} and \mathbf{D}_{M_T} defined in Section II-A, respectively.

For the fine estimation we combine the demodulated pilot signal in κ_t training frames to obtain an effective received 4-D

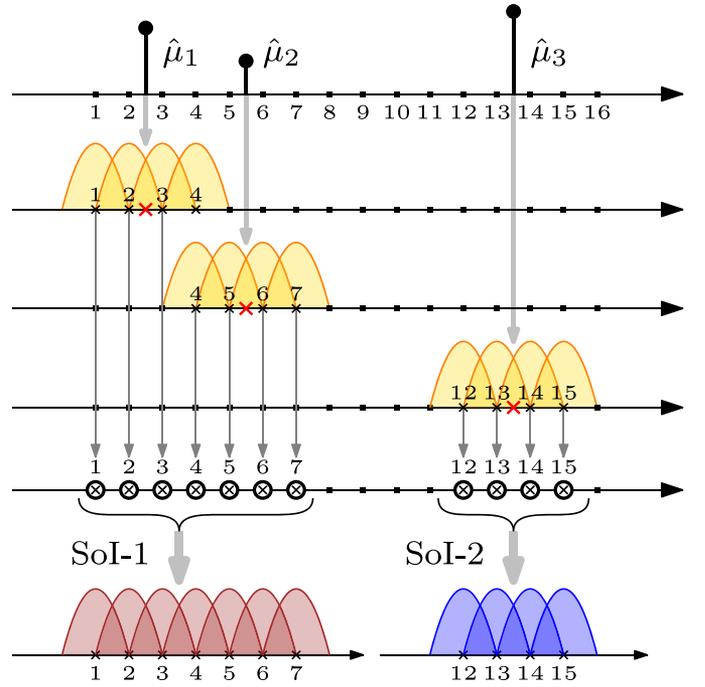


Fig. 4. Sliding window based sectorization for the transmit side. The three red crosses are positions of coarse estimates of spatial frequencies and the number stands for the beam index. We have $W_T = 3$, $\zeta_T = 1$, $L = 3$, and $M_T = 16$.

pilot signal tensor $\bar{\mathbf{Y}} \in \mathbb{C}^{\bar{N}^{(R)} \times \bar{N}^{(T)} \times N_f \times 1}$, i.e.,

$$\bar{\mathbf{Y}} = \mathcal{I}_{4,L} \times_1 (\bar{\mathbf{W}}^H \mathbf{A}_R) \times_2 (\bar{\mathbf{F}}^T \mathbf{A}_T) \times_3 \mathbf{A}_f \times_4 \mathbf{d}^T + \bar{\mathbf{Z}}. \quad (45)$$

where $\bar{\mathbf{Z}} \in \mathbb{C}^{\bar{N}^{(R)} \times \bar{N}^{(T)} \times N_f \times 1}$ is the effective noise tensor.

To obtain (45), we need to sweep through both the transmit dimension and the receive dimension in a sequential order. Let us define $N_t^{(T)}$ and $N_t^{(R)}$ as the number of training frames in the transmit direction and the receive direction, respectively. In general we use all the N_T transmit RF chains and N_R receive RF chains in every training frame to minimize the required training OFDM symbols. Hence, $N_t^{(T)} = \lceil \frac{\bar{N}^{(T)}}{N_T} \rceil$ and $N_t^{(R)} = \lceil \frac{\bar{N}^{(R)}}{N_R} \rceil$. Let us define the $N_t^{(T)}$ unique hybrid precoding matrices as $\mathbf{F}_{\kappa_1} = \bar{\mathbf{F}}(:, (\kappa_1 - 1)N_T + 1 : \kappa_1 N_T)$ for $1 \leq \kappa_1 < N_t^{(T)}$ and $\mathbf{F}_{\kappa_1} = \bar{\mathbf{F}}(:, \bar{N}^{(T)} - N_T + 1 : \bar{N}^{(T)})$ for $\kappa_1 = N_t^{(T)}$. Note that $\mathbf{F}_{N_t^{(T)}-1}$ and $\mathbf{F}_{N_t^{(T)}}$ will have overlapping columns if $\bar{N}^{(T)}$ is not a multiple of N_T . Similarly, we define the $N_t^{(R)}$ unique hybrid decoding matrices as $\mathbf{W}_{\kappa_2} = \bar{\mathbf{W}}(:, (\kappa_2 - 1)N_R + 1 : \kappa_2 N_R)$ for $1 \leq \kappa_2 < N_t^{(R)}$ and $\mathbf{W}_{\kappa_2} = \bar{\mathbf{W}}(:, \bar{N}^{(R)} - N_R + 1 : \bar{N}^{(R)})$ for $\kappa_2 = N_t^{(R)}$. Finally, we define $\mathbf{Y}_\kappa = \mathbf{Y}_{\kappa_1, \kappa_2}$ such that $\kappa = (\kappa_1 - 1)N_t^{(R)} + \kappa_2$. Therefore, starting from the first training frame, we fix the hybrid precoding matrix and then sweep through the $N_t^{(R)}$ hybrid decoding matrices. Then we change the hybrid precoding matrix and sweep through all the hybrid decoding matrices again. The procedure continues until all the hybrid precoding matrices are used. Therefore, we need

$\kappa_t = N_t^{(T)} \cdot N_t^{(R)}$ training frames. We have

$$\bar{\mathcal{Y}} = \left[\mathcal{Y}_1^{(T)} \sqcup_2 \cdots \sqcup_2 \mathcal{Y}_{N_t^{(T)}-2}^{(T)} \sqcup_2 \mathcal{Y}_{\kappa_2, L} \right], \quad (46)$$

where $\mathcal{Y}_{\kappa_1}^{(T)} = [\mathcal{Y}_{\kappa_1, 1} \sqcup_1 \cdots \sqcup_1 \mathcal{Y}_{\kappa_1, N_t^{(R)}-2} \sqcup_1 \mathcal{Y}_{\kappa_1, L}]$ for $\kappa_1 \in \{1, \dots, N_t^{(T)}\}$. We define $\mathcal{Y}_{\kappa_1, L} = [\mathcal{Y}_{\kappa_1, 1}^{(U)} \sqcup_1 (\mathcal{Y}_{\kappa_1, 1}^{(O)} + \mathcal{Y}_{\kappa_1, 2}^{(O)})/2 \sqcup_1 \mathcal{Y}_{\kappa_1, 2}^{(U)}]$, where $\mathcal{Y}_{\kappa_1, 1}^{(U)} = \mathcal{Y}_{\kappa_1, N_t^{(R)}-1}(1 : N_R - o_2, :, :)$, $\mathcal{Y}_{\kappa_1, 1}^{(O)} = \mathcal{Y}_{\kappa_1, N_t^{(R)}-1}(N_R - o_2 + 1 : N_R, :, :)$, $\mathcal{Y}_{\kappa_1, 2}^{(U)} = \mathcal{Y}_{\kappa_1, N_t^{(R)}}(o_2 + 1 : N_R, :, :)$, and $\mathcal{Y}_{\kappa_1, 2}^{(O)} = \mathcal{Y}_{\kappa_1, N_t^{(R)}}(1 : o_2, :, :)$. We have $o_2 = \max(0, N_t^{(R)} N_R - \bar{N}^{(R)})$. Moreover, $\mathcal{Y}_{\kappa_2, L} = [\mathcal{Y}_{\kappa_2, 1}^{(U)} \sqcup_1 (\mathcal{Y}_{\kappa_2, 1}^{(O)} + \mathcal{Y}_{\kappa_2, 2}^{(O)})/2 \sqcup_1 \mathcal{Y}_{\kappa_2, 2}^{(U)}]$, where $\mathcal{Y}_{\kappa_2, 1}^{(U)} = \mathcal{Y}_{N_t^{(T)}-1}^{(T)}(:, 1 : N_T - o_1, :)$, $\mathcal{Y}_{\kappa_2, 1}^{(O)} = \mathcal{Y}_{N_t^{(T)}-1}^{(T)}(:, N_T - o_1 + 1 : N_T, :)$, $\mathcal{Y}_{\kappa_2, 2}^{(U)} = \mathcal{Y}_{N_t^{(T)}}^{(T)}(:, o_1 + 1 : N_T, :)$, and $\mathcal{Y}_{\kappa_2, 2}^{(O)} = \mathcal{Y}_{N_t^{(T)}}^{(T)}(:, 1 : o_1, :)$. We have $o_1 = \max(0, N_t^{(T)} N_T - \bar{N}^{(T)})$.

Equation (45) corresponds to a single snapshot. To increase the number of snapshots, we smooth $\bar{\mathcal{Y}}$ on its 3-mode by using the same smoothing procedure as in Section IV-A. Let $\bar{\mathcal{Y}}_{(ss)} \in \mathbb{C}^{\bar{N}_R \times \bar{N}_T \times M_{\text{sub}, f} \times L_f}$ denote the tensor after smoothing. To compute the estimated spatial frequencies, $\hat{\mu}_{f, \ell}$, $\hat{\mu}_{T, \ell}$, and $\hat{\mu}_{R, \ell}$, $\forall \ell$, we apply hybrid 3-D UTE algorithm in Section II-C onto $\bar{\mathcal{Y}}_{(ss)}$, which is Algorithm 1 with some modifications.

The estimated channel gain vector $\hat{\mathbf{d}}$ is calculated by using LS, i.e.,

$$\hat{\mathbf{d}} = (\hat{\mathbf{A}}_f \diamond (\bar{\mathbf{F}}^T \hat{\mathbf{A}}_T) \diamond (\bar{\mathbf{W}}^H \hat{\mathbf{A}}_R))^+ \cdot \text{vec}(\bar{\mathcal{Y}}), \quad (47)$$

where the estimated array steering matrices are constructed by using the estimates $\hat{\mu}_{f, \ell}$, $\hat{\mu}_{T, \ell}$, and $\hat{\mu}_{R, \ell}$. Finally, the estimated vectorized channel is given by

$$\hat{\mathbf{h}}_{\text{vec}} = (\hat{\mathbf{A}}_f \diamond \hat{\mathbf{A}}_T \diamond \hat{\mathbf{A}}_R) \cdot \hat{\mathbf{d}}. \quad (48)$$

The proposed three-step channel estimation algorithm is summarized in Algorithm 2.

V. SIMULATION RESULTS

The proposed R -D Tensor-ESPRIT in DFT beamspace framework and its application in gridless mmWave channel estimation are evaluated by using Monte-Carlo simulations. We set $N_{\text{fft}} = 512$. The pilot spacing is set to $\Delta_{\text{pilot}} = 1$. In each OFDM symbol $N_f = 36$ consecutive subcarriers, starting from the first subcarrier, are used as pilot subcarriers. We consider $L_f = 24$ virtual subarrays in the frequency dimension. The inter-element spacing of the ULA is equal to half of the wavelength and isotropic antenna elements are used. To achieve a fair comparison especially when compared to other algorithms in the literature, the transmit power P_{pilot} in each training frame is normalized such that $P_{\text{pilot}} = (P_0 - P_{\text{ini}})/\kappa_t$, where P_0 denotes a total transmit power and P_{ini} denotes the transmit power in the coarse estimation step, and κ_t denotes the total number of required training frames when algorithms in the literature are used or the number of training frames in the fine estimation step when our proposed algorithms are used. There is no coarse estimation for algorithms in the literature. Therefore, $P_{\text{ini}} = 0$. The pilot transmit power is equally allocated

Algorithm 2: The Three-Step Channel Estimation Algorithm.

- 1: Inputs: number of antennas M_T and M_R , and number of RF chains N_T , and N_R , number of frequency subblocks L_f , number of paths L , window lengths W_T and W_R , maximum distances to center of SoI ζ_T and ζ_R .
 - 2: **Coarse estimation:**
 - Use N_T transmit antennas and N_R receive antennas that correspond to the number of RF chains at the transmitter and at the receiver to perform frame-wise training by using all the beams.
 - After the demodulation of pilot symbols, we re-arrange the received signal to obtain \mathcal{Y}_{ss} in (39). Then the coarse estimates of $(\mu_{T, \ell}, \mu_{R, \ell})$, $\forall \ell$, are computed by using 3-D UTE as in [26].
 - 3: **Sectorization:**
 - Find optimal transmit DFT beam indices $k_F^{(1)} \cdots k_L^{(1)} \cdots k_F^{(G_T)} \cdots k_L^{(G_T)}$ and receive beam indices $k_F^{(1)} \cdots k_L^{(1)} \cdots k_F^{(G_R)} \cdots k_L^{(G_R)}$ by using the sliding window based approach of Section IV-B.
 - 4: **Fine estimation:**
 - Perform frame-wise training by using the selected transmit and receive beams.
 - After demodulating the pilot signals, we re-arrange the received signal to obtain the smoothed tensor \mathcal{Y} in (46). Then we apply the hybrid 3-D UTE in Algorithm 1 to compute the fine estimates of the (spatial) frequencies $(\hat{\mu}_{f, \ell}, \hat{\mu}_{T, \ell}, \hat{\mu}_{R, \ell})$, $\forall \ell$.
 - Compute the LS estimate of the channel gain $\hat{\mathbf{d}}$ as in (47).
 - Construct the channel estimate as in (48).
 - 5: Outputs: frequencies $(\hat{\mu}_{f, \ell}, \hat{\mu}_{T, \ell}, \hat{\mu}_{R, \ell})$, $\forall \ell$, complex gain $\hat{\mathbf{d}}$, and the channel $\hat{\mathbf{h}}_{\text{vec}}$.
-

to N_f pilot tones in one OFDM symbol. We define the SNR in the simulations as $\text{SNR} = P_0/\sigma_n^2$, where P_0 is normalized to unity. When the estimation performance of the spatial frequencies is evaluated, the RMSE is computed, i.e., $\text{RMSE} = \sqrt{\mathbb{E}\{\sum_{r=1}^3 \sum_{\ell=1}^L (\mu_{x(r), \ell} - \hat{\mu}_{x(r), \ell})^2\}}$. When the estimation performance of the channel is evaluated, the normalized mean squared error (NMSE) is calculated, i.e., $\text{NMSE} = \mathbb{E}\{\|\mathbf{h}_{\text{vec}} - \hat{\mathbf{h}}_{\text{vec}}\|^2 / \|\mathbf{h}_{\text{vec}}\|^2\}$, where \mathbf{h}_{vec} is the effective channel defined in (36) and $\hat{\mathbf{h}}_{\text{vec}}$ is defined in (48). We use ‘‘SE-DFT,’’ ‘‘STE-DFT,’’ ‘‘UE-DFT,’’ and ‘‘UTE-DFT’’ to denote matrix based Standard ESPRIT in DFT beamspace, Standard Tensor-ESPRIT in DFT beamspace, matrix based Unitary ESPRIT in DFT beamspace, and Unitary Tensor-ESPRIT in DFT beamspace, respectively. Our derived Cramer-Rao bound inspired by [16] is used as the lower bound for different estimation algorithms and is denoted as ‘‘CRB’’. All the simulation results are obtained by averaging over 10 000 noise realizations.

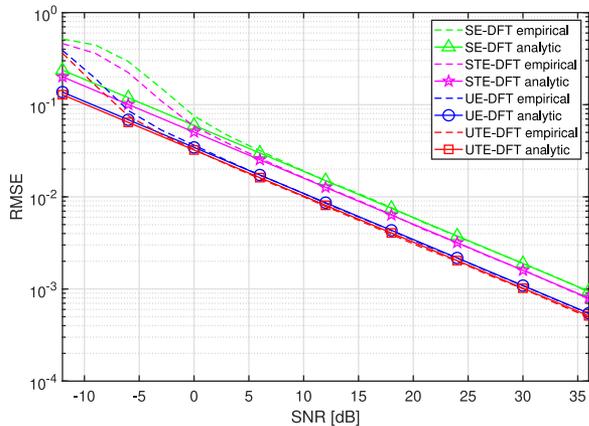


Fig. 5. RMSE vs. SNR for a 2-D array with $M_1 = 5$, $M_2 = 6$, $B_1 = B_2 = 3$, $N = 20$, $d = 2$.

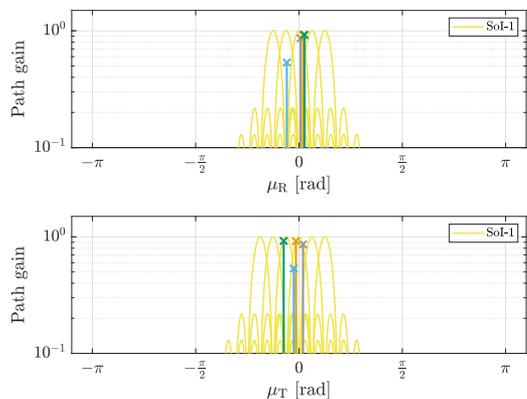


Fig. 6. The illustration of the configured (spatial) frequencies and the ideal sectorization result for $M_T = M_R = 32$, $N_T = N_R = 6$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$. We have $L = 4$, $\boldsymbol{\mu}_f = [-0.3759, -0.1906, -0.3143, -0.0557]$, $\boldsymbol{\mu}_R = [0.0254, 0.0642, -0.1852, 0.0844]$, $\boldsymbol{\mu}_T = [0.0612, -0.0462, -0.0798, -0.2326]$, and $\mathbf{d} = [-0.5808 - 0.6371j, 0.7512 + 0.5282j, -0.0953 + 0.5267j, -0.8834 - 0.2680j]$. The same path at the transmitter and the receiver is denoted by the same color.

A. Verification of the Analytic Performance

The derived analytical performance in Section II-D is verified in Fig. 5. We consider two correlated sources such that $\mathbb{E}\{s_{1,t}s_{2,t'}^*\} = \rho$, $\forall t \neq t' \in \{1, \dots, N\}$, where ρ is the correlation coefficient between two sources. In particular, we specify $\mu_1^{(1)} = \mu_2^{(2)} = 1$, $\mu_1^{(2)} = \mu_2^{(1)} = -0.5$, and $\rho = 0.9$. For both modes, i.e., $r = 1, 2$, the compression matrix $\mathbf{F}^{(r)}$ comprises the first two rows and the last row of the (phase-shifted) DFT matrix. Numerical results show that the empirical estimation errors agree with the analytical ones especially in the high SNR regime.

B. Performance of Proposed Channel Estimation Algorithms

In Fig. 6 the configured spatial frequencies and the corresponding ideal sectorization result are illustrated, i.e., the ideal spatial frequencies are used to perform the sectorization. In our performance simulations, coarse estimates are used to perform

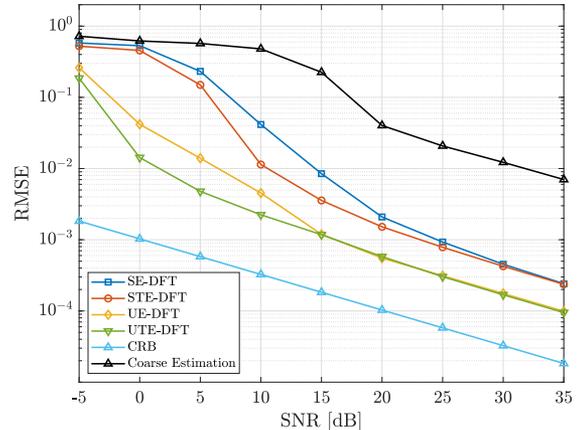


Fig. 7. RMSE vs. SNR for $M_T = M_R = 32$, $N_T = N_R = 6$, $L = 4$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$.

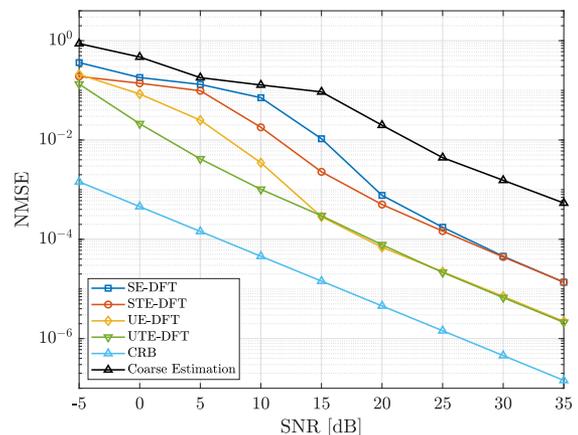


Fig. 8. NMSE vs. SNR for $M_T = M_R = 32$, $N_T = N_R = 6$, $L = 4$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$.

the sectorization. Fig. 7 and Fig. 8 show the RMSE of spatial frequencies and the NMSE of the channel in (36) for a scenario with $L = 4$ paths and closely spaced DoAs at the receiver side. The channel parameters are in the caption part of Fig. 6. We set $P_{\text{ini}} = 0.15P_0$. The proposed 3-D hybrid ESPRIT based channel estimation methods provide much better refined spatial estimates compared to the initial estimates. The most significant performance gain is obtained in both spatial frequency estimation and the subsequent channel estimation when the hybrid 3-D UTE algorithm is used. In the high SNR regime, the matrix based algorithm SE-DFT and UE-DFT converge to the corresponding tensor based algorithm STE-DFT and UTE-DFT.

In Figs. 9, 10, 11, 12, and 13 we compare the proposed hybrid 3-D UTE based channel estimation algorithm with the state of the art tensor based channel estimation algorithms in [16], [18], and [34]. We use ‘‘LRTD,’’ ‘‘TDBE-HOOI,’’ and ‘‘TDBE-CPD’’ to denote the algorithms in [16], [18], and [34], respectively. LRTD uses the CP decomposition to directly estimate the steering vectors $\mathbf{a}_T(\mu_{T,\ell})$, $\mathbf{a}_R(\mu_{R,\ell})$, and $\mathbf{a}_f(\mu_{f,\ell})$. Afterwards, the spatial frequencies are estimated by using a correlation based search algorithm. Therefore, the resolution will be limited by

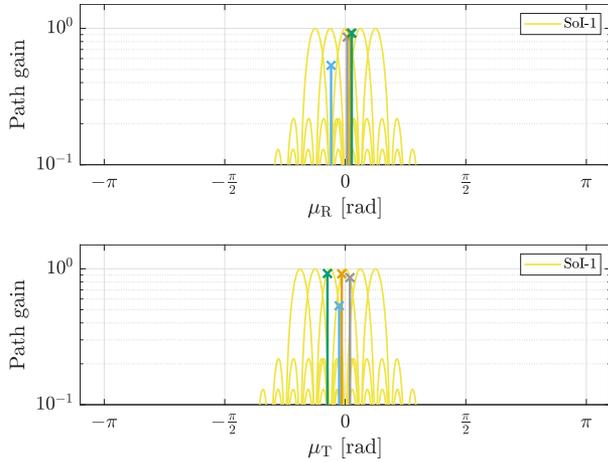


Fig. 9. The illustration of the configured (spatial) frequencies and the ideal sectorization result for $M_T = M_R = 32$, $N_T = N_R = 6$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$. We have $L = 4$, $\boldsymbol{\mu}_R = [0.0254, 0.0642, -0.1852, 0.0844]$, $\boldsymbol{\mu}_T = [0.0612, -0.0462, -0.0798, -0.2326]$, $\boldsymbol{\mu}_f = [-0.3759, -0.1906, -0.3143, -0.0557]$, and $\boldsymbol{d} = [-0.5808 - 0.6371j, 0.7512 + 0.5282j, -0.0953 + 0.5267j, -0.8834 - 0.2680j]$. The same path at the transmitter and the receiver is denoted by the same color.

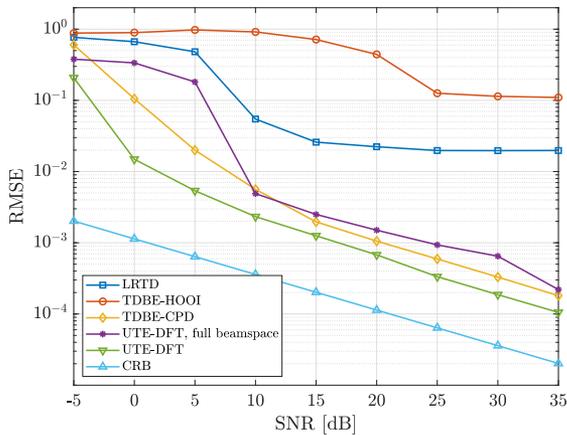


Fig. 10. RMSE vs. SNR for different algorithms for $M_T = M_R = 32$, $N_T = N_R = 6$, $L = 4$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$.

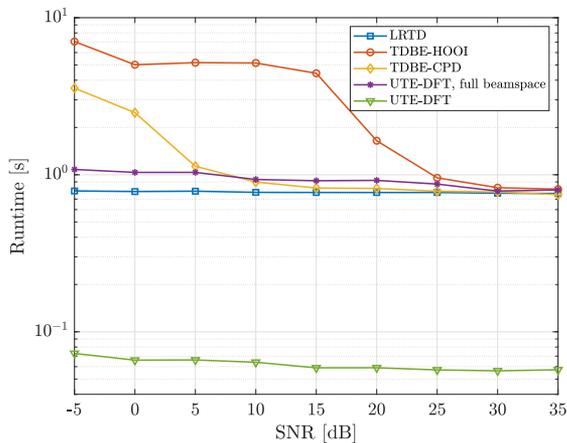


Fig. 11. Run time comparison for different algorithms for $M_T = M_R = 32$, $N_T = N_R = 6$, $L = 4$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$.

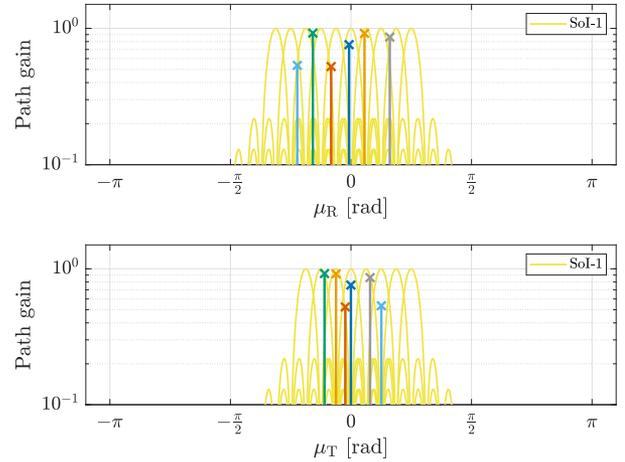


Fig. 12. The illustration of the configured (spatial) frequencies and the ideal sectorization result for $M_T = M_R = 32$, $N_T = N_R = 6$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$. We have $L = 4$, $\boldsymbol{\mu}_R = [0.5056, 0.1764, -0.6997, -0.4968, -0.025, -0.2582]$, $\boldsymbol{\mu}_T = [0.2487, -0.1948, 0.3961, -0.3452, -0.0009, -0.074]$, $\boldsymbol{\mu}_f = [-0.3734, -0.2837, -0.1571, -0.3267, -0.0528, -0.0237]$, and $\boldsymbol{d} = [-0.1535 + 0.8483j, -0.8787 - 0.2666j, 0.5163 - 0.1412j, 0.9008 - 0.2020j, 0.4167 + 0.6353j, 0.5155 - 0.0963j]$. The same path at the transmitter and the receiver is denoted by the same color.

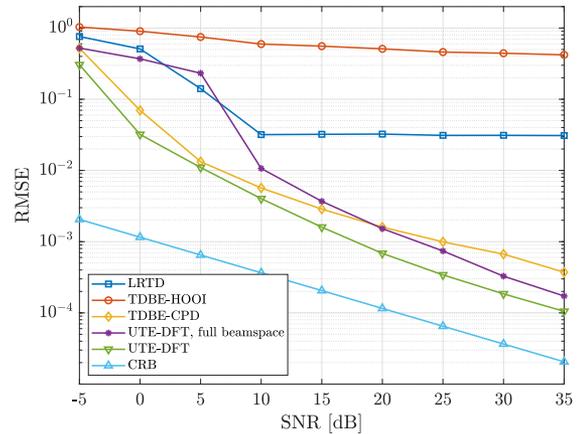


Fig. 13. RMSE vs. SNR for different algorithms for $M_T = M_R = 32$, $N_T = N_R = 6$, $L = 6$, $W_T = W_R = 3$, and $\zeta_T = \zeta_R = 1$.

the granularity of the search grid. In our simulations a search grid of size 3200 is used in all dimensions. Moreover, the number of random beams is all set to 6, i.e., $M = T = K = 6$ as defined in [16]. For TDBE-HOOI and TDBE-CPD we use the full beamspace as in the references [18] and [34]. This setting is also used in the original paper. The difference between TDBE-HOOI and TDBE-CPD lies in the signal subspace computation. The former one uses the higher-order orthogonal iteration (HOOI) algorithm while the latter one uses the an approximate CP decomposition based on “Nonlinear least squares by Gauss-Newton with dogleg trust region” as described in [35]. In the simulations we set $P_{\text{ini}} = 0.3P_0$. The run time of different algorithms is obtained by using a computer cluster, where each node of the cluster has 180 GB memory and a CPU of Intel Xeon E5-2650V4. We were using one physical core for each runtime of an algorithm.

Fig. 9 shows the path locations and the corresponding SoI for our hybrid 3-D UTE-DFT algorithm. In this scenario we have closely spaced paths at both the transmitter and the receiver. If the ideal sectorization is used, only one SoI is formed at both the transmitter and the receiver. We observe that the average required number of training frames for our proposed hybrid 3-D UTE-DFT algorithm is 2 when $\text{SNR} \leq 10$ dB and is 2-3 when $\text{SNR} < 10$ dB. For LRTD, TDBE-HOOI, and TDBE-CPD more training frames are needed especially when the full DFT beamspace is used. In Fig. 10 we compare the achieved RMSE of the (spatial) frequencies using different algorithms. The proposed hybrid 3-D UTE-DFT algorithm with a reduced-dimensional DFT beamspace provides the best estimation performance. The gains come from the focused power and the fact that UTE can better handle closely spaced paths. When the full DFT beamspace is used, the performance degrades because the transmit power is distributed equally to all beams. Due to the use of the CP decomposition, TDBE-CPD provides a better performance than TDBE-HOOI. LRTD suffers from the grid based search. Therefore, it yields an error floor in the high SNR regime. In Fig. 11 the comparison of the computational complexity is performed by comparing the run time of different algorithms. The run time is chosen because iterative algorithms are used to achieve a tensor decomposition in LRTD, TDBE-HOOI, and TDBE-CPD. As seen from the figure, the proposed UTE-DFT based channel estimation algorithm has the lowest complexity. TDBE-HOOI and TDBE-CPD yield a much higher complexity especially in the low to medium SNR regime because the tensor decomposition converges slowly. Fig. 12 and Fig. 13 provide the illustration of the configured (spatial) frequencies and the RMSE comparison for a scenario with $L = 6$ paths, respectively. In this simulation a perfect sectorization is used. It can be seen from Fig. 13 that even the hybrid 3-D UTE-DFT algorithm in the full DFT beamspace has achieved a better RMSE performance than the algorithms from the literature in the high SNR regime. Lastly, for our proposed algorithm we could also use the CP decomposition to obtain the initial estimates of the factor matrices instead of the HOSVD that provides a tensor based signal subspace estimate. Notice, however, that the computational complexity will be significantly increased.

VI. CONCLUSION

In this paper we have developed a gridless channel estimation algorithm for a hybrid mmWave MIMO-OFDM system with a frequency selective channel. The proposed algorithm is based on tensor-based ESPRIT-type algorithms in DFT beamspace. To this end, an R -D Standard/Unitary Tensor-ESPRIT in DFT beamspace framework has been developed. The corresponding analytic performance has also been derived. We have proven that for a single source the R -D ESPRIT in DFT beamspace provides a better performance than R -D ESPRIT in element space. Based on the derived hybrid 3-D Tensor-ESPRIT algorithm, a gridless channel estimation has been achieved in two stages. Numerical results show that the proposed algorithm can provide accurate channel estimates with only a few training resources.

APPENDIX A PROOF OF LEMMA 3

The proof is achieved by extending the concept in [26] to the R -D DFT beamspace case. In order to find the real-valued transformation, we first transform \mathcal{X} into a centro-Hermitian tensor. A centro-Hermitian tensor is obtained by using the forward-backward averaging (FBA) pre-processing. More precisely, we define the FBA of the tensor \mathcal{X} in the same way as in [26], i.e.,

$$\mathcal{X}^{(\text{fba})} = [\mathcal{X} \lrcorner_{R+1} \mathcal{X}^* \times_1 \mathbf{\Pi}_{M_1} \times_2 \mathbf{\Pi}_{M_2} \cdots \times_{R+1} \mathbf{\Pi}_N], \quad (49)$$

where $\mathbf{\Pi}_q$ is the $q \times q$ exchange matrix. According to the definition of centro-Hermitian tensor in [26], the tensor $\mathcal{X}^{(\text{fba})} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R \times 2N}$ is centro-Hermitian because

$$\mathcal{X}^{(\text{fba})} \equiv \mathcal{X}^{(\text{fba})*} \times_1 \mathbf{\Pi}_{M_1} \times_2 \mathbf{\Pi}_{M_2} \cdots \times_{R+1} \mathbf{\Pi}_{2N}. \quad (50)$$

According to Lemma 4 in [26], the centro-Hermitian tensor $\mathcal{X}^{(\text{fba})}$ can be mapped to a real-valued tensor by multiplying each of its modes with $\mathbf{Q}^{(r)\text{H}}$, where $\mathbf{Q}^{(r)}$ should be a unitary and left- $\mathbf{\Pi}$ -real matrix satisfying $\mathbf{\Pi}\mathbf{Q}^{(r)*} = \mathbf{Q}^{(r)}$. Since $\tilde{\mathbf{F}}^{(r)\text{H}}$ are left- $\mathbf{\Pi}$ -real matrices, we set $\mathbf{Q}^{(r)} = \tilde{\mathbf{F}}^{(r)\text{H}}$ for $r \in \{1, \dots, R\}$. Moreover, we define

$$\mathbf{Q}^{(R+1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_N & j\mathbf{I}_N \\ \mathbf{\Pi}_N & -j\mathbf{\Pi}_N \end{bmatrix}. \quad (51)$$

Then the real-valued transformation is calculated by

$$\begin{aligned} \mathcal{Z} &= \mathcal{X}^{(\text{fba})} \times_1 \mathbf{Q}^{(1)\text{H}} \times_2 \mathbf{Q}^{(2)\text{H}} \cdots \times_{R+1} \mathbf{Q}^{(R+1)\text{H}} \\ &= [\mathcal{X} \times_1 \tilde{\mathbf{F}}^{(1)} \times_2 \tilde{\mathbf{F}}^{(2)} \cdots \tilde{\mathbf{F}}^{(R)} \lrcorner_{R+1} \\ &\quad \mathcal{X}^* \times_1 \tilde{\mathbf{F}}^{(1)} \mathbf{\Pi}_{M_1} \times_2 \tilde{\mathbf{F}}^{(2)} \mathbf{\Pi}_{M_2} \cdots \times_R \tilde{\mathbf{F}}^{(R)} \mathbf{\Pi}_{M_R} \times_{R+1} \mathbf{\Pi}_N] \\ &\quad \times_{R+1} \mathbf{Q}^{(R+1)\text{H}} \\ &= \frac{1}{\sqrt{2}} [\tilde{\mathcal{X}} \lrcorner_{R+1} \tilde{\mathcal{X}}^* \times_{R+1} \mathbf{\Pi}_N] \times_{R+1} \begin{bmatrix} \mathbf{I}_N & \mathbf{\Pi}_N \\ -j\mathbf{I}_N & j\mathbf{\Pi}_N \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \left(\tilde{\mathcal{X}} \times_{R+1} \begin{bmatrix} \mathbf{I}_N \\ -j\mathbf{I}_N \end{bmatrix} + \tilde{\mathcal{X}}^* \times_{R+1} \begin{bmatrix} \mathbf{I}_N \\ j\mathbf{I}_N \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{2}} \left((\tilde{\mathcal{X}} \lrcorner_{R+1} (-j\tilde{\mathcal{X}})) + (\tilde{\mathcal{X}}^* \lrcorner_{R+1} j\tilde{\mathcal{X}}^*) \right) \\ &= \frac{1}{\sqrt{2}} \left((\tilde{\mathcal{X}} + \tilde{\mathcal{X}}^*) \lrcorner_{R+1} (-j(\tilde{\mathcal{X}} - \tilde{\mathcal{X}}^*)) \right) \\ &= \sqrt{2} \left[\text{Re}\{\tilde{\mathcal{X}}\} \lrcorner_{R+1} \text{Im}\{\tilde{\mathcal{X}}\} \right]. \quad (52) \end{aligned}$$

In the second step, the following property is used in the derivation.

$$[\mathcal{A} \lrcorner_p \mathcal{B}] \times_q \mathbf{U} = [\mathcal{A} \times_q \mathbf{U} \lrcorner_p \mathcal{B} \times_q \mathbf{U}], \quad p \neq q. \quad (53)$$

In the third step, the definition of $\tilde{\mathcal{X}}$ as in (14) and the property $\tilde{\mathbf{F}}^{(r)} \mathbf{\Pi}_{M_r} = \tilde{\mathbf{F}}^{(r)*}$ are used. In the fourth step, we use the property that

$$[\mathcal{A} \lrcorner_p \mathcal{B}] \times_p \begin{bmatrix} \mathbf{U} & \mathbf{W} \end{bmatrix} = \mathcal{A} \times_p \mathbf{U} + \mathcal{B} \times_p \mathbf{W}. \quad (54)$$

In the fifth step, we use the property that $\mathcal{A} \times_p [\mathbf{U} \mathbf{W}] = (\mathcal{A} \times_p \mathbf{U}) \sqcup_p (\mathcal{A} \times_p \mathbf{W})$. Note that the scalar $\sqrt{2}$ is not needed for the computation of subspace estimates.

APPENDIX B PROOF OF THEOREM 2

Let us define the SVD of the $(R+1)$ -th mode unfolding of \mathcal{Z}_0 as $[\mathcal{Z}_0]_{(R+1)}^T = \mathbf{E}_s \tilde{\Sigma}_s \tilde{\mathbf{V}}_s^H + \mathbf{E}_n \tilde{\Sigma}_n \tilde{\mathbf{V}}_n^H$, where \mathcal{Z}_0 is the noiseless version of \mathcal{Z} . The first-order perturbation analysis of the HOSVD from [28] can be extended such that the perturbed signal subspace estimate $\hat{\mathcal{E}}_s$ is given by $[\hat{\mathcal{E}}_s]_{(R+1)}^T = \mathbf{E}_s + [\Delta \mathcal{E}_s]_{(R+1)}^T$, where $\Delta \mathcal{E}_s$ is similar as $\Delta \mathbf{U}_s$ in [28] except that all the corresponding terms in $\Delta \mathbf{U}_s$ are replaced by the estimates from \mathcal{Z}_0 .

By following the derivation in [36], the tensor version of the first-order perturbation of the i -th eigenvalue in the r -th mode is given by

$$\Delta \bar{\lambda}_i^{(r)} \approx \tilde{\mathbf{p}}_i^T \cdot \left(\tilde{\mathbf{K}}_1^{(r)} \cdot \mathbf{E}_s \right)^+ \cdot \left[\tilde{\mathbf{K}}_2^{(r)} - \tilde{\lambda}_i^{(r)} \tilde{\mathbf{K}}_1^{(r)} \right] \cdot [\Delta \mathcal{E}_s]_{(R+1)}^T \cdot \tilde{\mathbf{q}}_i. \quad (55)$$

By using the first order Taylor expansion and the fact that $\frac{\partial \tan(x)}{\partial x} = 1 + \tan^2(x)$, we can conclude that

$$\Delta \bar{\lambda}_i^{(r)} \approx \Delta \mu_i^{(r)} (\bar{\lambda}_i^{(r)2} + 1)/2. \quad (56)$$

By inserting (56) into (55), we obtain (28). The computation of the corresponding MSE is straightforward.

APPENDIX C PROOF OF LEMMA 5

When $d = 1$, we drop the index i . The effective noiseless signal matrix $[\mathcal{X}_0]_{(R+1)}^T$ in (24) simplifies to $[\mathcal{X}_0]_{(R+1)}^T = \mathbf{b} \mathbf{s}^T$, where $\mathbf{b} = [\mathcal{B}]_{(R+1)}^T = \mathbf{b}^{(1)} \otimes \dots \otimes \mathbf{b}^{(R)}$, $\mathbf{b}^{(r)} = \mathbf{F}^{(r)} \mathbf{a}^{(r)}$, and $\mathbf{a}^{(r)} = \mathbf{a}(\mu^{(r)})$, $\forall r$. The rank-1 economy-size SVD of $[\mathcal{X}_0]_{(R+1)}^T$ is explicitly computed as

$$[\mathcal{X}_0]_{(R+1)}^T = \mathbf{u}_s \cdot \sigma_s \cdot \mathbf{v}_s^H, \quad (57)$$

where $\mathbf{u}_s = \mathbf{b}/\|\mathbf{b}\|$, $\sigma_s = \|\mathbf{b}\| \|\mathbf{s}\|$, and $\mathbf{v}_s = \mathbf{s}^*/\|\mathbf{s}\|$. Moreover, we have $\Psi^{(r)} = \lambda^{(r)}$ and therefore $\mathbf{p} = \mathbf{q} = \mathbf{1}$. Then the MSE of $\mu^{(r)}$ simplifies to

$$\mathbb{E}\{(\Delta \mu^{(r)})^2\} = \frac{1}{2} M \sigma_n^2 \cdot \|\beta\|^2 \cdot \|\alpha^{(r)}\|^2, \quad (58)$$

where $\beta = \mathbf{v}_s \sigma_s^{-1}$ and $\alpha^{(r)} = (\mathbf{U}_n^* \cdot \mathbf{U}_n^T) \cdot ((\tilde{\mathbf{J}}_1^{(r)} \cdot \mathbf{u}_s)^+ \cdot [\tilde{\mathbf{J}}_2^{(r)}/\lambda_i^{(r)} - \tilde{\mathbf{J}}_1^{(r)}])^T$. To obtain equation (58), we replace \mathbf{W}_{ten} in (27) by its matrix counterpart $\mathbf{W}_{\text{mat}} = (\sigma_s^{-1} \mathbf{v}_s^T) \otimes (\mathbf{U}_n \cdot \mathbf{U}_n^H)$.

We can further simplify $\alpha^{(r)T}$ by using (57), i.e.,

$$\alpha^{(r)T} = \|\mathbf{b}\| \cdot \left(\tilde{\mathbf{J}}_1^{(r)} \cdot \mathbf{b} \right)^+ \cdot \left[\tilde{\mathbf{J}}_2^{(r)}/\lambda^{(r)} - \tilde{\mathbf{J}}_1^{(r)} \right], \quad (59)$$

where the fact that $\mathbf{U}_n \cdot \mathbf{U}_n^H = \mathbf{I}_M - \frac{\mathbf{b} \mathbf{b}^H}{\|\mathbf{b}\|^2}$ and $\tilde{\mathbf{J}}_2^{(r)}/\lambda^{(r)} \mathbf{b} - \tilde{\mathbf{J}}_1^{(r)} \mathbf{b} = \mathbf{0}$ are used. Given that $\mathbb{E}\{\|\mathbf{s}\|^2\} = P_T N$, we can calculate $\|\beta\|^2$ as

$$\|\beta\|^2 = \|\mathbf{v}_s \sigma_s^{-1}\|^2 = \sigma_s^{-2} = \frac{1}{P_T N \|\mathbf{b}\|^2}. \quad (60)$$

By using the facts that $\mathbf{c}^+ = \frac{\mathbf{c}^H}{\|\mathbf{c}\|^2}$ and $\tilde{\mathbf{J}}_1^{(r)} \mathbf{b} = \tilde{\mathbf{J}}_2^{(r)} \mathbf{b}/\lambda^{(r)}$ we can express $\|\alpha^{(r)}\|^2$ as

$$\|\alpha^{(r)}\|^2 = \|\mathbf{b}\|^2 \cdot \frac{1}{\|\tilde{\mathbf{J}}_2^{(r)} \mathbf{b}\|^4} \cdot \left(\left\| \tilde{\mathbf{J}}_2^{(r)H} \tilde{\mathbf{J}}_2^{(r)} \mathbf{b} \right\|^2 + \left\| \tilde{\mathbf{J}}_1^{(r)H} \tilde{\mathbf{J}}_1^{(r)} \mathbf{b} \right\|^2 - 2 \cdot \text{Re} \left\{ \mathbf{b}^H \tilde{\mathbf{J}}_2^{(r)H} \tilde{\mathbf{J}}_2^{(r)} \tilde{\mathbf{J}}_1^{(r)H} \tilde{\mathbf{J}}_1^{(r)} \mathbf{b} \right\} \right). \quad (61)$$

To get more insights into (61), we study two special cases, i.e., $B_r = 2$ and $B_r = M_r$, $\forall r$.

When $B_r = M_r$, $\forall r$, matrices $\mathbf{J}_2^{(r)H} \mathbf{J}_2^{(r)}$ and $\mathbf{J}_1^{(r)H} \mathbf{J}_1^{(r)}$ are circulant matrices. The product of two circulant matrices is also a circulant matrix [37]. It is known that a circulant matrix can be diagonalized by using a DFT matrix. Let \mathbf{C} denote a circulant matrix. The following property holds, i.e.,

$$\mathbf{b}^{(r)H} \mathbf{C} \mathbf{b}^{(r)} = \mathbf{a}^{(r)H} \mathbf{D}^H \mathbf{C} \mathbf{D} \mathbf{a}^{(r)} = M_r^2 c_{1,1}, \quad (62)$$

where $c_{1,1}$ is the (1,1)-th element of \mathbf{C} and the fact that $\|\alpha^{(r)}\|^2 = M_r$ is used. Then it is straightforward to compute

$$\|\tilde{\mathbf{J}}_2^{(r)} \mathbf{b}\|^2 = \prod_{p=1, p \neq r}^R \|\mathbf{b}^{(p)}\|^2 \|\mathbf{J}_2^{(r)} \mathbf{b}^{(r)}\|^2 = 2M_r^2 \prod_{p=1, p \neq r}^R M_p^2.$$

Similarly, we have

$$\|\tilde{\mathbf{J}}_2^{(r)H} \tilde{\mathbf{J}}_2^{(r)} \mathbf{b}\|^2 = \|\tilde{\mathbf{J}}_1^{(r)H} \tilde{\mathbf{J}}_1^{(r)} \mathbf{b}\|^2 = 6 \prod_{p=1, p \neq r}^R M_p^2 M_r^2,$$

$$\mathbf{b}^H \tilde{\mathbf{J}}_2^{(r)H} \tilde{\mathbf{J}}_2^{(r)} \tilde{\mathbf{J}}_1^{(r)H} \tilde{\mathbf{J}}_1^{(r)} \mathbf{b}$$

$$= \prod_{p=1, p \neq r}^R M_p^2 M_r^2 \left(4 + 2 \cos \left(\frac{2\pi}{M_r} \right) \right).$$

Subsequently, $\|\alpha^{(r)}\|^2$ is simplified as

$$\|\alpha^{(r)}\|^2 = \left(1 - \cos \left(\frac{2\pi}{M_r} \right) \right) \prod_{\substack{p=1 \\ p \neq r}}^R \frac{1}{M_p^2}. \quad (63)$$

Finally, the MSE for the spatial frequency in the r -th mode is given by

$$\mathbb{E}\{(\Delta \mu^{(r)})^2\} = \frac{\sigma_n^2}{P_T N} \cdot \frac{\sin^2 \left(\frac{\pi}{M_r} \right)}{M} \quad (64)$$

When $B_r = 2$, $\forall r$, only two consecutive rows from the DFT matrix are selected. Then by using basic trigonometric identities and after some algebraic manipulation we obtain (31).

REFERENCES

- [1] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [2] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.
- [3] J. Zhang, A. Wiesel, and M. Haardt, "Low rank approximation based hybrid precoding schemes for multi-carrier single-user massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2016, pp. 3281–3285.
- [4] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

- [5] T. E. Bogale, L. B. Le, and X. Wang, "Hybrid analog-digital channel estimation and beamforming: Training-throughput tradeoff," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5235–5249, Dec. 2015.
- [6] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation and hybrid combining for mmwave: Phase shifters or switches," in *Proc. IEEE Inf. Theory Appl. Workshop*, 2015, pp. 90–97.
- [7] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Jun. 2016.
- [8] Z. Xiao, P. Xia, and X.-G. Xia, "Channel estimation and hybrid precoding for millimeter-wave mimo systems: A low-complexity overall solution," *IEEE Access*, vol. 5, pp. 16100–16110, Jul. 2017.
- [9] L. N. Ribeiro, A. L. F. de Almeida, N. J. Myers, and R. W. Heath, "Tensor-based estimation of mmWave MIMO channels with carrier frequency offset," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2019, pp. 4155–4159.
- [10] J. Zhang, I. Podkurkov, M. Haardt, and A. Nadeev, "Channel estimation and training design for hybrid analog-digital multi-carrier single-user massive MIMO systems," in *Proc. 20th Int. ITG Workshop Smart Antennas*, Munich, Germany, Mar. 2016, pp. 1–8.
- [11] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [12] J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.
- [13] J. P. Gonzalez-Coma, J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, L. Castedo, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for frequency selective multiuser mmWave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 353–367, May 2018.
- [14] A. Ali, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, Feb. 2018.
- [15] D. Zhu, J. Choi, and R. W. Heath, "Two-dimensional AoD and AoA acquisition for wideband mmWave systems with dual-polarized MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7890–7905, Dec. 2017.
- [16] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, 2017.
- [17] L. Cheng, G. Yue, X. Xiong, Y. Liang, and S. Li, "Tensor decomposition-aided time-varying channel estimation for millimeter wave MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1216–1219, Aug. 2019.
- [18] F. Wen, N. Garcia, J. Kulmer, K. Witrisal, and H. Wymeersch, "Tensor decomposition based beamspace ESPRIT for millimeter wave MIMO channel estimation," in *Proc. IEEE Glob. Commun. Conf.*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–7.
- [19] J. Zhang and M. Haardt, "Channel estimation and training design for hybrid multi-carrier mmWave massive MIMO systems: The beamspace ESPRIT approach," in *Proc. 25th Eur. Signal Process. Conf.*, Kos, Greece, Sep. 2017, pp. 385–389.
- [20] J. Zhang and M. Haardt, "Channel estimation for hybrid multi-carrier mmWave MIMO systems using three-dimensional unitary ESPRIT in DFT beamspace," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, Curacao, Dutch Antilles, Dec. 2017, pp. 1–5.
- [21] D. Rakhimov, J. Zhang, A. de Almeida, A. Nadeev, and M. Haardt, "Channel estimation for hybrid multi-carrier mmWave MIMO systems using 3-D unitary tensor-ESPRIT in DFT beamspace," in *Proc. 53th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2019, pp. 447–451.
- [22] G. Xu, S. D. Silverstein, R. H. Roy, and T. Kailath, "Beamspace ESPRIT," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 349–356, Feb. 1994.
- [23] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-form 2D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316–328, Feb. 1996.
- [24] C. P. Mathews, M. Haardt, and M. D. Zoltowski, "Performance analysis of closed-form, ESPRIT based 2-D angle estimator for rectangular arrays," *IEEE Signal Process. Lett.*, vol. 3, no. 4, pp. 124–126, Apr. 1996.
- [25] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [26] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, Jul. 2008.
- [27] M. Haardt and J. A. Nosssek, "Simultaneous schur decomposition of several non-symmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.
- [28] F. Roemer, M. Haardt, and G. Del Galdo, "Analytical performance assessment of multi-dimensional matrix- and tensor-based ESPRIT-type algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2611–2625, May 2014.
- [29] M. Haardt and J. A. Nosssek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, May 1995.
- [30] J. Steinwandt, F. Roemer, M. Haardt, and G. Del Galdo, "Performance analysis of multi-dimensional ESPRIT-type algorithms for arbitrary and strictly non-circular sources with spatial smoothing," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2262–2276, May 2017.
- [31] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [32] J. Zhang, I. Podkurkov, M. Haardt, and A. Nadeev, "Efficient multidimensional parameter estimation for joint wideband radar and communication systems based on OFDM," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, New Orleans, LA, Mar. 2017, pp. 3096–3100.
- [33] A. Ali, N. González-Prelcic, and A. Ghosh, "Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14903–14917, Dec. 2020.
- [34] F. Wen, H. C. So, and H. Wymeersch, "Tensor decomposition-based beamspace esprit algorithm for multidimensional harmonic retrieval," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Barcelona, Spain, 2020, pp. 4572–4576.
- [35] L. Sorber, M. Van Barel, and L. De Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank- $(L_r, L_r, 1)$ terms, and a new generalization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 695–720, 2013.
- [36] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.
- [37] H. Lütkepohl, *Handbook of Matrices*. Hoboken, NJ, USA: Wiley, 1996.