Channel Estimation for Hybrid Multi-Carrier mmWave MIMO Systems Using 3-D Unitary Tensor-ESPRIT in DFT beamspace

Damir Rakhimov [*], *Jianshu Zhang* [*], *André de Almeida* [†], *Adel Nadeev* [‡], *Martin Haardt* [*]
[*] Communications Research Laboratory, Ilmenau University of Technology, P. O. Box 100565, D-98684 Ilmenau, Germany
Email: {damir.rakhimov, jianshu.zhang, martin.haardt}@tu-ilmenau.de
[†] Department of Teleinformatics Engineering, Federal University of Ceara (UFC), Fortaleza, Brazil
[‡] Inst. of Radio Electronics and Telecommunications, Kazan National Research Technical University, Kazan, Russia

***Abstract*—In this work we propose a gridless channel estimation scheme for MIMO systems in the millimeter wave (mmWave) band which is based on R-D Unitary Tensor-ESPRIT in DFT beamspace. Compared to conventional ESPRIT based algorithms in element space, the beamspace approach can be applied to MIMO systems with hybrid architectures. Moreover, the proposed scheme can significantly reduce the training overhead for communication systems operating in the mmWave band. The proposed algorithm involves coarse and fine estimation steps. During the coarse estimation step, Unitary Tensor-ESPRIT in element-space is applied to the array with a reduced size aperture to obtain initial information about the angular profiles, such as directions of arrival (DoAs), directions of departure (DoDs), and propagation delays. Based on these estimates, in a second step the fine estimation of angular profiles, propagation delays and channel gains is performed applying 3-D Unitary Tensor-ESPRIT in DFT beamspace. We explain how to combine received signals from different sectors of interest (SoIs) and to perform joint processing. Simulation results show the tensor gain for the proposed algorithm in addition to the improved channel estimation.**

***Index Terms*—3-D Unitary Tensor-ESPRIT in DFT beamspace, MIMO-OFDM, semi-blind estimation.**

I. Introduction

Hybrid MIMO systems require additional analog precoders and decoders to overcome high pathloss penality during propagation. The task of designing optimal analog precoders and decoders is challenging [1]. In addition, for multicarrier systems, the analog precoding and decoding matrices will be identical for all subcarriers [2]. For this reason, the design of accurate and fast channel estimation algorithms for hybrid MIMO systems is still an open issue. Many authors use compressed sensing (CS) methods for channel estimation in millimeter wave (mmWave) frequencies. For example, in [3] a compressed sensing based channel estimation algorithm for hybrid MIMO in a frequency-flat channel is considered. The authors of [4] propose a channel estimation scheme based on CS for single carrier waveforms. In [5], an adaptive multi-grid sparse recovery approach for channel estimation in Hybrid MIMO communication systems is proposed. The authors in [6] develop a channel estimation algorithm for a frequency-flat channels by involving multiple measurement vectors (MMV) to improve the channel estimation accuracy. There also exist extensions for a multipath channel. For example, in [7] the authors propose a two stage CS based channel estimation algorithm for a frequency-selective channel that provides a reduction of the involved computational complexity. In general, a drawback of CS based channel estimation is its higher pilot overhead. Moreover, all the aforementioned algorithms depend on the on-grid assumption of the channel parameters, which will require an additional grid-offset estimation procedure for practical usage, which additionally complicates a solution.

To overcome the problem, the authors in [8] have proposed a channel estimation algorithm that is based on estimating the dominant directions of arrival (DoAs) and directions of departure (DoDs). In [9] authors have proposed a two-stage gridless channel estimation algorithm, which estimates angular profiles and channel gains. The core of the solution is represented by a three-dimensional (3-D) Standard ESPRIT in DFT beamspace algorithm. The algorithm provides high resolution estimates of DoAs and DoDs, when all pairs of DoAs and DoDs fall in a specific sector of interest (SoI). An improved algorithm was introduced in [10], where the dominant DoAs and DoDs are estimated in a first step and then based on these values more accurate parameter estimates are obtained in a second step. The disadvantage of the algorithm is the separate processing of signals from different SoIs and as a result it requires additional interference cancellers. The mentioned drawbacks of algorithms motivate us to develop novel algorithms that will have an improved performance and require a smaller training overhead.

In this paper, we propose a gridless channel estimation algorithm for a hybrid point-to-point mmWave multi-carrier massive MIMO system based on 3-D Unitary Tensor-ESPRIT in DFT beamspace. We use a similar approach as in [10] for the coarse estimation, where an array with a reduced size aperture and wider beams is used. Such an architecture can be obtained by switching off antennas or phase shifters. For the fine estimation step, we use all available antennas of the array and perform joint processing of information from different SoIs. The proposed algorithm is based on a tensor version [11] of the 3-D Unitary ESPRIT algorithm in DFT beamspace. The 2-D matrix version was already presented in [12], where phase shifted columns or rows of the DFT matrices are used for analog precoding and decoding. The advantages of the algorithm are that it does not require *a priori* knowledge of SoIs, pairs of DoAs and DoDs can belong to different SoIs, and the number of required training frames is relatively small.

NOTATION: We use lowercase (a), bold lowercase ($\boldsymbol{a}$) and bold capital ($\boldsymbol{A}$) letters for scalars, vectors, and matrices, respectively. Calligraphic letters ($\boldsymbol{\mathcal{T}}$) are used for tensors. Symbols $(.)^{\mathrm{T}}$, $(.)^{+}$ and $(.)^{\mathrm{H}}$ denote transpose, Moore-Penrose pseudo-inverse, and Hermitian transposition. The Kronecker product and Khatri-Rao product are denoted as $\otimes$ and $\diamond$, $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix, and $[\boldsymbol{\mathcal{Y}}]_{(k)}$ represents the k-mode unfolding of the tensor $\boldsymbol{\mathcal{Y}}$. The expression $\boldsymbol{\mathcal{A}} \times_n \boldsymbol{B}$ defines an n-mode product between a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{\boldsymbol{I}_1 \times \boldsymbol{I}_2 \times \ldots \times \boldsymbol{I}_N}$ and a matrix $\boldsymbol{B} \in \mathbb{C}^{\boldsymbol{I}_n \times \boldsymbol{J}}$, for $n \in \{1, ..., N\}$. We use $\mathrm{diag}\{\boldsymbol{x}\}$ to represent diagonal matrix with elements of the vector $\boldsymbol{x}$ on the main diagonal, $\mathbb{E}\{\}$ denotes expectation operator.

II. Problem formulation

A. System model

We consider a point-to-point massive MIMO-OFDM mmWave communication system. The transmitter has M_{T} antennas with N_{T} RF chains, while the receiver has M_{R} antennas with N_{R} RF chains. We assume $M_{\mathrm{T}} \gg N_{\mathrm{T}}$ and $M_{\mathrm{R}} \gg N_{\mathrm{R}}$. A CP-OFDM based modulation scheme is applied to combat the multipath effect [13]. The corresponding FFT size that equals to the number of subcarriers per OFDM symbol is N_{fft}, while the total number of pilots per OFDM symbol is N_{p}, and they are transmitted on subcarriers with the indices n_{p}, $p \in \{1, \cdots, N_{\mathrm{p}}\}$. Let N_t be the total number of the transmitted OFDM symbols. At the transmitter and the receiver side we assume a uniform linear array (ULA) with an antenna spacing of $\Delta = \lambda/2$. Perfect time and frequency synchronization are assumed. The maximum delay is less than the length of the cycle prefix (CP).

The received signal on the n-th subcarrier in the m-th OFDM symbol is given by [9]

$$\boldsymbol{y}_n[m] = \boldsymbol{W}_n^{\mathrm{H}}[m](\boldsymbol{H}_n\boldsymbol{F}_n[m]\boldsymbol{s}_n[m] + \boldsymbol{\eta}_n[m]) \in \mathbb{C}^{N_{\mathrm{R}}}, \quad (1)$$

where $\boldsymbol{H}_n \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}}$ is the channel matrix on the n-th subcarrier, $\boldsymbol{s}_n[m] \in \mathbb{C}^{N_{\mathrm{T}}}$ denotes the transmitted data vector, and $\boldsymbol{\eta}_n[m] \in \mathbb{C}^{M_{\mathrm{R}}}$ denotes the zero mean circularly symmetric complex Gaussian (ZMCSCG) noise with covariance matrix $\mathbb{E}\{\boldsymbol{\eta}_n[m]\boldsymbol{\eta}_n^{\mathrm{H}}[m]\} = \boldsymbol{I}_{M_{\mathrm{R}}}\sigma_n^2$ for all $n \in \{1,...,N_{\mathrm{fft}}\}$ and $m \in \{1,\cdots,N_{\mathrm{t}}\}$. The decoding matrix $\boldsymbol{W}_n^{\mathrm{H}}[m] = \boldsymbol{W}_{\mathrm{D},n}^{\mathrm{H}}[m]\boldsymbol{W}_{\mathrm{A}}^{\mathrm{H}}[m] \in \mathbb{C}^{N_{\mathrm{R}}\times M_{\mathrm{R}}}$, as well as the precoding matrix $\boldsymbol{F}_n[m] = \boldsymbol{F}_{\mathrm{A}}[m]\boldsymbol{F}_{\mathrm{D},n}[m] \in \mathbb{C}^{M_{\mathrm{T}}\times N_{\mathrm{T}}}$ include digital and analog parts. We assume that the analog precoding $\boldsymbol{F}_{\mathrm{A}}[m] \in \mathbb{C}^{M_{\mathrm{T}}\times N_{\mathrm{T}}}$ and decoding $\boldsymbol{W}_{\mathrm{A}}^{\mathrm{H}}[m] \in \mathbb{C}^{N_{\mathrm{R}}\times M_{\mathrm{R}}}$ are implemented by using fully-connected networks of phase-shifters. Thereby, the analog precoding and decoding matrices have unit modulus entries. In the channel estimation step, we do not use digital precoders and decoders on the pilot subcarriers n_{p}, i.e., $\boldsymbol{F}_{\mathrm{D},n_{\mathrm{p}}}[m] = \boldsymbol{I}_{N_{\mathrm{T}}}$ and $\boldsymbol{W}_{\mathrm{D},n_{\mathrm{p}}}^{\mathrm{H}}[m] = \boldsymbol{I}_{N_{\mathrm{R}}}$. The power constraint for all pilots within one symbol is set to be $\sum_{n_{\mathrm{p}}} \|\boldsymbol{F}_{n_{\mathrm{p}}}[m]\boldsymbol{s}_{n_{\mathrm{p}}}[m]\|_2^2 = P_{pilot},\ \forall m$.

B. Channel model

Let $\Delta_{\mathrm{f}} = 1/(N_{\mathrm{fft}}\cdot T_{\mathrm{s}})$ represent the subcarrier spacing and T_{s} denote the sampling interval. The channel is modeled by a L-path model as in [14] and [15]. In this paper, we assume that the number of dominant paths is known. On the n-th subcarrier the channel is given by

$$\boldsymbol{H}_n = \sum_{\ell=1}^{L} \alpha_\ell \cdot \boldsymbol{a}_{\mathrm{R}}(\theta_\ell)\cdot \boldsymbol{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_\ell)\cdot e^{j(n-1)\mu_{\mathrm{f},\ell}} \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}}, \quad (2)$$

where $n \in \{1,...,N_{\mathrm{fft}}\}$ and $\ell \in \{1,...,L\}$. Moreover, α_ℓ denotes the complex-valued channel gain of the ℓ-th path, $\mu_{\mathrm{f},\ell} = -2\pi\tau_\ell\Delta_{\mathrm{f}}$ is angular frequency corresponding to the the propagation delay τ_ℓ, $\boldsymbol{a}_{\mathrm{R}}(\theta_\ell) \in \mathbb{C}^{M_{\mathrm{R}}\times 1}$ and $\boldsymbol{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_\ell) \in \mathbb{C}^{M_{\mathrm{T}}\times 1}$ are the array steering vectors with a Vandermonde structure (due to the ULA assumption). These array steering vectors $\boldsymbol{a}(\theta)$ have the following structure

$$\boldsymbol{a}(\theta) = \begin{bmatrix}1 & e^{j\mu} & \cdots & e^{j(M-1)\mu}\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{M\times 1}, \quad (3)$$

where μ is a spatial frequency that is given by $\mu = -\frac{2\pi}{\lambda}\Delta\sin\theta$. Note that we also have the Vandermonde structure in the frequency domain on the N_{fft} subcarriers as in (3).

If we stack the channel realizations $\boldsymbol{H}_n$ for different subcarriers $n \in \{1,...,N_{\mathrm{fft}}\}$ along the frequency dimension, we obtain a three dimensional channel tensor $\boldsymbol{\mathcal{H}} = \begin{bmatrix}\boldsymbol{H}_1 \llcorner_3 \cdots \llcorner_3 \boldsymbol{H}_{N_{\mathrm{fft}}}\end{bmatrix} \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}\times N_{\mathrm{fft}}}$. Taking into account L-path model (2), the channel tensor $\boldsymbol{\mathcal{H}}$ incorporating all subcarriers can also be rewritten in the following form

$$\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{4,L} \times_1 \boldsymbol{A}_{\mathrm{R}} \times_2 \boldsymbol{A}_{\mathrm{T}}^{*} \times_3 \boldsymbol{A}_{\mathrm{f}} \times_4 \boldsymbol{d}^{\mathrm{T}} \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}\times N_{\mathrm{fft}}\times 1}, \quad (4)$$

where $\boldsymbol{A}_{\mathrm{R}} = \begin{bmatrix}\boldsymbol{a}_{\mathrm{R}}(\theta_1),\ \boldsymbol{a}_{\mathrm{R}}(\theta_2),\ \cdots \boldsymbol{a}_{\mathrm{R}}(\theta_L)\end{bmatrix} \in \mathbb{C}^{M_{\mathrm{R}}\times L}$, $\boldsymbol{A}_{\mathrm{T}} = \begin{bmatrix}\boldsymbol{a}_{\mathrm{T}}(\phi_1),\ \boldsymbol{a}_{\mathrm{T}}(\phi_2),\ \cdots \boldsymbol{a}_{\mathrm{T}}(\phi_L)\end{bmatrix} \in \mathbb{C}^{M_{\mathrm{T}}\times L}$ are receive and transmit steering matrices, $\boldsymbol{A}_{\mathrm{f}} = \begin{bmatrix}\boldsymbol{a}_{\mathrm{f}}(\tau_1),\ \boldsymbol{a}_{\mathrm{f}}(\tau_2),\ \cdots \boldsymbol{a}_{\mathrm{f}}(\tau_L)\end{bmatrix} \in \mathbb{C}^{N_{\mathrm{p}}\times L}$ is a matrix with propagation delays, and $\boldsymbol{d} = \begin{bmatrix}\alpha_1,\ \cdots,\ \alpha_L\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{L}$ is a vector with channel gains.

C. Received signal

The training phase consists of training frames. Each training frame comprises N_{T} OFDM symbols. Identical pilot positions for all subcarriers and frames, as well as identical analog precoding and decoding matrices are used in each frame [10]. The received signal on the n_{p}-th pilot subcarrier and in the k-th frame is expressed as

$$\boldsymbol{Y}_{n_{\mathrm{p}},k} = \boldsymbol{W}_k^{\mathrm{H}}\boldsymbol{H}_{n_{\mathrm{p}}}\boldsymbol{F}_k\boldsymbol{S}_{n_{\mathrm{p}},k} + \boldsymbol{Z}_{n_{\mathrm{p}},k} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}, \quad (5)$$

where $\boldsymbol{Y}_{n_{\mathrm{p}},k} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$ is the received signal after stacking N_{T} consecutive vectors $\boldsymbol{y}_{n_{\mathrm{p}}}[m] \in \mathbb{C}^{N_{\mathrm{R}}}$ next to each other, $\boldsymbol{F}_k \in \mathbb{C}^{M_{\mathrm{T}}\times N_{\mathrm{T}}}$ and $\boldsymbol{W}_k^{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{R}}\times M_{\mathrm{R}}}$ are precoding and decoding matrices. Here, $\boldsymbol{W}_k$ and $\boldsymbol{F}_k$ contain N_{T} and N_{R} phase-shifted columns of a DFT matrix [12], respectively, and $\boldsymbol{H}_{n_{\mathrm{p}}} \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}}$ denotes the channel between the transmitter and the receiver on the n_{p}-th subcarrier. The noise matrix is given by $\boldsymbol{Z}_{n_{\mathrm{p}},k} = \boldsymbol{W}_k^{\mathrm{H}}\boldsymbol{N}_{n_{\mathrm{p}},k}$, where $\boldsymbol{N}_{n_{\mathrm{p}},k} \in \mathbb{C}^{M_{\mathrm{R}}\times N_{\mathrm{T}}}$ is a matrix with samples of ZMCSCG noise. Moreover, $\boldsymbol{S}_{n_{\mathrm{p}},k} \in \mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}}$ is the accumulated matrix on the pilot subcarrier n_{p} and the training frame k. We assume that the matrix $\boldsymbol{S}_{n_{\mathrm{p}},k}$ has orthogonal rows for the subcarriers with pilots, i.e., $\boldsymbol{S}_{n_{\mathrm{p}},k}\boldsymbol{S}_{n_{\mathrm{p}},k}^{\mathrm{H}} = \gamma\boldsymbol{I}_{N_{\mathrm{T}}}$ for $n_{\mathrm{p}} \in \{n_1,...,n_{N_{\mathrm{p}}}\}$.

Multiplying the received signal on the n_{p}-th subcarrier and in the k-th frame by $\boldsymbol{S}_{n_{\mathrm{p}},k}^{\mathrm{H}}$ from the right side and normalizing the result by γ, the vectorization of the received signal can be represented as

$$\boldsymbol{y}_{n_{\mathrm{p}},k} = \mathrm{vec}\left\{\boldsymbol{Y}_{n_{\mathrm{p}},k}\cdot\frac{1}{\gamma}\boldsymbol{S}_{n_{\mathrm{p}},k}^{\mathrm{H}}\right\} \in \mathbb{C}^{N_{\mathrm{R}}N_{\mathrm{T}}} \quad (6)$$

By stacking N_{p} consecutive received signals $\boldsymbol{y}_{n_{\mathrm{p}},k}$, we construct the combined received signal vector $\boldsymbol{y}_k = \begin{bmatrix}\boldsymbol{y}_{n_1,k}^{\mathrm{T}} & \cdots & \boldsymbol{y}_{n_{N_{\mathrm{p}}},k}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{R}}N_{\mathrm{T}}N_{\mathrm{p}}}$. Taking into account the channel model (4), the combined received signals for all pilot subcarriers can be represented as

$$\boldsymbol{y}_k = (\boldsymbol{\Phi}^{\mathrm{T}}\otimes\boldsymbol{F}_k^{\mathrm{T}}\otimes\boldsymbol{W}_k^{\mathrm{H}})\cdot(\boldsymbol{A}_{\mathrm{f}}\diamond\boldsymbol{A}_{\mathrm{T}}^{*}\diamond\boldsymbol{A}_{\mathrm{R}})\cdot\boldsymbol{d} + \boldsymbol{z}_k, \quad (7)$$

where $\boldsymbol{z}_k \in \mathbb{C}^{N_{\mathrm{R}}N_{\mathrm{T}}N_{\mathrm{p}}}$ is the effective noise vector, $\boldsymbol{\Phi} \in \mathbb{R}^{N_{\mathrm{fft}}\times N_{\mathrm{p}}}$ is a pilot selection matrix that selects N_{p} out of N_{fft} subcarriers corresponding to the transmitted pilots[1]. We can consider the vector $\boldsymbol{y}_k$ as the 4-mode unfolding of the tensor $\boldsymbol{\mathcal{Y}}_k \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}\times N_{\mathrm{p}}\times 1}$, which can be defined as:

$$\boldsymbol{\mathcal{Y}}_k = \boldsymbol{\mathcal{I}}_{4,L}\times_1(\boldsymbol{W}_k^{\mathrm{H}}\boldsymbol{A}_{\mathrm{R}})\times_2(\boldsymbol{F}_k^{\mathrm{T}}\boldsymbol{A}_{\mathrm{T}}^{*})\times_3(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{A}_{\mathrm{f}})\times_4\boldsymbol{d}^{\mathrm{T}} + \boldsymbol{\mathcal{Z}}_k, \quad (8)$$

where $\boldsymbol{\mathcal{Z}}_\kappa = \boldsymbol{\mathcal{N}}_k\times_1\boldsymbol{W}_k^{\mathrm{H}}\times_2\boldsymbol{S}_{n_{\mathrm{p}},k}^{\mathrm{H}}$ is the effective noise tensor, $\boldsymbol{\mathcal{N}}_k$ is the tensor with samples of ZMCSCG noise.

III. Proposed Channel Estimation Algorithm

In this section, we describe the proposed channel estimation algorithm.

A. Pilot based channel estimation

The proposed algorithm starts with an initial estimation of the channel using pilots on each OFDM symbol. To avoid interference between different transmitters, each OFDM symbol is transmitted sequentionally from different RF chains, i.e., at each time, only one RF chain is active. In the training phase, the N_{t} training OFDM symbols are divided into κ_{t} training frames with identical structure, i.e., $N_{\mathrm{t}} = \kappa_{\mathrm{t}}\cdot N_{\mathrm{T}}$. The digital precoders and decoders on the pilot subcarriers are set to identity matrices $\boldsymbol{F}_{\mathrm{D},n_{\mathrm{p}},k} = \boldsymbol{I}_{N_{\mathrm{T}}}$, $\boldsymbol{W}_{\mathrm{D},n_{\mathrm{p}},k} = \boldsymbol{I}_{N_{\mathrm{R}}},\ \forall n_{\mathrm{p}} \in \{n_1,...,n_{N_{\mathrm{p}}}\},\ \kappa \in \{1,...,\kappa_{\mathrm{t}}\}$

We perform sector-wise estimation of the channel, so the analog precoder or the decoder change only for different training frames. Then, the received signals for different training frames can be combined coherently and processed jointly, while in [10] the received training data in different frames are also used separately. We have $\kappa_{\mathrm{t}} = N_{\mathrm{prec,T}}\cdot N_{\mathrm{prec,R}}$, where $N_{\mathrm{prec,T}}$ and $N_{\mathrm{prec,R}}$ are the number of used analog precoders and decoders, respectively.

B. Coarse estimation

In the first step of the algorithm, we perform a coarse estimation of the angular and delay information of the channel without an analog beamforming network. The goal is to obtain initial information about possible SoIs for further analysis. To obtain rough estimates, we propose to form a wide-beam by switching off some antennas of the arrays. We keep only $M_{\mathrm{T}} = N_{\mathrm{T}}$ and $M_{\mathrm{R}} = N_{\mathrm{R}}$ antennas corresponding to the number of transceiver chains. In this case, the tensor of the combined received signals $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}\times N_{\mathrm{p}}\times 1}$ for one training frame can be represented as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{I}}_{4,L}\times_1\boldsymbol{A}_{\mathrm{R}}^{(\mathrm{eff})}\times_2\boldsymbol{A}_{\mathrm{T}}^{(\mathrm{eff})^{*}}\times_3(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{A}_{\mathrm{f}})\times_4\boldsymbol{d}^{\mathrm{T}} + \boldsymbol{\mathcal{Z}}, \quad (9)$$

[1]Control data can be transmitted on the subcarriers that are not used for pilots.

Algorithm 1 3-D Unitary Tensor-ESPRIT in element space

1: **Pre-processing**
- **Smoothing**
  - Obtain $\boldsymbol{\mathcal{Y}}_{\mathrm{ss}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}} \times N_{\mathrm{sub,p}} \times L_{\mathrm{f}}}$ by applying smoothing to $\boldsymbol{\mathcal{Y}}$
- **Forward-backward averaging**
  - $\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{ss}} = \boldsymbol{\mathcal{Y}}_{\mathrm{ss}}^{*} \times_1 \boldsymbol{\Pi}_{N_{\mathrm{R}}} \times_2 \boldsymbol{\Pi}_{N_{\mathrm{T}}} \times_3 \boldsymbol{\Pi}_{N_{\mathrm{sub,p}}} \times_4 \boldsymbol{\Pi}_{2L_{\mathrm{f}}}$,
  - $\boldsymbol{\mathcal{Z}} = \left[\boldsymbol{\mathcal{Y}}_{\mathrm{ss}} \hookrightarrow_4 \tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{ss}}\right] \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}} \times N_{\mathrm{sub,p}} \times 2L_{\mathrm{f}}}$.
- **Real-valued transformation**
  - $\varphi(\boldsymbol{\mathcal{Z}}) = \boldsymbol{\mathcal{Z}} \times_1 \boldsymbol{Q}_{N_{\mathrm{R}}}^{\mathrm{H}} \times_2 \boldsymbol{Q}_{N_{\mathrm{T}}}^{\mathrm{H}} \times_3 \boldsymbol{Q}_{N_{\mathrm{sub,p}}}^{\mathrm{H}} \times_4 \boldsymbol{Q}_{2L_{\mathrm{f}}}^{\mathrm{H}}$.

2: **Tensor signal subspace estimation**
- Compute HOSVD of $\varphi(\boldsymbol{\mathcal{Z}})$ and truncate it to rank L
  - $\varphi(\boldsymbol{\mathcal{Z}}) = \boldsymbol{\mathcal{S}}_Z \times_1 \boldsymbol{E}_1 \times_2 \boldsymbol{E}_2 \times_3 \boldsymbol{E}_3 \times_4 \boldsymbol{E}_4 \approx \boldsymbol{\mathcal{S}}_Z^{[s]} \times_1 \boldsymbol{E}_1^{[s]} \times_2 \boldsymbol{E}_2^{[s]} \times_3 \boldsymbol{E}_3^{[s]} \times_4 \boldsymbol{E}_4^{[s]}$.
- Determine $\boldsymbol{\mathcal{E}}^{[s]} \in \mathbb{R}^{N_{\mathrm{R}} \times N_{\mathrm{T}} \times N_{\mathrm{sub,p}} \times L}$
  - $\boldsymbol{\mathcal{E}}^{[s]} = \boldsymbol{\mathcal{S}}_Z^{[s]} \times_1 \boldsymbol{E}_1^{[s]} \times_2 \boldsymbol{E}_2^{[s]} \times_3 \boldsymbol{E}_3^{[s]}$.

3: **Solution of the invariance equation**
- Construct the selection matrices
  - $\boldsymbol{K}_1^{(r)} \in \mathbb{C}^{(B^{(r)}-1) \times B^{(r)}}$ and $\boldsymbol{K}_2^{(r)} \in \mathbb{C}^{(B^{(r)}-1) \times B^{(r)}}$, $B^{(r)} \in \{N_{\mathrm{R}}, N_{\mathrm{T}}, N_{\mathrm{sub,p}}\}$, $r = 1, 2, 3$ [11].
- Construct 3 shift invariance equations from [11]
  - $\boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{K}_1^{(r)} \times_4 \boldsymbol{\Upsilon}^{(r)} \approx \boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{K}_2^{(r)}$, $r = 1, 2, 3$.
- Solve the equations for the matrices $\boldsymbol{\Upsilon}^{(r)} \in \mathbb{R}^{L \times L}$, $\forall r$, e.g., via LS.

4: **Spatial frequency estimation**
- Obtain the estimates of $\hat{\boldsymbol{\Omega}}^{(r)}$ by computing the joint EVD of $\boldsymbol{\Upsilon}^{(r)} = \boldsymbol{T} \cdot \hat{\boldsymbol{\Omega}}^{(r)} \cdot \boldsymbol{T}^{-1}$, $\forall r$.
- Compute $\hat{\mu}_i^{(r)} = 2 \arctan(\hat{\boldsymbol{\Omega}}_{i,i}^{(r)})$, $\forall i, r$.

where $\boldsymbol{A}_{\mathrm{R}}^{(\mathrm{eff})} \in \mathbb{C}^{N_{\mathrm{R}} \times L}$, $\boldsymbol{A}_{\mathrm{T}}^{(\mathrm{eff})} \in \mathbb{C}^{N_{\mathrm{T}} \times L}$ are the steering matrices of the shortened antenna arrays at the receiver and the transmitter.

We apply 3-D Unitary Tensor-ESPRIT in element space (Algorithm 1) to the reduced aperture array to obtain initial estimates of dominant spatial frequencies (DoAs, DoDs, and delays), where $\boldsymbol{\Pi}_q \in \mathbb{R}^{q \times q}$ is the exchange matrix with ones on the anti-diagonal and zeros elsewhere $\boldsymbol{Q}_q \in \mathbb{C}^{q \times q}$ is a left-$\boldsymbol{\Pi}$-real and unitary matrix, i.e., $\boldsymbol{\Pi} \cdot \boldsymbol{Q}_q^{*} = \boldsymbol{Q}_q$ [11]. Since only a single snapshot is available, smoothing with L_{f} subarrays is performed along the frequency dimension, yielding

$$\boldsymbol{\mathcal{Y}}_{\mathrm{ss}} = \left[\boldsymbol{\mathcal{Y}}_{1,1,1} \hookrightarrow_4 \boldsymbol{\mathcal{Y}}_{1,1,2} \hookrightarrow_4 \cdots \hookrightarrow_4 \boldsymbol{\mathcal{Y}}_{1,1,L_{\mathrm{f}}}\right], \tag{10}$$

where $\boldsymbol{\mathcal{Y}}_{\mathrm{ss}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}} \times N_{\mathrm{sub,p}} \times L_{\mathrm{f}}}$ is a tensor with the received data after smoothing and $\boldsymbol{\mathcal{Y}}_{1,1,\ell_{\mathrm{f}}} = \boldsymbol{\mathcal{Y}} \times_3 \boldsymbol{J}_{\mathrm{ss},\ell_{\mathrm{f}}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}} \times N_{\mathrm{sub,p}} \times 1}$ for $\ell_{\mathrm{f}} = 1, \cdots, L_{\mathrm{f}}$. The selection matrix $\boldsymbol{J}_{\mathrm{ss},\ell_{\mathrm{f}}}$ is defined as

$$\boldsymbol{J}_{\mathrm{ss},\ell_{\mathrm{f}}} = \left[\boldsymbol{0}_{N_{\mathrm{sub,p}} \times (\ell_{\mathrm{f}}-1)} \quad \boldsymbol{I}_{N_{\mathrm{sub,p}}} \quad \boldsymbol{0}_{N_{\mathrm{sub,p}} \times (L_{\mathrm{f}}-\ell_{\mathrm{f}})}\right], \tag{11}$$

where L_{f} is the number of frequency subblocks for the smoothing, and $N_{\mathrm{sub,p}}$ is the number of subcarriers per subblock.

The coarse estimation step requires only one training frame. The maximum number of resolvable paths using 3-D Unitary Tensor-ESPRIT is equal to:

$$L_{\max} = \min\Big\{N_{\mathrm{sub,p}}(N_{\mathrm{T}}-1)N_{\mathrm{R}},\ N_{\mathrm{sub,p}}N_{\mathrm{T}}(N_{\mathrm{R}}-1), (N_{\mathrm{sub,p}}-1)N_{\mathrm{T}}N_{\mathrm{R}},\ 2L_{\mathrm{f}}\Big\}. \tag{12}$$

C. Sectorization

After the coarse estimation step, we obtain an initial estimate of $\boldsymbol{\mu}_\ell$, $\forall \ell$. Based on these estimates, we perform sectorization by selecting B consecutive rows of a phase-shifted DFT matrix, starting from the $k_{\min}$-th row, to form a beam pattern encompassing the dominant paths estimated in the first step, namely, the sector of interest (SoI). We perform the sectorization for the transmitter and the receiver sides independently. It starts by selecting the D nearest beams for each dominant path, from the predefined DFT grid of beams. Depending on the value of the parameter D, we can control the width of the resulting sector of interest. Numerical results show that for a low SNR range, a better strategy is to construct wider SoIs corresponding to larger values of D, while for a high SNR range, a better performance is achieved for smaller values of D.

After the assignment of beams to each dominant path, the information about all the chosen beams are combined by taking into account any possible overlap. Then, we divide the formed set into different SoIs. The number of beams assigned to each SoI depends on the number of available RF chains. A preliminary analysis shows that a better performance is achieved by choosing the angular locations close to the center of the associated sector of interest.

D. Fine estimation

At the fine estimation step, we scan the relevant SoIs to obtain high-resolution estimates of the dominant paths. The high-resolution parameter estimation procedure requires one training frame for each SoI. For this reason, the total number of required training frames is equal to the overall number of SoIs, and can be found as

$$\kappa_{\mathrm{t}} = N_{\mathrm{prec,T}} \cdot N_{\mathrm{prec,R}} = \left\lceil \frac{N_{\mathrm{B,T}}}{N_{\mathrm{T}}} \right\rceil \cdot \left\lceil \frac{N_{\mathrm{B,R}}}{N_{\mathrm{R}}} \right\rceil \leq \frac{M_{\mathrm{T}}}{N_{\mathrm{T}}} \cdot \frac{M_{\mathrm{R}}}{N_{\mathrm{R}}}, \tag{13}$$

where $N_{\mathrm{B,T}}$ and $N_{\mathrm{B,R}}$ are the total number of selected beams from the DFT grid at the transmitter and the receiver after the sectorization procedure. As we can see from (13), the required number of frames for channel estimation is significantly smaller than for the full DFT beamspace.

At the receiver, the received signal tensors are combined to obtain the final effective signal tensor.

Let $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ denote the tensors that contain the same hybrid precoder, i.e., $\boldsymbol{F}_1 = \boldsymbol{F}_2$, but different hybrid decoders $\boldsymbol{W}_1^{\mathrm{H}}$ and $\boldsymbol{W}_2^{\mathrm{H}}$. If $\boldsymbol{W}$ and $\boldsymbol{W}$ have non-overlapping columns, then we concatenate $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ along the 1-modes as $[\boldsymbol{\mathcal{Y}}_1 \hookrightarrow_1 \boldsymbol{\mathcal{Y}}_2]$. If $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ have r_1 overlapping columns, then we average the parts of the tensors that correspond to the overlapping columns. Let us define $\boldsymbol{\mathcal{Y}}_{11} = \boldsymbol{\mathcal{Y}}_1(1:(N_{\mathrm{R}}-r_1),:,:)$, $\boldsymbol{\mathcal{Y}}_{12} = \boldsymbol{\mathcal{Y}}_1((N_{\mathrm{R}}-r_1+1):N_{\mathrm{R}},:,:)$, $\boldsymbol{\mathcal{Y}}_{21} = \boldsymbol{\mathcal{Y}}_2(1:r_1,:,:)$, $\boldsymbol{\mathcal{Y}}_{22} = \boldsymbol{\mathcal{Y}}_2((r_1+1):N_{\mathrm{R}},:,:)$, where $\boldsymbol{\mathcal{Y}}_{12}$ and $\boldsymbol{\mathcal{Y}}_{21}$ correspond to the same hybrid precoder and decoder. The combination of $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ is achieved via $[\boldsymbol{\mathcal{Y}}_{11} \hookrightarrow_1 (\boldsymbol{\mathcal{Y}}_{12} + \boldsymbol{\mathcal{Y}}_{21})/2 \hookrightarrow_1 \boldsymbol{\mathcal{Y}}_{22}]$. After all the training frames are combined according to the aforementioned method, we obtain the final effective combined received signal tensor $\bar{\boldsymbol{\mathcal{Y}}} \in \mathbb{C}^{N_{\mathrm{B,R}} \times N_{\mathrm{B,T}} \times N_{\mathrm{p}} \times 1}$, i.e.,

$$\bar{\boldsymbol{\mathcal{Y}}} = \boldsymbol{\mathcal{I}}_{4,L} \times_1 (\bar{\boldsymbol{W}}^{\mathrm{H}} \boldsymbol{A}_{\mathrm{R}}) \times_2 (\bar{\boldsymbol{F}}^{\mathrm{T}} \boldsymbol{A}_{\mathrm{T}}^{*}) \times_3 (\boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{A}_{\mathrm{f}}) \times_4 \boldsymbol{d}^{\mathrm{T}} + \bar{\boldsymbol{\mathcal{Z}}}, \tag{14}$$

where $\bar{\boldsymbol{\mathcal{Z}}}$ is the effective noise tensor, $\bar{\boldsymbol{W}} \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{B,R}}}$ and $\bar{\boldsymbol{F}} \in \mathbb{C}^{M_{\mathrm{T}} \times N_{\mathrm{B,T}}}$ comprise the non-overlapped phase-shifted columns [12] of DFT matrices that are used to form the SoIs. Then we apply Unitary Tensor-ESPRIT in DFT beamspace (Algorithm 2) to obtain fine estimates of the spatial frequencies (DoAs, DoDs, and delays) from the tensor $\bar{\boldsymbol{\mathcal{Y}}}$ of the combined received training frames. The advantages of the algorithm are the lower computational complexity and the real-valued signal processing for the signal subspace estimation and the subsequent spatial frequencies estimation step. Again, we utilize smoothing as a preprocessing step that is carried out along the frequency dimension. After the high-resolution estimation of the spatial frequencies $\left[\mu_{\mathrm{R},\ell} \quad \mu_{\mathrm{T},\ell} \quad \mu_{\mathrm{f},\ell}\right]^{\mathrm{T}}$, $\forall \ell$, we obtain a least squares estimate of the channel gains as

$$\hat{\boldsymbol{d}} = (\hat{\boldsymbol{A}}_{\mathrm{f}}^{(\mathrm{eff})} \diamond \hat{\boldsymbol{A}}_{\mathrm{T}}^{(\mathrm{eff})*} \diamond \hat{\boldsymbol{A}}_{\mathrm{R}}^{(\mathrm{eff})})^{+} \cdot \bar{\boldsymbol{y}} \in \mathbb{C}^{L}, \tag{15}$$

where $\bar{\boldsymbol{y}} = \mathrm{vec}\{\bar{\boldsymbol{\mathcal{Y}}}\} \in \mathbb{C}^{N_{\mathrm{B,R}} N_{\mathrm{B,T}} N_{\mathrm{p}}}$, while $\hat{\boldsymbol{A}}_{\mathrm{R}}^{(\mathrm{eff})} \in \mathbb{C}^{N_{\mathrm{B,R}} \times L}$, $\hat{\boldsymbol{A}}_{\mathrm{T}}^{(\mathrm{eff})} \in \mathbb{C}^{N_{\mathrm{B,T}} \times L}$, and $\hat{\boldsymbol{A}}_{\mathrm{f}}^{(\mathrm{eff})} \in \mathbb{C}^{N_{\mathrm{p}} \times L}$ are reconstructed versions of the effective receive steering matrix, the effective transmit steering matrix and the effective matrix with propagation delays, respectively. The maximum number of resolvable uncorrelated sources using 3-D Unitary Tensor-ESPRIT in DFT beamspace equals to:

Algorithm 2 3-D Unitary Tensor-ESPRIT in DFT beamspace

1: **Pre-processing :**
- **Beamspace transformation** (analog precoding and decoding) and pilot subcarrier selection
  - $\bar{\boldsymbol{\mathcal{Y}}} = \boldsymbol{\mathcal{X}} \times_1 \bar{\boldsymbol{W}}^{\mathrm{H}} \times_2 \bar{\boldsymbol{F}}^{\mathrm{T}} \times_3 \boldsymbol{\Phi} \in \mathbb{C}^{N_{\mathrm{B,R}} \times N_{\mathrm{B,T}} \times N_{\mathrm{p}} \times 1}$.
- **Smoothing**
  - Obtain $\boldsymbol{\mathcal{Y}}_{\mathrm{ss}} \in \mathbb{C}^{N_{\mathrm{B,R}} \times N_{\mathrm{B,T}} \times N_{\mathrm{sub,p}} \times L_{\mathrm{f}}}$ by applying smoothing to $\bar{\boldsymbol{\mathcal{Y}}}$.
- **Real-valued transformation**
  - $\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{ss}} = \boldsymbol{\mathcal{Y}}_{\mathrm{ss}} \times_3 \boldsymbol{Q}^{\mathrm{H}}_{N_{\mathrm{sub,p}}} \in \mathbb{C}^{N_{\mathrm{B,R}} \times N_{\mathrm{B,T}} \times N_{\mathrm{sub,p}} \times L_{\mathrm{f}}}$,
  - $\boldsymbol{\mathcal{Z}} = \left[\mathrm{Re}\left\{\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{ss}}\right\} \sqcup_4 \mathrm{Im}\left\{\tilde{\boldsymbol{\mathcal{Y}}}_{\mathrm{ss}}\right\}\right]$.

2: **Tensor signal subspace estimation**
- Compute HOSVD of $\boldsymbol{\mathcal{Z}}$ and truncate it to rank L.
  - $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{S}}_Z \times_1 \boldsymbol{E}_1 \times_2 \boldsymbol{E}_2 \times_3 \boldsymbol{E}_3 \times_4 \boldsymbol{E}_4 \approx \boldsymbol{\mathcal{S}}_Z^{[s]} \times_1 \boldsymbol{E}_1^{[s]} \times_2 \boldsymbol{E}_2^{[s]} \times_3 \boldsymbol{E}_3^{[s]} \times_4 \boldsymbol{E}_4^{[s]}$.
- Determine the signal space $\boldsymbol{\mathcal{E}}^{[s]} \in \mathbb{R}^{N_{\mathrm{B,R}} \times N_{\mathrm{B,T}} \times N_{\mathrm{sub,p}} \times L}$
  - $\boldsymbol{\mathcal{E}}^{[s]} = \boldsymbol{\mathcal{S}}_Z^{[s]} \times_1 \boldsymbol{E}_1^{[s]} \times_2 \boldsymbol{E}_2^{[s]} \times_3 \boldsymbol{E}_3^{[s]}$.

3: **Solution of the invariance equations**
- Construct the selection matrices
  - $\boldsymbol{\Gamma}_1^{(r)} \in \mathbb{C}^{(B^{(r)}-1) \times B^{(r)}}$ and $\boldsymbol{\Gamma}_2^{(r)} \in \mathbb{C}^{(B^{(r)}-1) \times B^{(r)}}$, $B^{(r)} \in \{N_{\mathrm{B,R}}, N_{\mathrm{B,T}}\}$, $r = 1, 2$ [12].
  - $\boldsymbol{K}_1^{(r)} \in \mathbb{C}^{(N_{\mathrm{sub,p}}-1) \times N_{\mathrm{sub,p}}}$ and $\boldsymbol{K}_2^{(r)} \in \mathbb{C}^{(N_{\mathrm{sub,p}}-1) \times N_{\mathrm{sub,p}}}$, $r = 3$ [11].
- Construct 2 shift invariance equations in DFT beamspace
  - $\boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{\Gamma}_1^{(r)} \times_4 \boldsymbol{\Upsilon}^{(r)} = \boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{\Gamma}_2^{(r)}, r = 1, 2.$
- Construct 1 shift invariance equation from [11]
  - $\boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{K}_1^{(r)} \times_4 \boldsymbol{\Upsilon}^{(r)} \approx \boldsymbol{\mathcal{E}}^{[s]} \times_r \boldsymbol{K}_2^{(r)}, \ r = 3.$
- Solve the equations for the matrices $\boldsymbol{\Upsilon}^{(r)} \in \mathbb{R}^{L \times L}$, $\forall r$, e.g., via LS.

4: **Spatial frequency estimation**
- Obtain the estimates of $\hat{\boldsymbol{\Omega}}^{(r)}$ by computing the joint EVD of $\boldsymbol{\Upsilon}^{(r)} = \boldsymbol{T} \cdot \hat{\boldsymbol{\Omega}}^{(r)} \cdot \boldsymbol{T}^{-1}$, $\forall r$.
- Compute $\hat{\mu}_i^{(r)} = 2 \arctan(\hat{\boldsymbol{\Omega}}_{i,i}^{(r)})$, $\forall i, r$.

$$L_{\max} = \min\Big\{N_{\mathrm{sub,p}}(N_{\mathrm{B,T}}-1)N_{\mathrm{B,R}},\ N_{\mathrm{sub,p}}N_{\mathrm{B,T}}(N_{\mathrm{B,R}}-1), (N_{\mathrm{sub,p}}-1)N_{\mathrm{B,T}}N_{\mathrm{B,R}},\ 2L_{\mathrm{f}}\Big\}. \tag{16}$$

IV. Simulation Results

We provide some numerical results to show the performance of the proposed algorithm. For all scenarios, a MIMO-OFDM communication system with QPSK modulation is assumed. The total number of subcarriers is set to $N_{\mathrm{fft}} = 64$ and the size of the cyclic prefix is $N_{\mathrm{cp}} = 16$. Moreover, the number of transmit and receive antennas is given by $M_{\mathrm{T}} = M_{\mathrm{R}} = 24$, and the number of scatterers is varied in our simulation experiments. The receiver and transmitter antennas are connected to RF chains via fully-connected networks of phase shifters, with $N_{\mathrm{T}} = N_{\mathrm{R}} = 6$. The total number of pilots per OFDM symbol is $N_{\mathrm{p}} = 25$, which are transmitted on subcarriers with the following indexes $n_{\mathrm{p}} \in \{2, \cdots, 26\}$. Regarding the smoothing, we divide the total number of subcarriers with pilots into $L_{\mathrm{f}} = 18$ blocks. For each transmission, the total power of one OFDM symbol is set to unity. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} = 10 \log \frac{1}{N_{\mathrm{p}} \sigma_n^2}$. In addition, we assume a L-path channel model (4) during simulations. For each path ℓ, we fix the direction of arrival $\mu_{\mathrm{R},\ell}$, direction of departure $\mu_{\mathrm{T},\ell}$, delay τ_ℓ, and propagation gain α_ℓ during all Monte Carlo runs. We assume that the channel stays constant during the channel estimation procedure.

We simulate all algorithms for $D = \{2, 4\}$ to evaluate the influence of the width of a beam that is formed in the direction of one dominant path. The performance of the proposed algorithm is evaluated in terms of the mean square error (MSE) and the normalized mean square error (NMSE) of the angular frequencies for three different scenarios: 1) Two overlapping SoIs, 2) Single SoI, and 3) Two non-overlapping SoIs. Each point on the plots of the MSE and the NMSE corresponds to an average of 2000 Monte Carlo runs. We calculate the MSE via

$$\mathrm{MSE} = \frac{1}{3L}\ \mathbb{E}\left\{\sum_{x \in \{\mathrm{f,T,R}\}} \sum_{\ell=1}^{L} \left(\mu_{x,\ell} - \hat{\mu}_{x,\ell}\right)^2\right\}, \tag{17}$$

and the NMSE of the vectorized channel $\boldsymbol{h}_{\mathrm{vec}} = (\boldsymbol{A}_{\mathrm{f}} \diamond \boldsymbol{A}_{\mathrm{T}}^* \diamond \boldsymbol{A}_{\mathrm{R}}) \cdot \boldsymbol{d}$ as

$$\mathrm{NMSE} = \mathbb{E}\left\{\frac{\|\boldsymbol{h}_{\mathrm{vec}} - \hat{\boldsymbol{h}}_{\mathrm{vec}}\|_F^2}{\|\boldsymbol{h}_{\mathrm{vec}}\|_F^2}\right\}. \tag{18}$$

TABLE I. Simulation parameters

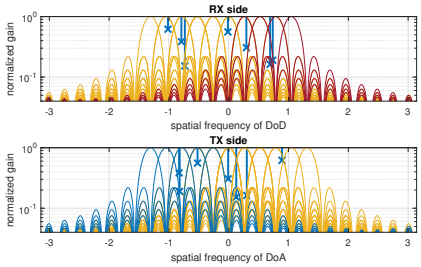
#	$\{\mu_{\mathrm{f},\ell}, \mu_{\mathrm{R},\ell}, \mu_{\mathrm{T},\ell}\}$	α_ℓ
	Scenario #1	
1	$[-2.2492, 0.6920, 0.3020]^{\mathrm{T}}$	$-0.1201 + 0.1094i$
2	$[-2.0844, 0.2947, -0.0099]^{\mathrm{T}}$	$-0.0225 - 0.3053i$
3	$[-1.9085, -0.0142, -0.5217]^{\mathrm{T}}$	$-0.5324 - 0.1670i$
4	$[-1.1468, -0.7918, -0.8300]^{\mathrm{T}}$	$0.2248 - 0.3124i$
5	$[-1.1534, -0.7340, 0.1257]^{\mathrm{T}}$	$-0.1108 + 0.1049i$
6	$[-1.7746, -1.0161, 0.8910]^{\mathrm{T}}$	$0.5111 + 0.3479i$
7	$[-1.5643, 0.7358, -0.8211]^{\mathrm{T}}$	$-0.1045 - 0.1587i$
	Scenario #2	
1	$[-2.0844, -0.2788, -0.1017]^{\mathrm{T}}$	$0.9211 - 0.4053i$
2	$[-1.9085, -0.2613, -0.2961]^{\mathrm{T}}$	$0.3438 + 0.1936i$
3	$[-1.1468, 0.0748, -0.2817]^{\mathrm{T}}$	$0.0909 + 0.5533i$
4	$[-1.1534, -0.0245, -0.3522]^{\mathrm{T}}$	$-0.3444 - 0.5987i$
	Scenario #3	
1	$[-1.5388, -1.2507, -1.5674]^{\mathrm{T}}$	$-0.0512 - 0.2322i$
2	$[0.0374, -1.2563, -1.9265]^{\mathrm{T}}$	$-0.0566 + 0.2011i$
3	$[1.2508, -1.3129, -1.6231]^{\mathrm{T}}$	$0.4097 + 0.2411i$
4	$[2.4561, 1.6183, 1.5428]^{\mathrm{T}}$	$0.0842 - 0.0704i$
5	$[2.8858, 1.3112, 1.5714]^{\mathrm{T}}$	$0.0571 + 0.0623i$
6	$[0.2967, 1.6045, 1.7625]^{\mathrm{T}}$	$0.4583 - 0.3366i$

For the first simulation scenario, we consider two overlapping SoIs (Fig. 1a) with $L = 7$ paths. For the second scenario (Fig. 1d) we consider the case of one cluster with multiple paths, $L = 4$. All paths are located close together to be in the range of one SoI. For the third simulation scenario (Fig. 1g) we evaluate the performance of the algorithm in the case of two separate clusters with $L = 6$ paths. The values for all scenarios used during the simulations are presented in Table I. Fig. 1a, Fig. 1d, and Fig. 1g illustrate the SoIs for the different scenarios, as well as DoAs, DoDs of the paths and their amplitudes (in blue).

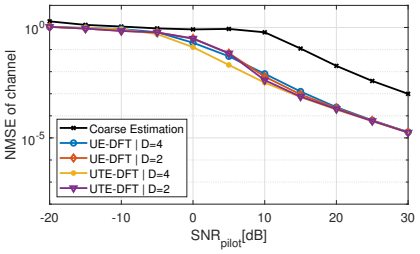
We use "UE-DFT" and "UTE-DFT" to denote Unitary ESPRIT in DFT beamspace and Unitary Tensor-ESPRIT in DFT beamspace, respectively. Fig. 1b, Fig. 1e, and Fig.1h show that the proposed 3-D Unitary Tensor-ESPRIT in DFT beamspace based channel estimation algorithm provides significantly better refined spatial estimates compared to the initial estimates. For all scenarios the proposed algorithm outperforms the other candidates. The influence of the width of the beams changes depending on the value of the SNR. As we can see from Fig. 1i, at low SNRs a better strategy is to use wider beams $D = 4$, while at high SNR narrower beams $D = 2$ yield more accurate estimates of the parameters. In addition to improved parameter estimates, we observe a reduction in the required training overhead. In the second scenario, the channel estimation procedure (coarse and fine steps together) requires only two training frames, which is significantly less than the number of frames for the full scan of the DFT beamspace, which would require 24 frames for this simulation setup.

V. Conclusion

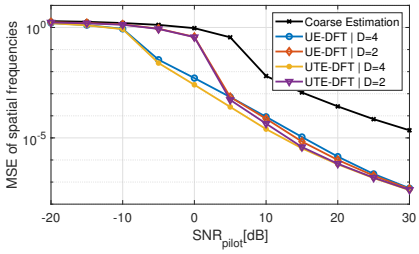
In this paper, we have proposed a two-step gridless channel estimation scheme based on 3-D Unitary Tensor-ESPRIT in DFT beamspace for mmWave hybrid analog-digital massive MIMO systems. The training design uses the columns and rows of phase shifted DFT matrices as the hybrid precoding and decoding matrices together with orthogonal and identical pilot symbols on all pilot subcarriers. 3-D Unitary Tensor-ESPRIT in DFT beamspace is used to provide high



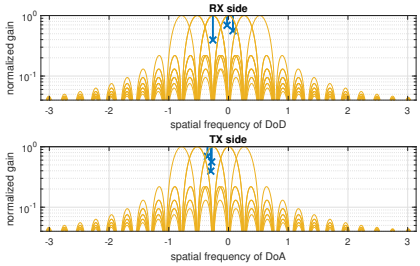
(a) Sectorization of the dominant spatial frequencies for scenario 1, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 7$, $D = 2$, $N_{\mathrm{prec,R}} = N_{\mathrm{prec,T}} = 2$.



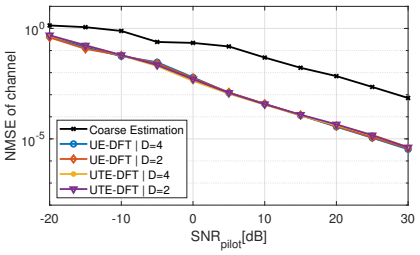
(b) Estimation NMSE of equivalent channel vs. SNR for scenario 1, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 7$.



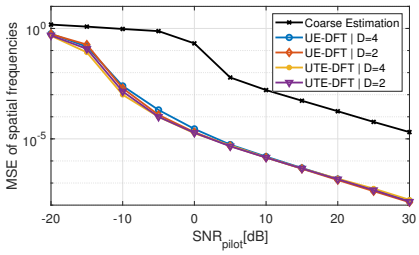
(c) Estimation MSE of spatial frequencies vs. SNR for scenario 1, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 7$.



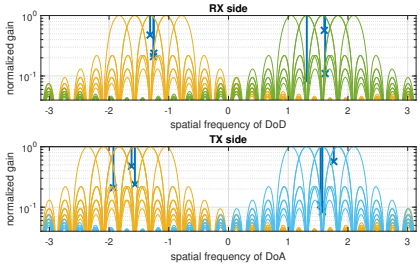
(d) Sectorization of the dominant spatial frequencies for scenario 2 , $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 3$, $D = 2$, $N_{\mathrm{prec,R}} = N_{\mathrm{prec,T}} = 1$.



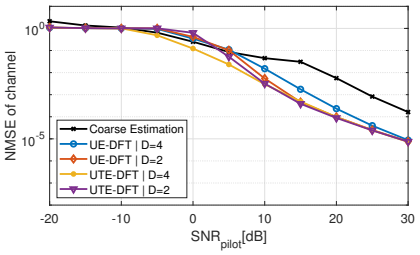
(e) Estimation NMSE of equivalent channel vs. SNR for scenario 2, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 3$.



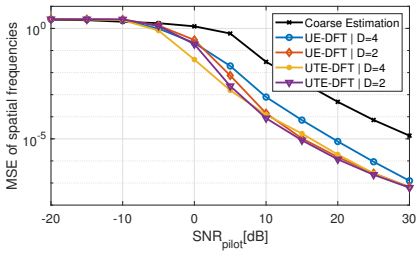
(f) Estimation MSE of spatial frequencies vs. SNR for scenario 2, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 3$.



(g) Sectorization of the dominant spatial frequencies, scenario 3, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 6$, $D = 2$, $N_{\mathrm{prec,R}} = N_{\mathrm{prec,T}} = 2$.



(h) Estimation NMSE of equivalent channel vs. SNR for scenario 3, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 6$.



(i) Estimation MSE of spatial frequencies vs. SNR for scenario 3, $M_{\mathrm{R}} = M_{\mathrm{T}} = 24$, $N_{\mathrm{R}} = N_{\mathrm{T}} = 6$, $L = 6$.

Fig. 1: Numerical results of performance of the 3-D Unitary Tensor-ESPRIT in DFT beamspace for different scenarios

resolution estimates of the spatial frequencies of the dominant paths, which comprises a coarse initial estimation and a fine estimation. Numerical results confirm the high resolution estimates of the spatial frequencies and the channel using only a few OFDM symbols for training.

ACKNOWLEDGMENT

This work has been supported by CAPES/PROBRAL Proc. no. 88887.144009/2017-00 and CAPES/PRINT Proc. no. 88887.311965/2018-00. The authors gratefully acknowlegde the support of DFG under project no. 402834619.

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 Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[2] 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 International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3281–3285.

[3] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[4] K. Venugopal, A. Alkhateeb, and R. W. Heath, "Channel Estimation for Hybrid Architecture-Based Wideband Millimeter Wave Systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.

[5] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel Estimation via Orthogonal Matching Pursuit for Hybrid MIMO Systems in Millimeter Wave Communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, 2016.

[6] R. M. Rial, C. Rusu, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel Estimation and Hybrid Combining for mmWave: Phase shifters or switches?" in *Proc. Information Theory and Applications Workshop (ITA)*, 2015, pp. 90–97.

[7] 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 International ITG Workshop on Smart Antennas*, 2016, pp. 1–8.

[8] D. Zhu, J. Choi, and R. W. Heath, "Two-Dimensional AoD and AoA Acquisition for Wideband Millimeter-Wave Systems With Dual-Polarized MIMO," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 7890–7905, 2017.

[9] 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 European Signal Processing Conference (EUSIPCO)*, 2017.

[10] ——, "Channel Estimation for Hybrid Multi-carrier mmWave MIMO Systems Using Three-dimensional Unitary ESPRIT in DFT beamspace," in *Proc. IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2017, pp. 1–5.

[11] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-Order SVD-Based Subspace Estimation to Improve the Parameter Estimation Accuracy in Multidimensional Harmonic Retrieval Problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, 2008.

[12] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-form 2-D Angle Estimation with Rectangular Arrays in Element Space or Beamspace via Unitary ESPRIT," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 316–328, 1996.

[13] R. Prasad, *OFDM for wireless communications systems*, ser. Artech House universal personal communications series. Artech House, 2004.

[14] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[15] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, A. F. Molisch, C. Oestges, and H. Özcelik, "Survey of Channel and Radio Propagation Models for Wireless MIMO Systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, 2007.