Channel Estimation and Training Design for Hybrid Multi-carrier MmWave Massive MIMO Systems: The Beamspace ESPRIT Approach

Jianshu Zhang and Martin Haardt
Communications Research Laboratory, Ilmenau University of Technology, Germany
Email: {jianshu.zhang, martin.haardt}@tu-ilmenau.de

***Abstract*—In this paper, we study the channel estimation problem for a cyclic prefix OFDM (CP-OFDM) based millimeter wave (mmWave) hybrid analog-digital MIMO system, where the analog processing is achieved using only phase shift networks. A three-dimensional (3-D) Standard ESPRIT in discrete Fourier transform (DFT) beamspace approach is developed to estimate the unknown frequency-selective channel. The required training protocol, analog precoding and decoding matrices, as well as pilot patterns are also discussed. Simulation results show that the proposed 3-D Standard ESPRIT in DFT beamspace based channel estimation algorithm provides accurate channel estimates when there is a sufficient number of snapshots.**

Keywords*—MmWave Massive MIMO, hybrid precoding and decoding, multi-dimensional harmonic retrieval, Standard ESPRIT in DFT beamspace.

I. Introduction

Millimeter wave (mmWave) massive MIMO techniques will not only gain from large chunks of underutilized spectrum in the mmWave band [1] but will also benefit from a significantly reduced form factor of the massive MIMO array [2]. However, one RF chain per antenna is impractical because the involved power consumption and the hardware cost are too high. To exploit the MIMO gain under a reasonable cost, one promising solution is to deploy hybrid analog-digital precoding schemes, realized using phase shifters or switches in the RF domain [3], and digital precoding schemes, implemented in the digital baseband domain as in conventional MIMO. If analog precoding is achieved using phase shifters only, the analog precoding matrix should have only constant modulus entries [4], [5], [6]. Furthermore, when a wideband multi-carrier system is considered, we have to use the same phase shifts for all subcarriers [7]. These two constraints are stringent so that they lead to significant challenges not only for the precoding of the transmitted data but also for the required channel estimation tasks [4], [5], [8]. When a frequency-flat channel is considered, in [5] an adaptive compressed sensing (CS) based channel estimation algorithm is proposed to estimate the channel of a hybrid analog-digital massive MIMO system. This CS based channel estimation algorithm has been further extended in [9] by involving multiple measurement vectors (MMV) to improve the channel estimation accuracy. An adaptive multi-grid sparse recovery approach is applied in [10]. For a frequency-selective channel we have proposed multi-stage CS based channel estimation methods in [11] to reduce the involved computational complexity. Unfortunately, all the above CS based methods depend on the on-grid assumption of the channel parameters, which requires a grid-offset estimation for practical use. The grid-offset estimation itself is already a challenging task. Although in [12] a gridless CS mmWave channel estimation algorithm is developed based on the gridless CS theory [13], the proposed algorithm is not aware of the hardware constraint in a hybrid MIMO architecture and is only suitable for a one-dimensional estimation problem, e.g., single-antenna terminals. Hence, this motivates us to seek a gridless mmWave channel estimation method for hybrid mmWave massive MIMO systems.

In this paper we develop a gridless channel estimation algorithm for a hybrid point-to-point mmWave multi-carrier massive MIMO system. A cyclic prefix OFDM (CP-OFDM) based multi-carrier modulation scheme is used and training using pilot tones is considered. The resulting channel estimation problem is formulated as a three-dimensional (3-D) harmonic retrieval problem with compressed measurements. Inspired by the Unitary ESPRIT in DFT beamspace algorithm in [14], we develop a 3-D standard ESPRIT in DFT beamspace based channel estimation algorithm. When combined with our proposed training design, which involves the design of a training protocol, the analog precoding and decoding matrices, and the pilot patterns, the proposed algorithm provides high-resolution estimates of the spatial frequencies and subsequently accurate channel estimates without the on-grid assumption. Only a few training symbols are used during the estimation. Simulation results also show that by using the proposed channel estimation algorithm the achievable system sum rate is only slightly affected in the low SNR regime.

Notation: Upper-case and lower-case bold-faced letters denote matrices and vectors, respectively. The expectation, transpose, conjugate, Hermitian transpose, inverse, and Moore-Penrose pseudo inverse are denoted by $\mathbb{E}\{\cdot\}$, $\{\cdot\}^{\mathrm{T}}$, $\{\cdot\}^{*}$, $\{\cdot\}^{\mathrm{H}}$, $\{\cdot\}^{-1}$, 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 the 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

A. System Model

We study a point-to-point mmWave massive MIMO system as in [11], where a base station (BS) transmits data

ISBN 978-0-9928626-7-1 © EURASIP 2017

to a multi-antenna user equipment (UE). The BS has M_{T} transmit antennas and N_{T} RF chains. The UE has M_{R} receive antennas and N_{R} RF chains. The number of RF chains is assumed to be much smaller than the number of antenna elements, i.e., $M_{\mathrm{T}} \gg N_{\mathrm{T}}$ 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}. Let $\boldsymbol{s}_n[m] \in \mathbb{C}^{N_{\mathrm{T}}}$ represent the transmitted pilot vector on the n-th pilot subcarrier in the m-th OFDM symbol over N_{T} RF chains ($n \in \{k_1, \cdots, k_{N_{\mathrm{f}}}\} \subset \{1, \cdots, N_{\mathrm{fft}}\}$, $p \in \{1, \cdots, N_{\mathrm{f}}\}$, $m \in \{1, \cdots, N_{\mathrm{t}}\}$). Thereby, the training procedure consists of N_{t} OFDM symbols each with N_{f} pilot tones. The pilot tones and the data tones are interleaved on all subcarriers and then passed through the IFFT filter. A CP of length N_{CP} symbols is added, followed by an RF precoder $\boldsymbol{F}[m] \in \mathbb{C}^{M_{\mathrm{T}} \times N_{\mathrm{T}}}$ using analog circuitry. The transmit power of the training vectors satisfies $\sum_{n=k_1}^{k_{N_{\mathrm{f}}}} \|\boldsymbol{F}[m]\boldsymbol{s}_n[m]\|^2 = P_{\mathrm{pilot}}$ for all m.

We consider a frequency selective quasi-static block fading channel. Assume that N_{CP} has the same length as the maximum excess delay of the channel such that the inter-symbol interference is avoided. After passing through the channel, first, an RF decoder $\boldsymbol{W}^{\mathrm{H}}[m] \in \mathbb{C}^{N_{\mathrm{R}} \times M_{\mathrm{R}}}$ is used at the UE. Afterwards, the CP is removed from the received signal. By using the FFT filter the time domain signal is transformed into the frequency domain. Let $\boldsymbol{H}_n[m] \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}}}$ denote the discrete channel transfer function (CTF) on the n-th pilot subcarrier in the m-th OFDM symbol of the UE. The received training signal on the n-th pilot subcarrier in the m-th OFDM symbol is given by [11]

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

where $\boldsymbol{z}_n[m]$ 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 and m. We assume that the RF precoder and decoder are implemented using a network of phase shifters. Hence, the matrices $\boldsymbol{F}[m]$ and $\boldsymbol{W}[m]$ contain only unit modulus entries.

Our goal is to design $\boldsymbol{W}[m]$, $\boldsymbol{F}[m]$, and $\boldsymbol{s}_n[m]$, $\forall n, m$, such that the channel can be accurately estimated at the receiver.

B. Channel Model

We consider uniform linear arrays (ULAs) at both ends. Therefore, the array steering vector is a Vandermonde vector. We define an M-element Vandermonde vector $\boldsymbol{a}(\mu)$ as

$$\boldsymbol{a}(\mu) = \begin{bmatrix} 1 & e^{j\mu} & \cdots & e^{j(M-1)\mu} \end{bmatrix}^{\mathrm{T}}. \tag{1}$$

Due to the lack of diffraction, a mmWave massive MIMO channel is modeled using a finite number of scatterers, e.g., L scatterers [15]. Each scatterer contributes to a single propagation path between the BS and the UE, which accounts for one time delay τ_ℓ and one pair of spatial frequencies $(\mu_{\mathrm{T},\ell}, \mu_{\mathrm{R},\ell})$ for $\ell \in \{0, \cdots, L-1\}$. The frequency domain representation of the channel is given by [16]

$$\boldsymbol{H}(f) = \sum_{\ell=0}^{L-1} \alpha_\ell \boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},\ell})\boldsymbol{a}_{\mathrm{T}}^{\mathrm{T}}(\mu_{\mathrm{T},\ell})e^{-j2\pi\tau_\ell f}, \tag{2}$$

where α_ℓ denote the complex gain of the ℓ-th path. The vectors $\boldsymbol{a}_{\mathrm{T}}(\mu_{\mathrm{T},\ell}) \in \mathbb{C}^{M_{\mathrm{T}}}$ and $\boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},\ell}) \in \mathbb{C}^{M_{\mathrm{R}}}$ denote the array steering vectors of the BS and the UE, respectively.

Let $T_{\mathrm{s}} = 1/(N_{\mathrm{fft}} \cdot \Delta_{\mathrm{f}})$ represent the sampling period and Δ_{f} denote the subcarrier spacing. We model the sampled CTF on the n-th subcarrier in the m-th OFDM symbol as

$$\boldsymbol{H}_n[m] = \sum_{\ell=0}^{L-1} \alpha_\ell[m]\boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},\ell})\boldsymbol{a}_{\mathrm{T}}^{\mathrm{T}}(\mu_{\mathrm{T},\ell})e^{-j2\pi\tau_\ell(n-1)\Delta_{\mathrm{f}}}, \tag{3}$$

Note that equation (3) implies that the angular-delay profile remains unchanged while the complex channel gain might vary during different OFDM symbol periods.

C. Proposed Training Protocol

In our training protocol we divide the total number of N_{t} training OFDM symbols into κ_{m} training frames, each consisting of N_{T} OFDM symbols. That is, $N_{\mathrm{t}} = \kappa_{\mathrm{m}} \cdot N_{\mathrm{T}}$. We assume that the channel gains $\alpha_\ell[m]$, $\forall \ell$ are approximately the same in each training frame whereas they are different in different training frames, e.g., [12]. Under this assumption, we will use the notation $\alpha_{\kappa,\ell}$ instead of $\alpha_\ell[m]$ in the rest of the paper. Furthermore, the same analog precoding and decoding matrices are used in all κ_{m} training frames. Thus, the indices m of $\boldsymbol{F}[m]$ and $\boldsymbol{W}[m]$ are dropped. The training vector $\boldsymbol{s}_n[m] \in \mathbb{C}^{N_{\mathrm{T}}}$ in each frame is designed such that $\boldsymbol{S}_n = [\boldsymbol{s}_n[1] \quad \cdots \quad \boldsymbol{s}_n[N_{\mathrm{T}}]] \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ is a scaled orthogonal matrix for all n. Note that more than N_{T} OFDM symbols in one frame can achieve the same goal, but at a cost of wasting more training resources. The same training matrix $\boldsymbol{S}_n$ is used in different frames. Therefore, by pre-multiplying $\boldsymbol{S}_n^{\mathrm{H}}$ from the right-hand side of the received signal matrix of N_{T} consecutive OFDM symbols in each frame we obtain

$$\boldsymbol{Y}_{n,\kappa} = \boldsymbol{W}^{\mathrm{H}}\boldsymbol{H}_{n,\kappa}\boldsymbol{F} + \bar{\boldsymbol{Z}}_{n,\kappa},$$

where $\kappa \in \{1, \cdots, \kappa_{\mathrm{m}}\}$ and $\bar{\boldsymbol{Z}}_{n,\kappa} = \boldsymbol{W}^{\mathrm{H}}\boldsymbol{Z}_{n,\kappa}\boldsymbol{S}_n^{\mathrm{H}}$ denotes the effective noise, where $\boldsymbol{Z}_{n,\kappa} \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{T}}}$ comprises N_{T} consecutive $\boldsymbol{z}_n[m]$ in the κ-th frame. By vectorizing $\boldsymbol{Y}_{n,\kappa}$ and stacking them on top of each other for $n \in \{k_1, \cdots, k_{N_{\mathrm{f}}}\}$ we obtain

$$\boldsymbol{y}_\kappa = (\boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{F}^{\mathrm{T}} \otimes \boldsymbol{W}^{\mathrm{H}}) \cdot (\boldsymbol{A}_{\mathrm{f}} \diamond \boldsymbol{A}_{\mathrm{T}} \diamond \boldsymbol{A}_{\mathrm{R}}) \cdot \boldsymbol{\alpha}_\kappa + \bar{\boldsymbol{z}}_\kappa, \tag{4}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{N_{\mathrm{fft}} \times N_{\mathrm{f}}}$ is a column selection matrix with ones on the (k_p, p)-th elements and zeros otherwise, and

$$\begin{aligned}
\boldsymbol{A}_{\mathrm{R}} &= [\boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},0}) \quad \cdots \quad \boldsymbol{a}_{\mathrm{R}}(\mu_{\mathrm{R},L-1})] \in \mathbb{C}^{M_{\mathrm{R}} \times L} \\
\boldsymbol{A}_{\mathrm{T}} &= [\boldsymbol{a}_{\mathrm{T}}(\mu_{\mathrm{T},0}) \quad \cdots \quad \boldsymbol{a}_{\mathrm{T}}(\mu_{\mathrm{T},L-1})] \in \mathbb{C}^{M_{\mathrm{T}} \times L} \\
\boldsymbol{A}_{\mathrm{f}} &= [\boldsymbol{a}_{\mathrm{f}}(\mu_{\mathrm{f},0}) \quad \cdots \quad \boldsymbol{a}_{\mathrm{f}}(\mu_{\mathrm{f},L-1})] \in \mathbb{C}^{N_{\mathrm{fft}} \times L} \\
\boldsymbol{\alpha}_\kappa &= [\alpha_{\kappa,0} \quad \cdots \quad \alpha_{\kappa,L-1}]^{\mathrm{T}} \in \mathbb{C}^{L}
\end{aligned}$$

with $\mu_{\mathrm{f},\ell} = -2\pi\tau_\ell\Delta f$. The vector $\bar{\boldsymbol{z}}_\kappa$ denotes the vectorized version of the effective noise matrix $\begin{bmatrix} \bar{\boldsymbol{Z}}_{k_1,\kappa} & \cdots & \bar{\boldsymbol{Z}}_{k_{N_{\mathrm{f}}},\kappa} \end{bmatrix}$. Borrowing the array signal processing terminology [17], we see equation (4) corresponds to a single snapshot model.

We increase the number of snapshots by placing the $\boldsymbol{y}_\kappa$ next to each other, $\forall \kappa$. Finally, we get

$$\bar{\boldsymbol{Y}} = (\boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{F}^{\mathrm{T}} \otimes \boldsymbol{W}^{\mathrm{H}}) \cdot (\boldsymbol{A}_{\mathrm{f}} \diamond \boldsymbol{A}_{\mathrm{T}} \diamond \boldsymbol{A}_{\mathrm{R}}) \cdot \boldsymbol{\Gamma} + \bar{\boldsymbol{Z}}, \tag{5}$$

ISBN 978-0-9928626-7-1 © EURASIP 2017

where $\boldsymbol{\Gamma} = [\boldsymbol{\alpha}_1 \ \cdots \ \boldsymbol{\alpha}_{\kappa_\mathrm{m}}] \in \mathbb{C}^{L\times\kappa_\mathrm{m}}$ and $\bar{\boldsymbol{Z}} = [\bar{\boldsymbol{z}}_1 \ \cdots \ \bar{\boldsymbol{z}}_{\kappa_\mathrm{m}}] \in \mathbb{C}^{N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}\times\kappa_\mathrm{m}}$. We set $\boldsymbol{B}_\mathrm{R} = \boldsymbol{F}^\mathrm{T}\boldsymbol{A}_\mathrm{R} \in \mathbb{C}^{N_\mathrm{R}\times L}$, $\boldsymbol{B}_\mathrm{T} = \boldsymbol{W}^\mathrm{H}\boldsymbol{A}_\mathrm{T} \in \mathbb{C}^{N_\mathrm{T}\times L}$, and $\boldsymbol{B}_\mathrm{f} = \boldsymbol{\Phi}^\mathrm{T}\boldsymbol{A}_\mathrm{f} \in \mathbb{C}^{N_\mathrm{f}\times L}$. Then equation (5) is rewritten as

$$\bar{\boldsymbol{Y}} = \boldsymbol{B}\cdot\boldsymbol{\Gamma} + \bar{\boldsymbol{Z}} \in \mathbb{C}^{N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}\times\kappa_\mathrm{m}}, \tag{6}$$

where $\boldsymbol{B} = \boldsymbol{B}_\mathrm{f} \diamond \boldsymbol{B}_\mathrm{T} \diamond \boldsymbol{B}_\mathrm{R} \in \mathbb{C}^{N_\mathrm{f}N_\mathrm{t}N_\mathrm{R}\times L}$ denotes a virtual array steering matrix and equation (6) corresponds to a κ_m-snapshot model.

III. Standard ESPRIT in Beamspace Based Channel Estimation

Let $\boldsymbol{\mu}_\ell = [\mu_{\mathrm{R},\ell}, \mu_{\mathrm{T},\ell}, \mu_{\mathrm{f},\ell}]^\mathrm{T} \in \mathbb{R}^3$ denote a tuple of 3-D frequencies. Using the reformulation in (6) the channel estimation task can be accomplished by firstly estimating the 3-D frequency vector $\boldsymbol{\mu}_\ell$ and then the complex gain matrix $\boldsymbol{\Gamma}$, e.g., the LS estimate of $\boldsymbol{\Gamma}$ is given by $\boldsymbol{\Gamma} = \boldsymbol{B}^+\bar{\boldsymbol{Y}}$. Without the compressed measurements in (6), the spatial frequency estimation problem yields a 3-D harmonic retrieval problem as in [17] problem and thus can be solved using the algorithm in the same paper. Nevertheless, inspired by the Unitary ESPRIT in DFT beamspace algorithm in [14], we develop a 3-D Standard ESPRIT in DFT beamspace to estimate spatial frequency vectors $\boldsymbol{\mu}_\ell$, $\forall\ell$. Similarly as other ESPRIT-type algorithms, e.g., [17], the involved two major steps in our algorithm are the estimation of signal subspace and solving shift invariance equations. More precisely, we aim at designing $\boldsymbol{\Phi}$, $\boldsymbol{F}$ and $\boldsymbol{W}$ as well as selection matrices $\boldsymbol{J}_{x,i} \in \mathbb{C}^{M_{\mathrm{sel},\mathrm{x}}\times N_x}$ for $i = 1, 2$ and $x \in \{\mathrm{f}, \mathrm{R}, \mathrm{T}\}$ such that the following three shift invariance equations exist

$$\boldsymbol{J}_{r,1}\boldsymbol{B}\boldsymbol{\Phi}_x = \boldsymbol{J}_{r,2}\boldsymbol{B}, r = 1, 2, 3, x \in \{\mathrm{f}, \mathrm{R}, \mathrm{T}\}. \tag{7}$$

The spatial frequencies to be estimated are contained in the diagonal matrices $\boldsymbol{\Phi}_x = \mathrm{diag}\left\{\begin{bmatrix} e^{j\mu_{x,0}} & \cdots & e^{j\mu_{x,L-1}}\end{bmatrix}\right\}$, $\forall x$. Similarly as in [17], the compound selection matrices $\boldsymbol{J}_{r,i}$, $\forall r, i$, have Kronecker structures and are expressed as

$$\begin{aligned}\boldsymbol{J}_{1,i} &= \boldsymbol{I}_{N_\mathrm{f}} \otimes \boldsymbol{I}_{N_\mathrm{t}} \otimes \boldsymbol{J}_{\mathrm{R},i},\\ \boldsymbol{J}_{2,i} &= \boldsymbol{I}_{N_\mathrm{f}} \otimes \boldsymbol{J}_{\mathrm{T},i} \otimes \boldsymbol{I}_{N_\mathrm{R}},\\ \boldsymbol{J}_{3,i} &= \boldsymbol{J}_{\mathrm{f},i} \otimes \boldsymbol{I}_{N_\mathrm{T}} \otimes \boldsymbol{I}_{N_\mathrm{R}}.\end{aligned} \tag{8}$$

After the shift invariance equations are created, we compute an orthonormal basis of the column space of $\boldsymbol{B}$, which is denoted by $\boldsymbol{U}_\mathrm{s} \in \mathbb{C}^{N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}\times L}$ from the measurement matrix $\bar{\boldsymbol{Y}}$. Then we replace the unknown matrix $\boldsymbol{B}$ in (7) by $\boldsymbol{U}_\mathrm{s}$, which yields

$$\boldsymbol{J}_{r,1}\boldsymbol{U}_\mathrm{s}\boldsymbol{\Psi}_x \approx \boldsymbol{J}_{r,2}\boldsymbol{U}_\mathrm{s}, , r = 1, 2, 3, x \in \{\mathrm{f}, \mathrm{R}, \mathrm{T}\}, \tag{9}$$

where an eigendecomposition satisfies $\boldsymbol{\Psi}_x = \boldsymbol{T}^{-1}\boldsymbol{\Phi}_x\boldsymbol{T}$, $\forall x$. Consequently, the eigenvalues of $\boldsymbol{\Psi}_x \approx (\boldsymbol{J}_{r,1}\boldsymbol{U}_\mathrm{s})^+\boldsymbol{J}_{r,2}\boldsymbol{U}_\mathrm{s}$ are estimates of $e^{j\mu_{x,\ell}}$, $\forall x, \ell$. In this way the spatial frequency vectors $\boldsymbol{\mu}_\ell$, $\forall\ell$ can be computed using the algorithm in [17] and an automatic pairing of the frequency estimates is achieved. In Section III-A we discuss how to build the shift invariance equations. In Section III-B we explain how to obtain the orthonormal basis $\boldsymbol{U}_\mathrm{s}$.

A. Building Shift Invariance Equations

Thanks to the Kronecker structure in (8), we can design $\boldsymbol{J}_{x,i}$ for all dimensions separately such that

$$\boldsymbol{J}_{x,1}\boldsymbol{B}_x\boldsymbol{\Phi}_x = \boldsymbol{J}_{x,2}\boldsymbol{B}_x, x \in \{\mathrm{f}, \mathrm{R}, \mathrm{T}\}. \tag{10}$$

When $x = \mathrm{f}$, equation (10) can be achieved by setting $k_p = p$, i.e., $\boldsymbol{\Phi}$ contains the first N_f columns of $\boldsymbol{I}_{N_\mathrm{fft}}$. Then $\boldsymbol{B}_\mathrm{f}$ becomes a reduced-dimensional version of $\boldsymbol{A}_\mathrm{f}$. We can use $\boldsymbol{J}_{\mathrm{f},1} = \begin{bmatrix}\boldsymbol{I}_{N_\mathrm{f}-1} & \boldsymbol{0}_{(N_\mathrm{f}-1)\times 1}\end{bmatrix}$ and $\boldsymbol{J}_{\mathrm{f},2} = \begin{bmatrix}\boldsymbol{0}_{(N_\mathrm{f}-1)\times 1} & \boldsymbol{I}_{N_\mathrm{f}-1}\end{bmatrix}$, when a maximum overlap is used [17].

The same design cannot be applied when $x = \mathrm{R}, \mathrm{T}$ because $\boldsymbol{F}$ and $\boldsymbol{W}$ must have constant modulus entries. Inspired by the Unitary ESPRIT in DFT beamspace in [14], we propose to construct $\boldsymbol{F}^\mathrm{T}$ and $\boldsymbol{W}^\mathrm{H}$ using N_T and N_R successive rows of $M_\mathrm{T} \times M_\mathrm{T}$ and $M_\mathrm{R} \times M_\mathrm{R}$ DFT matrices, respectively. By applying a similar manipulation as in [14] we have the subsequent statement.

Lemma 1. *When $x = \mathrm{T}, \mathrm{R}$ and $i = 1, 2$, the shift invariance equation* (10) *is fulfilled if $\boldsymbol{J}_{x,i} \in \mathbb{C}^{(N_x-1)\times N_x}$, $\forall x, i$, is a sub-matrix of $\boldsymbol{G}_{x,i}$. Let $1 \le \gamma_{\min,x} \le M_x$ denote the starting index. Then $\boldsymbol{J}_{x,i}$ comprises $N_x - 1$ consecutive rows and N_x columns of $\boldsymbol{G}_{x,i}$ starting from $\gamma_{\min,x}$, $\forall x, i$. The matrices $\boldsymbol{G}_{x,i} \in \mathbb{C}^{M_x\times M_x}$ are defined by*

$$\boldsymbol{G}_{x,1} = \begin{bmatrix} 1 & -e^{-j\frac{2\pi}{M_x}} & \cdots & 0 & 0 & 0\\ 0 & 1 & -e^{-j\frac{2\pi}{M_x}} & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 1 & -e^{-j\frac{2\pi}{M_x}}\\ 1 & 0 & 0 & \cdots & 0 & -e^{j\frac{2\pi}{M_x}}\end{bmatrix}$$

and

$$\boldsymbol{G}_{x,2} = \begin{bmatrix} 1 & 1 & \cdots & 0 & 0 & 0\\ 0 & e^{j\frac{2\pi}{M_x}} & -e^{j\frac{2\pi}{M_x}} & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & e^{j\frac{(M_x-2)2\pi}{M_x}} & -e^{j\frac{(M_x-2)2\pi}{M_x}}\\ 1 & 0 & 0 & \cdots & 0 & e^{j\frac{(M_x-1)2\pi}{M_x}}\end{bmatrix}$$

Proof: The proof is straightforward by following [14]. ■

When $N_x < M_x$ for $x = \mathrm{T}, \mathrm{R}$, the scope of the search, i.e., the spatial sector specified by $\mu_{x,\ell}$, will be narrowed down because each row of $\boldsymbol{J}_{x,1}$ and $\boldsymbol{J}_{x,2}$ relates to two successive components of the DFT beamspace array steering vectors. Following the discussion in [18], it can be derived that the steering angles should fall in the region

$$\mu_{x,\ell} \in [\gamma_{\min,x} - 1, \gamma_{\min,x} + N_x - 2]\cdot\frac{2\pi}{M_x}, \forall\ell.$$

B. Estimation of the signal subspace

When $\kappa_\mathrm{m} \ge L$, by directly applying the SVD on $\bar{\boldsymbol{Y}}$ we can construct $\boldsymbol{U}_\mathrm{s}$, i.e., the orthonormal basis of $\boldsymbol{B}$, by using the first L left singular vectors of $\bar{\boldsymbol{Y}}$. However, a large κ_m also implies a large number of training frames. To reduce the number of training symbols, we propose to use the spatial smoothing algorithm in [17] to extend the number of snapshots. The spatial smoothing technique can be applied over three dimensions separately. It increases the number of snapshots by reducing the dimension of effective measurements. When a DFT beamspace is considered, i.e., $x = \mathrm{R}, \mathrm{T}$, this also leads to a reduced effective aperture of the spatial domain. Therefore, we only apply the spatial smoothing along the frequency domain.

ISBN 978-0-9928626-7-1 © EURASIP 2017

Similarly as described in [17], let us divide N_f *virtual* sensors in the frequency domain into L_f subarrays, each contains $M_{\mathrm{sub,f}} = N_\mathrm{f} - L_\mathrm{f} + 1$ elements. We define selection matrices $\boldsymbol{J}_{\ell_\mathrm{f},1,1} = \bar{\boldsymbol{J}}_{\ell_\mathrm{f}}^{(N_\mathrm{f})} \otimes \boldsymbol{I}_{N_\mathrm{T}} \otimes \boldsymbol{I}_{N_\mathrm{R}}$, where $\ell_\mathrm{f} \in \{1, \cdots, L_\mathrm{f}\}$ and $\bar{\boldsymbol{J}}_{\ell_\mathrm{f}}^{(N_\mathrm{f})} = \begin{bmatrix} \boldsymbol{0}_{M_{\mathrm{sub,f}} \times (\ell_\mathrm{f}-1)} & \boldsymbol{I}_{M_{\mathrm{sub,f}}} & \boldsymbol{0}_{M_{\mathrm{sub,f}} \times (L_\mathrm{f}-\ell_\mathrm{f})} \end{bmatrix} \in \mathbb{C}^{M_{\mathrm{sub,f}} \times N_\mathrm{f}}$. Using the definitions above we express the final spatially smoothed data matrix as

$$\bar{\boldsymbol{Y}}_{\mathrm{ss}} = \begin{bmatrix} \boldsymbol{J}_{1,1,1}\bar{\boldsymbol{Y}} & \cdots & \boldsymbol{J}_{L_\mathrm{f},1,1}\bar{\boldsymbol{Y}} \end{bmatrix} \in \mathbb{C}^{M_\mathrm{sub} \times L_\mathrm{v}} \tag{11}$$

where $M_\mathrm{sub} = M_{\mathrm{sub,f}} N_\mathrm{T} N_\mathrm{R}$ and $L_\mathrm{v} = L_\mathrm{f} \kappa_\mathrm{m}$.

Lastly, we compute the SVD of $\bar{\boldsymbol{Y}}_{\mathrm{ss}}$ to obtain a new orthonormal basis $\boldsymbol{U}_\mathrm{s} \in \mathbb{C}^{M_\mathrm{sub} \times L}$, which consists of the first L left singular vectors of $\boldsymbol{J}_{1,1,1}\bar{\boldsymbol{Y}}$. The selection matrix $\boldsymbol{J}_{\mathrm{f},i}$ in (8) should be adjusted to yield a dimension of $(M_{\mathrm{sub,f}} - 1) \times M_{\mathrm{sub,f}}$ for $i = 1, 2$, when a maximum overlap is considered. By using the proposed 3-D Standard ESPRIT in DFT beamspace algorithm, the maximum number of resolvable scatterers is

$$\begin{aligned} L_\mathrm{max} = \min((M_{\mathrm{sub,f}} - 1)N_\mathrm{T}N_\mathrm{R}, M_{\mathrm{sub,f}} \\ \cdot (N_\mathrm{T} - 1)N_\mathrm{R}, M_{\mathrm{sub,f}}N_\mathrm{T}(N_\mathrm{R} - 1), L_\mathrm{v}). \end{aligned} \tag{12}$$

Remark 1. *Note that it is also possible to extend other high-resolution line-spectra estimation tools based on our DFT beamspace principle, e.g., the MUSIC algorithm [19]. However, compared to ESPRIT-type algorithms, the MUSIC algorithm yields a 3-D search and thus is computationally much more expensive.*

IV. Simulation Results

The proposed 3-D Standard ESPRIT in DFT beamspace based channel estimation algorithm is evaluated using Monte-Carlo simulations. A total transmit power over all subcarriers in one OFDM symbol is set to unity. The SNR is defined as $\mathrm{SNR} = 1/(N_\mathrm{fft}\sigma_n^2)$. The overall transmit power of the training vectors is defined as $P_\mathrm{pilot} = \frac{N_\mathrm{f}}{N_\mathrm{fft}}$ and it is equally allocated to N_f pilot tones in one OFDM symbol. We set $N_\mathrm{fft} = 64$ and $L = 2$. The channel gain α_ℓ, $\forall \ell$, is ZMCSCG distributed with $\sum_{\ell=1}^{L} \mathbb{E}\{|\alpha_\ell|^2\} = 1$. Two training frames are used, i.e., $\kappa_\mathrm{m} = 2$. When the sum rate of the system is evaluated, we consider only $N_\mathrm{data} = 2$ active data subcarriers to be computationally efficient. The inter-element spacing of the ULA is equal to half of the wavelength. We set $M_\mathrm{T} = 64$, $M_\mathrm{R} = 32$, $N_\mathrm{T} = 8$, $N_\mathrm{R} = 4$, and $N_\mathrm{f} = 8$ in the simulations. According to the discussion in Section III-A, this setup corresponds to a search region of approximately $\pi/4$ for both $\mu_{\mathrm{R},\ell}$ and $\mu_{\mathrm{T},\ell}$, $\forall \ell$. Therefore, in our simulations we set $\mu_{x,\ell} \in (0, \pi/4)$ for $x = \mathrm{R}, \mathrm{T}, \forall \ell$. 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.,

$$\mathrm{MSE} = \frac{1}{3L}\mathbb{E}\left\{\sum_{r=1}^{3}\sum_{\ell=0}^{L-1}(\mu_{x(r),\ell} - \hat{\mu}_{x(r),\ell})^2\right\}.$$

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

$$\mathrm{NMSE} = \frac{1}{\kappa_\mathrm{m}}\mathbb{E}\left\{\sum_{\kappa=1}^{\kappa_\mathrm{m}}\frac{\|\boldsymbol{h}_\kappa - \hat{\boldsymbol{h}}_\kappa\|^2}{\|\boldsymbol{h}_\kappa\|^2}\right\}.$$

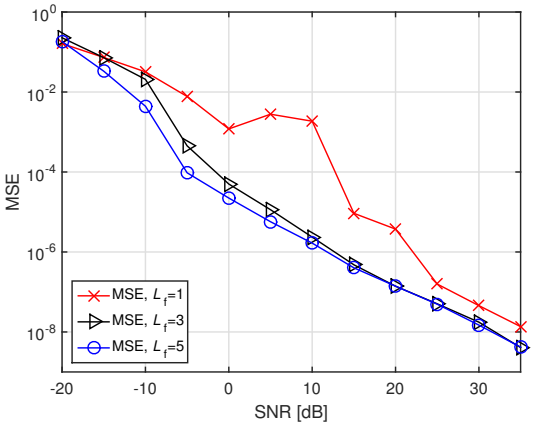

Figure 1. MSE of the spatial frequencies vs. SNR for different number of virtual subarrays in the frequency domain.

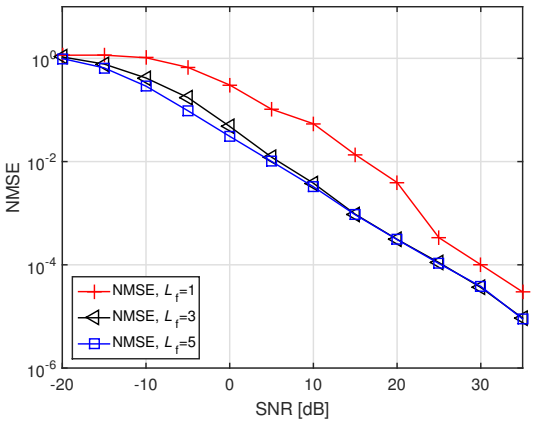

Figure 2. NMSE of the channel vs. SNR for different number of virtual subarrays in the frequency domain.

The proposed 3-D Standard ESPRIT in DFT beamspace algorithm is denoted as "B-ESPRIT". The simulation results are obtained by averaging over 500 channel realizations.

As depicted in Figures 1 and 2 both the MSE and the NMSE reduce when the number of virtual subarrays in the frequency domain increases. Moreover, when the number of virtual subarrays is fixed, the 3-D Standard ESPRIT in DFT beamspace algorithm provides more accurate estimates of the spatial frequencies than those of the channel. This advantage might be exploited when the hybrid analog-digital design is based on the spatial frequencies, e.g., [4]. Note that when $L_\mathrm{f} = 1$, i.e., the number of effective snapshots is equal to that of spatial frequencies, the performance curve of the proposed algorithm is not smooth. This phenomenon does not appear when the number of effective snapshots increases. This implies that to achieve a better performance we should have more snapshots than the number of spatial frequencies.

To evaluate the achievable sum rate of the system with estimated channel state information (CSI), we apply the channel matching based hybrid design as described in [11] and [6]. Since our focus is on the effects of imperfect CSI rather than the overall system spectral efficiency, the demonstrated sum rate in Fig. 3 accounts only for the data transmission

ISBN 978-0-9928626-7-1 © EURASIP 2017

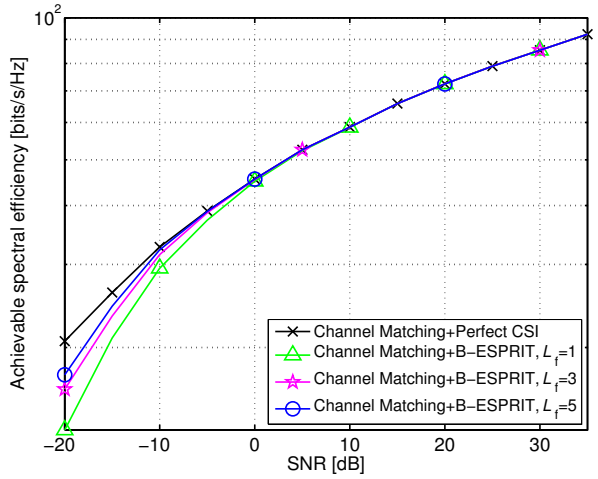

Figure 3. Achievable sum rate vs. SNR when the channel matching based design in [6] is used and there are $N_{\mathrm{ss}} = 2$ spatial streams.

phase, during which the analog as well as digital precoding and decoding schemes are modified according to the channel matching based hybrid design and the estimated CSI. The achievable sum rate increases when the number of virtual subarrays increases. But the performance difference under different simulation settings is obvious only in the low to medium SNR regime.

V. Conclusion

A gridless 3-D Standard ESPRIT in beamspace based channel estimation algorithm has been proposed to estimate the CSI of a hybrid mmWave massive MIMO system. The required training protocol, analog processing as well as pilot patterns have been discussed and developed. Simulation results show that the proposed 3-D Standard ESPRIT in beamspace algorithm provides accurate estimates of both the CSI and the spatial frequencies using only a few training symbols especially when combined with spatial smoothing technique. Moreover, the channel estimation error only slightly affects the achievable sum rate when the channel matching based hybrid design is applied.

Acknowledgments

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

References

[1] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Communications Magazine*, pp. 106–113, Feb. 2014.

[2] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?," *IEEE Communications Magazine*, pp. 56–62, Sept. 2014.

[3] R. Mendez-Rial, C. Rusu, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, Jr., "Channel estimation and hybrid combining for mmWave: phase shifters or switches," in *Proc. of Information Theory and Applications Workshop (ITA 2015)*, UCSD, USA, Feb. 2015.

[4] 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.

[5] 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.

[6] J. Zhang, M. Haardt, I. Soloveychik, and A. Wiesel, "A channel matching based hybrid analog-digital strategy for massive multi-user MIMO downlink systems," in *Proc. 9th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2016)*, Rio de Janeiro, Brazil, July 2016.

[7] 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.

[8] T. E. Bogale, L. B. Le, and X. Wang, "Hybrid analog-digital channel estimation and beamforming: Training-throughput tradeoff," *submitted to IEEE Transactions on Communications, available at http://http://arxiv.org/abs/1509.0509*, pp. 1–15, 2015.

[9] 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.

[10] 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.

[11] 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.

[12] S. Haghighatshoar and G. Caire, "Enhancing the estimation of mm-Wave large array channels by exploiting spatio-temporal correlation and sparse scattering," in *Proc. IEEE Int. Workshop on Smart Antennas (WSA)*, Munich, Germany, Mar. 2016.

[13] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.

[14] 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.

[15] 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.

[16] 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*, Jan. 2007.

[17] 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.

[18] C. P. Mathews, M. Haardt, and M. D. Zoltowski, "Implementation and performance analysis of 2D DFT beamspace ESPRIT," in *Proc. of the 48th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1995.

[19] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 2, pp. 331–341, Apr. 1986.