Subspace Tracking Algorithms for Millimeter Wave MIMO Channel Estimation with Hybrid Beamforming

Stefano Buzzi, *Senior Member*, *IEEE*, and Carmen D'Andrea

Abstract—This paper proposes the use of subspace tracking algorithms for performing MIMO channel estimation at millimeter wave (mmWave) frequencies. Using a subspace approach, we develop a protocol enabling the estimation of the right (resp. left) singular vectors at the transmitter (resp. receiver) side; then, we adapt the projection approximation subspace tracking with deflation (PASTd) and the orthogonal Oja (OOJA) algorithms to our framework and obtain two channel estimation algorithms. The hybrid analog/digital nature of the beamformer is also explicitly taken into account at the algorithm design stage. Numerical results show that the proposed estimation algorithms are effective, and that they perform better than two relevant competing alternatives available in the open literature.

Index Terms—MIMO Channel Estimation, mmWave, clustered channel model

I. Introduction

The use of frequency bands in the range $10-100$ GHz, a.k.a. millimeter waves (mmWaves), for cellular communications, is among the most striking technological innovations brought by fifth generation (5G) wireless networks [1]. Indeed, the scarcity of available frequency bands in the sub-6 GHz spectrum has been the main thrust for considering the use of higher frequencies for cellular applications, and indeed recent research [2] has shown that mmWaves, despite increased path-loss and atmospheric absorption phenomena, can be actually used for cellular communications over short-range distances (up to 100-200 meters), provided that multiple antennas are used at both sides of the communication link: MIMO processing, thus, is one distinguishing and key feature of mmWave systems. When considering mmWave links, one challenging task is that of MIMO channel estimation. This paper tackles this issue, and in particular the contribution of this paper can be summarized as follows.

1) Using a subspace approach, we develop a protocol enabling the estimation of the right (resp. left) singular vectors at the transmitter (resp. receiver) side; then, we adapt the PASTd algorithm [3], and the orthogonal Oja (OOJA) algorithm [4] to our framework and obtain two subspace-based channel estimation algorithms.

2) We adapt the proposed algorithms in order to take into account the hybrid analog/digital beamforming structure usually employed in mmWave wireless links. In particular, we assume that both at the transmitter and

at the receiver the front-end RF modules are made by an analog combining matrix whose columns correspond to beam-steering array response vectors, and show that the proposed channel estimation algorithms based on subspace tracking can be applied in this scenario too.

3) We compare the proposed algorithms with two competing alternatives available in the open literature, namely the *Approximate Maximum Likelihood* (AML) algorithm in [5], [6], and the *Subspace Estimation using Arnoldi iteration* (SE-ARN) in [7]. In this context, we also generalize the AML algorithm – presented in [5], [6] for the case in which the mobile station (MS) has just one antenna – to the case in which the MS is equipped with multiple antennas.

4) We also show how the proposed channel estimation schemes can be used to implement a simple pilot-less differential modulation scheme in the case of multiplexing order one, and the corresponding symbol error probability (SER) is numerically evaluated.

This paper is organized as follows. Next Section contains the description of the system model, while in Section III the protocol to enable estimation of the left and right dominant singular vectors of the channel matrix using the subspace tracking algorithms is described. Section IV is devoted to the description of other estimation schemes, including the extension of the AML algorithm [5] to the case in which the MS is equipped with multiple antennas, and the SE-ARN. Section V is devoted to the illustration of the numerical results, while, finally, concluding remarks are given in Section VI.

Notation: We denote vectors by boldface lowercase letters (e.g. $\mathbf{x}$), matrices by boldface capital letters (e.g. $\mathbf{X}$), scalar constant by nonboldface letters (e.g x or X). The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate-transpose, respectively and the identity matrix of order k is expressed with $\mathbf{I}_k$.

II. System model

We consider a single-cell and single-user MIMO link that can be representative of the Base Station (BS) - Mobile Station (MS) link in a wireless cellular system using an orthogonal multiple access scheme. The number of parallel streams, i.e. the multiplexing order, is denoted by M. We consider a time-division-duplex (TDD) scenario, so that the BS-to-MS channel is the conjugate transpose of the MS-to-BS

The authors are with the Department of Electrical and Information Engineering, University of Cassino and Lazio Meridionale, I-03043 Cassino, Italy (buzzi@unicas.it, carmen.dandrea@unicas.it).
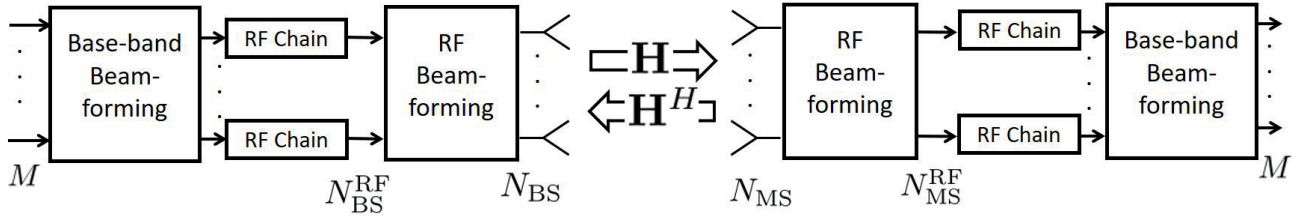
Figure 1. Block scheme of the considered transceiver architecture.

channel, provided that the transmission time does not exceed the channel coherence interval. We denote by N_{BS} the number of antennas at the BS, and by N_{MS} the number of antennas at the MS, and assume for simplicity a bi-dimensional model, i.e. both the BS and the MS are equipped with a uniform linear array (ULA). A schematic representation of the considered system is reported in Fig. 1. In particular, beamforming at both sides of the link is of the hybrid type, in the sense that, in order to lower hardware complexity, analog signals combining is performed at RF in the transceiver front-end in order to reduce the dimensionality of the signals, and, then fully digital baseband combining is performed. We will denote by $N_{\text{BS}}^{\text{RF}} < N_{\text{BS}}$ and $N_{\text{MS}}^{\text{RF}} < N_{\text{MS}}$ the number of RF chains at the BS and at the MS, respectively.

A. The clustered channel model

We denote by $\mathbf{H}$ the $(N_{\text{MS}} \times N_{\text{BS}})$-dimensional matrix representing the BS-to-MS channel. Due to TDD operation the reverse-link propagation channel is expressed as $\mathbf{H}^H$. According to the popular narrowband clustered mmWave channel model [8]–[10], the channel matrix is expressed as:

$$\mathbf{H} = \gamma \sum_{i=1}^{N_{\text{cl}}} \sum_{l=1}^{N_{\text{ray},i}} \alpha_{i,l} \sqrt{L(r_{i,l})} \mathbf{a}_{\text{MS}}(\phi_{i,l}^{\text{MS}}) \mathbf{a}_{\text{BS}}^H(\phi_{i,l}^{\text{BS}}) + \mathbf{H}_{\text{LOS}} .$$
$$(1)$$

In Eq. (1), we are implicitly assuming that the propagation environment is made of N_{cl} scattering clusters, each of which contributes with $N_{\text{ray},i}$ propagation paths, $i = 1, \ldots, N_{\text{cl}}$, plus a possibly present LOS component. We denote by $\phi_{i,l}^{\text{MS}}$ and $\phi_{i,l}^{\text{BS}}$ the downlink angle of arrival (that coincides with the uplink angle of departure) at the MS and the downlink angle of departure (that coincides with the uplink angle of departure) at the BS of the lth ray in the ith scattering cluster, respectively. The quantities $\alpha_{i,l}$ and $L(r_{i,l})$ are the complex path gain and the attenuation associated to the (i, l)-th propagation path. The complex gain $\alpha_{i,l} \sim \mathcal{CN}(0, \sigma_{\alpha,i}^2)$, with $\sigma_{\alpha,i}^2 = 1$ [8]. The factors $\mathbf{a}_{\text{MS}}(\phi_{i,l}^{\text{MS}})$ and $\mathbf{a}_{\text{BS}}(\phi_{i,l}^{\text{BS}})$ represent the normalized receive and transmit array response vectors evaluated at the corresponding angles of arrival and departure; for a ULA with half-wavelength inter-element spacing we have
$$\mathbf{a}_{\text{BS}}(\phi_{i,l}^{\text{BS}}) = \frac{1}{\sqrt{N_{\text{BS}}}} [1 \ e^{-j\pi \sin \phi_{i,l}^{\text{BS}}} \ \cdots \ e^{-j\pi(N_{\text{BS}}-1) \sin \phi_{i,l}^{\text{BS}}}]^T .$$
A similar expression can be also given for $\mathbf{a}_{\text{MS}}(\phi_{i,l}^{\text{MS}})$. Finally,
$$\gamma = \sqrt{\frac{N_{\text{BS}} N_{\text{MS}}}{\sum_{i=1}^{N_{\text{cl}}} N_{\text{ray},i}}}$$ is a normalization factor ensuring that the received signal power scales linearly with the product

$N_{\text{BS}} N_{\text{MS}}$. Regarding the LOS component, denoting by $\phi_{\text{LOS}}^{\text{MS}}$, $\phi_{\text{LOS}}^{\text{BS}}$, the downlink arrival and departure angles corresponding to the LOS link, we assume that

$$\mathbf{H}_{\text{LOS}} = I_{\text{LOS}}(d) \sqrt{N_{\text{MS}} N_{\text{BS}} L(d)} e^{j\theta} \mathbf{a}_{\text{MS}}(\phi_{\text{LOS}}^{\text{MS}}) \mathbf{a}_{\text{BS}}^H(\phi_{\text{LOS}}^{\text{BS}}) .$$
$$(2)$$

In the above equation, $\theta \sim \mathcal{U}(0, 2\pi)$, d is the link length, while $I_{\text{LOS}}(d)$ is a random variate indicating if a LOS link exists between transmitter and receiver, with p the probability that $I_{\text{LOS}}(d) = 1$. A detailed description of all the parameters needed for the generation of sample realizations for the channel model of Eq. (1) is reported in [11], and we refer the reader to this reference for further details on the channel model.

III. SUBSPACE-BASED CHANNEL EIGENDIRECTIONS ESTIMATION

Given the structured (parametric) channel model in (1), differently from what usually happens for MIMO channels at conventional sub-6 GHz frequencies where all the entries of the channel matrix are to be estimated, here we are actually interested only to the dominant left and right singular vectors of the channel matrix itself; it is also quite easy to realize that these singular vectors tend to coincide, in the limit of large number of antennas, with the ULA array responses at the angles corresponding to the rays with the complex gain with the largest norm.

A. Subspace-based channel estimation with fully-digital (FD) beamforming architecture

We start by considering the case of FD beamforming, i.e. no analog beamforming is performed and the number of RF chains coincides with the number of antennas, both at the BS and at the MS. The proposed protocol for channel estimation consists of two successive phases. In phase (a), the BS transmits a suitable probing signal and the MS estimates the dominant left eigenvectors of the channel matrix; then, in phase (b), the MS transmits a suitable signal and the BS estimates the dominant right eigenvectors of the channel matrix.

With regard to phase (a), let $\mathbf{s}_{\text{BS}}(n)$, with $n = 1, \ldots, P_{\text{BS}}$, be a sequence of N_{BS}-dimensional random column vectors with identity covariance matrix[1]. These vectors are transmitted by the BS at (discrete) time $n = 1, \ldots, P_{\text{BS}}$; the signal

---

[1]As an example, a sequence of random uniform binary-valued antipodal symbols can be used.

received at the MS at time n is expressed as the following $(N_{\text{MS}} \times 1)$-dimensional vector:

$$\mathbf{r}_{\text{MS}}(n) = \mathbf{H}\mathbf{s}_{\text{BS}}(n) + \mathbf{w}_{\text{MS}}(n) , \qquad (3)$$

where $\mathbf{w}_{\text{MS}}(n)$ is the N_{MS}-dimensional AWGN vector, modeled as $\mathcal{CN}(0, \sigma_n^2)$ independent RVs. Letting $\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H$ denote the singular value decomposition of the channel matrix, the covariance matrix of the received signal is expressed as

$$\mathbf{R}_{\text{MS}} = E\left[\mathbf{r}_{\text{MS}}(n)\mathbf{r}_{\text{MS}}^H(n)\right] = \mathbf{U}\boldsymbol{\Lambda}^2\mathbf{U}^H + \sigma_n^2\mathbf{I}_{N_{\text{MS}}} . \quad (4)$$

Given (4), it is thus easily seen that we can estimate the M dominant left singular vectors of the channel matrix by estimating the M dominant directions of the subspace spanned by the received vectors $\mathbf{r}_{\text{MS}}(n)$, with $n = 1, \ldots, P_{\text{BS}}$. The signal processing literature is rich of adaptive subspace tracking algorithms that can be straightforwardly applied at the MS to obtain an estimate of the principal left singular vectors of the channel matrix. Deferring later the specification of the adopted subspace tracking algorithms, let $\mathbf{D}_{MS}$ be the $(N_{\text{MS}} \times M)$-dimensional matrix containing the estimate of the M dominant singular vectors of $\mathbf{R}_{\text{MS}}$; the matrix $\mathbf{D}_{\text{MS}}$ will be used at the MS as a precoder during data transmission and as a combiner when receiving data from the BS. After P_{BS} symbol intervals, phase (a) is over and phase (b) starts. The MS now transmits random independent vectors with identity covariance matrix in order to enable estimation at the BS of the dominant right eigenvectors of the channel matrix $\mathbf{H}$. More precisely, let $\mathbf{s}_{\text{MS}}(n)$, with $n = P_{\text{BS}} + 1, \ldots, P_{\text{BS}} + P_{\text{MS}}$, be a sequence of M-dimensional random vectors with identity covariance matrix and assume that the MS transmits the following N_{MS}-dimensional data vectors

$$\mathbf{x}_{\text{MS}}(n) = \mathbf{D}_{\text{MS}}\mathbf{s}_{\text{MS}}(n), \qquad (5)$$

where $\mathbf{s}_{\text{MS}}(n)$ is the $(M \times 1)$-dimensional vector whose entries take value in $\{\pm 1\}$. The received discrete-time signal at the BS is represented by the following N_{BS}-dimensional vector:

$$\mathbf{r}_{\text{BS}}(n) = \mathbf{H}^H\mathbf{x}_{\text{MS}}(n) + \mathbf{w}_{\text{BS}}(n) , \qquad (6)$$

where $\mathbf{w}_{\text{BS}}(n)$ is the N_{BS}-dimensional AWGN vector, modeled as $\mathcal{CN}(0, \sigma_n^2)$ independent RVs[2]. Similarly to phase (a), under the assumption of negligible errors in the estimation of the left singular vectors of the channel matrix, the covariance matrix of the received signal at the BS is expressed as

$$\mathbf{R}_{\text{BS}} = E\left[\mathbf{r}_{\text{BS}}(n)\mathbf{r}_{\text{BS}}^H(n)\right] \approx \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^H + \sigma_n^2\mathbf{I}_{N_{\text{BS}}} , \quad (7)$$

thus implying that the the M dominant right singular vectors of the channel matrix can be estimated by running adaptive subspace tracking algorithms at the BS. We will denote by $\mathbf{D}_{\text{BS}}$ is the $(N_{\text{BS}} \times M)$-dimensional matrix containing the estimates of the M dominant singular vectors of $\mathbf{R}_{\text{MS}}$.

[2]Note that in phase (b) the MS is already using the estimated precoder $\mathbf{D}_{\text{MS}}$; the illustrated procedure also works if the MS does not use the precoder and sends N_{MS}-dimensional random vectors.

B. Subspace-based channel estimation with hybrid (HY) beamforming architecture

In the previous section a FD beamforming structure has been assumed. We now examine the case in which, for complexity reasons, a HY beamforming architecture is adopted. The front-end processing consists of an analog RF combining matrix aimed at reducing the number of RF chains needed to implement the base-band processing. From a mathematical point of view, the beamforming matrices at the MS and at the BS can be expressed as

$$\begin{aligned}
\mathbf{D}_{\text{MS}} &= \mathbf{D}_{\text{MS,RF}}\mathbf{D}_{\text{MS,BB}} , \qquad \text{and} \\
\mathbf{D}_{\text{BS}} &= \mathbf{D}_{\text{BS,RF}}\mathbf{D}_{\text{BS,BB}} ,
\end{aligned} \qquad (8)$$

respectively. In (8), $\mathbf{D}_{\text{MS,RF}}$ is an $(N_{\text{MS}} \times N_{\text{MS}}^{\text{RF}})$-dimensional matrix with unit-norm entries, while $\mathbf{D}_{\text{MS,BB}}$ is an $(N_{\text{MS}}^{\text{RF}} \times M)$-dimensional matrix with no constraint on its entries. Similarly, $\mathbf{D}_{\text{BS,RF}}$ is an $(N_{\text{BS}} \times N_{\text{BS}}^{\text{RF}})$-dimensional matrix with unit-norm entries, and $\mathbf{D}_{\text{BS,BB}}$ is an $(N_{\text{BS}}^{\text{RF}} \times M)$-dimensional baseband combining matrix. The design of HY analog/digital combiners is a vastly explored research topic; most papers try to find the hybrid combiner that best approximates, according to some criterion, the optimal FD combiner. In this paper, we use a simpler and different approach. We assume that $\mathbf{D}_{\text{MS,RF}}$ and $\mathbf{D}_{\text{BS,RF}}$ have a fixed structure and in particular contain on their column the ULA array responses corresponding to a grid of discrete angles spanning the range $[-\pi/2, \pi/2]$. In particular, letting

$$\begin{aligned}
\theta_{\text{MS}}(i) &= \left(-\frac{\pi}{2} + \frac{\pi(i-1)}{N_{\text{MS}}^{\text{RF}}}\right) , \quad i = 1, \ldots, N_{\text{MS}}^{\text{RF}} , \\
\theta_{\text{BS}}(i) &= \left(-\frac{\pi}{2} + \frac{\pi(i-1)}{N_{\text{BS}}^{\text{RF}}}\right) , \quad i = 1, \ldots, N_{\text{BS}}^{\text{RF}} ,
\end{aligned} \qquad (9)$$

the RF combiners have the following structure:

$$\begin{aligned}
\mathbf{D}_{\text{BS,RF}} &= \left[\mathbf{a}_{\text{BS}}\left(\theta_{\text{BS}}(1)\right), \ldots \mathbf{a}_{\text{BS}}\left(\theta_{\text{BS}}(N_{\text{BS}}^{\text{RF}})\right)\right] , \\
\mathbf{D}_{\text{MS,RF}} &= \left[\mathbf{a}_{\text{MS}}\left(\theta_{\text{MS}}(1)\right), \ldots \mathbf{a}_{\text{MS}}\left(\theta_{\text{MS}}(N_{\text{MS}}^{\text{RF}})\right)\right] .
\end{aligned} \qquad (10)$$

Now, focus on the scheme in Fig. 1 and consider the cascade of the BS analog beamformer, the channel $\mathbf{H}$ and the MS analog beamformer; it is straightforward to show that this cascade can be modeled through the matrix $\widetilde{\mathbf{H}} = \mathbf{D}_{\text{MS,RF}}^H\mathbf{H}\mathbf{D}_{\text{BS,RF}}$, of dimension $N_{\text{MS}}^{\text{RF}} \times N_{\text{BS}}^{\text{RF}}$. As a consequence, the channel estimation scheme outlined in the previous section under the assumption of FD beamforming can be now re-applied on the (reduced-dimension) composite channel $\widetilde{\mathbf{H}}$.

C. PASTd algorithm

The PASTd algorithm was introduced in [3]; one of its most popular applications is reported in the highly-cited paper [12]. In order to illustrate this algorithm, let $\mathbf{r}$ be an $(N \times 1)$-dimensional random vector with autocorrelation matrix $\mathbf{C} = E\left[\mathbf{r}\mathbf{r}^H\right]$. Consider the scalar function

$$\begin{aligned}
J(\mathbf{W}) &= E\left[\|\mathbf{r} - \mathbf{W}\mathbf{W}^T\mathbf{r}\|^2\right] \\
&= \text{tr}(\mathbf{C}) - 2\text{tr}(\mathbf{W}^T\mathbf{C}\mathbf{W}) + \text{tr}(\mathbf{W}^T\mathbf{C}\mathbf{W}\mathbf{W}^T\mathbf{W}),
\end{aligned} \qquad (11)$$

with a $(N \times M)$-dimensional matrix argument $\mathbf{W}$, with $M < N$. It is shown in [3] that

- the matrix $\mathbf{W}$ is a stationary point of $J(\mathbf{W})$ if and only if $\mathbf{W} = \mathbf{T}_M\mathbf{Q}$, where $\mathbf{T}_M$ is a $(N \times M)$-dimensional

---

Algorithm 1 The PASTd Algorithm

1: $\mathbf{x}_1(n) = \mathbf{r}(n)$
2: **for** $m = 1 : M$ **do**
3:     $y_m(n) = \mathbf{u}_m^H(n-1)\mathbf{x}_m(n)$
4:     $\lambda_m(n) = \beta\lambda_m(n-1) + |y_m(n)|^2$
5:     $\mathbf{u}_m(n) = \mathbf{u}_m(n-1) + [\mathbf{x}_m(n) - \mathbf{u}_m(n-1)y_m(n)]\frac{y_m(n)^*}{\lambda_m(n)}$
6:     $\mathbf{x}_{m+1}(n) = \mathbf{x}_m(n) - \mathbf{u}_m(n)y_m(n)$
7: **end for**

Algorithm 2 The OOJA Algorithm

1: $\mathbf{v}(n) = \mathbf{W}^T(n)\mathbf{r}(n)$
2: $\mathbf{z}(n) = \mathbf{W}(n)\mathbf{v}(n)$
3: $\mathbf{p}(n) = \mathbf{r}(n) - \mathbf{z}(n)$
4: $\varphi(n) = \frac{1}{\sqrt{1 + \delta^2\|\mathbf{p}(n)\|^2\|\mathbf{v}(n)\|^2}}$
5: $\tau(n) = \frac{\phi(n)-1}{\|\mathbf{v}(n)\|^2}$
6: $\bar{\mathbf{p}}(n) = -\frac{\tau(n)\mathbf{z}(n)}{\delta} + \phi(n)\mathbf{p}(n)$
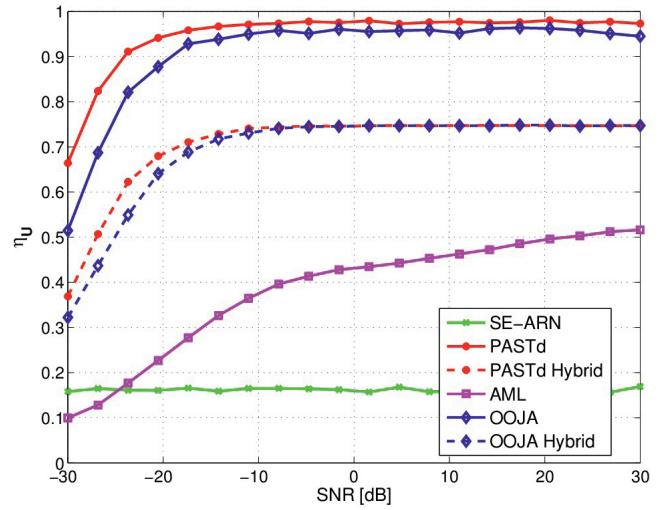7: $\mathbf{W}(n+1) = \mathbf{W}(n) - \delta\bar{\mathbf{p}}(n)\mathbf{v}^T(n)$

---

matrix contains any M distinct eigenvectors of $\mathbf{C}$ and $\mathbf{Q}$ is any $(M \times M)$-dimensional unitary matrix.

- All stationary points of $J(\mathbf{W})$ are saddle points except when $\mathbf{T}_M$ contains the M dominant eigenvectors of $\mathbf{C}$. In that case, $J(\mathbf{W})$ attains the global minimum.

Therefore, for $M = 1$ the solution of minimizing $J(\mathbf{W})$ is given by the most dominant eigenvector of $\mathbf{C}$. In practical applications, only sample vectors $\mathbf{r}(i)$ are available and so the statistical average in (11) is replaced by an exponentially-windowed time-average, i.e.:

$$J[\mathbf{W}(n)] = \sum_{i=1}^{n} \beta^{n-i}\|\mathbf{r}(i) - \mathbf{W}(n)\mathbf{W}(n)^T\mathbf{r}(i)\|^2, \quad (12)$$

where β is the forgetting factor. The key trick of the PASTd approach is to approximate $\mathbf{W}(n)^T\mathbf{r}(i)$ in (12), the unknown projection of $\mathbf{r}(i)$ onto the columns of $\mathbf{W}(n)$, by $\mathbf{y}(i) = \mathbf{W}(i-1)^T\mathbf{r}(i)$, which can be calculated for $i = 1, \ldots n$ at sampling time n. This results in a modified cost function

$$\tilde{J}[\mathbf{W}(n)] = \sum_{i=1}^{n} \beta^{n-i}\|\mathbf{r}(i) - \mathbf{W}(n)\mathbf{y}(n)\|^2, \quad (13)$$

The recursive least squares (RLS) algorithm can then be used to solve for $\mathbf{W}(n)$ that minimizes the exponentially weighted least square criterion (13). The PASTd algorithm for tracking the eigenvalues and eigenvectors of the signal subspace is based on deflaction technique and its basic idea is as follows. For $M = 1$, by minimizing $\tilde{J}[\mathbf{W}(n)]$ in (13) the most dominant eigenvector is updated; then, the contrib-uton from this estimated eigenvector is removed from $\mathbf{r}(n)$ itself, and the second most dominant eigenvector can be now extracted from the data. This procedure is applied repeatedly until the M dominant eigenvectors are sequentially estimated. The complete PASTd procedure is reported in Algorithm 1. *Computational Complexity*: The PASTd algorithm costs $4NM + O(N)$ flops per iteration.

D. OOJA algorithm

The OOJA algorithm was introduced in [4]; it builds upon the minor subspace extraction algorithm of Oja in [13]. Ac-cording to Oja's algorithm the iteration is:

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \delta\left[\mathbf{r}(n)\mathbf{v}^T(n) - \mathbf{W}(n)\mathbf{v}(n)\mathbf{v}^T(n)\right]$$
$$= \mathbf{W}(n) - \delta\mathbf{p}(n)\mathbf{v}^T(n), \quad (14)$$

where $\mathbf{W}$ is a $(N \times M)$-dimensional matrix, with $M < N$, $\mathbf{v}(n) = \mathbf{W}^T(n)\mathbf{r}(n)$, $\mathbf{p}(n) = \mathbf{r}(n) - \mathbf{W}(n)\mathbf{v}(n)$ and δ is a learning parameter. The OOJA algorithm consists of iteration

Figure 2. η_U versus the received SNR for a system with $N_{\text{MS}} \times N_{\text{BS}} = 30 \times 100$. For the HY implementations we have used $N_{\text{MS}}^{\text{RF}} = 10$ and $N_{\text{BS}}^{\text{RF}} = 20$.

(14), proposed in [13], plus the following orthogonalization step of the weight matrix:

$$\mathbf{W}(n+1) = \mathbf{W}(n+1)\left(\mathbf{W}^T(n+1)\mathbf{W}(n+1)\right)^{-1/2}, \quad (15)$$

where $\left(\mathbf{W}^T(n+1)\mathbf{W}(n+1)\right)^{-1/2}$ denotes the inverse square root of $\mathbf{W}^T(n+1)\mathbf{W}(n+1)$. In [4] the authors write (14) as:

$$\mathbf{W}(n+1) = \left[\mathbf{W}(n) - \delta\mathbf{p}(n)\mathbf{v}^T(n)\right]\left(\mathbf{I}_M + \tau(n)\mathbf{v}(n)\mathbf{y}^T(n)\right)$$
$$= \mathbf{W}(n) - \delta\bar{\mathbf{p}}(n)\mathbf{v}^T(n), \quad (16)$$

where

$$\tau(n) = \frac{1}{\|\mathbf{v}(n)\|^2}\left[\frac{1}{\sqrt{1 + \delta^2\|\mathbf{p}(n)\|^2\|\mathbf{v}(n)\|^2}} - 1\right] \quad (17)$$

and

$$\bar{\mathbf{p}}(n) = -\frac{\tau(n)\mathbf{W}(n)\mathbf{v}(n)}{\delta} + \left[1 + \tau(n)\|\mathbf{v}\|^2\right]\mathbf{p}(n). \quad (18)$$

The complete OOJA procedure is reported in Algorithm 2. *Computational Complexity*: The OOJA algorithm costs $3NM + O(N)$ flops per iteration.

IV. NUMERICAL SIMULATIONS

We now provide some numerical results. The first perfor-mance measure that we are considering is the normalized correlation between the true channel eigenvectors and the
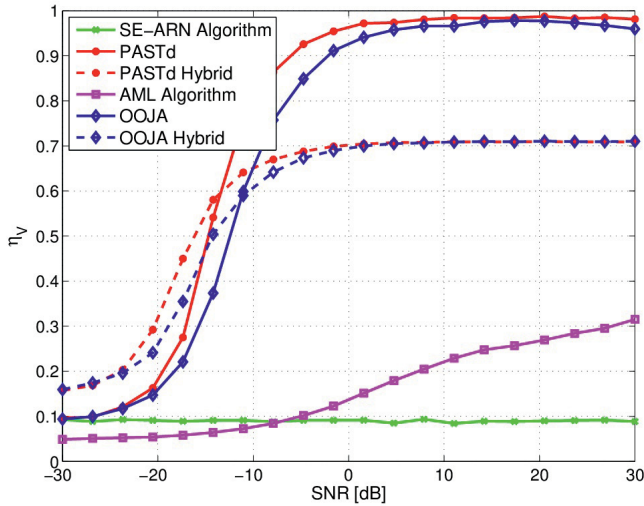
374

Figure 3. η_U versus the received SNR for a system with $N_{\mathrm{MS}} \times N_{\mathrm{BS}} = 30 \times 100$. For the HY implementations we have used $N_{\mathrm{MS}}^{\mathrm{RF}} = 10$ and $N_{\mathrm{BS}}^{\mathrm{RF}} = 20$.

Figure 4. CDF of η_U for a system with $N_{\mathrm{MS}} \times N_{\mathrm{BS}} = 30 \times 100$. For the HY implementations we have used $N_{\mathrm{MS}}^{\mathrm{RF}} = 10$ and $N_{\mathrm{BS}}^{\mathrm{RF}} = 20$.

estimated ones. In particular, we denote by $\mathbf{u}$ and $\mathbf{v}$ the left and the right eigenvector of the channel matrix corresponding to the dominant eigenvalue respectively, and by $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ the estimates of left and right dominant eigenvector of the channel matrix, respectively. The normalized correlations are defined as

$$\eta_U = \frac{\left|\mathbf{u}^H \hat{\mathbf{u}}\right|}{\|\mathbf{u}\|\|\hat{\mathbf{u}}\|}, \qquad (19)$$

and as

$$\eta_V = \frac{\left|\mathbf{v}^H \hat{\mathbf{v}}\right|}{\|\mathbf{v}\|\|\hat{\mathbf{v}}\|}. \qquad (20)$$

We will also report the achievable rate and the error probability assuming differential phase shift keying signaling and differential non-coherent detection - notice indeed that the proposed channel estimation methods do not rely on known training symbols, and the dominant eigenvector is estimated with no information on the signal phase.

In our simulation setup, we consider a communication bandwidth of $W = 500$ MHz centered over the carrier frequency $f_0 = 73$ GHz. The distance between the transmitter and the receiver is 50 m; the additive thermal noise is assumed to have a power spectral density of -174 dBm/Hz, while the front-end receiver at the BS and at the MS is assumed to have a noise figure of 3 dB. The shown results come from an average over 500 random scenario realizations with independent channels.

The length of the training phase for all the algorithms is $P_{\mathrm{BS}} = P_{\mathrm{MS}} = 30$; of these training samples, the first ten are used to perform an SVD of the sample covariance matrix of the data, and the corresponding dominant eigenvector is used to initialize the PASTd and OOJA algorithms. In Figs. 2 and 3 we report η_U and η_V versus the received SNR for the considered algorithms and also for the SE-ARN and AML algorithms. In Figs. 4 and 5 we report instead the cumulative distribution function (CDF) of η_U and η_V at an SNR of 10 dB, respectively – the remaining parameters are the same as the ones taken in Figs. 2-3. First of all, it is seen that the
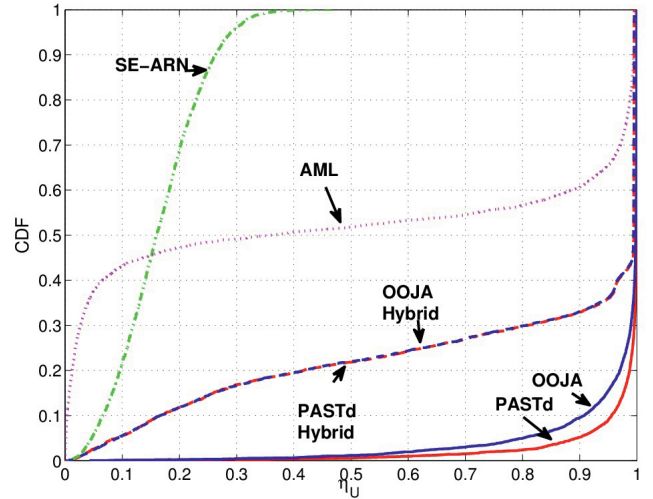
proposed subspace tracking algorithms achieve good performance and outperform competing alternatives. As expected, FD implementations achieves better performance than their HY suboptimal implementations. Further investigations are actually needed here to test the system performance with different types of analog front-end beamformers. Fig. 6 reports the achievable spectral efficiency for the BS-to-MS link versus SNR, for multiplexing order $M = 1$ (in subplot (a)) and $M = 3$ (in subplot (b)). In particular, the plotted performance measure is the following

$$\mathcal{R}_{\mathrm{MS}} = \log_2 \det \left[\mathbf{I}_M + P_{\mathrm{T,BS}} \left(\sigma_n^2 \mathbf{D}_{\mathrm{MS}}^H \mathbf{D}_{\mathrm{MS}} \right)^{-1} \times \right. \qquad (21)$$

$$\left. \mathbf{D}_{\mathrm{MS}}^H \mathbf{H} \mathbf{D}_{\mathrm{BS}} \mathbf{D}_{\mathrm{BS}}^H \mathbf{H}^H \mathbf{D}_{\mathrm{MS}} \right] .$$

with $P_{\mathrm{T,BS}}$ the transmitted power at the BS. Results again show that the proposed algorithms have good performance levels. Finally, Fig. 7 we report the symbol error probability versus SNR for the BS-to-MS link and assuming multiplexing order $M = 1$. Subplot (a) refers to the case in which the training phase length is 10 (with two symbols used to perform the SVD to initialize the PASTd and OOJA algorithms), while subplot (b) refers to the case in which the training phase length is 50 (with 10 symbols used for initialization). A differential 16-PSK modulation has been assumed. Overall, results confirm again the good performance of the proposed estimation techniques, even for the case of short training length.

V. CONCLUSION

This paper has been focused on the problem of channel estimation for wireless single-user MIMO links at mmWave frequencies. Exploiting the clustered propagation channel model, it is shown that subspace tracking algorithms can be used in order to track the principal eigenvectors of the channel matrix. Results have shown that the proposed estimation algorithms are effective and capable of attaining good performance levels
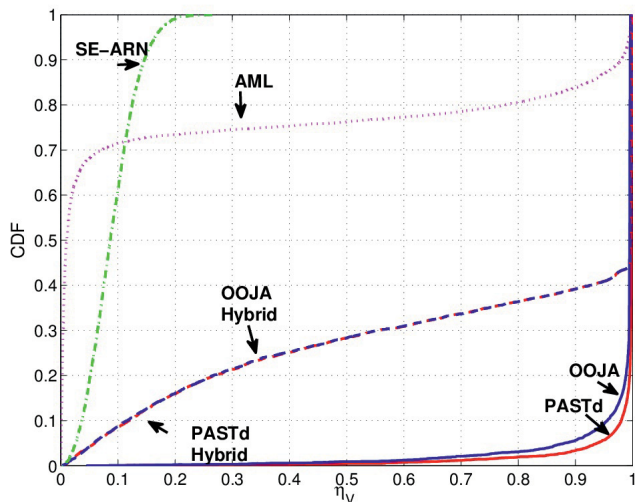
375

Figure 5. CDF of η_V for a system with $N_{\mathrm{MS}} \times N_{\mathrm{BS}} = 30 \times 100$. For the HY implementations we have used $N_{\mathrm{MS}}^{\mathrm{RF}} = 10$ and $N_{\mathrm{BS}}^{\mathrm{RF}} = 20$.

Figure 7. Error probability for the BS-to-MS link versus received SNR. Parameters: 16-PSK differential modulation, $N_{\mathrm{MS}} \times N_{\mathrm{BS}} = 30 \times 100$, $N_{\mathrm{MS}}^{\mathrm{RF}} = 10$, $N_{\mathrm{BS}}^{\mathrm{RF}} = 20$, 10 training symbols in subplot (a), and 50 training symbols in subplot (b).

Figure 6. Spectral efficiency for the BS-to-MS link versus received SNR. Parameters: $N_{\mathrm{MS}} \times N_{\mathrm{BS}} = 30 \times 100$, $N_{\mathrm{MS}}^{\mathrm{RF}} = 10$, $N_{\mathrm{BS}}^{\mathrm{RF}} = 20$, multiplexing order $M = 1$ in subplot (a) and $M = 3$ in subplot (b).

also for very short training phases. This research can be extended in several directions. First of all, for the HY solutions, we have used a fixed analog beamformer. Further research is needed in order to properly design more effective analog beamformers, possibly to be tuned via an adaptive algorithm. Additionally, note that we have been proposing here a single-user estimation technique; multiuser joint channel learning schemes are thus to be investigated. Finally, it would be of great interest to investigate the performance of the proposed adaptive algorithms in a dynamic environment, wherein channel impulse response changes due to, e.g., MS mobility.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
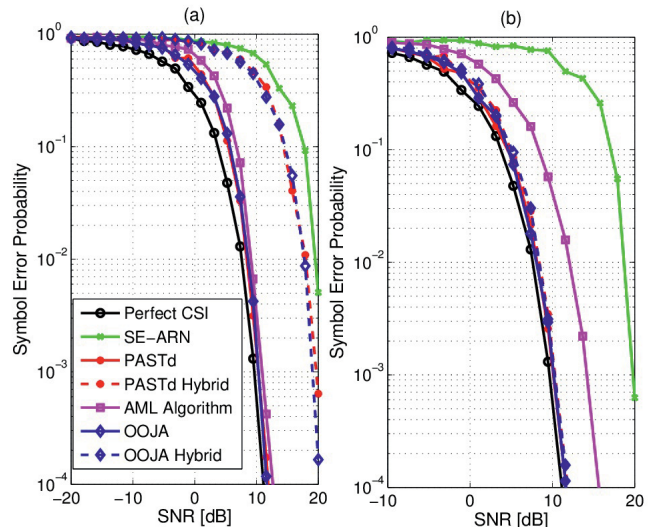
[2] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[3] B. Yang, "Projection approximation subspace tracking," *IEEE Transactions on Signal processing*, vol. 43, no. 1, pp. 95–107, Jan. 1995.

[4] K. Abed-Meraim, S. Attallah, A. Chkeif, and Y. Hua, "Orthogonal OJA algorithm," *IEEE Signal processing letters*, vol. 7, no. 5, pp. 116–119, May 2000.

[5] S. Haghighatshoar and G. Caire, "Low-complexity massive MIMO subspace estimation and tracking from low-dimensional projections," *arXiv preprint arXiv:1608.02477*, Aug. 2016.

[6] ——, "Massive MIMO channel subspace estimation from low-dimensional projections," *IEEE Transactions on Signal Processing*, Oct. 2016.

[7] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 528–542, Apr. 2016.

[8] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[9] S. Haghighatshoar and G. Caire, "Enhancing the estimation of mm-Wave large array channels by exploiting spatio-temporal correlation and sparse scattering," in *Proc. of 20th International ITG Workshop on Smart Antennas (WSA 2016)*, 2016.

[10] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *2014 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2014, pp. 3326–3331.

[11] S. Buzzi and C. D'Andrea, "On clustered statistical MIMO millimeter wave channel simulation," *arXiv preprint arXiv:1604.00648*, May 2016.

[12] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.

[13] E. Oja, "Principal components, minor components, and linear neural networks," *Neural networks*, vol. 5, no. 6, pp. 927–935, Dec. 1992.