Full Rank Spatial Channel Estimation at Millimeter Wave Systems

Hsiao-Lan Chiang, Wolfgang Rave, Tobias Kadur, and Gerhard Fettweis
Vodafone Chair for Mobile Communications
Technische Universität Dresden
Helmholtzstrasse 18, Dresden, Germany
Email: {hsiao-lan.chiang, rave, tobias.kadur, gerhard.fettweis}@ifn.et.tu-dresden.de

***Abstract*—Spatial channel estimation (CE) for hybrid beamforming architecture is one of the crucial challenges for millimeter wave (mmW) links. The limited beamwidth makes the CE incapable of resolving multiple channel paths from the same cluster when the angular spread is less than or equal to the beamwidth. Therefore, how to efficiently obtain reliable links which provide maximum spatial degree-of-freedom gain for data transmission is the key technology. To this end, a novel CE method to estimate full rank array response matrices is presented. The full rank array response matrices lead to the full rank reconstructed channel matrix, which provides the maximum spatial degree-of-freedom gain of the sparse mmW channels. Compared with the previously proposed CE methods which use transmit or receive diversity rather than spatial degrees of freedom, the proposed one shows significant increase of data rates as SNR is getting larger.**

I. Introduction

With the rapid increase of data rates in wireless communication, bandwidth shortage is getting more critical. As a result, there is a growing interest in using millimeter wave (mmW) for future wireless communications taking advantage of the enormous amount of available spectrum [1]. Measurements of most large and small scale parameters for mmW backhaul and access links have been presented in [2][3]. It seems that path loss in such environment is very severe, and in order to improve capacity and service quality, mmW together with large antenna array is a promising approach [4].

At mmW systems, a combination of analog beamforming (ABF, operating in passband) [5][6] and digital beamforming (DBF, operating in baseband) is one of the low-cost solutions to higher data rate transmission. This combination is commonly denoted as hybrid beamforming (HBF) [7]. ABF is typically implemented at each antenna using a finite set of possible phase shifts, where one entry can be selected from a codebook to produce a beam pattern with a specific main lobe direction. A hybrid beam can be regarded as a linear combination of multiple analog beams with the coefficients defined in the DBF, which are to be optimally combined according to some criteria, e.g., to maximize mutual information.

The problem of spatial channel estimation (CE) is complicated by the fact that only steering angles can be changed by phase shifters equipped in ABF structure. For this reason, to obtain full channel information means to obtain the coupling of ABFs across the channel and therefore exhaustive analog beam training is required in the beginning [8]. Typically, the problem of CE can be formulated as an l_0-norm optimization problem [9] since mmW channels consist of a quite small number of significant paths, meaning that angles of departure and arrival (AoDs and AoAs) are sparse in the angular domain.

Orthogonal matching pursuit (OMP) [10] offers a solution to l_0-norm optimization problems of CE. The columns in the codebook used in OMP is widely assumed to be mutual incoherent or orthogonal, and each column is the coupling of one predefined departure array response vector and one predefined arrival array response vector. In general, it is difficult for these predefined array response vectors to realize multiple paths from the same cluster if the angular spreads are small, for example, less than the half-power beamwidth (HPBW). In such a case, either the estimated departure or the estimated arrival array response vectors could be repeated. Therefore, the reconstructed channel matrix may only provide transmit or receive diversity. However, neither provides fewer data rates than *spatial degrees of freedom* [11], which exploits multiple resolvable channel paths for data transmission.

This paper presents a novel CE method to obtain maximum *spatial degree-of-freedom gain*. By enforcing the rank of the estimated array response matrices, the reconstructed channel by these estimated matrices can ensure to be full rank. Consequently, the reconstructed channel matrix can offer the maximum spatial degree-of-freedom gain. Compared with the non-full rank channel matrices, the simulation results by the proposed full rank CE show significant increase of data rates.

This paper is organized as follows: Section II describes the systems and the mmW channel models. Section III introduces the non-full rank CE, the motivation of the full rank CE, and how to obtain it. Simulation results are presented in Section IV, and we conclude our work in Section V.

We use the following notations throughout this paper. a is a scalar, $\mathbf{a}$ is a vector, and $\mathbf{A}$ is a matrix. a_i is the i^{th} entry of $\mathbf{a}$; $a_{i,j}$ or $(\mathbf{A})_{i,j}$ is the $(i,j)^{\text{th}}$ entry of $\mathbf{A}$; $\mathbf{a}(i)$ is the i^{th} column vector of $\mathbf{A}$; and $\mathbf{A}(1:N)$ denotes the first N column vectors $[\mathbf{a}(1), \cdots, \mathbf{a}(N)]$ of $\mathbf{A}$. $\mathbf{A}^*$, $\mathbf{A}^H$, and $\mathbf{A}^T$ denote complex conjugate, Hermitian transpose and transpose of $\mathbf{A}$, respectively. $\|\mathbf{a}\|_0$ is the l_0-norm of $\mathbf{a}$ [12]; $\|\mathbf{a}\|_2$ is the 2-norm of $\mathbf{a}$; $|\mathbf{A}|$ and $\|\mathbf{A}\|_F$ are the determinant and Frobenius norm of $\mathbf{A}$, respectively. $\text{diag}(\mathbf{A})$ is the vector formed by the diagonal elements of $\mathbf{A}$; $\text{vec}(\mathbf{A})$ is vectorization

978-1-5090-2061-4/16/$31.00 ©2016 European Union

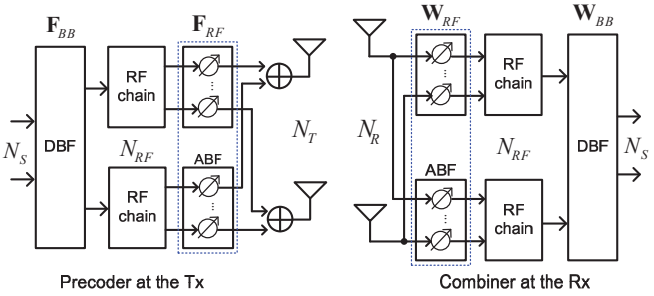

Figure 1. The beamforming system diagram, where each RF chain includes digital-to-analog converter (DAC) at the transmitter and analog-to-digital converter (ADC) at the receiver, and each ABF consists of multiple phase shifters connecting to one RF chain.

of $\mathbf{A}$; rank($\mathbf{A}$) is the (column) rank of $\mathbf{A}$; and $[\mathbf{A} \,|\, \mathbf{B}]$ denotes horizontal concatenation. $\mathbf{I}_N$ is the $N \times N$ identity matrix.

II. System Models

In the assumed mmW systems [13], uplink and downlink transmission timing between a transmitter and a receiver is synchronized. Fig. 1 shows a single link where the transmitter with N_T antennas communicates N_S data streams to the receiver with N_R antennas. The transmitter is equipped with a precoder, which consists of an $N_T \times N_{RF}$ ABF matrix $\mathbf{F}_{RF}$ and an $N_{RF} \times N_S$ DBF matrix $\mathbf{F}_{BB}$ [7]. The precoder steers N_S hybrid beams, and each hybrid beam is formed by a weighted combination (defined in DBF) of the N_{RF} analog beams. The N_{RF} analog beams at the transmitter are selected from a codebook $\mathcal{F} = \{\tilde{\mathbf{f}}_{RF}(l) = \text{the } l^{\text{th}} \text{ column of } \tilde{\mathbf{F}}_{RF}, l = 1, \cdots, N_F\}$, where $\tilde{\mathbf{F}}_{RF} \in \mathbb{C}^{N_T \times N_F}$ is given by

$$\tilde{\mathbf{F}}_{RF} = \frac{1}{\sqrt{N_T}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j2\pi \frac{1}{N_F}} & \cdots & e^{j2\pi \frac{N_F-1}{N_F}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j2\pi \frac{N_T-1}{N_F}} & \cdots & e^{j2\pi \frac{(N_F-1)(N_T-1)}{N_F}} \end{bmatrix}. \tag{1}$$

The power constraint on each hybrid beam is enforced by $\|\mathbf{F}_{RF}\mathbf{f}_{BB}(n_s)\|_2^2 = 1, n_s = 1, \cdots, N_S$, where $\mathbf{f}_{BB}(n_s)$ is the n_s^{th} column of $\mathbf{F}_{BB}$. At the receiver, the codebook $\mathcal{W} = \{\tilde{\mathbf{w}}_{RF}(k) = \text{the } k^{\text{th}} \text{ column of } \tilde{\mathbf{W}}_{RF}, k = 1, \cdots, N_W\}$ is used in the ABF, and $\tilde{\mathbf{W}}_{RF} \in \mathbb{C}^{N_R \times N_W}$ has the similar structure as (1). The same number N_{RF} of the RF devices and the same number N_S of the data streams are assumed at the receiver.

The received signal after the combiner $\mathbf{W}_{RF}\mathbf{W}_{BB}$ can be written as

$$\mathbf{r} = \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s} + \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{z}, \tag{2}$$

where $\mathbf{r} \in \mathbb{C}^{N_S \times 1}$ is the combined received data vector, $\mathbf{s} \in \mathbb{C}^{N_S \times 1}$ is the transmitted data vector satisfying $\mathrm{E}[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_S}\mathbf{I}_{N_S}$, and $\mathbf{z}$ is an $N_R \times 1$ additive white Gaussian noise (AWGN) vector, where the elements are complex Gaussian random variables, $z_{n_r} \sim \mathcal{CN}(0, \sigma_z^2)$, $n_r = 1, \cdots, N_R$.

mmW channels are different to Rayleigh or Rician fading channel models, which are often assumed for centimeter wave communications. One key difference is its sparsity in spatial frequency domain [3][4]. Thus, the channel matrix $\mathbf{H}$ in (2) can be modeled as the sum of outer products of the array response vectors associated with the N_P paths [11]

$$\begin{aligned} \mathbf{H} &= \sum_{p=1}^{N_P} \alpha_p \mathbf{a}_A(p) \mathbf{a}_D^H(p) \\ &= \underbrace{[\mathbf{a}_A(1), \cdots, \mathbf{a}_A(N_P)]}_{\mathbf{A}_A} \underbrace{\begin{bmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_{N_P} \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \mathbf{a}_D(1)^H \\ \vdots \\ \mathbf{a}_D(N_P)^H \end{bmatrix}}_{\mathbf{A}_D^H} \\ &= \mathbf{A}_A \mathbf{D} \mathbf{A}_D^H, \end{aligned} \tag{3}$$

which consists of three factors: the arrival array response matrix $\mathbf{A}_A \in \mathbb{C}^{N_R \times N_P}$, the diagonal matrix $\mathbf{D} \in \mathbb{C}^{N_P \times N_P}$ with N_P complex attenuation coefficients $\{\alpha_p\}$, and the departure array response matrix $\mathbf{A}_D \in \mathbb{C}^{N_T \times N_P}$. Each array response vector in $\mathbf{A}_D$ can be expressed as [6]

$$\mathbf{a}_D(p) = \frac{1}{\sqrt{N_T}} \left[1, e^{j2\pi\lambda_0^{-1}\sin\phi_{D,p}\Delta_d}, \cdots, e^{j2\pi\lambda_0^{-1}\sin\phi_{D,p}(N_T-1)\Delta_d}\right]^T, \tag{4}$$

where $\phi_{D,p}$ stands for the p^{th} AoD, $\Delta_d = \frac{\lambda_0}{2}$ is the distance between two antennas, and λ_0 is the wavelength at the carrier frequency. Each array response vector in $\mathbf{A}_A$ has the similar form as (4). Assume Gaussian signaling [14], given $\mathbf{F}_{BB}$, $\mathbf{F}_{RF}$, $\mathbf{W}_{RF}$, $\mathbf{W}_{BB}$, the system throughput is shown as

$$C = \log_2 \left| \mathbf{I}_{N_S} + \frac{1}{N_S} \cdot \mathbf{R}_z^{-1} \cdot \left(\mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB}\right) \cdot \left(\mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB}\right)^H \right|, \tag{5}$$

where $\mathbf{R}_z = \sigma_z^2 \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{W}_{RF} \mathbf{W}_{BB}$ is the noise covariance matrix after combining.

III. Spatial Channel Estimation

To simplify the CE problem, let us assume that the DBFs at the transmitter and the receiver initially operates with $\mathbf{F}_{BB} = \mathbf{I}$ and $\mathbf{W}_{BB} = \mathbf{I}$. To obtain the observations for CE, all the combinations of the columns in the codebooks $\mathcal{F}$ and $\mathcal{W}$ are trained by using a known training sequence $\{s_1, \ldots, s_T\}$ of length T (the coherence time of the channel is assumed to be sufficiently long). After removing the training signals, the average of the noisy observations for the coupling

coefficient with a single combination of $\tilde{\mathbf{f}}_{RF}(l)$ and $\tilde{\mathbf{w}}_{RF}(k)$ can be written as

$$\begin{aligned} y_{k,l} &= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{s_t}\left(\tilde{\mathbf{w}}_{RF}^H(k)\mathbf{H}\tilde{\mathbf{f}}_{RF}(l)\,s_t + \tilde{\mathbf{w}}_{RF}^H(k)\mathbf{z}(t)\right) \\ &= \tilde{\mathbf{w}}_{RF}^H(k)\mathbf{H}\tilde{\mathbf{f}}_{RF}(l) + \frac{1}{T}\sum_{t=1}^{T}\frac{1}{s_t}\tilde{\mathbf{w}}_{RF}^H(k)\mathbf{z}(t) \qquad (6) \\ &= \tilde{\mathbf{w}}_{RF}^H(k)\mathbf{H}\tilde{\mathbf{f}}_{RF}(l) + z'_k, \\ k &= 1,\cdots,N_W, l = 1,\cdots,N_F, \end{aligned}$$

where z'_k is the averaged noise coupling with $\tilde{\mathbf{w}}_{RF}(k)$. Collecting the observations associated with all the combinations of $\tilde{\mathbf{w}}_{RF}(k)$ and $\tilde{\mathbf{f}}_{RF}(l)$ in a matrix, we obtain

$$\begin{aligned} \mathbf{Y} &= \begin{bmatrix} y_{1,1} & \cdots & y_{1,N_F} \\ \vdots & \ddots & \vdots \\ y_{N_W,1} & \cdots & y_{N_W,N_F} \end{bmatrix} \qquad (7) \\ &= \tilde{\mathbf{W}}_{RF}^H\mathbf{H}\tilde{\mathbf{F}}_{RF} + \mathbf{Z}', \end{aligned}$$

where $\mathbf{Z}'$ is the $N_W \times N_F$ AWGN matrix coupling with $\tilde{\mathbf{W}}_{RF}$. Vectorizing (7) and using rules for Kronecker product, it becomes

$$\begin{aligned} \mathbf{y}_V &= \mathrm{vec}(\mathbf{Y}) \\ &= (\tilde{\mathbf{F}}_{RF}^T \otimes \tilde{\mathbf{W}}_{RF}^H)\mathrm{vec}(\mathbf{H}) + \mathrm{vec}(\mathbf{Z}') \\ &= (\tilde{\mathbf{F}}_{RF}^T \otimes \tilde{\mathbf{W}}_{RF}^H)\mathrm{vec}(\mathbf{A}_A\mathbf{D}\mathbf{A}_D^H) + \mathrm{vec}(\mathbf{Z}') \qquad (8) \\ &= \underbrace{(\tilde{\mathbf{F}}_{RF}^T \otimes \tilde{\mathbf{W}}_{RF}^H)}_{\mathbf{\Phi}}(\mathbf{A}_D^* \otimes \mathbf{A}_A)\,\mathrm{vec}(\mathbf{D}) + \mathrm{vec}(\mathbf{Z}') \\ &= \mathbf{\Phi}(\mathbf{A}_D^* \otimes \mathbf{A}_A)\,\mathrm{vec}(\mathbf{D}) + \mathrm{vec}(\mathbf{Z}'), \end{aligned}$$

where $\mathbf{y}_V \in \mathbb{C}^{N_F N_W \times 1}$ is the measurement vector.

A. Non-Full Rank Channel Estimation

The CE problem in (8) can be formulated as a least squares problem subject to a sparsity constraint as [8]

$$\begin{aligned} (\hat{\mathbf{A}}_D, \hat{\mathbf{A}}_A, \hat{\mathbf{D}}) &= \underset{\underline{\mathbf{A}}_D, \underline{\mathbf{A}}_A, \underline{\mathbf{D}}}{\arg\min} \|\mathbf{y}_V - \mathbf{\Phi}(\underline{\mathbf{A}}_D^* \otimes \underline{\mathbf{A}}_A)\,\mathrm{vec}(\underline{\mathbf{D}})\|_2^2, \\ \text{s.t. } &\begin{cases} \underline{\mathbf{a}}_D(p') \in \mathcal{A}_\mathcal{D}, p' = 1,\cdots.N'_P, \\ \underline{\mathbf{a}}_A(p') \in \mathcal{A}_\mathcal{A}, \\ \|\mathrm{vec}(\underline{\mathbf{D}})\|_0 = N'_P, \end{cases} \qquad (9) \end{aligned}$$

where N'_P ($N'_P \le N_P$) is the number of the estimated channel paths. $\mathcal{A}_\mathcal{D} = \{\tilde{\mathbf{a}}_D(i_d) = \text{the } i_d^{\text{th}} \text{ column of } \tilde{\mathbf{A}}_D, i_d = 1,\cdots,N_D\}$ is the codebook consisting of N_D candidates of the departure array response vectors, and $\tilde{\mathbf{A}}_D \in \mathbb{C}^{N_T \times N_D}$ is presented in a given realization

$$\tilde{\mathbf{A}}_D = \frac{1}{\sqrt{N_T}}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j2\pi\frac{1}{N_D}} & \cdots & e^{j2\pi\frac{N_D-1}{N_D}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j2\pi\frac{N_T-1}{N_D}} & \cdots & e^{j2\pi\frac{(N_D-1)(N_T-1)}{N_D}} \end{bmatrix}. \qquad (10)$$

The codebook $\mathcal{A}_\mathcal{A} = \{\tilde{\mathbf{a}}_A(i_a) = \text{the } i_a^{\text{th}} \text{ column of } \tilde{\mathbf{A}}_A, i_a = 1,\cdots,N_A\}$ consists of N_A candidates of the arrival array response vectors, and $\tilde{\mathbf{A}}_A \in \mathbb{C}^{N_R \times N_A}$ has the similar structure as (10). The diagonal matrix $\underline{\mathbf{D}} \in \mathbb{C}^{N'_P \times N'_P}$ has N'_P non-zero entries. We assume that $N_D = N_T$ and $N_A = N_R$, so that $\tilde{\mathbf{A}}_D$ and $\tilde{\mathbf{A}}_A$ are IDFT matrices. (9) can be solved by OMP [8][10], and the pseudocode is shown in **Algorithm 1**. OMP is an iterative greedy algorithm that selects the column of the codebook which is most correlated with the current residual ($\mathbf{y}_R$ in **Algorithm 1**) at each step.

Algorithm 1: Non-Full Rank CE by OMP [8]
Input: $\mathbf{y}_V$, $\tilde{\mathbf{A}}_D$, $\tilde{\mathbf{A}}_A$
Output: $\hat{\mathbf{A}}_D$, $\hat{\mathbf{A}}_A$, $\hat{\mathbf{D}}$

1. $\hat{\mathbf{A}}_D$ = empty matrix, $\hat{\mathbf{A}}_A$ = empty matrix, $\hat{\mathbf{\Psi}}$ = empty matrix
2. $\mathbf{y}_R = \mathbf{y}_V$
3. $\mathbf{\Psi} = \mathbf{\Phi}(\tilde{\mathbf{A}}_D^* \otimes \tilde{\mathbf{A}}_A)$
4. for $p' = 1 : N'_P$
5. $\quad \mathbf{g} = \mathbf{\Psi}^H\mathbf{y}_R$
6. $\quad (\hat{i}_D, \hat{i}_A) = \underset{\substack{i_D \in \{1,\cdots,N_D\} \\ i_A \in \{1,\cdots,N_A\}}}{\arg\max} \left(\mathbf{g}\mathbf{g}^H\right)_{i,i}$ where $i = (i_D - 1)\cdot N_A + i_A$
7. $\quad \hat{\mathbf{A}}_D = [\hat{\mathbf{A}}_D \,|\, \tilde{\mathbf{a}}_D(\hat{i}_D)]$ and $\hat{\mathbf{A}}_A = [\hat{\mathbf{A}}_A \,|\, \tilde{\mathbf{a}}_A(\hat{i}_A)]$
8. $\quad \hat{\mathbf{\Psi}} = [\hat{\mathbf{\Psi}} \,|\, \mathbf{\Phi}(\tilde{\mathbf{a}}_D(\hat{i}_D)^* \otimes \tilde{\mathbf{a}}_A(\hat{i}_A))]$
9. $\quad \mathbf{y}_R = (\mathbf{I}_{N_F N_W} - \hat{\mathbf{\Psi}}(\hat{\mathbf{\Psi}}^H\hat{\mathbf{\Psi}})^{-1}\hat{\mathbf{\Psi}}^H)\mathbf{y}_V$
10. end
11. $\mathrm{diag}(\hat{\mathbf{D}}) = (\hat{\mathbf{\Psi}}^H\hat{\mathbf{\Psi}})^{-1}\hat{\mathbf{\Psi}}^H\mathbf{y}_V$

In **Algorithm 1** Step 8, the OMP leads the current selected column orthogonal to the previous selected ones, which means that any two columns of $\hat{\mathbf{\Psi}}$, $\hat{\psi}(i)$ and $\hat{\psi}(j), i \ne j$, are distinct. However, it cannot guarantee that the corresponding array response vectors are different as well. Let $\hat{\psi}(i) = \mathbf{\Phi}(\tilde{\mathbf{a}}_D(\hat{i}_D)^* \otimes \tilde{\mathbf{a}}_A(\hat{i}_A))$ and $\hat{\psi}(j) = \mathbf{\Phi}(\tilde{\mathbf{a}}_D(\hat{j}_D)^* \otimes \tilde{\mathbf{a}}_A(\hat{j}_A))$. We can find that $\hat{\psi}(i) \ne \hat{\psi}(j)$ holds if one of the following three conditions happens

$$\begin{aligned} &(1)\quad \tilde{\mathbf{a}}_D(\hat{i}_D) \ne \tilde{\mathbf{a}}_D(\hat{j}_D) \text{ and } \tilde{\mathbf{a}}_A(\hat{i}_A) = \tilde{\mathbf{a}}_A(\hat{j}_A), \\ &(2)\quad \tilde{\mathbf{a}}_D(\hat{i}_D) = \tilde{\mathbf{a}}_D(\hat{j}_D) \text{ and } \tilde{\mathbf{a}}_A(\hat{i}_A) \ne \tilde{\mathbf{a}}_A(\hat{j}_A), \\ &(3)\quad \tilde{\mathbf{a}}_D(\hat{i}_D) \ne \tilde{\mathbf{a}}_D(\hat{j}_D) \text{ and } \tilde{\mathbf{a}}_A(\hat{i}_A) \ne \tilde{\mathbf{a}}_A(\hat{j}_A), \end{aligned}$$

where the first two conditions include the repeated array propagation vectors. Therefore, the estimated array response matrix could be rank deficient, i.e., $\mathrm{rank}(\hat{\mathbf{A}}_D) < N'_P$ or $\mathrm{rank}(\hat{\mathbf{A}}_A) < N'_P$, which leads to a rank deficient channel matrix,

$$\mathrm{rank}(\hat{\mathbf{H}}) \le \min(\mathrm{rank}(\hat{\mathbf{A}}_D), \mathrm{rank}(\hat{\mathbf{A}}_A)) < N'_P.$$

B. Proposed Full Rank Channel Estimation

We hope that $\mathrm{rank}(\hat{\mathbf{H}}) = N'_P$ for the reason that the reconstructed channel matrix can provide N'_P links for parallel data transmission in order for higher data rates. Accordingly, two constraints, $\mathrm{rank}(\underline{\mathbf{A}}_D) = N'_P$ and $\mathrm{rank}(\underline{\mathbf{A}}_A) = N'_P$, are considered in the CE problem formulation, which becomes a

Table I
AN EXAMPLE OF THE GIVEN AND THE SELECTED AODS AND AOAS WITH $N_P = 3$ AND $N'_P = 2$ BY THE NON-FULL RANK AND THE FULL RANK CE.

	The given/selected angle w.r.t. the 1st strongest path	The given/selected angle w.r.t. the 2nd strongest path	The given/selected angle w.r.t. the3rd strongest path	*Note*
Given AoDs	$\phi_{D,1} = 48^o$	$\phi_{D,2} = 51^o$	$\phi_{D,3} = -30^o$	-
Estimated AoDs by the non-full rank CE (**Algorithm 1**)	$\hat{\phi}_{D,1} = 48.59^o$	$\hat{\phi}_{D,2} = 48.59^o$	-	$\hat{\mathbf{a}}_D(1) = \hat{\mathbf{a}}_D(2)$
Estimated AoDs by the full rank CE (**Algorithm 2**)	$\hat{\phi}_{D,1} = 48.59^o$	$\hat{\phi}_{D,2} = -30^o$ (close to the 3rd given AoD)	-	$\hat{\mathbf{a}}_D(1) \neq \hat{\mathbf{a}}_D(2)$
Given AoAs	$\phi_{A,1} = -29^o$	$\phi_{A,2} = -24^o$	$\phi_{A,3} = 25^o$	-
Estimated AoAs by the non-full rank CE (**Algorithm 1**)	$\hat{\phi}_{A,1} = -30^o$	$\hat{\phi}_{A,2} = -22.02^o$	-	$\hat{\mathbf{a}}_A(1) \neq \hat{\mathbf{a}}_A(2)$
Estimated AoAs by the full rank CE (**Algorithm 2**)	$\hat{\phi}_{A,1} = -30^o$	$\hat{\phi}_{A,2} = 25.94^o$ (close to the 3rd given AoA)	-	$\hat{\mathbf{a}}_A(1) \neq \hat{\mathbf{a}}_A(2)$

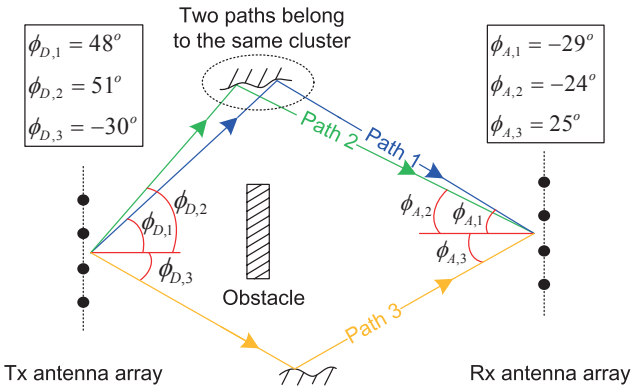

Figure 2. A typical example of mmW channels with $N_P = 3$, where Path 1 and 2 belong to the same cluster so that $\phi_{D,1} \approx \phi_{D,2}$ and $\phi_{A,1} \approx \phi_{A,2}$.

Algorithm 2: Proposed Full Rank CE by OMP

Input: $\mathbf{y}_V$, $\tilde{\mathbf{A}}_D$, $\tilde{\mathbf{A}}_A$

Output: $\hat{\mathbf{A}}_D$, $\hat{\mathbf{A}}_A$, $\hat{\mathbf{D}}$

1. $\hat{\mathbf{A}}_D$ = empty matrix, $\hat{\mathbf{A}}_A$ = empty matrix, $\hat{\mathbf{\Psi}}$ = empty matrix, $\mathcal{I}_\mathcal{D} = \varnothing$, $\mathcal{I}_\mathcal{A} = \varnothing$
2. $\mathbf{y}_R = \mathbf{y}_V$
3. $\mathbf{\Psi} = \mathbf{\Phi}(\tilde{\mathbf{A}}_D^* \otimes \tilde{\mathbf{A}}_A)$
4. for $p' = 1 : N'_P$
5. $\quad \mathbf{g} = \mathbf{\Psi}^H \mathbf{y}_R$
6. $\quad (\hat{\imath}_D, \hat{\imath}_A) = \underset{\substack{i_D \in \{1,\cdots,N_D\}\setminus\mathcal{I}_\mathcal{D} \\ i_A \in \{1,\cdots,N_A\}\setminus\mathcal{I}_\mathcal{A}}}{\arg\max} \left(\mathbf{g}\mathbf{g}^H\right)_{i,i}$ where $i = (i_D - 1) \cdot N_A + i_A$
7. $\quad \hat{\mathbf{A}}_D = [\hat{\mathbf{A}}_D \,|\, \tilde{\mathbf{a}}_D(\hat{\imath}_D)]$ and $\hat{\mathbf{A}}_A = [\hat{\mathbf{A}}_A \,|\, \tilde{\mathbf{a}}_A(\hat{\imath}_A)]$
8. $\quad \hat{\mathbf{\Psi}} = [\hat{\mathbf{\Psi}} \,|\, \mathbf{\Phi}(\tilde{\mathbf{a}}_D(\hat{\imath}_D)^* \otimes \tilde{\mathbf{a}}_A(\hat{\imath}_A))]$
9. $\quad \mathbf{y}_R = (\mathbf{I}_{N_F N_W} - \hat{\mathbf{\Psi}}(\hat{\mathbf{\Psi}}^H \hat{\mathbf{\Psi}})^{-1} \hat{\mathbf{\Psi}}^H) \mathbf{y}_V$
10. $\quad \mathcal{I}_\mathcal{D} = \mathcal{I}_\mathcal{D} \cup \{\hat{\imath}_D\}$ and $\mathcal{I}_\mathcal{A} = \mathcal{I}_\mathcal{A} \cup \{\hat{\imath}_A\}$
11. end
12. $\text{diag}(\hat{\mathbf{D}}) = (\hat{\mathbf{\Psi}}^H \hat{\mathbf{\Psi}})^{-1} \hat{\mathbf{\Psi}}^H \mathbf{y}_V$

least squares problem subject to one sparsity and two *full rank* constraints as

$$(\hat{\mathbf{A}}_D, \hat{\mathbf{A}}_A) = \underset{\underline{\mathbf{A}}_D, \underline{\mathbf{A}}_A}{\arg\min} \|\mathbf{y}_V - \mathbf{\Phi}\left(\underline{\mathbf{A}}_D^* \otimes \underline{\mathbf{A}}_A\right) \text{vec}(\underline{\mathbf{D}})\|_2^2,$$
$$\text{s.t.} \begin{cases} \underline{\mathbf{a}}_D(p') \in \mathcal{A}_\mathcal{D}, \\ \underline{\mathbf{a}}_A(p') \in \mathcal{A}_\mathcal{A}, \\ \|\text{vec}(\underline{\mathbf{D}})\|_0 = N'_P, \\ \text{rank}(\underline{\mathbf{A}}_D) = N'_P, \\ \text{rank}(\underline{\mathbf{A}}_A) = N'_P. \end{cases} \tag{11}$$

(11) can also be solved by OMP with some modifications, and the pseudocode is shown in **Algorithm 2**. Compared with **Algorithm 1** Step 6, the different part is that, in **Algorithm 2** Step 6, the indices in the sets $\mathcal{I}_\mathcal{D}$ and $\mathcal{I}_\mathcal{A}$ (consisting of all the previous selected indices) will not be selected again to ensure that there are no repeated vectors in $\hat{\mathbf{A}}_D$ and $\hat{\mathbf{A}}_A$.

A simple example of the selected AoDs and AoAs by **Algorithm 1** and **Algorithm 2** is shown in Fig. 2 and Table I with $N_P = 3$, $N'_P = 2$, $N_T = N_R = N_D = N_A = 32$, $T = 512$, SNR $= 0$ dB, and $|\alpha_1|^2 > |\alpha_2|^2 > |\alpha_3|^2$, where the two strongest paths from the same cluster have nearly the same AoDs and AoAs[1]. The capacity of the channels by these two algorithms is introduced in terms of the rank of the channel.

1) rank($\hat{\mathbf{H}}$) by the non-full rank CE (Algorithm 1): From Table I, we can find that the two selected vectors $\hat{\mathbf{a}}_D(1)$ and $\hat{\mathbf{a}}_D(2)$ are repeated. Since $\hat{\mathbf{a}}_D(1) = \hat{\mathbf{a}}_D(2)$ and $\hat{\mathbf{a}}_A(1) \neq \hat{\mathbf{a}}_A(2)$, the reconstructed channel becomes

$$\hat{\mathbf{H}} = \hat{\mathbf{A}}_A \hat{\mathbf{D}} \hat{\mathbf{A}}_D^H = [\hat{\mathbf{a}}_A(1), \hat{\mathbf{a}}_A(2)] \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} \hat{\mathbf{a}}_D(1)^H,$$

and because $\hat{\mathbf{A}}_A$ has full column rank,

$$\text{rank}(\hat{\mathbf{H}}) = \text{rank}(\hat{\mathbf{A}}_D) = 1,$$

which provides receive diversity but no spatial degree-of-freedom gain [11].

[1]The predefined 32 angles in the codebooks $\mathcal{A}_\mathcal{D}$ and $\mathcal{A}_\mathcal{A}$ are $\{0^o, \pm 3.58^o, \pm 7.18^o, \pm 10.81^o, \pm 14.48^o, \pm 18.21^o, \pm 22.02^o, \pm 25.94^o, \pm 30^o, \pm 34.23^o, \pm 38.68^o, \pm 43.43^o, \pm 48.59^o, \pm 54.34^o, \pm 61.05^o, \pm 69.54^o, 90^o$ (or -90^o)$\}$.

Table II
SIMULATION PARAMETERS.

Parameter	Value
Carrier frequency [GHz]	60
Training time (T in (6))	512, 1024, 1536
Number of paths (N_P)	5
Number of the estimated paths (N'_P)	2
Number of RF chains (N_{RF})	2
Number of data streams (N_S)	2
Number of Tx antennas (N_T)	32
Number of Rx antennas (N_R)	32
Number of columns in $\tilde{\mathbf{A}}_D$ (N_D)	32
Number of columns in $\tilde{\mathbf{A}}_A$ (N_A)	32
Path attenuation coefficient (α_p)	$\alpha_p \in \mathbb{C}$, $\sum_{p=1}^{N_P} \lvert\alpha_p\rvert^2 = 1$ and $\lvert\alpha_1\rvert^2 > \cdots > \lvert\alpha_{N_P}\rvert^2$

2) rank($\hat{\mathbf{H}}$) by the full rank CE (Algorithm 2): Since $\hat{\mathbf{a}}_D(1) \neq \hat{\mathbf{a}}_D(2)$ and $\hat{\mathbf{a}}_A(1) \neq \hat{\mathbf{a}}_A(2)$, the reconstructed channel matrix can be written as

$$\hat{\mathbf{H}} = \hat{\mathbf{A}}_A \hat{\mathbf{D}} \hat{\mathbf{A}}_D^H = [\hat{\mathbf{a}}_A(1), \hat{\mathbf{a}}_A(2)] \begin{bmatrix} \hat{\alpha}_1 & 0 \\ 0 & \hat{\alpha}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}}_D(1)^H \\ \hat{\mathbf{a}}_D(2)^H \end{bmatrix},$$

and due to the fact that rank($\hat{\mathbf{A}}_A$) = 2 and rank($\hat{\mathbf{A}}_D$) = 2, we have rank($\hat{\mathbf{H}}$) = 2, which provides two spatial degrees of freedom [11], and each link supports one data stream transmission and reception.

IV. NUMERICAL RESULTS

Table II lists all the simulation parameters. Assume that the two strongest paths are from the same cluster and the angular spread is less than or equal to HPBW ($\approx \frac{\pi}{N_T}$(or $\frac{\pi}{N_R}$) = 5.6^o), therefore the first two AoDs and AoAs are defined as

$$\begin{cases} \phi_{D,1} \sim \mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2}) \\ \phi_{D,2} = \phi_{D,1} + \Delta\phi \end{cases} \text{and} \begin{cases} \phi_{A,1} \sim \mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2}) \\ \phi_{A,2} = \phi_{A,1} + \Delta\phi \end{cases}, \quad (12)$$

where $\Delta\phi_D$ is the angular spread modeled as a uniform distribution, $\Delta\phi \sim \mathcal{U}(-\frac{5.6^o}{2}, \frac{5.6^o}{2})$. Except the angles defined in (12), others are uniformly distributed over the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$. Two CE methods are simulated: the non-full rank CE, see **Algorithm 1**, and the proposed full rank CE, see **Algorithm 2**.

A. Mean Absolute Error of the CE

As introduced in Subsection III-B, when $\phi_{D,1} \approx \phi_{D,2}$ and/or $\phi_{A,1} \approx \phi_{A,2}$, the second selected angles $\hat{\phi}_{D,2}$ and $\hat{\phi}_{A,2}$ by **Algorithm 2** are mostly close to $\phi_{D,3}$ and $\phi_{A,3}$, respectively. Hence, the mean absolute error (MAE) between the given angles and the selected ones by the non-full rank and the full rank CE methods shown in Fig. 3 is given by

$$\text{MAE}_{\text{ non-full rank}} = \mathrm{E}\left[\frac{\sum_{p=1}^{N'_P} \left(\left|\hat{\phi}_{D,p} - \phi_{D,p}\right| + \left|\hat{\phi}_{A,p} - \phi_{A,p}\right|\right)}{2N'_P}\right] \quad (13)$$

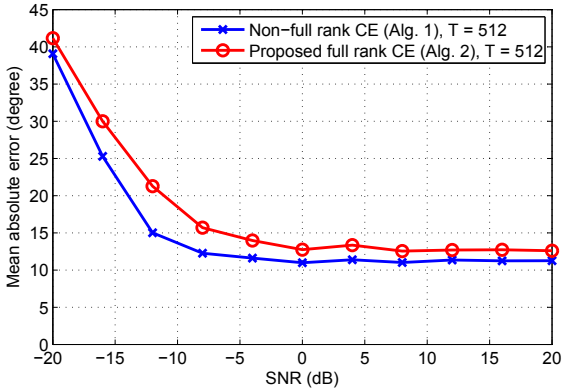

Figure 3. The MAE between the given AoDs/AoAs and the selected ones by the non-full rank and the full rank CE methods with $T = 512$.

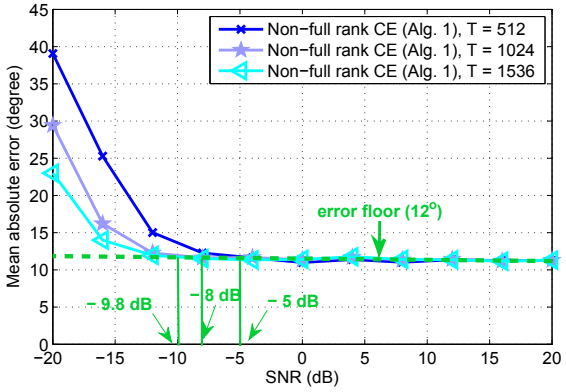

Figure 4. The MAE between the given AoDs/AoAs and the selected ones by the non-full rank CE method with $T = 512, 1024, 1536$.

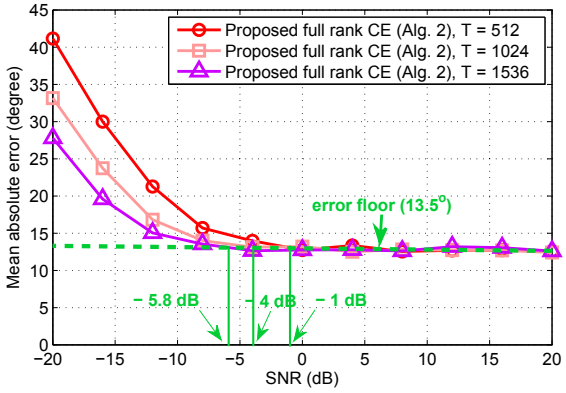

Figure 5. The MAE between the given AoDs/AoAs and the selected ones by the full rank CE method with $T = 512, 1024, 1536$.

and

$$\text{MAE}_{\text{ full rank}} = \mathrm{E}\left[\frac{1}{2N'_P}\left(\left|\hat{\phi}_{D,1} - \phi_{D,1}\right| + \left|\hat{\phi}_{A,1} - \phi_{A,1}\right| + \left|\hat{\phi}_{D,2} - \phi_{D,3}\right| + \left|\hat{\phi}_{A,2} - \phi_{A,3}\right|\right)\right]. \quad (14)$$

Note that only two estimated paths ($N'_P = 2$) are considered in (14).

In Fig. 3, **Algorithm 2** shows worse MAE because the two selected paths are mainly the first and the third paths ($p = 1, 3$) and the path loss of the third path is larger than the first two paths. When the number T of the observations increases, see Fig. 4 and 5, the MAE by these two methods decrease significantly because the averaged noise power becomes smaller. In Fig. 4, the error floor of the non-full rank CE method by $T = 512, 1024, 1536$ occurs when SNR $= -5, -8, -9.8$ dB, and we can find that when the length of the training signals are doubled and tripled ($T = 1024, 1536$), there are 3 and 4.8 dB improvements in SNR. Moreover, in Fig. 5, the error floor of the proposed full rank CE method by $T = 512, 1024, 1536$ occurs when SNR $= -1, -4, -5.8$ dB. Similar to the cases in Fig. 4, when the length of the training signals are doubled and tripled, there are 3 and 4.8 dB improvements in SNR as well.

B. Achievable Data Rate

Based on the estimated $\hat{\mathbf{A}}_A$, $\hat{\mathbf{A}}_D$, and $\hat{\mathbf{D}}$ obtained in (9) and (11), the reconstructed channel $\hat{\mathbf{H}} = \hat{\mathbf{A}}_A \hat{\mathbf{D}} \hat{\mathbf{A}}_D^H$ has a singular value decomposition (SVD):

$$\hat{\mathbf{H}} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{15}$$

where the columns of $\mathbf{V} \in \mathbb{C}^{N_T \times N'_P}$ and $\mathbf{U} \in \mathbb{C}^{N_R \times N'_P}$ are, respectively, the right and the left singular vectors of $\hat{\mathbf{H}}$, and the diagonal entries of $\mathbf{\Lambda} = diag(\lambda_1, \cdots, \lambda_{N'_P}) \in \mathbb{R}^{N'_P \times N'_P}$, $\lambda_1 \geq \cdots \geq \lambda_{N'_P} > 0$, are the singular values of $\hat{\mathbf{H}}$. The design criterion that can be used for the precoder reconstruction is to minimize the Frobenius norm of the error between the precoder and the right singular vectors of $\hat{\mathbf{H}}$ [7]

$$(\hat{\mathbf{F}}_{RF}, \hat{\mathbf{F}}_{BB}) = \underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}}{\arg \min} \|\mathbf{V}(1:N_S) - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2,$$
$$\text{s.t. } \mathbf{f}_{RF}(n_{rf}) \in \{\tilde{\mathbf{f}}_{RF}(l), l = 1, \cdots, N_F\}, n_{rf} = 1, \cdots, N_{RF},$$
$$\|\mathbf{F}_{RF}\mathbf{f}_{BB}(n_s)\|_2 = \|\mathbf{v}(n_s)\|_2, n_s = 1, \cdots, N_S, N_S \leq N'_P. \tag{16}$$

(16) can be solved by OMP [10]. The same design criterion leads to the desired combiner. Two cases of the given AoDs and AoAs simulated in Fig. 6 are listed as follows:

- Case 1: all the AoDs and AoAs are independent and uniformly distributed over the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ except the first two strongest paths whose AoDs and AoAs belong to the same cluster and the values are defined in (12).
- Case 2: all the AoDs and AoAs are independent and uniformly distributed over the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

In Fig. 6 Case 1, **Algorithm 2** shows better results than **Algorithm 1** because it can always obtain spatial degree-of-freedom gain even if its MAE is worse, while **Algorithm 1** may provide transmit or receive diversity mostly. When the number T of the observations increases, the data rates do not show obvious improvement since the decrease of MAE by more observations occurs when SNR < 0 dB. In Case 2, the data rates by **Algorithm 1** approach the results by **Algorithm 2** because the given AoDs and AoAs in this case are all from different clusters; therefore, **Algorithm 1** provides not only transmit or receive diversity but also spatial degree-of-freedom gain as **Algorithm 2**.

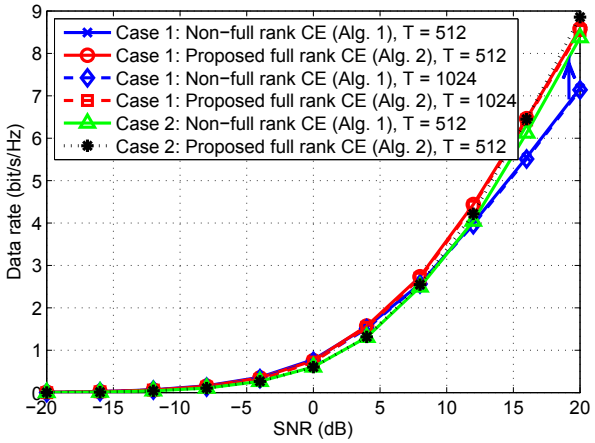

Figure 6. The data rates by two different CE methods and the same HBF reconstruction method.

V. Conclusion

Owing to sparse mmW channels and limited beamwidth, exploiting the maximum spatial degrees of freedom is an essential technology in order to increase the data rates. This paper presents a solution to obtain it by the estimated full rank array response matrices. From the simulation results, it shows that the proposed method achieves higher data rates than the previously proposed ones especially when SNR increases.

Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°619563 (MiWaveS).

References

[1] T. Rappaport, R. W. Heath Jr., T. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*, Prentice Hall, 2014.

[2] T. A. Thomas, et al., "3D mmWave Channel Model Proposal," *IEEE VTC Fall*, pp. 1–6, 2014.

[3] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, Sept 2015.

[4] T. S. Rappaport, et al., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, Vol. 1, pp. 335–349, May 2013.

[5] A. Hajimiri, et al., "Integrated Phased Array Systems in Silicon," *IEEE Proceedings*, Vol. 93, pp. 1637–1655, Sep. 2005.

[6] J. C. Liberti, et al., *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice Hall, NJ, 1999.

[7] O. E. Ayach, et al., "Low Complexity Precoding for Large Millimeter Wave MIMO Systems," *IEEE International Conference on Communications (ICC)*, pp. 3724–3729, Jun. 2012.

[8] R. Mendez-Rial, C. Rusu, et al, "Channel estimation and hybrid combining for mmWave: Phase shifters or switches?" *Information Theory and Applications Workshop (ITA)*, pp. 90–97, 2015.

[9] E. Candes and T. Tao, "Decoding by Linear Programming," *IEEE Transactions on Information Theory*, Vol. 51, pp. 4203–4215, Dec. 2005.
[10] T. T. Cai and L. Wang, "Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise," *IEEE Transactions on Information Theory*, Vol. 57, No. 7, pp. 4680–4688, Jul. 2011.
[11] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Chapter 7, 2005.
[12] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhauser, 2013.
[13] V. Frascolla, et al., "Challenges and Opportunities for Millimeter-wave Mobile Access Standardization" *IEEE Workshop on Globecom*, pp. 553–558, 2014.
[14] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.