IJICTR
International Journal of Information & Communication Technology Research
Volume 9- Number 4 – Autumn 2017 (1-10)

Subspace-Based Approaches for Hybrid Millimeter-Wave Channel Estimation

M.Shakhsi Dastgahian*
Electrical Engineering Department
Ferdowsi University of Mashhad
Mashhad, Iran
Email: majid.shakhsidastgahian@mail.um.ac.ir

H.Khoshbin Ghomash
Electrical Engineering Department
Ferdowsi University of Mashhad
Mashhad, Iran
Email: khoshbin@um.ac.ir

Received: March 1, 2017 - Accepted: September 3, 2017

Abstract—**Millimeter wave communication (mmWC) is a promising volunteer for 5G communication systems with high data rates. To subdue the channel propagation characteristics in this frequency band, high dimensional antenna arrays need to be deployed at transceiver. Employing such a deployment, prevents to use of ADC or RF chain in each branch of MIMO system because of power constraints. Thus, Such systems impose to have a hybrid analog/digital precoding/combining architecture. Hence, channel estimation revision seems to be essential. This paper propose new algorithms to estimate the mmW channel by exploiting the sparse nature of the channel and finding the subspace of received signal vectors based on MUSIC. By combining the multiple measurement vector (MMV) concept, MISIC , subspace augmentation (SA) and two-stage orthogonal subspace matching pursuit (TOSMP) approaches, we try to recover the indices of non-zero elements of an unknown channel matrix accurately even under the defective- rank condition. These indices are called support in the context. Simulation results indicate MUSIC-based approaches offer lower estimation error and higher sum rates compared with conventional MMV solutions.**

Keywords—Millimeter wave MIMO systems; sparse channel estimation; support; multiple measurement vectors (MMV); subspace augmentation (SA).

I. INTRODUCTION

Thanks to prominent feature of huge unlicensed spectral frequency between 30GHz to 300GHz in millimeter wave (mmW) systems, modern wireless communication tends to develop standards to achieve high-throughput [1].

One of the challenges in outdoor ambience is severe path loss and shadowing phenomena due to oxygen

* Corresponding Author

absorption, humidity fades, and reflective outdoor materials in millimetre-wave band. Fortunately, progression in RF circuits combined with very small wavelengths of mmW signals (between 1 to 10mm) makes it possible to pack a miniaturized large number of antennas into transceivers thereby providing high beamforming gains that can compensate path loss and even out-of-cell interference [2]. Utilizing of the massive MIMO in mmW system helps to provide pencil-shaped beams so that it focuses energy into a favorable direction [3]. Large scale Moreover, adaptive beamforming may make systems less vulnerable to unfavorable shadowing effects. In contrast to conventional lower frequency systems, millimeter wave (mmW) systems, avoid to dedicate one complete RF chain and one high-resolution ADC or DAC to each branch of antenna due to a high cost and power consumption of such components. For this reason, low complexity sub-optimal analog beamforming based on beam training is proposed to be used instead of fully baseband solutions to support only single stream MIMO communication [4].

To send several data streams simultaneously and achieve more accurate beamforming gain, a hybrid architecture has been proposed in [5] where the processing is divided across the analog stage with number of RF chains much lower than the number of antennas and baseband stage. Baseband or digital stage is for correction of limitation of analog RF section. In [6] the sparse nature of the poor scattering mmW channel is exploited to develop low-complexity channel estimation. Hybrid beamformers designing problem based on estimation of a few eigenmodes according to echoing between transmitter and receiver has been investigated in [7] based on subspace estimation rather than estimation of the whole channel by utilizing the concept of the reciprocity of the channel in TDD MIMO systems.

In this paper, we consider a hybrid beamforming model for downlink single-user mmW systems. We assume to have an known sensing matrix in several times of training mode as well as to know the geometry of the arrays in source and destination. Thus, we utilize the multiple measurement vectors (MMV) model of sparse for millimeter channel and propose different approaches for solving the channel estimation problem [8]. The main contribution of the paper is developing the MUSIC-based methods [9] rather than the existing simultaneous algorithms for solving the joint sparse channel recovery. One of the challenges in millimeter wave channel estimation can be rank-deficiency. A matrix is rank defective when rank is smaller than the dimension of the matrix. For rank-defective channel estimation by simultaneous algorithms we can use contents of [8, 10-12].

We use the following notations throughout this paper.

The bold upper-case letters denotes matrices, and bold lower-cases represent vectors. Furthermore, $\|\mathbf{A}\|_F$ is Frobenius norm, whereas $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^*$ and $\mathbf{A}^\dagger$ are its transpose, conjugate transpose (Hermitian), conjugate and Moore-Penrose pseudo-inverse, accordingly. $\mathbf{A}\otimes\mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\mathbb{E}[.]$, is used to designate expectation. $\mathbf{A}_\chi$ is a submatrix of $\mathbf{A}$ composed of columns indicated by set χ . $R(\mathbf{\Psi})$ is abbreviated to range of the matrix $\mathbf{\Psi}$. An $n\times n$ unitary matrix is represented by $\mathbf{I}_n$.

II. SYSTEM MODEL

Assume a single user downlink Millimetre-Wave MIMO system with N_t transmit antennas at the base station (BS) and N_r receive antennas at the Mobile station (MS) whereas each side is equipped by N_{rf}^t and N_{rf}^r RF chains. N_s data streams are considered to send into sparse channel. In our model, the number of component of the transmitter array is more than the receiver and also the number of RF chains is satisfied to $N_s \le N_{rf} \le \min(N_t, N_r)$ Fig.1 depicts a hybrid single user MIMO mmW transceiver with spatial multiplexing gain and phase shifter as an analog beamformer.

The downlink signal at the receiver side after filtering on baseband is given by,

$$\mathbf{y}_r = \mathbf{G}_r^H \mathbf{F}_r^H \mathbf{H} \mathbf{F}_t \mathbf{G}_t \mathbf{t} + \mathbf{G}_r^H \mathbf{F}_r^H \mathbf{n} = \mathbf{C}^H \mathbf{H} \mathbf{P} \mathbf{t} + \mathbf{C}^H \mathbf{n} \quad (1)$$

where $\mathbf{H} \in C^{N_r \times N_t}$ is the complex sparse channel assumed to be slowly block-fading, $\mathbf{F}_t \in C^{N_t \times N_{rf}^t}$ is the analog or *RF* precoder, $\mathbf{G}_t \in C^{N_{rf}^t \times N_t}$ the baseband precoder, $\mathbf{t} \in C^{N_t \times 1}$ is transmit signal vector with covariance matrix $\mathbb{E}[\mathbf{t}\mathbf{t}^H] = (P_s / N_s)\mathbf{I}_{N_s}$ and $\mathbf{n} \in C^{N_r \times 1}$ is the additive Gaussian noise at the receiver with $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \sigma_n^2 \mathbf{I}_{N_r}$. Similarly, $\mathbf{F}_r \in C^{N_r \times N_{rf}^r}$ and $\mathbf{G}_r \in C^{N_{rf}^r \times N_r}$ are the RF band and baseband combiners, respectively. matrix $\mathbf{C}$ is defined as $\mathbf{F}_r \mathbf{G}_r \in C^{N_r \times N_r}$.

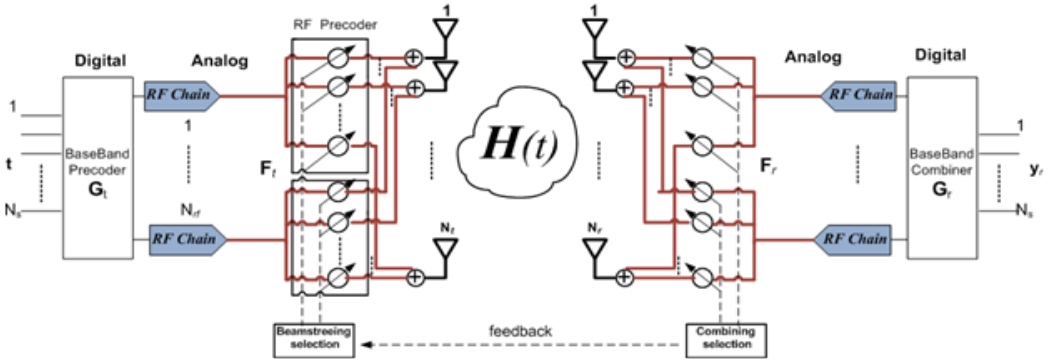


Figure1. Hybrid Model of millimetre wave channel

In Fig.1, the block of phase shifters as an analog

precoder/combiner can be chosen from predefined codebooks or as random matrices with variable phase and constant amplitude. Thus, a possible value set for μth phase shifter and νth RF chain in matrix $\mathbf{F}$ is

$$[\mathbf{F}]_{\mu,\nu} = 1/\sqrt{N_t}\, e^{j\vartheta_{\mu,\nu}} \quad (2)$$

where $\vartheta_{\mu,\nu}$ as a M-bits quantized angle is chosen from uniform distribution in range of $[0, 2\pi)$. The total power constraint is compelled by normalizing $\mathbf{G}_t$ such that $\|\mathbf{F}_t\mathbf{G}_t\|_F^2 = N_s$.

Based on parametric physical model of channel with L scatterers and assumption that each scatterer contributes a single propagation path between the BS and MS, the nonlinear channel $\mathbf{H}$ in spatial angles (but linear in the path gains) can be indicated as

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L}} \sum_{l=1}^{L} \beta_l \mathbf{v}_r(\theta_l)\mathbf{v}_t^H(\varphi_l) = \sqrt{\frac{N_t N_r}{L}} \mathbf{V}_r(\theta)\boldsymbol{\beta}\mathbf{V}_t^H(\varphi) \quad (3)$$

where $\boldsymbol{\beta} = diag([\beta_1, \beta_2, ..., \beta_L]) \in C^{L\times L}$ is the L dimensional propagation path gain diagonal matrix with independently and identically distributed complex Gaussian diagonal entries with zero mean and variance $1/L$. The $\mathbf{V}_r(\theta) \in C^{N_r \times L}$ and $\mathbf{V}_t(\varphi) \in C^{N_t \times L}$ in (3) represent array response matrices at the BS and MS, respectively. Such matrices are given by

$$\mathbf{V}_r(\theta) = [\mathbf{v}_r(\theta_1), \mathbf{v}_r(\theta_2), ..., \mathbf{v}_r(\theta_L)]/\sqrt{N_r} \quad (4)$$

$$\mathbf{V}_t(\varphi) = [\mathbf{v}_t(\varphi_1), \mathbf{v}_t(\varphi_2), ..., \mathbf{v}_t(\varphi_L)]/\sqrt{N_t} \quad (5)$$

The φ_i and θ_i denote the Angle of Departure (AoDs) and Angle of Arrival (AoAs) of ith independent path from L total path. By assuming of an uniform linear arrays (ULA) model, $\mathbf{v}_r(\theta_l)$ and $\mathbf{v}_t(\varphi_l)$ can be defined as

$$\mathbf{v}_r(\theta_l) = \left[1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta_l)}, ..., e^{-j\frac{2\pi}{\lambda}(N_r-1)d\sin(\theta_l)}\right]^T/\sqrt{N_r} \quad (6)$$

$$\mathbf{v}_t(\varphi_l) = \left[1, e^{-j\frac{2\pi}{\lambda}d\sin(\varphi_l)}, ..., e^{-j\frac{2\pi}{\lambda}(N_t-1)d\sin(\varphi_l)}\right]^T/\sqrt{N_t} \quad (7)$$

where λ is the signal wavelength and d is the inter-antenna distance and set to $\frac{\lambda}{2}$ at both the BS and MS. It is assumed $\mathbf{v}_r(\theta_l)$ and $\mathbf{v}_t(\varphi_l)$ vary slowly and they can be well estimated at both sides.

III. PROBLEM FORMULATION

In this section, we take the advantages of the sparse nature of the mmW channel and formulize channel estimation as a compressive sensing problem. Contrary to [6] where training analogue vectors are obtained from a multi-level hierarchically, here they are found as random vectors with fluctuating phase by using one RF chain in training step. As a result, the received signal can be written as

$$\upsilon_k = \mathbf{f}_{r,k}^H \mathbf{H}\mathbf{f}_{t,k}\, t_k + \mathbf{f}_{r,k}^H \mathbf{n}_k \quad (8)$$

where υ_k and t_k are received and transmitted symbol, $\mathbf{f}_{r,k}$ and $\mathbf{f}_{t,k}$ are training analog beamformer at the MS and BS and $\mathbf{n}_k$ is additive received vector noise at the kth instant. For representing the sparse characteristics of the channel, we can apply lemma $vec(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})vec(\mathbf{B})$ to (8) from [13] . Thus we can rewrite (8) as

$$\upsilon_k = (\mathbf{f}_{t,k}^T \otimes \mathbf{f}_{r,k}^H)vec(\mathbf{H})\, t_k + \mathbf{f}_{r,k}^H \mathbf{n}_k \quad (9)$$

We can assume that BS sends equal symbols t in separate M time slot with different precoder vectors and also MS received in M different combiners. With such an assumption, MS stacks the M measurements in a vector as

$$\mathbf{y} = \boldsymbol{\Theta}\mathbf{h}t + \boldsymbol{\zeta} \quad (10)$$

where $\mathbf{y} = [\upsilon_1, ..., \upsilon_M]^T$, $\boldsymbol{\Theta} = \left[(\mathbf{f}_{t,1}^T \otimes \mathbf{f}_{r,1}^H), ..., (\mathbf{f}_{t,M}^T \otimes \mathbf{f}_{r,M}^H)\right]^T$, $\mathbf{h} = vec(\mathbf{H})$ and $\boldsymbol{\zeta} = [\mathbf{f}_{r,1}^H \mathbf{n}_1, ..., \mathbf{f}_{r,M}^H \mathbf{n}_M]^T$.

Under the consideration of virtual model of the channel, we can characterize physical channels by joint spatial beams in fixed virtual transmit and receive directions determined by resolution of the arrays [14]. Using this linear model of channel, user channel $\mathbf{H}$ can be modelled as

$$\mathbf{H} = \sum_{m=1}^{N_r}\sum_{n=1}^{N_t} \mathbf{H}_v(m,n)\mathbf{v}_r(\theta_m)\mathbf{v}_t^H(\varphi_n) = \mathbf{U}_r\mathbf{H}_v\mathbf{U}_t^H \quad (11)$$

$\mathbf{U}_r = [\mathbf{v}_r(\hat{\theta}_1), \mathbf{v}_r(\hat{\theta}_2), ..., \mathbf{v}_r(\hat{\theta}_{N_r})]$ is an $N_r \times N_r$ array response matrix Similar to (3) excepting that instead of spatial frequencies $\frac{2\pi}{\lambda}d\sin(\theta_l), l = 1, ..., L$, we substitute the virtual spatial frequencies $\frac{2\pi k}{N_r}, k = 1, 2, ..., N_r$. Similarly, $\mathbf{U}_t = [\mathbf{v}_t(\hat{\varphi}_1), \mathbf{v}_t(\hat{\varphi}_2), ..., \mathbf{v}_t(\hat{\varphi}_{N_t})]$ is an $N_t \times N_t$ array response matrix with virtual spatial frequencies $\frac{2\pi i}{N_t}, i = 1, 2, ..., N_t$. Thanks to these spatial virtual directions, the matrices $\mathbf{U}_r$ and $\mathbf{U}_t$ are full-rank DFT matrices. Therefore, $\mathbf{H}_v$ is unitarily equivalent to $\mathbf{H}$ and captures all of channel information. $\mathbf{H}_v \in C^{N_r \times N_t}$ represents the virtual complex channel matrix and is not generally diagonal. For virtual angles where there is no scattering, the corresponding entries are approximately zero. Note that, the virtual representation does not distinguish between scatterers that are within the spatial resolution.

By vectorising the channel matrix in equation (11), we have

$$\mathbf{h} = vec(\mathbf{H}) = (\mathbf{U}_t^T \otimes \mathbf{U}_r)vec(\mathbf{H}_v) \quad (12)$$

$$= (\mathbf{U}_t^T \otimes \mathbf{U}_r)\mathbf{h}_v = \mathbf{W}\mathbf{h}_v \quad (13)$$

where $\mathbf{W} \in C^{N_t N_r \times N_t N_r}$ is defined as a complex dictionary matrix of the channel and $\mathbf{h}_v \in C^{N_t N_r \times 1}$ represents a sparse vector with L non zero entries as $L << N_r N_t$.

Replacing (13) in the stacked measurement vector in (10) and assuming $t = 1$, we can write,

$$\mathbf{y} = \mathbf{\Theta W h}_v + \zeta \quad (14)$$

$$= \mathbf{\Psi h}_v + \zeta \quad (15)$$

where $\mathbf{\Psi} \in C^{M \times N_t N_r}$ is sensing matrix with the constraint of $M < N_t N_r$. Equation (15) can be seen as a single measurement vector (SMV) compressive sensing problem due to L-level sparse vector $\mathbf{h}_v$. Some of important greedy techniques such as orthogonal Matching Pursuit (OMP) and its derivaions, Hard Iterative Thresholding (IHT) and its extensions have been offered to resolve the SMV sparse problems [15].

When the SNR is very low, which is typical case at mmW systems, we need to enhance the number of measurements comparable to the dimension of unknown sparse vector. To prevent large stacking of the measurements, exploiting Multiple Measurement Vector (MMV) or joint sparsity is proposed. Rather than recovering the K unknown vector, one attempts to simultaneously recover all vectors by finding the row support of the unknown $\mathbf{H}_v$ from the matrix formulation as

$$\mathbf{Y} = \mathbf{\Psi H}_v + \mathbf{E} \quad (16)$$

Where, $\mathbf{H}_v = [\mathbf{h}_{v,1}, ..., \mathbf{h}_{v,K}] \in C^{N_r N_t \times K}$, $\mathbf{E} = [\zeta_1, ..., \zeta_K] \in C^{M \times K}$ and thus $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_K] \in C^{M \times K}$. When predominant nonzero entries of $\mathbf{h}_{v,k}$ are shared in the same locations, MMV algorithms can lead to computational promotion [16]. In this paper, we assume that $\mathbf{H}_v$ is L-row-sparse, i.e., that its row-support, which is defined as $\text{supp}(\mathbf{H}_v) := \{i : \mathbf{H}_{v,(i)} \neq 0\}$, has cardinality at most L.

One of the famous greedy algorithms for solving the MMV problem is Simultaneously Orthogonal Matching Pursuit (SOMP) [17]. However, SOMP is a rank blind approach. Namely, it does not allow for perfect recovery in the full rank case with small number of measurements. Another greedy blind rank algorithms such as Simultaneous Iterative Hard Thresholding (SIHT), Simultaneous Hard Thresholding Pursuit (SHTP) are proposed in [8]. In contrast to high computational complexity of rank-blind methods, MUSIC (Multiple Signal Classification) approach, where is a rank-aware, provides guaranteed recovery in the full row rank cases with the mild complexity.

IV. MUSIC-BASED METHODS

One of the simplest approaches evolves the rank of observations is MUSIC, a popular algorithm in array signal processing which is identical to rank aware Thresholding techniques in the full rank case [18]. However, one of the main disadvantages of the MUSIC technique is its tending to failure under the condition of $\text{rank}(\mathbf{H}_{v,(\chi)}) < L$. To compensate this limitation, one can use a greedy selection algorithm to find $s = L - r$ atoms of dictionary (or equivalently supports) and then apply MUSIC to an augmented data matrix to recognize the rest supports. Motivated by these facts, we investigate rank aware algorithms to improve estimation of unknown channel matrix.

A. mmW channel estimation by MUSIC

Inspired by using of the MUSIC for joint sparse recovery in [19], signal subspace is needed to estimate by eigenvalue decomposition (EVD) of the approximated covariance matrix $\mathbf{YY}^H / K$ or singular value decomposition (SVD) of observation matrix $\mathbf{Y}$. To distinguish between the signal and noise subspace, it is better to truncate the number of eigenvalues to sparsity level of the channel, L. The resulted subspace matrix is constituted of L eigenvectors proportional to L dominant eigenvalues. But MUSIC is tending to failure when $\mathbf{\Psi}_\chi \mathbf{H}_{v,(\chi)}$ is ill-conditioned in the presence of the noise, or when $\mathbf{H}_{v,(\chi)}$ does not have full row rank. If the condition number, defined as the ratio of the largest to the smallest singular value, is large enough then the matrix is said to be ill-conditioned. Also whereas K, the number of snapshots, is smaller than the sparsity level L, then no more than K rows can be linearly independent, and the nonzero rows of unknown matrix turns into rank defective. Correlation between sources or multi-path propagation is another reason that caused to rank defective matrix. Therefore, in such conditions we choose r large eigenvalues, where r is the rank of $\mathbf{\Psi H}_v$, instead of L large eigenvalues where $r < L$. But rank is usually unknown in receiver. To estimate rank, difference between (i)*th* and (i+1)*th* descended sorted eigenvalue of M eigenvalues of $\mathbf{YY}^H / K$ is calculated. Note that calculation is begun to (M-1)*th* and M*th* eigenvalues. If the result of dividing of subtracted value to the largest eigenvalue is less than a predefined threshold then rank is decreased one unit and rank finding procedure will be persist until M-1 computations. After the identification of the rank, signal subspace estimation is performed. Given the signal subspace $\hat{S}$, MUSIC for solving the joint sparse recovery problem accepts each columns of $\mathbf{\Psi}$ if that column is posed on the estimated subspace and continues until L columns is selected. MUSIC method is summarized in algorithm 1 to find the unknown virtual channel matrix. In algorithm 1, function *Eigen* refers to eigenvalue decomposition and returns back a diagonal matrix included eigenvalues.

Algorithm 1 MUSIC

1. EVs =sort(*Eigen*($\mathbf{YY}^H / K$)) degenerately;
2. *if* K<L
3. r=whole of number of EVs larger than zero;
4. *else*
5. ρ = rank estimation from M EVs as explained in section IV.A;
6. r= min (ρ , L);

7. *construct matrix* $\mathbf{U}_{\hat{s}}$ from r eigenvector proportional to r largest eigenvalues;
8. $\mathbf{P}_{\hat{s}} = \mathbf{U}_{\hat{s}}\mathbf{U}_{\hat{s}}^{H}$;
9. $\chi = \{\varnothing\}$;
10. *for* j=1: $N_t N_r$
11. $\kappa_j = \left\|\mathbf{P}_{\hat{s}}\boldsymbol{\Psi}_j\right\|_2 / \left\|\boldsymbol{\Psi}_j\right\|_2$
12. *end for*
13. $\chi = \{$ select L largest elements of vector κ $\}$;
14. $\hat{\mathbf{H}}_{\mathrm{v},(\chi)} = \boldsymbol{\Psi}_{(\chi)}^{\dagger}\mathbf{Y}$;

B. Subspace augmentation methods

The range of the matrix $\boldsymbol{\Psi}$ is nothing more than the space spanned by set of all possible linear combinations columns of $\boldsymbol{\Psi}$ and is denoted by $R(\boldsymbol{\Psi})$. Practically speaking, Signal subspace, $\hat{S}$, is estimated from the finitely many snapshot matrix $\mathbf{Y} = \boldsymbol{\Psi}\mathbf{H}_{\mathrm{v}} + \mathbf{E}$ and exactly computed from $\mathbf{Y}\mathbf{Y}^{\mathrm{H}} / \mathrm{K}$ when $K \to \infty$. Mathematically speaking, signal subspace is defined by

$$S \overset{\Delta}{=} R(\boldsymbol{\Psi}\mathbf{H}_{\mathrm{v}}) = R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)}) \tag{17}$$

When $\mathbf{H}_{\mathrm{v},(\chi)}$ is a full row rank matrix, the range of signal subspace of $\mathbf{Y} = \boldsymbol{\Psi}\mathbf{H}_{\mathrm{v}} + \mathbf{E}$, i.e. $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$, is coincided with $R(\boldsymbol{\Psi}_{\chi})$. If we assume that $\mathbf{H}_{\mathrm{v},(\chi)}$ is a full row rank matrix and SNR is high, then $R(\boldsymbol{\Psi}_{\chi}) = R(\mathbf{Y})$ .Namely, only the columns of $\boldsymbol{\Psi}$ that selected by elements of χ , lie within $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$. Consequently, one can find the value of elements of support by projecting the columns of $\boldsymbol{\Psi}$ into orthogonal subspace of $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$ [9].

However, in practical issue of estimating the rank, due to limitations in SNR, finite number of K or truncation of eigenvalues, estimation of signal subspace is inaccurate. Thus, relation $\left\|\mathbf{P}_{\hat{s}}^{\perp}\boldsymbol{\Psi}_k\right\| = 0$ is not satisfied and we have to minimize $\mathbf{P}_{\hat{s}}^{\perp}\boldsymbol{\Psi}_k$ or equivalently maximize $\mathbf{P}_{\hat{s}}\boldsymbol{\Psi}_k$ in algorithm 1. Also, in the rank defective case, i.e., $\mathrm{rank}(\mathbf{H}_{\mathrm{v},(\chi)}) = r < L$, in spite of exact rank estimation, it may happen that some of columns of sensing matrix as an element of the support not belong to the signal subspace S and thus MUSIC makes a mistake to select a true component of support set. To prohibit the wrong estimating of the support, one can estimate a subspace spanned by $L - r$ columns of $\boldsymbol{\Psi}$ as an augmentation subspace by conventional MMV algorithms like SOMP in a probabilistic way and attach this subspace to the *r*-dimensional subspace obtained by MUSIC method deterministically. The details will be discussed in next section.

B.1 Subspace Augmentation MUSIC-MMV

Suppose that $\mathbf{H}_{\mathrm{v}}$ has L nonzero rows within support $\chi \subset \{1,...,N_t N_r\}$ and also $\mathbf{H}_{\mathrm{v}}$ is rank defective, i.e. $r < L$. Let δ be an arbitrary subset of χ with $L - r$ elements. If we estimate *r*-dimensional subspace from $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$ then we can write,

$$R(\boldsymbol{\Psi}_{\chi}) = \hat{S} + R(\boldsymbol{\Psi}_{\delta}) \tag{18}$$

where $\hat{S}$ is estimated subspace of $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$ which obtains by applying the EVD over $\mathbf{Y}\mathbf{Y}^{\mathrm{H}} / \mathrm{N}$. The goal is to find an augmented signal subspace.

Theory 1: Suppose that $\tilde{S}$ is an augmented signal subspace within $R(\boldsymbol{\Psi}_{\chi})$. If we have equation (18) as $\bar{S} = \hat{S} + R(\boldsymbol{\Psi}_{\delta})$ where $\hat{S}$ is estimated subspace, then projection matrix on augmented signal subspace attains as

$$\mathbf{P}_{\bar{S}} = \mathbf{P}_{\hat{S}} + (\mathbf{P}_{\hat{S}}^{\perp}\boldsymbol{\Psi}_{\delta})(\mathbf{P}_{\hat{S}}^{\perp}\boldsymbol{\Psi}_{\delta})^{\dagger} \tag{19}$$

Proof is in appendix A.

Remember that In SOMP algorithm as a conventional solution for MMV, the selection rule is given by, $l = \arg\max_{k \in [N_t N_r] \setminus \delta} \left\|\mathbf{P}_{R(\boldsymbol{\Psi}_{\delta})}^{\perp}\mathbf{Y}\boldsymbol{\Psi}_{:,k}\right\|_2$.

The key point in algorithm2 is replacing stacked data matrix $\mathbf{Y}$ by the estimated orthonormal signal subspace $\mathbf{U}_{\hat{s}}$ in step 5. That is, SA-MUSIC algorithm in step 5 incrementally updates the partially support by the following selection rule,

$$l = \arg\max_{k \in [N_t N_r] \setminus \delta} \left\|\mathbf{P}_{R(\boldsymbol{\Psi}_{\delta})}^{\perp}\mathbf{U}_{\hat{s}}\boldsymbol{\Psi}_{:,k}\right\|_2 \tag{20}$$

where δ is $\{\varnothing\}$ at the outset of iterations.

The implementation on (19) and (20) seems impractical. We know that the projection onto $R(\boldsymbol{\Psi}_{\chi})$ is achieved by $\mathbf{P}_{R(\boldsymbol{\Psi}_{\chi})} = \boldsymbol{\Psi}_{\chi}\boldsymbol{\Psi}_{\chi}^{\dagger} = \mathbf{U}_1\mathbf{U}_1^{H}$ where $\mathbf{U}_1$ is attained by SVD on $\boldsymbol{\Psi}_{\chi} = [\mathbf{U}_1 \mid \mathbf{U}_2]\boldsymbol{\Sigma}[\mathbf{V}_1 \mid \mathbf{V}_2]^{H}$ [13]. A proper substitution to construct the projection onto $\hat{S} + R(\boldsymbol{\Psi}_{\delta})$ is exploiting the orthonormal basis from $[\mathbf{U}_{\hat{s}} \ \boldsymbol{\Psi}_{\delta}]$ by SVD or QR decompositions. However, SVD doubles the computation time rather than QR factorization, but provides more reliable and consistent rank determination.

Proposition 1: The selection rule of augmented support by SA-MUSIC-MMV is attained as

$\kappa_j = \left\|\bar{\mathbf{U}}_{aug}^{H}\boldsymbol{\Psi}_{\gamma(j)}\right\|_2 / \left\|\boldsymbol{\Psi}_{\gamma(j)}\right\|_2$ where $\gamma = \{1: N_t N_r\} \setminus \delta$ and observed matrix is noisy.

Proof. According to MUSIC algorithm, every column of $\boldsymbol{\Psi}$ belong to support index, lies within the range of $\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)}$. Provided that $\mathbf{U}_s$ is error-free orthonormal basis of $R(\boldsymbol{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$, it can be orthonormal basis of $R(\boldsymbol{\Psi}_{\chi})$. Thus each of columns of $\boldsymbol{\Psi}_{\chi}$ certainly lies within this subspace. In the other hand, we have $\mathbf{P}_s^{\perp} = \mathbf{I} - \mathbf{U}_s\mathbf{U}_s^{H}$ and thus none of selected columns would not lie within $\mathbf{P}_s^{\perp}$ .Therefore one can compute $\left\|\mathbf{P}_s^{\perp}\boldsymbol{\Psi}_k\right\| = \left\|\boldsymbol{\Psi}_k - \mathbf{U}_s\mathbf{U}_s^{H}\boldsymbol{\Psi}_k\right\| = 0$ if and only if $k \in \chi$. Let δ be an index subset of χ with cardinality of

$L-r$ chosen by (20), then $rank([\hat{\mathbf{U}}_s{}^H \boldsymbol{\Psi}_\delta]) = L-r$ where $\hat{\mathbf{U}}_s$ is noise subspace resulted by SVD on $\mathbf{Y}$ and consists of orthonormal columns such that $\hat{\mathbf{U}}_s{}^H \mathbf{Y} = \mathbf{0}$. If we suppose the noisy model of $\mathbf{Y} = \boldsymbol{\Psi}\mathbf{H}_v + \mathbf{E}$, then for every $k \in \gamma = \{1:N_t N_r\} \setminus \delta$ the term of $rank(\hat{\mathbf{U}}_s{}^H [\boldsymbol{\Psi}_\delta, \boldsymbol{\Psi}_{:,k}]) = L-r$ is satisfied if and only if $k \in \chi$. Thus the term of $min \left\| \left[\hat{\mathbf{U}}_s, \boldsymbol{\Psi}_\delta \right]^H \boldsymbol{\Psi}_{\gamma(j)} \right\|_2$ can be appropriate selection rule.

Algorithm 2 describes Subspace Augmentation MUSIC-MMV (SA-MUSIC) as follows,

1. $\delta = \{\varnothing\}$;
2. Continue by step 1 to 8 from algorithm1, exploit $\mathbf{U}_{\hat{s}}$ and r ;
3. $\gamma = \{1:N_t N_r\} \setminus \delta$
4. *for* j=1 : length(γ)
5. $\eta = \arg \max_{k \in [N_t N_r] \setminus \delta} \left\| \mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)} \mathbf{U}_{\hat{s}} \boldsymbol{\Psi}_{:,k} \right\|_2$
6. $\delta = \delta \bigcup \eta$;
7. *end for*
8. $\mathbf{U}_{aug} = [\mathbf{U}_{\hat{s}}; \boldsymbol{\Psi}_\delta]$;find augmentation subspace
9. $\bar{\mathbf{U}}_{aug} = \text{orth}(\mathbf{U}_{aug})$; find orthonormal basis for the range of $\mathbf{U}_{aug}$ by SVD or QR .
10. $\gamma = \{1:N_t N_r\} \setminus \delta$
11. *for* j=1 : length(γ)
12. $\kappa_j = \left\| \bar{\mathbf{U}}^H_{aug} \boldsymbol{\Psi}_{\gamma(j)} \right\|_2 / \left\| \boldsymbol{\Psi}_{\gamma(j)} \right\|_2$
13. *end for*
14. $\chi = \delta \bigcup \{$ r largest elements of κ $\}$;
15. $\hat{\mathbf{H}}_{v,(\chi)} = \boldsymbol{\Psi}^{\dagger}_{(\chi)} \mathbf{Y}$;

$\mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)}$ is perpendicular complementary projection on range of $\boldsymbol{\Psi}_\delta$.

B.2 Two stage orthogonal subspace matching pursuit (TOSMP)

In this section, we present an mmW channel estimation method based on subspace and MMV concepts no need to augmentation. At the primary stage, the algorithm finds $M-1$ the most effective columns of $\boldsymbol{\Psi}$ by partial orthogonal subspace matching pursuit (POSMP) method introduced in [11] and then calculates the primary estimation of $\mathbf{H}_v$. POSMP is a modified version of [20], and its criterion is based on minimization the angle between spaces $\mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)} \mathbf{U}_{\hat{s}}$ and $\mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)} R(\boldsymbol{\Psi}_\delta)$.

Definition 1 [21]: the angle between two subspace S_1 and S_2 is as follows $\angle(S_1, S_2) = \sin^{-1}(\min\{\|P^{\perp}_{S_1} P_{S_2}\|, \|P^{\perp}_{S_2} P_{S_1}\|\})$.

In the secondary stage, the L nonzero rows of estimated channel matrix $\mathbf{H}_v$ via the best row-norm, is recognized. When the rank is deficient, i.e. $rank(\mathbf{H}_v) = r \le L$, it is proven that the TOSMP outperforms the SA-MUSIC specially, under the condition $K < r$.

Algorithm 3 represents the TOSMP method.

1. $\delta = \{\varnothing\}$;
2. Continue by step 1 to 8 from algorithm1, exploit $\mathbf{U}_{\hat{s}}$ and r ;
3. $\gamma = \{1:N_t N_r\} \setminus \delta$
4. *for* j=1 : M-1
5. $\eta = \arg \max_{k \in [N_t N_r] \setminus \delta} \left\| \mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)} \mathbf{U}_{\hat{s}} \mathbf{P}^{\perp}_{R(\boldsymbol{\Psi}_\delta)} \boldsymbol{\Psi}_{:,k} \right\|_2$
6. $\delta = \delta \bigcup \eta$;
7. *end for*
8. $\hat{\mathbf{H}}_{v,(\delta)} = \boldsymbol{\Psi}^{\dagger}_{(\delta)} \mathbf{Y}$
9. *for* j=1 : length(δ)
10. $\kappa_j = \left\| \hat{\mathbf{H}}_{v,\delta(j),:} \right\|_2$
11. *end for*
12. $\chi = \{$ L largest elements of κ $\}$;
13. $\hat{\mathbf{H}}_{v,(\chi)} = \boldsymbol{\Psi}^{\dagger}_{(\chi)} \mathbf{Y}$;

C. The Spectral efficiency

After estimation of virtual channel and thus extracting estimation of mmW channel, thanks to assumption of the Gaussian signalling over the link in (2), we are able to achieve the estimated spectral efficiency as follow,

$$\hat{R} = \log_2 \left(\left| \mathbf{I}_{N_s} + \frac{P_s}{N_s} \hat{\mathbf{R}}_n^{-1} \hat{\mathbf{H}}_{eff} \hat{\mathbf{H}}^H_{eff} \right| \right) \qquad (21)$$

where $\hat{\mathbf{H}}_{eff} = \hat{\mathbf{C}}^H \hat{\mathbf{H}} \hat{\mathbf{P}}$ is effective estimated channel with $\hat{\mathbf{C}} = \hat{\mathbf{F}}_r \hat{\mathbf{G}}_r$ and $\hat{\mathbf{P}} = \hat{\mathbf{F}}_t \hat{\mathbf{G}}_t$ as estimated combiner and precoder matrices respectively. Also, $\hat{\mathbf{R}}_n = \sigma_n^2 \hat{\mathbf{C}} \hat{\mathbf{C}}^H$ is the estimated combined noise covariance matrix in the downlink. Considering the estimated virtual channel from one of the proposed algorithms and using (11), the estimated mmW channel model can be written

$$\hat{\mathbf{H}} = \mathbf{U}_r \hat{\mathbf{H}}_{v,(\chi)} \mathbf{U}_t^H \qquad (22)$$

By applying the singular value decomposition on $\hat{\mathbf{H}}$ in (22) and choosing the first N_s columns of left- handed unitary matrix, i.e., $\hat{\mathbf{U}}$, and choosing the first N_s columns of right-handed unitary matrix, i.e., $\hat{\mathbf{V}}$, and finally replacing them rather than $\hat{\mathbf{C}}$ and $\hat{\mathbf{P}}$ respectively, we can design $\hat{\mathbf{F}}_r$, $\hat{\mathbf{G}}_r$, $\hat{\mathbf{F}}_t$ and $\hat{\mathbf{G}}_t$ by

Fig.2 Normalized Minimum Mean Square Error of Estimated Channel

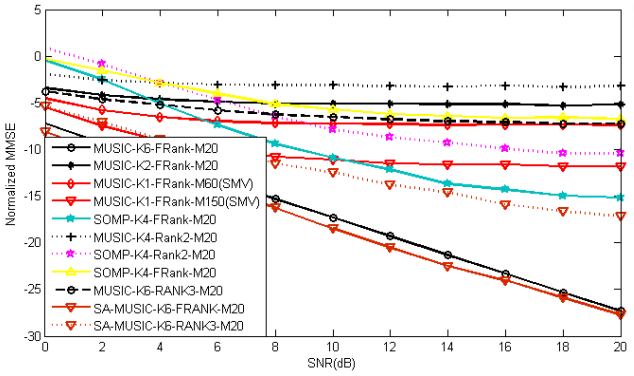


Fig.3. Spectral Efficiency of Estimated mmW-Channel

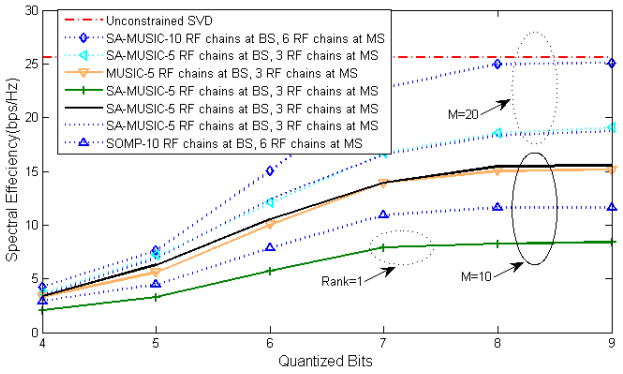


Fig.4 NMSE versus SNR for
N_t =32,N_r =8,L=5,M=40,different-rank and K

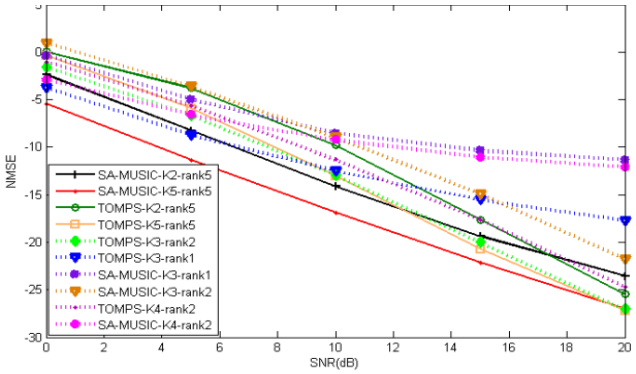
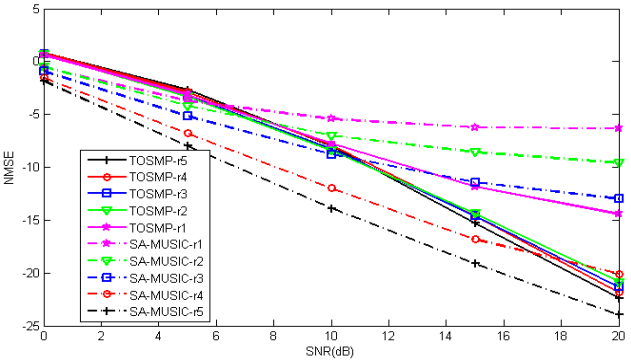


Fig.5 NMSE versus SNR for
N_t =32,N_r =8,L=8,M=40,K=5,different-rank



solving the general sparse optimization problem as follow,

$$\left(\hat{\mathbf{F}}_x,\hat{\mathbf{G}}_x\right)=\arg\min\ \left\|\hat{\mathbf{W}}_x-\hat{\mathbf{F}}_x\hat{\mathbf{G}}_x\right\|_F,\qquad \text{st.}$$

$$\left[\hat{\mathbf{F}}_x\right]_{:,l}\ and\ \left[\hat{\mathbf{G}}_x\right]_{:,l}\in\left\{\left[\mathbf{C}_x\right]_{:,i}\ 1\le i\le \mathrm{N}_{\mathrm{x,cod}}\right\},l=1,2,...,\mathrm{N}_{\mathrm{x,rf}}$$

$$\left\|\hat{\mathbf{F}}_x\hat{\mathbf{G}}_x\right\|_F^2=\mathrm{N_s}\qquad(23)$$

where subscript X can be substituted for t or r, $\hat{\mathbf{W}}_x$ can be replaced by $\hat{\mathbf{U}}$ or $\hat{\mathbf{V}}$ dependent on X and $\mathbf{C}_x$ is a general codebook included of quantized steering vectors for transmitter or receiver and is chosen from

$$\left\{\begin{array}{l}\mathbf{v}(\frac{k\pi}{\mathrm{N}_Q}),k=0,1,...,\mathrm{N}_{\mathrm{x,cod}}-1,\mathrm{N}_Q=2^Q\\ \mathbf{v}_x(\vartheta)=\left[1,\mathrm{e}^{j\pi\cos(\vartheta)},...,\mathrm{e}^{j\pi(N_x-1)\cos(\vartheta)}\right]^{\mathrm{T}}/\sqrt{\mathrm{N}_x}\end{array}\right\}\qquad(24)$$

Q in (23) is number of bits for controlling the phase shifters and $\mathrm{N}_{\mathrm{x,cod}}$ is number of steering vectors existing in transmitter/receiver codebook.

V. SIMULATION RESULTS

In this section, we evaluate numerical results of proposed algorithms to estimate a typical millimetre-Wave MIMO channel with hybrid precoding structure and compare their performance to conventional MMV and SMV problems.

BS and MS are equipped with 64 and 32 miniature antenna and 10 and 6 RF chains respectively. We assume that scatterer number is 6 independent on indoor or outdoor environment.

Each scatterer as a cluster further assumed to contribute a single propagation path between the BS and MS. RF phase shifters in analog parts of precoder and combiner are able to be controlled with 7 quantization bits. The operational carrier frequency of the system is 28GHz with consideration of bandwidth of 100MHz. the path-loss exponent is assumed to be $\beta_{loss}=3.5$. The angles of arrival and departure are selected randomly with a uniform distribution from range of $[0,2\pi]$.

In Fig.2, we evaluate the NMSE behaviour of virtual channel estimation in the proposed algorithms for the cases $L=6$ as multipath number or sparsity level of channel and set $K=2$ to 6 while number of training beam vectors, i.e. M is held to constant to 20. The impact of K on the performance of the proposed algorithms especially on measurement in an environment with constant AoDs and AoAs and variable channel gain, has incremental growth. When K=1, the algorithm turns into the SMV and needs to increase the number of measurements, i.e., M, for better results. However, enlarging of the measurement vector increases the time of solving the problem exhaustingly.

The performance of proposed SA-MUSIC (MMV is selected here to be SOMP) with deployments of M=20 and decreasing rank to 3 outperformances than the Full-Rank MUSIC and SOMP.

Fig.6 Impact of M,K,rank over NMSE

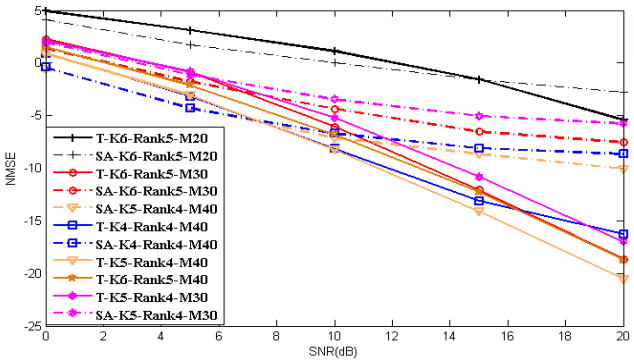


Fig.7 Success rate of support extraction

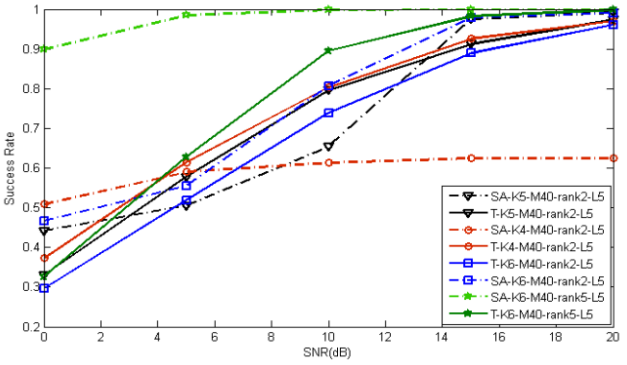


Fig.8 Spectral efficiency versus Quantized bit numbers of phase shifter

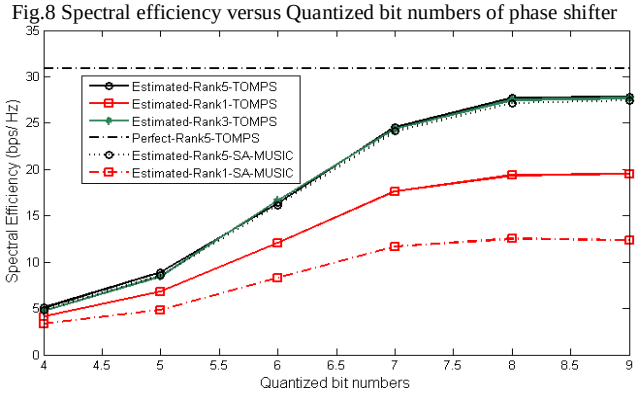
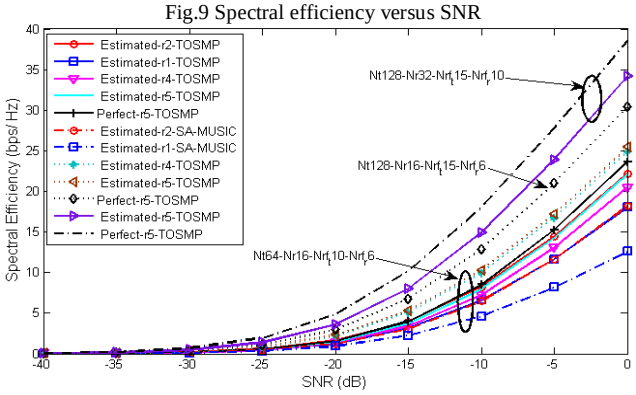


Fig.9 Spectral efficiency versus SNR



The performance of MUSIC has more improvement when K, i.e. number of columns of In Fig.2. The MS antenna number with different RF chains and quantized bits in phase shifters, and compared with the spectral efficiency of the perfect channel.

In Fig.4, performance of the algorithms TOSMP and SA-MUSIC under the metric of NMSE is examined. As is clear of figure, TOMSP outperforms SA-MUSIC chiefly in low rank condition. The arrangement for this simulation is N_t =32,N_r =8,L=5-M40 for different rank and different measurement K. As shown in this figure, SA-MUSIC fails when $rank(\mathbf{H}_v) < K$. It means that TOSMP is having less sensitive to rank-deficiency. Another simulation is tested for N_t =32,N_r =8,L=5,M40,K=5 and different ranks in Fig.5. Since K<L, $rank(\mathbf{H}_v)$ is limited to K. When $rank(\mathbf{H}_v) < K$, TOMPS outperformance SA-MUSIC. However in condition of $K \le rank(\mathbf{H}_v) < L$ SA-MUSIC can estimate augmented subspace properly. Impact of variations over M, K and rank investigate in Fig.6. in the case K=6 and rank=5, increase of M from 20 to 40 improves performance both of SA-MUSIC and TOSMP but, TOSMP outperforms SA-MUSIC specially in moderate and high SNRs.

Fig.7 shows success rate comparison of two mentioned algorithms for fixed deployments of N_t =32,N_r =8,L=5-M40 . Success rate is a metric for recognizing the percentage of true support extraction. For the full rank case with sufficient measurements, SA-MUSIC has a better behaviour. But for the rank-defective case, TOSMP outperforms for middle to high SNRs.

spectral efficiency in terms of quantized bit number of phase shifters is represented in Fig.3. The desired

We also evaluate the performance of spectral efficiency versus Quantized bit numbers of phase shifter in Fig.8. By adopting the structure of N_t =64,N_r =32,L=8,M40 , K=5,N_{rf}^t =10 and $N_{rf}^r = 6$, comparison between two algorithm for different quantized bit numbers of phase shifters have been done. The simulation result represents that TOSMP is robust under rank-defective conditions.

Fig.9 shows the performance of spectral efficiency of hybrid precoder and combiner system versus SNR under the condition of L=8,M40,K=5,N_{rf}^t =10 and $N_{rf}^r = 6$ with different N_t , N_{rf}^t , N_r and N_{rf}^r . This figure clearly shows that TOMPS is more efficient than SA-MUSIC under the condition of $rank < K$. It is also shown the role of transceiver antenna gain and size in improvement of spectral efficiency.

VI. CONCLUSION

We have explored the ability of MUSIC algorithm based on subspace augmentation for Rank-defective mmW-channel estimation with large antenna array. We introduce robust TOSMP algorithm based on orthogonal matching pursuit inside of SA-algorithm when number of measurements is lower than the rank. However, TOSMP has higher computational complexity than SA-MUSIC. Numerical results show that SA-MUSIC provides good a spectral efficiency and normalized

MSE than the MUSIC and conventional MMV methods specially, when the rank is deficient. for compensating of such shortage, we proposed TOSMP algorithm that outperforms SA-MUSIC when the measurement numbers is smaller than the rank. It would be interesting to extract the hybrid precoding/combining of rank-defective multi-user mmW according to some studied articles such as [22] for the future works.

Appendix A. Proof OF Theory 1

By assuming $\mathbf{U}_{\bar{S}} = orth[\mathbf{U}_{\hat{S}}, \mathbf{\Psi}_{\delta}]$ as an enhanced basis matrix with L columns within $\Omega(\mathbf{\Psi}_{\chi})$ such that $\mathbf{U}_{\hat{S}}$ as is a $r \times r$ estimated signal subspace within $\Omega(\mathbf{\Psi}\mathbf{H}_{\mathrm{v}}) = \Omega(\mathbf{\Psi}_{\chi}\mathbf{H}_{\mathrm{v},(\chi)})$ we have,

$$\Omega(\mathbf{U}_{\bar{S}}) = \Omega[\mathbf{U}_{\hat{S}}, \mathbf{\Psi}_{\delta}] \qquad \text{(A.1)}$$

On the other hand, $\Omega(\mathbf{U}_{\hat{S}})^{\perp} = \Omega(\mathbf{Q}_{\hat{S}})$ where $\mathbf{Q}_{\hat{S}}$ is an estimation of noise subspace of $\mathbf{Y}\mathbf{Y}^{\mathrm{H}}/\mathrm{K}$. By applying the projection update rule on (A.1) we have,

$$\Omega(\mathbf{U}_{\bar{S}}) = \Omega(\mathbf{U}_{\hat{S}}) + \Omega(\mathbf{P}_{\mathbf{U}_{\hat{S}}}^{\perp}\mathbf{\Psi}_{\delta}) = \Omega(\mathbf{\Psi}_{\delta}) + \Omega(\mathbf{P}_{\mathbf{\Psi}_{\delta}}^{\perp}\mathbf{U}_{\hat{S}}) \qquad \text{(A.2)}$$

Since $\mathbf{P}_{\mathbf{Q}_{\hat{S}}}$ is equivalent to $\mathbf{I} - \mathbf{P}_{\mathbf{U}_{\hat{S}}} = \mathbf{P}_{\mathbf{U}_{\hat{S}}}^{\perp}$ then we have,

$$\Omega(\mathbf{U}_{\bar{S}}) = \Omega(\mathbf{U}_{\hat{S}}) + \Omega(\mathbf{P}_{\mathbf{Q}_{\hat{S}}}\mathbf{\Psi}_{\delta}) \qquad \text{(A.3)}$$

By applying the projection operator in both of equation (A.2) and considering $\bar{S}$ and $\hat{S}$ instead of $\mathbf{U}_{\bar{S}}$ and $\mathbf{U}_{\hat{S}}$ respectively and knowing that $\mathbf{P}_{\mathbf{G}} = \mathbf{G}\mathbf{G}^{\dagger}$ for an arbitrary matrix of $\mathbf{G}$, we have

$$\mathbf{P}_{\bar{S}} = \mathbf{P}_{\Omega[\mathbf{U}_{\hat{S}}\mathbf{\Psi}_{\delta}]} = \mathbf{P}_{\Omega(\mathbf{U}_{\hat{S}})} + \mathbf{P}_{\Omega(\mathbf{P}_{\Omega(\mathbf{U}_{\hat{S}})}^{\perp}\mathbf{\Psi}_{\delta})} \qquad \text{(A.4)}$$
$$= \mathbf{P}_{\hat{S}} + (\mathbf{P}_{\hat{S}}^{\perp}\mathbf{\Psi}_{\delta})(\mathbf{P}_{\hat{S}}^{\perp}\mathbf{\Psi}_{\delta})^{\dagger}$$

In practical issues, one can calculate $\mathbf{P}_{\bar{S}}$ by considering of the unitary matrix part of QR decomposition on $[\mathbf{U}_{\hat{S}}\mathbf{\Psi}_{\delta}]$ since the columns of orthogonal projector matrix can be obtained from any set of orthonormal vectors onto $\Omega[\mathbf{U}_{\hat{S}}, \mathbf{\Psi}_{\delta}]$.

REFERENCES

[1] R. C. Daniels, J. N. Murdock, T. S. Rappaport, and R. W. Heath, "60 GHz wireless: Up close and personal," *IEEE Microwave Magazine*, vol. 11, pp. 44-50, 2010.

[2] M. R. Akdeniz, L. Yuanpeng, M. K. Samimi, S. Shu, S. Rangan, T. S. Rappaport, *et al.*, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *Selected Areas in Communications, IEEE Journal on*, vol. 32, pp. 1164-1179, 2014.

[3] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *Communications Magazine, IEEE*, vol. 52, pp. 186-195, 2014.

[4] H. Sooyoung, K. Taejoon, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter Wave Beamforming for Wireless Backhaul and Access in Small Cell Networks," *Communications, IEEE Transactions on*, vol. 61, pp. 4391-4403, 2013.

[5] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 1499-1513, 2014.

[6] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, pp. 831-846, 2014.

[7] H. Ghauch, M. Bengtsson, T. Kim, and M. Skoglund, "Subspace estimation and decomposition for hybrid analog-digital millimetre-wave MIMO systems," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2015, pp. 395-399.

[8] J. D. Blanchard, M. Cermak, D. Hanle, and J. Yirong, "Greedy Algorithms for Joint Sparse Recovery," *Signal Processing, IEEE Transactions on*, vol. 62, pp. 1694-1704, 2014.

[9] K. Jong Min, L. Ok Kyun, and Y. Jong Chul, "Compressive MUSIC: Revisiting the Link Between Compressive Sensing and Array Signal Processing," *Information Theory, IEEE Transactions on*, vol. 58, pp. 278-301, 2012.

[10] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *Information Theory, IEEE Transactions on*, vol. 58, pp. 1135-1146, 2012.

[11] K. Lee, Y. Bresler, and M. Junge, "Subspace methods for joint sparse recovery," *Information Theory, IEEE Transactions on*, vol. 58, pp. 3613-3641, 2012.

[12] M. S. Dastgahian and H. Khoshbin, "Rank-defective Millimeter-Wave Channel Estimation Based on Subspace-Compressive Sensing," *Digital Communications and Networks*, 2016.

[13] C. D. Meyer, *Matrix analysis and applied linear algebra*: Siam, 2000.

[14] A. M. Sayeed, "Deconstructing multiantenna fading channels," *Signal Processing, IEEE Transactions on*, vol. 50, pp. 2563-2579, 2002.

[15] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*: Cambridge University Press, 2012.

[16] S. Foucart, "Recovering jointly sparse vectors via hard thresholding pursuit," *Proc. Sampling Theory and Applications (SampTA)],(May 2-6 2011),* 2011.

[17] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing,* vol. 86, pp. 589-602, 2006.

[18] R. Gribonval, H. Rauhut, K. Schnass, and P. Vandergheynst, "Atoms of all channels, unite! Average case analysis of multi-channel sparse recovery using greedy algorithms," *Journal of Fourier analysis and Applications,* vol. 14, pp. 655-687, 2008.

[19] P. Feng, "Universal minimum-rate sampling and spectrum-blind reconstruction for multiband signals," University of Illinois at Urbana-Champaign, 1998.

[20] J. D. Blanchard and M. E. Davies, "Recovery Guarantees for Rank Aware Pursuits," *Signal Processing Letters, IEEE,* vol. 19, pp. 427-430, 2012.

[21] P. Å. Wedin, "On angles between subspaces of a finite dimensional inner product space," in *Matrix Pencils*, ed: Springer, 1983, pp. 263-285.

[22] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and S. Li, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Transactions on Wireless Communications,* vol. 15, pp. 7501-7516, 2016.

Majid Shakhsi Dastgahian received the B.S. degree in electronic and completed M.S. degree in communication from Ferdowsi University of Mashhad, Iran, in 2003 and 2011 respectively. He is currently pursuing the Ph.D. degree with the school of communication Engineering, Ferdowsi University of Mashhad, Iran. He has been working as a manager of electrical laboratories and a lecturer at Imam Reza International University since 2005. His current research interests are focused on the area of 5G, application of compressed sensing and Matrix completion in next-generation signal processing for mobile and wireless communications, e.g., indoor and outdoor millimeter wave channel estimation.

Hossein Khoshbin received the B.Sc. degree in electronics engineering and the M.Sc. degree in communications engineering in 1985 and 1987, respectively, both from Isfahan University of Technology, Isfahan, Iran. He received the Ph.D. degree in communications engineering from the University of Bath, United Kingdom, in 2000. He is currently an Associative Professor at the Department of Electrical and Computer Engineering, Ferdowsi University, Mashhad, Iran. His research interests include communication theory, digital and wireless communications.