CHANNEL ESTIMATION FOR HYBRID MULTI-CARRIER MMWAVE MIMO SYSTEMS USING THREE-DIMENSIONAL UNITARY ESPRIT IN DFT BEAMSPACE

Jianshu Zhang and Martin Haardt

Communications Research Laboratory, Ilmenau University of Technology, Germany

ABSTRACT

In this paper we study the channel estimation problem for a CP-OFDM based mmWave hybrid analog-digital MIMO system, where the analog processing is achieved using only phase shift networks. A two-stage three-dimensional (3-D) Unitary ESPRIT in DFT beamspace based channel estimation algorithm is proposed to estimate the angular-delay profile and subsequently the unknown frequency-selective channel. The required training protocol, analog precoding and decoding matrices, as well as pilot patterns are discussed. Simulation results show that the proposed multi-stage 3-D Unitary ESPRIT in DFT beamspace based channel estimation algorithm provides high resolution channel estimates.

Index Terms— MmWave Massive MIMO, hybrid precoding and decoding, tensor, Unitary ESPRIT in DFT beamspace.

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. To this end, it is often assumed that the number of RF chains is significantly less than the number of antennas, and the analog processing is realized using only phase shifters. The latter assumption introduces non-convex constraints on the design of analog precoding and decoding matrices [1], [2]. When a broadband multi-carrier system is considered, we have to use the same phase shifts for all subcarriers [3]. These new constraints lead to significant challenges for the design of MIMO processing algorithms as well as the corresponding channel estimation schemes. Many efforts have been put on providing accurate channel estimates for hybrid MIMO processing by using compressed sensing (CS) methods. For instance, an adaptive CS based channel estimation algorithm is proposed in [2] for a frequency-flat channel. This CS based channel estimation algorithm has been further extended in [4] by involving multiple measurement vectors (MMV) to improve the channel estimation accuracy. An adaptive multi-grid sparse recovery approach is applied in [5]. For a frequency-selective channel we have proposed two-stage CS based channel estimation methods in [6] to reduce the involved computational complexity. In [7] CS based channel estimation methods have also been studied for systems based on a single-carrier waveform. Nevertheless, all the above CS based methods depend on the on-grid assumption of the channel parameters, which will require an grid-offset estimation for practical use. The grid-offset estimation itself is already a challenging task. To circumvent this challenging problem, in [8] the angular profiles, i.e., direction of arrivals (DoAs) and the direction of departures (DoDs), are estimated. In [9] we have also proposed a gridless channel estimation algorithm, which first estimates the angular profiles and then the channel gains. The proposed algorithm is based on three-dimensional (3-D) Standard ESPRIT in DFT beamspace algorithm, which can provide high resolution estimates of DoAs and DoDs when all pairs of DoAs and DoDs fall in a specific sector of interest that is known a priori. In practice, it is difficult to ensure that all the DoAs and DoDs lie in a limited sector of interest. This motivates us to develop novel algorithms that still work when some pairs of DoA and DoD are outside the sector of interest.

In this paper we develop a gridless channel estimation algorithm for a hybrid point-to-point mmWave multi-carrier massive MIMO system. The proposed channel estimation algorithm involves a coarse estimation step, where coarse estimates of spatial frequencies and complex channel gains are provided with a wide beam and reduced effective aperture by switching off antennas or phase shifters. These coarse estimates are then used to form an enlarged effective sector of interest or to help suppressing the interference within a sector of interest. Thereby, the proposed algorithm requires neither a priori knowledge of the sector of the interest nor the assumption that all pairs of DoAs and DoDs are within a specific sector of interest. The developed algorithm is based on the 3-D Unitary ESPRIT algorithm in discrete Fourier transform (DFT) beamspace [10], i.e., the analog precoding and decoding matrices comprise phase shifted columns or rows of DFT matrices. Simulation results show that the proposed channel estimation algorithms provide high resolution channel estimates and spatial frequency estimates.

Notation: Upper-case and lower-case bold-faced letters denote matrices and vectors, respectively. The expectation, transpose, conjugate, Hermitian transpose, and Moore-Penrose pseudo-inverse are denoted by $\mathbb{E}\{\cdot\}$, $\{\cdot\}^{\mathrm{T}}$, $\{\cdot\}^{*}$, $\{\cdot\}^{\mathrm{H}}$, and $\{\cdot\}^{+}$, 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 and column exchange matrix are $\boldsymbol{I}_m$, and $\boldsymbol{\Pi}_m$, respectively. The $m \times n$ matrix with all zero elements is $\mathbf{0}_{m\times n}$.

II. PROBLEM FORMULATION

II-A. System Model and Channel Model

We study a point-to-point mmWave massive MIMO system. The transmitter has M_{T} transmit antennas and N_{T} RF chains. The receiver has M_{R} receive antennas and N_{R} RF chains. We have $M_{\mathrm{T}} \gg N_{\mathrm{T}}$ and $M_{\mathrm{R}} \gg N_{\mathrm{R}}$. A CP-OFDM based modulation scheme is applied to combat the multipath effect. The corresponding FFT size is N_{fft}. The N_{R}-dimensional received training signal on the n-th pilot subcarrier in the m-th OFDM symbol is given by [9]

$$\boldsymbol{y}_n[m] = \boldsymbol{W}^{\mathrm{H}}[m](\boldsymbol{H}_n\boldsymbol{F}[m]\boldsymbol{s}_n[m] + \boldsymbol{z}_n[m]), \tag{1}$$

where $\boldsymbol{H}_n \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{T}}}$ denotes the time-invariant channel transfer function (CTF) on the n-th pilot subcarrier and $\boldsymbol{z}_n[m]$ represents zero mean circularly symmetric complex Gaussian (ZMCSCG) noise with covariance matrix $\mathbb{E}\{\boldsymbol{z}_n[m]\boldsymbol{z}_n^{\mathrm{H}}[m]\} = \sigma_{\mathrm{n}}^2\boldsymbol{I}_{M_{\mathrm{R}}}$ for all $n \in \{k_1, \cdots, k_{N_{\mathrm{f}}}\} \subset \{1, \cdots, N_{\mathrm{fft}}\}$ and $m \in \{1, \cdots, N_{\mathrm{t}}\}$. The tuples (m, n) represent discrete time-frequency indices for pilot symbols. The matrices $\boldsymbol{W}^{\mathrm{H}}[m] \in \mathbb{C}^{N_{\mathrm{R}}\times M_{\mathrm{R}}}$ and $\boldsymbol{F}[m] \in \mathbb{C}^{M_{\mathrm{T}}\times N_{\mathrm{T}}}$ denote the analog decoder and precoder, respectively. The analog matrices have only unit modulus entries because we

assume that the analog circuitry is implemented using only phase shift networks. Lastly, the vector $\boldsymbol{s}_n[m] \in \mathbb{C}^{N_\mathrm{T}}$ represents the transmitted pilot symbols on all the transmitting RF chains at a certain time-frequency bin (m,n). The transmit power constraint for the training symbols is set to be $\sum_{n=k_1}^{k_{N_\mathrm{f}}} \|\boldsymbol{F}[m]\boldsymbol{s}_n[m]\|^2 = P_{\mathrm{pilot}}$ for all m. Note that we do not use digital precoders and decoders on pilot subcarriers.

The mmWave massive MIMO channel is modeled using a finite number of scatterers, e.g., L scatterers [11]. Each scatterer contributes to a single propagation path between the transmitter and the receiver, which accounts for one time delay τ_ℓ and one pair of DoA $\theta_{\mathrm{R},\ell} \in [0, 2\pi)$ and DoD $\theta_{\mathrm{T},\ell} \in [0, 2\pi)$. The continuous time-domain channel impulse response (CIR) is given by

$$\boldsymbol{H}(t) = \sum_{\ell=1}^{L} d_\ell \boldsymbol{a}_\mathrm{R}(\theta_{\mathrm{R},\ell})\boldsymbol{a}_\mathrm{T}^\mathrm{T}(\theta_{\mathrm{T},\ell})\delta(t-\tau_\ell), \tag{2}$$

where d_ℓ denotes the random complex gain of the ℓ-th path. The vectors $\boldsymbol{a}_\mathrm{T}(\theta_{\mathrm{T},\ell})$ and $\boldsymbol{a}_\mathrm{R}(\theta_{\mathrm{R},\ell})$ denote the array steering vectors of the transmitter and the receiver, respectively. In this paper we consider a uniform linear array (ULA) [1]. We define the array steering vector as a function of a spatial frequency μ, i.e.,

$$\boldsymbol{a}(\mu) = a_0(\mu)\begin{bmatrix}1 & e^{j\mu} & \cdots & e^{j(M-1)\mu}\end{bmatrix}^\mathrm{T} \in \mathbb{C}^M, \tag{3}$$

where $a_0(\mu)$ denotes the antenna element gain pattern that is identical for all antennas. The transformation between μ and its corresponding spatial angle θ is given by $\mu = -2\pi\sin(\theta)\Delta_\mathrm{d}/\lambda$, where Δ_d denotes the identical antenna spacing and λ is the wavelength of the signal.

We consider an OFDM implementation that is described in [12], i.e., an additional CP, which is known as the window part of the symbol, is added and is pulse shaped to reduce the out of band emission during the transmission. At the receiver both the window part and the CP are discarded and the rectangular receiver filter is implemented by the DFT. Thereby, the pulse shaping function has no influence on the effective channel. Let $T_\mathrm{s} = 1/(N_\mathrm{fft} \cdot \Delta_\mathrm{f})$ represent the sampling period and Δ_f denote the subcarrier spacing. The sampled CTF $\boldsymbol{H}_n$ in (1) is given by

$$\boldsymbol{H}_n = \sum_{\ell=1}^{L} d_\ell \boldsymbol{a}_\mathrm{R}(\mu_{\mathrm{R},\ell})\boldsymbol{a}_\mathrm{T}^\mathrm{T}(\mu_{\mathrm{T},\ell})e^{j(n-1)\mu_{\mathrm{f},\ell}}, \tag{4}$$

where $\mu_{\mathrm{f},\ell} = -2\pi\tau_\ell\Delta_\mathrm{f}$. It is assumed that $\tau_\ell\Delta_\mathrm{f} < 1$, $\forall\ell$.

Our goal is to design $\boldsymbol{W}[m]$, $\boldsymbol{F}[m]$, and $\boldsymbol{s}_n[m]$, $\forall n, m$, such that the channel parameters $\mu_{x,\ell}$, $\forall\ell, x \in \{\mathrm{f}, \mathrm{T}, \mathrm{R}\}$ and d_ℓ, $\forall\ell$ are accurately estimated at the receiver.

II-B. The Proposed Training Procedure

We divide the total N_t training OFDM symbols into κ_t training frames, each consisting of N_T OFDM symbols [2]. That is, $N_\mathrm{t} = \kappa_\mathrm{t} \cdot N_\mathrm{T}$. The analog precoding and decoding matrices remain constant in one training frame whereas they might change in different frames. The same time-frequency pilot patterns and training symbols are used in all frames and on different pilot subcarriers. The accumulated pilot matrix $\boldsymbol{S} = [\boldsymbol{s}[1] \quad \cdots \quad \boldsymbol{s}[N_\mathrm{T}]] \in \mathbb{C}^{N_\mathrm{T}\times N_\mathrm{T}}$ in each frame is a scaled orthogonal matrix, i.e., $\boldsymbol{S}\boldsymbol{S}^\mathrm{H} = \gamma\boldsymbol{I}_{N_\mathrm{T}}$. After stacking N_T consecutive $\boldsymbol{y}_n[m]$ next to each other, multiplying the result by $\boldsymbol{S}^\mathrm{H}$ from the right-hand side, dividing by γ on each subcarrier, and then stacking the vectorized version of the processed signal on top of each other, the obtained signal is expressed as

$$\boldsymbol{y}_\kappa = (\boldsymbol{B}_\mathrm{f} \diamond \boldsymbol{B}_{\mathrm{T},\kappa} \diamond \boldsymbol{B}_{\mathrm{R},\kappa}) \cdot \boldsymbol{d} + \boldsymbol{z}_\kappa, \kappa \in \{1, \cdots, \kappa_\mathrm{t}\} \tag{5}$$

where $\boldsymbol{d} = [d_1 \quad \cdots \quad d_L]$, $\boldsymbol{B}_\mathrm{f} = \boldsymbol{\Phi}^\mathrm{T}\boldsymbol{A}_\mathrm{f} \in \mathbb{C}^{N_\mathrm{f}\times L}$, $\boldsymbol{B}_{\mathrm{T},\kappa} = \boldsymbol{F}_\kappa^\mathrm{T}\boldsymbol{A}_\mathrm{T} \in \mathbb{C}^{N_\mathrm{T}\times L}$, and $\boldsymbol{B}_{\mathrm{R},\kappa} = \boldsymbol{W}_\kappa^\mathrm{H}\boldsymbol{A}_\mathrm{R} \in \mathbb{C}^{N_\mathrm{R}\times L}$, where $\boldsymbol{\Phi} \in \mathbb{R}^{N_\mathrm{fft}\times N_\mathrm{f}}$ denote a column selection matrix with ones on the (n,p)-th elements and zeros otherwise, where $p \in \{1, \cdots, N_\mathrm{f}\}$ and $n = k_p$. The array steering matrices $\boldsymbol{A}_\mathrm{f} \in \mathbb{C}^{N_\mathrm{fft}\times L}$, $\boldsymbol{A}_\mathrm{T} \in \mathbb{C}^{M_\mathrm{T}\times L}$, and $\boldsymbol{A}_\mathrm{R} \in \mathbb{C}^{M_\mathrm{R}\times L}$ contain L array steering vectors on the frequency dimension, the transmit (Tx) dimension, and the receive (Rx) dimension, respectively. The effective noise $\boldsymbol{z}_\kappa$ is computed by $\boldsymbol{z}_\kappa = (\boldsymbol{S}^* \otimes \boldsymbol{W}_\kappa^\mathrm{H}) \cdot \begin{bmatrix}\boldsymbol{z}_{k_1,\kappa}^\mathrm{T}/\gamma & \cdots & \boldsymbol{z}_{k_{N_\mathrm{f}},\kappa}^\mathrm{T}/\gamma\end{bmatrix}^\mathrm{T}$, where $\boldsymbol{z}_{n,\kappa}$ comprises N_T consecutive noise vectors $\boldsymbol{z}_n[m]$ in the κ-the frame.

III. MULTI-STAGE 3-D UNITARY ESPRIT IN DFT BEAMSPACE BASED CHANNEL ESTIMATION

In this section, we discuss the proposed channel estimation methods using the 3-D Unitary ESPRIT algorithm in DFT beamspace. To this end, the matrix $\boldsymbol{W}_\kappa^\mathrm{H}$ comprises N_R consecutive phase shifted rows of a $M_\mathrm{R} \times M_\mathrm{R}$ DFT matrix starting from the $k_\kappa^{(\mathrm{R})}$-th row and $\boldsymbol{F}_\kappa$ comprises N_T consecutive phase shifted columns of a $M_\mathrm{T} \times M_\mathrm{T}$ DFT matrix starting from the $k_\kappa^{(\mathrm{T})}$-th column as defined in [10]. Note that once the estimates of the spatial frequencies $\mu_{x,\ell}$, $\forall x, \ell$, are obtained, the complex gain vector $\boldsymbol{d}$ can be estimated by using a least squares (LS) method, e.g., a LS estimate of $\boldsymbol{d}$ could be computed by $\boldsymbol{d} = (\boldsymbol{B}_\mathrm{f} \diamond \boldsymbol{B}_{\mathrm{T},\kappa} \diamond \boldsymbol{B}_{\mathrm{R},\kappa})^+\boldsymbol{y}_\kappa$ via (5). To estimate the spatial frequencies, we use the 3-D Unitary ESPRIT in DFT beamspace algorithm. Compared to Standard ESPRIT, the Unitary ESPRIT algorithm has a reduced computational complexity due to a real-valued implementation and provides an enhanced performance due to the inherent forward-backward averaging (FBA). However, similarly as the Standard ESPRIT in DFT beamspace algorithm in [9], the Unitary ESPRIT algorithm in DFT beamspace requires prior information of the sectors of interest for the spatial frequencies in order to reduce the number of required training frames. Therefore, we estimate the spatial frequencies and the complex path gains in two stages: **a**) a coarse estimation step, where coarse estimates of (dominant) spatial frequencies and corresponding path gains are obtained. The estimated spatial frequencies are used to define sectors of interest in the DFT beamspace. The term "sector of interest" is inherited from the context of array signal processing [10]. It comes from the fact that a beamforming matrix comprising of consecutive rows of a DFT matrix forms a beam pattern encompassing only a certain range of angular locations, namely, the sector of interest; **b**) a fine estimation step, in which high-resolution estimates of the spatial frequencies are computed by using the 3-D Unitary ESPRIT in DFT beamspace algorithm. In this step, the training is carried out based on the selected channel estimation strategies and the sector(s) of interest, which are defined in the coarse estimation step. The coarse estimation is discussed in Section III-A. The proposed fine channel estimation algorithms to obtain high resolution channel estimates and the corresponding training procedure are described in Sections III-B and III-C, respectively.

III-A. Coarse estimation

Coarse estimation of channel parameters: The coarse estimates can be obtained by using an advanced antenna / hardware architecture, e.g., a multi-band antenna together with a fully digital sub-6 GHz hardware, e.g., [13], or by using a pre-estimation, e.g.,

[1] Our proposed channel estimation algorithms can be used for any centro-symmetric antenna arrays, e.g., they can be extended to uniform rectangular arrays for a joint estimation of azimuth and elevation angles [10].

[2] The use of N_T OFDM symbols is sufficient to compute the inverse of the training matrix $\boldsymbol{S}$ in one frame.

via the on-grid CS based channel estimation method [6]. Note that in the former case the angular spread information in a sub-6 GHz band is different than that in the mmWave band [13]. The latter solution comes at a cost of more training resources. In contrast, we study another possibility, i.e., obtaining an effective ULA antenna system, which operates in the mmWave band but with the same number of active antennas and RF chains. This could be achieved by using switches on antennas or phase shifters. Taking the transmitter as an example, to obtain $M_{\mathrm{T}}^{(\mathrm{eff})} = N_{\mathrm{T}}$ consecutive active antennas at the transmitter, we switch off $(M_{\mathrm{T}} - N_{\mathrm{T}})$ transmit antennas or $(M_{\mathrm{T}} - N_{\mathrm{T}})N_{\mathrm{T}}$ phase shifters. Then the effective aperture of the antenna systems is reduced, but a wide beam is achieved. By using the same method at the receiver, we obtain an effective antenna system with $M_{\mathrm{R}}^{(\mathrm{eff})} = N_{\mathrm{R}}$. Equation (5) turns into

$$\boldsymbol{y}^{(\mathrm{ini})} = (\boldsymbol{B}_{\mathrm{f}} \diamond \boldsymbol{B}_{\mathrm{T}}^{(\mathrm{eff})} \diamond \boldsymbol{B}_{\mathrm{R}}^{(\mathrm{eff})}) \cdot \boldsymbol{d} + \boldsymbol{z}^{(\mathrm{eff})}. \tag{6}$$

Consequently, the square matrices $\boldsymbol{W} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ and $\boldsymbol{F} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ allow us to use the complete DFT beamspace, or the element space when digital precoders and decoders are applied. Note that the index κ is dropped because one traing frame is sufficient to cover the full beamspace.

Nonetheless, since $\boldsymbol{y}^{(\mathrm{ini})}$ is only a vector, to estimate more than $L_{\max} = 2$ scatters by using 3-D Unitary ESPRIT in DFT beamspace, we need to perform spatial smoothing on the frequency dimension to increase the number of snapshots [9]. To this end, we divide N_f virtual sensors in the frequency dimension into L_{f} virtual subarrays, each contains $M_{\mathrm{sub,f}} = N_{\mathrm{f}} - L_{\mathrm{f}} + 1$ elements. After spatial smoothing, we obtain

$$\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{ini})} = \left[\bar{\boldsymbol{J}}_1^{(\mathrm{ss})}\boldsymbol{y}^{(\mathrm{ini})} \quad \cdots \quad \bar{\boldsymbol{J}}_{L_{\mathrm{f}}}^{(\mathrm{ss})}\boldsymbol{y}^{(\mathrm{ini})}\right] \in \mathbb{C}^{M_{\mathrm{sub,f}}N_{\mathrm{R}}N_{\mathrm{T}} \times L_{\mathrm{f}}}, \tag{7}$$

where $\bar{\boldsymbol{J}}_{\ell_{\mathrm{f}}}^{(\mathrm{ss})} = \boldsymbol{J}_{\ell_{\mathrm{f}}} \otimes \boldsymbol{I}_{N_{\mathrm{T}}} \otimes \boldsymbol{I}_{N_{\mathrm{T}}}$ with $\boldsymbol{J}_{\ell_{\mathrm{f}}} = \left[\boldsymbol{0}_{M_{\mathrm{sub,f}} \times (\ell_{\mathrm{f}}-1)} \quad \boldsymbol{I}_{M_{\mathrm{sub,f}}} \quad \boldsymbol{0}_{M_{\mathrm{sub,f}} \times (L_{\mathrm{f}}-\ell_{\mathrm{f}})}\right] \in \mathbb{R}^{M_{\mathrm{sub,f}} \times N_f}$ and $\ell_{\mathrm{f}} \in \{1, \cdots, L_{\mathrm{f}}\}$. Furthermore, in order to obtain the estimate of the real-valued signal subspace, we need to use the following theorem.

Theorem 1. *Let $\boldsymbol{Q}_{M_{\mathrm{sub,f}}} \in \mathbb{C}^{M_{\mathrm{sub,f}} \times M_{\mathrm{sub,f}}}$ and $\boldsymbol{Q}_{2L_{\mathrm{f}}} \in \mathbb{C}^{2L_{\mathrm{f}} \times 2L_{\mathrm{f}}}$ represent the left-$\boldsymbol{\Pi}$-real matrices that are constructed by following Example 3.3 in [14]. Then the following matrix is real-valued*

$$\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{Re})} = \bar{\boldsymbol{Q}}\left[\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{ini})} \quad \bar{\boldsymbol{\Pi}}\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{ini})^*}\boldsymbol{\Pi}_{L_{\mathrm{f}}}\right]\boldsymbol{Q}_{2L_{\mathrm{f}}}, \tag{8}$$

where $\bar{\boldsymbol{Q}} = \boldsymbol{Q}_{M_{\mathrm{sub,f}}} \otimes \boldsymbol{I}_{N_{\mathrm{T}}} \otimes \boldsymbol{I}_{N_{\mathrm{R}}}$ and $\bar{\boldsymbol{\Pi}} = \boldsymbol{\Pi}_{M_{\mathrm{sub,f}}} \otimes \boldsymbol{I}_{N_{\mathrm{T}}} \otimes \boldsymbol{I}_{N_{\mathrm{R}}}$. The signal subspace of $\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{Re})}$ corresponds to the real-valued counterpart of the complex-valued signal subspace of $\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{ini})}$.

Proof: The proof is omitted due to the space limitation. ■

Finally, we apply the 3-D Unitary ESPRIT in DFT beamspace algorithm, which follows a similar procedure as the 3-D Standard ESPRIT in DFT beamspace algorithm in [9]. The major difference is that real-valued invariance equations and real-valued signal subspace are computed in order to achieve a real-valued implementation. More precisely, the real-valued estimate of the signal subspace is obtained by computing the SVD of $\boldsymbol{Y}_{\mathrm{ss}}^{(\mathrm{Re})}$. For the construction of the invariance equations, the selection matrices have the same Kronecker structure as in [15] and the selection matrices in [10] should be used for the Rx and the Tx dimension. Lastly, the shift invariance equations are solved via LS. Afterwards, the spatial frequencies are obtained by using the algorithm in [15] so that an automatic pairing of the spatial frequencies in the different dimensions is achieved. We can also use tensor-based signal processing in [16] to improve the performance. By using the proposed 3-D Unitary ESPRIT in DFT beamspace algorithm, the maximum number of resolvable scatterers in one training frame is given by

$$\begin{aligned} L_{\max} = \min((M_{\mathrm{sub,f}} - 1)N_{\mathrm{T}}N_{\mathrm{R}}, M_{\mathrm{sub,f}} \\ \cdot (N_{\mathrm{T}} - 1)N_{\mathrm{R}}, M_{\mathrm{sub,f}}N_{\mathrm{T}}(N_{\mathrm{R}} - 1), 2L_{\mathrm{f}}). \end{aligned} \tag{9}$$

Sectorization: Coarse estimates of the spatial frequencies are used to determine the sectors of interest to be scanned in the fine estimation step so that high-resolution estimates of the spatial frequencies can be obtained. The sectorization rules used in this paper are two-fold: a). Independent sectorization is used along the Rx and the Tx dimension; b). In the s_{R}-th sector ($s_{\mathrm{R}} \in \{1, \cdots N_{\mathrm{t,R}}\}$) along the Rx dimension and the s_{T}-th sector ($s_{\mathrm{T}} \in \{1, \cdots N_{\mathrm{t,T}}\}$) along the Tx dimension, we make sure that $\max_i \left|\mu_{\mathrm{R},i}^{(s_{\mathrm{R}})} - \frac{(2k^{(s_{\mathrm{R}})}+N_{\mathrm{R}})\pi}{M_{\mathrm{R}}}\right| \leq \zeta_{\mathrm{R}}$ and $\max_j \left|\mu_{\mathrm{T},j}^{(s_{\mathrm{T}})} - \frac{(2k^{(s_{\mathrm{T}})}+N_{\mathrm{T}})\pi}{M_{\mathrm{T}}}\right| \leq \zeta_{\mathrm{T}}$, where $\mu_{\mathrm{R},i}^{(s_{\mathrm{R}})}$ and $\mu_{\mathrm{T},j}^{(s_{\mathrm{T}})}$ denote the sptatial frequencies within the sector of interest, and $k^{(s_{\mathrm{R}})}$ and $k^{(s_{\mathrm{T}})}$ represent the starting row and column of the corresponding sectors, respectively. The scalars ζ_{R} and ζ_{T} denote the maximum offset from the center of the sector. Obviously, when ζ_{R} and ζ_{T} decrease, the number of sectors increases.

III-B. Fine estimation via a joint estimation of all paths

A joint estimation of all the spatial frequencies can be achieved by forming an enlarged effective sector of interest such that there are no spatial frequencies outside of the effective sector of interest. To this end, we use different precoding or decoding matrices in different training frames and should concatenate different $\boldsymbol{y}_\kappa$. Let us define $\bar{\kappa} \in \{1, \cdots, \kappa_{\mathrm{t}}\}$ and $\bar{\kappa} \neq \kappa$. According to the sectorization rule in Section III-A, $\boldsymbol{y}_\kappa$ and $\boldsymbol{y}_{\bar{\kappa}}$ might contain repeated samples due to the scan of overlapping sectors of interest, e.g., when $\boldsymbol{W}_\kappa^{\mathrm{H}}$ and $\boldsymbol{W}_{\bar{\kappa}}^{\mathrm{H}}$ have same rows. Therefore, during the concatenation we will average over repeated samples in $\boldsymbol{y}_\kappa$. At the end we obtain

$$\boldsymbol{y}^{(\mathrm{eff})} = (\boldsymbol{B}_{\mathrm{f}} \diamond \bar{\boldsymbol{B}}_{\mathrm{T}}^{(\mathrm{eff})} \diamond \bar{\boldsymbol{B}}_{\mathrm{R}}^{(\mathrm{eff})}) \cdot \boldsymbol{d} + \boldsymbol{z}^{(\mathrm{eff})}, \tag{10}$$

where $\bar{\boldsymbol{B}}_{\mathrm{T}}^{(\mathrm{eff})} = \boldsymbol{F}^{(\mathrm{eff})^{\mathrm{T}}}\boldsymbol{A}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{T}}^{(\mathrm{eff})} \times L}$, and $\bar{\boldsymbol{B}}_{\mathrm{R}}^{(\mathrm{eff})} = \boldsymbol{W}^{(\mathrm{eff})^{\mathrm{H}}}\boldsymbol{A}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{R}}^{(\mathrm{eff})} \times L}$ with $N_{\mathrm{R}} \leq N_{\mathrm{R}}^{(\mathrm{eff})} \leq N_{\mathrm{t,R}}N_{\mathrm{R}}$ and $N_{\mathrm{T}} \leq N_{\mathrm{T}}^{(\mathrm{eff})} \leq N_{\mathrm{t,T}}N_{\mathrm{T}}$. Let us define $\bar{\boldsymbol{W}} = [\boldsymbol{W}_1 \ \cdots \ \boldsymbol{W}_{N_{\mathrm{t,R}}}] \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{t,R}}N_{\mathrm{R}}}$ and $\bar{\boldsymbol{F}} = [\boldsymbol{F}_1 \ \cdots \ \boldsymbol{F}_{N_{\mathrm{t,R}}}] \in \mathbb{C}^{M_{\mathrm{T}} \times N_{\mathrm{t,T}}N_{\mathrm{T}}}$. Then the matrices $\boldsymbol{W}^{(\mathrm{eff})}$ and $\boldsymbol{F}^{(\mathrm{eff})}$ comprise non-overlapping columns of $\bar{\boldsymbol{W}}$ and $\bar{\boldsymbol{F}}$, respectively. Finally, we can apply the 3-D Unitary ESPRIT in DFT beamspace to obtain spatial frequency estimates with a higher resolution. Note that the Khatri-Rao product in (10) requires the use of $\kappa_{\mathrm{t}} = N_{\mathrm{t,T}} \cdot N_{\mathrm{t,R}}$ training frames in the fine estimation step.

III-C. Fine estimation via a sequential estimation of spatial frequencies in different sectors

Fine estimates of the spatial frequencies can also be obtained by first suppressing the interference caused by frequencies outside of the sector of interest, and then applying the 3-D Unitary ESPRIT in DFT beamspace. There are two methods to suppress the interference, i.e., subtracting the interference and projecting into the null space of the interference. Compared to the joint estimation, the number of training frames and the computational complexity can be significantly reduced by using sector-wise channel estimation because there are fewer channel parameters in each sector. When the sector-wise method is used, we make sure that the spatial frequencies to be estimated in each sector along the Tx dimension and along the Rx dimension correspond the same paths. Thereby, the number of required training frames in the fine estimation step

reduces to $\kappa_{\mathrm{t}} = \max(N_{\mathrm{t,T}}, N_{\mathrm{t,R}})$. Moreover, we have a one-to-one mapping between the κ-th training frame and the κ-th sector.

Successive sector-wise interference cancellation (SsIC) Let $\bar{\boldsymbol{A}}_{\mathrm{f},\kappa}$, $\bar{\boldsymbol{A}}_{\mathrm{T},\kappa}$, and $\bar{\boldsymbol{A}}_{\mathrm{R},\kappa}$ represent estimated array steering matrices, which comprise the steering vectors that do not lie in the κ-th sector. Let $\bar{\boldsymbol{d}}_{\kappa}$ contain the complex gains of the paths that are not in the κ-th sector, e.g., obtained from the coarse estimation step. When the SsIC method is applied, the effective signal to be used for channel estimation in the κ-th training frame is given by

$$\boldsymbol{y}_{\kappa}^{(\mathrm{eff})} = \boldsymbol{y}_{\kappa} - (\boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{F}_{\kappa}^{\mathrm{T}} \otimes \boldsymbol{W}_{\kappa}^{\mathrm{H}})(\bar{\boldsymbol{A}}_{\mathrm{f},\kappa} \diamond \bar{\boldsymbol{A}}_{\mathrm{T},\kappa} \diamond \bar{\boldsymbol{A}}_{\mathrm{T},\kappa}) \cdot \bar{\boldsymbol{d}}_{\kappa}.$$

The estimated spatial frequencies and the corresponding path gains will be used for parameter estimation in the next sector. Therefore, the estimation order will affect the final performance. In our design the estimation order is based on the path gain. Let $\boldsymbol{d}_{\kappa}$ contain complex gains of the paths that are in the κ-th sector. Then our algorithm starts with the sector that contains the largest channel gains, i.e., $\kappa = \arg\max_{\kappa} \|\boldsymbol{d}_{\kappa}\|^2$.

Successive null space projection (SNSP) Let $\mathcal{L}_{\kappa} \subset \{1, \cdots L\}$ be a subset that contains the paths of interest in the κ-th sector, where $\mathcal{L}_{\kappa} \cap \mathcal{L}_{\bar{\kappa}} = \emptyset$ and $\mathcal{L}_1 \cup \cdots \cup \mathcal{L}_{\kappa_{\mathrm{t}}} = \{1, \cdots, L\}$. When the SNSP method is applied, the effective signal to be used for channel estimation is given by

$$\boldsymbol{y}_{\kappa}^{(\mathrm{eff})} = \prod_{j \notin \mathcal{L}_{\kappa}} (\boldsymbol{I}_{N_{\mathrm{f}}} \otimes (\boldsymbol{I}_{N_{\mathrm{T}}} - \boldsymbol{B}_{\mathrm{T},\kappa,j}) \otimes (\boldsymbol{I}_{N_{\mathrm{R}}} - \boldsymbol{B}_{\mathrm{R},\kappa,j}))\boldsymbol{y}_{\kappa},$$

where $\boldsymbol{B}_{\mathrm{T},\kappa,j} = \boldsymbol{F}_{\kappa}^{\mathrm{T}} \boldsymbol{a}_{\mathrm{T}}(\mu_{\mathrm{T},j}) \boldsymbol{a}_{\mathrm{T}}^{\mathrm{H}}(\mu_{\mathrm{T},j}) \boldsymbol{F}_{\kappa}^{*} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ and $\boldsymbol{B}_{\mathrm{R},\kappa,j} = \boldsymbol{W}_{\kappa}^{\mathrm{H}} \boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},j}) \boldsymbol{a}_{\mathrm{R}}^{\mathrm{H}}(\mu_{\mathrm{R},j}) \boldsymbol{W}_{\kappa} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ for $j \in \{1, \cdots, L\}$ and $j \notin \mathcal{L}_{\kappa}$. The same estimation order used for the SsIC algorithm is applied.

IV. SIMULATION RESULTS

The proposed multi-stage 3-D Unitary ESPRIT in DFT beamspace based channel estimation algorithm is evaluated using Monte-Carlo simulations. The transmit power of the pilots P_{pilot} is set to unity and is equally allocated to N_{f} pilot tones in one OFDM symbol. We define the SNR of the pilots as $\mathrm{SNR}_{\mathrm{pilot}} = 1/(N_{\mathrm{f}}\sigma_n^2)$. We set $N_{\mathrm{fft}} = 64$ and $L = 6$. The random complex gain is ZMCSCG distributed with $\sum_{\ell=1}^{L} \mathbb{E}\{|d_{\ell}|^2\} = 1$. In each OFDM symbol $N_{\mathrm{f}} = 25$ consecutive subcarriers, starting from the first non-DC subcarrier are used as pilot subcarriers. The inter-element spacing of the ULA is equal to half of the wavelength and isotropic antenna elements are used. We assume $M_{\mathrm{T}} = M_{\mathrm{R}} = 16$, $N_{\mathrm{T}} = N_{\mathrm{R}} = 4$, and $L_{\mathrm{f}} = 18$ virtual subarrays in the frequency dimension, each with $M_{\mathrm{sub,f}} = 8$ elements. The training matrix $\boldsymbol{S}$ is a scaled DFT matrix with a scaling factor $\gamma = \sqrt{\frac{P_{\mathrm{pilot}}}{N_{\mathrm{T}}^2 N_{\mathrm{f}}}}$ during the coarse estimation stage and $\gamma = \sqrt{\frac{P_{\mathrm{pilot}}}{N_{\mathrm{T}} M_{\mathrm{T}} N_{\mathrm{f}}}}$ during the fine estimation stage. The minimum distance to the center of the sector is $\zeta_{\mathrm{R}} = 2\pi/M_{\mathrm{R}}$ and $\zeta_{\mathrm{T}} = 2\pi/M_{\mathrm{T}}$. We investigate the performance of the estimation algorithm for both spatial frequencies and the channel. For the former one the mean squared estimation error (MSE) is used as the criterion, i.e.,

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

Let us define $\boldsymbol{h}_{\mathrm{vec}} = (\boldsymbol{A}_{\mathrm{f}} \diamond \boldsymbol{A}_{\mathrm{T}} \diamond \boldsymbol{A}_{\mathrm{R}}) \cdot \boldsymbol{d}$. The channel estimation performance is evaluated using the normalized mean squared estimation error (NMSE), i.e.,

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

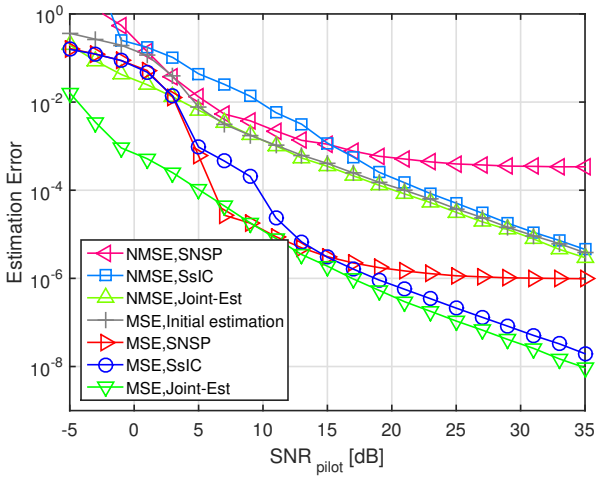

Fig. 1. Estimation error vs. SNR for $M_{\mathrm{T}} = M_{\mathrm{R}} = 16$, $N_{\mathrm{T}} = N_{\mathrm{R}} = 4$, $L = 6$, and $2L_{\mathrm{f}} = 36$ effective snapshots.

The proposed method in Section III-B is denoted as "Joint-Est". The proposed estimation methods in Section III-C, i.e., the SsIC algorithm and the SNSP algorithm, are denoted as "SsIC" and "SNSP", respectively. The 6 tuples of spatial frequencies used in our simulations are $[0.3555, 1.0178, 0.9469]^{\mathrm{T}}$, $[-0.8478, -0.9612, -0.4837]^{\mathrm{T}}$, $[0.8345, 0.0473, -0.5417]^{\mathrm{T}}$, $[0.0125, -1.0325, -0.1955]^{\mathrm{T}}$, $[0.2898, 0.5103, -0.7448]^{\mathrm{T}}$, and $[-1.0225, 0.3584, 0.6906]^{\mathrm{T}}$. The simulation results are obtained by averaging over 500 noise realizations.

Fig. 1 shows that the proposed two-stage 3-D Unitary ESPRIT in DFT beamspace based channel estimation methods provide improved estimates based on coarse estimates. A significant performance gain is obtained when the joint estimation method is used or when the SsIC and the SNSP algorithms are used in the high SNR regime. In general better initial estimates lead to better refined estimates. But the SNSP algorithm performs worse than the other two algorithms in the high SNR regime. This is due to the fact that the steering vectors are correlated. Thus, the orthogonal complements of them are also correlated. Projecting into the null space of the interfering steering sectors will also reduce the effective signal power.

V. CONCLUSION

A gridless two-stage 3-D Unitary ESPRIT in DFT beamspace based channel estimation algorithm has been proposed to estimate the CSI of a hybrid mmWave massive MIMO system. Based on a coarse estimate, the proposed algorithm provides high resolution estimates without scanning the whole DFT beamspace. The required training protocol, analog processing as well as pilot patterns have been discussed and developed. Simulation results show that the proposed channel estimation algorithm provides accurate estimates of both the CSI and the spatial frequencies using only a few training symbols.

ACKNOWLEDGMENTS

The authors acknowledge the financial support by the Carl-Zeiss-Foundation (http://carl-zeiss-stiftung.de/).

VI. REFERENCES

[1] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave

MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 1499–1513, Mar. 2014.
[2] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 831–846, Oct. 2014.
[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. on Acoustics, Speech, and Signal Processing (ICASSP)*, Shanghai, China, Mar. 2016.
[4] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González Prelcic, and R. W. Heath, Jr., "Channel estimation and hybrid combining for mmwave: phase shifters or switches," in *Proc. IEEE Information Theory and Applications Workshop (ITA)*, 2015.
[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, June 2016.
[6] 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. IEEE Int. Workshop on Smart Antennas (WSA)*, Munich, Germany, Mar. 2016.
[7] K. Venugopal, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "Channel estimation for hybrid architecture based wideband millimeter wave systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 35, pp. 1996–2009, Sept. 2017.
[8] D. Zhu, J. Choi, and R. W. Heath, Jr., "Two-dimensional AoD and AoA acquisition for wideband mmWave systems with cross-polarized MIMO," *submitted to IEEE Transactions on Wireless Communications*, pp. 1–31, 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)*, Kos, Greece, Sept. 2017.
[10] 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 Transactions on Signal Processing*, vol. 44, no. 2, pp. 316–328, Feb. 1996.
[11] 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, pp. 1058–1076, June 2010.
[12] R. Prasad, *OFDM for wireless communications systems*, Artech House, Inc., 2004.
[13] A. Ali, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "Millimeter wave beam-selection using out-of-band spatial information," *submitted to IEEE Transactions on Wireless Communications*, pp. 1–30, 2017.
[14] M. Haardt, *Efficient One-, Two-, and Multidimensional High-Resolution Array Signal Processing*, Shaker Verlag, 1997.
[15] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several non-symmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 161–169, Jan. 1998.
[16] 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 Transactions on Signal Processing*, vol. 56, pp. 3198–3213, July 2008.