

Coherence Optimized Channel Estimation for Mm-Wave Massive MIMO

Faisal AKRAM, Imran RASHID, Abdul GHAFOOR, Adil Masood SIDDIQUI

Dept. of Electrical Engineering, College of Signals, National University of Science and Technology (NUST), Pakistan

faisal.akram@mcs.edu.pk, irashid@mcs.edu.pk, abdulghafoor-mcs@nust.edu.pk, dradil@mcs.edu.pk

Submitted September 17, 2019 / Accepted June 19, 2020

Abstract. *Mm-wave MIMO communication makes a hybrid combination of analog RF and digital baseband processing more attractive, where digital baseband precoders/combiners able to adapt to the pre-defined analog (switch based) RF processors. Non-uniform two-dimensional quantized azimuth and elevation angle grid antenna array responses are suggested for uniform planar array (UPA) and are proven orthogonal. Training vectors (or sensing matrix) are designed for suggested antenna array response with unitary RF processing for UPA in mm-wave hybrid MIMO system. Proposed training vectors achieve minimized total coherence of the equivalent sensing matrix for hybrid MIMO system. Open-loop channel estimation of the mm-wave channel is done by using iterative re-weight based super resolution algorithm to exploit its sparse nature. Extensive simulations reveal the benefit of coherence optimization where normalized mean squared error is reduced and spectral efficiency is improved in comparison to existing methods.*

Keywords

Channel estimation, compressed sensing, hybrid MIMO, mm-wave communication, sparse channel

1. Introduction

Millimeter wave communication is a promising technology enabling solutions to the global bandwidth shortage in the frequency spectrum below 6 GHz [1]. Mm-wave systems have smaller wavelengths as compared to microwave frequencies. Therefore, larger and more compact antenna arrays are possible for mm-wave systems due to their smaller inter-antenna separation, which results in increased beamforming gains [1].

The channel for mm-wave spectrum is inherently sparse due to its peculiar propagation characteristic with limited diffraction, material penetration and small number of scatterers [2]. As a result, mm-wave channel matrices are rank deficient, depending on the few dominant singular values which tend to be much smaller than the number of antennas [3].

Consequently, a hybrid MIMO architecture with RF beamforming (analog) cascaded with digital baseband processing has been proposed in [4], [5]. Number of RF chains in hybrid MIMO architecture is determined by the effective rank of the channel.

In mm-wave MIMO systems, analog RF processors are often implemented using phase shifters with constant modulus condition [4]. In [6], space-time alignment allows mm-wave channel estimation during beam sweeping time. In [4], the designing of baseband precoding/combining and RF beamformer is jointly done. This technique outperforms the solution where firstly the RF beam is steered to the angle of arrival (AoA)/ angle of departure (AoD) and then subsequently optimization of baseband precoding/combining is done. In [7], channel estimation for time varying mm-wave channels is considered in two stages. Initially, the AoAs/AoDs estimation is done using adaptive angle estimation and then pilot beamforming is done based on estimated AoAs/AoDs. It is designed to maximize pilot power for efficient path gain recovery. Two-stage algorithm in [8] performs estimation of position and rotation angle for single transmitter. Initially, multiple measurement vectors matching pursuit is used for coarse estimation followed by a refinement stage based on the space-alternating (generalized) expectation maximization algorithm. In [9], correlation based adaptive compressed sensing using cosine beam pattern and hierarchical codebook are used for better angle resolution estimation (AoA/AoD), which improves spectral efficiency. In [10], hybrid mm-wave massive MIMO channel estimation is done in two stages. Initially, joint AoA/AoD estimation is transformed into two one dimensional sub-problems. Then subsequently, sparse signal reconstruction is done using the initial support set.

Spatial channel covariance estimation is used in [11], for designing of analog precoders for hybrid (MIMO) architecture. Matrix completion based channel estimation is done with generalized conditional gradient framework. In [12], alternating minimization approach which is robust to array-inherent impairments is developed. In [13], successive interference cancellation allows designing of near-optimal precoders for multiple sub-arrays at the transmitter/receiver resulting in higher energy efficiency. In [14], clustered sparse

Bayesian learning approach is used for hybrid MIMO channel estimation without knowledge of channel statistics and angular information of users. In [15], sparse Bayesian learning based block-sparse channel estimation is done for mm-wave hybrid MIMO system using multiple measurement vectors. Furthermore, a recursive hierarchical Bayesian Kalman filter is used for estimation of time-selective channel. In [16], Kalman based formulation for hybrid analog/digital precoding in multi-user environments is proposed with affordable complexity.

In [17], channel estimation using training sequence (TS) design based on inverse discrete Fourier transform (IDFT) of a frequency domain binary TS is done. It is followed by an investigation of genetic algorithm to optimize the auto-coherence and cross-coherence of the blocks of sensing matrix. In [18], randomly constructed pilot signals and beamforming vectors are considered and the mutual coherence parameter of the resulting sensing matrix is analysed. In [19], hybrid analog/digital precoding and combining matrices are designed based on minimum mean squared error (MMSE) criteria. In [20], open-loop channel estimator for hybrid MIMO using uniform linear array (ULA) is proposed. Baseband training vectors are proposed for non-uniform quantized angles grid. The antenna array response is based on one dimension with unitary RF-beamformers. However, baseband training vectors for two-dimensional non-uniform quantized angles, matching the fine angular resolution of mm-wave signals are required to be investigated. Coherence minimization achieved for the equivalent sensing matrix also shows room for further improvement.

In this paper, training vectors forming baseband precoder/combiner are designed for mm-wave hybrid MIMO system. The environment is based on parametric channel model with quantized AoAs/AoDs (azimuth & elevation angle) for UPA and the pilot vectors are a cascade of RF beamformer and baseband precoder/combiner. For channel estimation, iterative re-weight super resolution algorithm [21] is used. The contributions of the paper are mentioned below:

- A sparse channel recovery problem is formulated for mm-wave hybrid MIMO system. Redundant dictionary consisting of antenna array response vectors with non-uniformly distributed angular grid for azimuth and elevation in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is suggested. Such an antenna array response is shown to have full rank (orthogonal) which facilitates coherence minimization of the overall equivalent sensing matrix for the mm-wave hybrid MIMO system. The angular grids for azimuth angles θ_γ and elevation angles ϕ_γ are defined such that $\sin(\theta_\gamma)$ and $\sin(\phi_\gamma)$ are uniformly distributed in $[-1, 1]$. The azimuth and elevation angular resolution of the proposed scheme can be finer than the parametric channel model. The non-uniformly distributed azimuth and elevation grids are shown to reduce the mutual coherence of the redundant dictionary.

- Pilot beam patterns for unitary RF beamformers are designed to minimize the overall coherence of the equivalent sensing matrix. Baseband precoder/combiner are optimized for a given RF beamformer and orthogonal antenna array response to lower the overall coherence of the mm-wave hybrid MIMO system. Coherence minimization of baseband precoder/combiner is done by modifying Grassmannian codebook generation algorithm for dictionary adaptation. The overall coherence of the system is taken as the maximum normalized inner product of the effective sensing matrix for the transmitter/receiver. The effective sensing matrix is formed by the product of the RF beamformer, baseband precoder/combiner and the respective combined azimuth and elevation angular antenna array response for UPA. It is observed that reduced coherence of the system enables better estimation accuracy.
- A detailed analysis of the coherence achieved through proposed technique for overall mm-wave MIMO system with different fixed antenna array but varying resolution (finer) antenna array responses is provided.
- Normalized error performance and achievable spectral efficiency of the proposed approach is compared to existing techniques. Error performance is also analyzed with varying number of paths and antenna array grid resolutions that are equivalent to the specific antenna configuration and further finer resolutions.

2. System Model

2.1 Signal Model

Consider a hybrid mm-wave MIMO system model (Fig. 1) [20] with uniform planar antenna array having N_T transmit and N_R receive antennas which are assumed to be multiples of their respective RF chains i.e, N_T^{RF} and N_R^{RF} . The RF beamformers are realized using analog phase shifters. During the channel estimation, the transmitter uses beam patterns $\{\mathbf{x}_v \in \mathbb{C}^{N_T \times 1}\}$ and the receiver uses beam patterns $\mathbf{w}_q \in \mathbb{C}^{N_R \times 1}$, where $v = 1, 2, \dots, N_x$, $q = 1, 2, \dots, N_y$, $\|\mathbf{x}_v\|_2^2 = 1$, $\|\mathbf{w}_q\|_2^2 = 1$. The receive vector for the v -th transmit beam is given as:

$$\mathbf{y}_{v,\bar{q}} = \mathbf{W}_{\bar{q}}^H \mathbf{H} \mathbf{x}_{v,s_v} + \mathbf{W}_{\bar{q}}^H \mathbf{n}_{\bar{q},v} \quad (1)$$

where s_v is the pilot symbol being transmitted with average power P [22], $\bar{q} \in \{1, 2, \dots, \frac{N_y}{N_R^{RF}}\}$, $\mathbf{W}_{\bar{q}} \in \mathbb{C}^{N_R \times N_R^{RF}}$ represents receive beam pattern, $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix and $\mathbf{n}_{\bar{q},v} \in \mathbb{C}^{N_R \times 1}$ is the noise vector with normal distribution having variance σ_n^2 . Collectively for \bar{q} and v the receive matrix $\mathbf{Y} \in \mathbb{C}^{N_y \times N_x}$ is given as:

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H} \mathbf{X} \mathbf{S} + \mathbf{W}^H \mathbf{N} \quad (2)$$

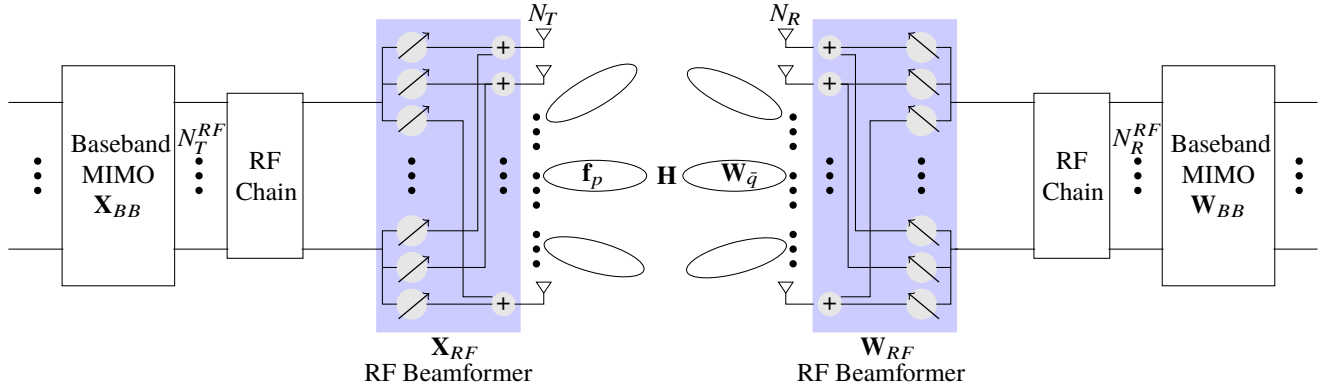


Fig. 1. The mm-wave MIMO system model [20].

where $\mathbf{X} \in \mathbb{C}^{N_T \times N_x}$ is the precoding matrix, $\mathbf{W} \in \mathbb{C}^{N_R \times N_y}$ is the combining matrix, $\mathbf{N} \in \mathbb{C}^{N_y \times N_x}$ is the noise matrix and $\mathbf{S} = \sqrt{P} \mathbf{I}_{N_x}$ with identity matrix $\mathbf{I}_{N_x} \in \mathbb{R}^{N_x \times N_x}$ is the pilot symbols matrix [22]. Therefore,

$$\mathbf{Y} = \sqrt{P} \mathbf{W}^H \mathbf{H} \mathbf{X} + \mathbf{W}^H \mathbf{N}. \quad (3)$$

In the hybrid MIMO framework [20], the precoding and combining matrices are decomposed into RF and Baseband (BB) components as $\mathbf{X} = \mathbf{X}_{RF} \mathbf{X}_{BB}$ and $\mathbf{W} = \mathbf{W}_{RF} \mathbf{W}_{BB}$, where $\mathbf{X}_{BB} \in \mathbb{C}^{N_T \times N_x}$, $\mathbf{W}_{BB} \in \mathbb{C}^{N_R \times N_y}$. RF components $\mathbf{X}_{RF} \in \mathbb{C}^{N_T \times N_T}$ and $\mathbf{W}_{RF} \in \mathbb{C}^{N_R \times N_R}$ are assumed to be unitary matrices.

2.2 Channel Model

Since mmWave channels are sparse in nature due to limited scattering [22], a geometric channel model is adopted. Let the channel have L scatterers and each one of them contributes to a propagation path between transmitter and receiver, then the channel model for UPA is given as:

$$\mathbf{H} = \sqrt{\frac{M_{tx} M_{ty} M_{rx} M_{ry}}{\rho}} \sum_{\gamma=1}^L \alpha_{\gamma} \mathbf{a}_R(\theta_{\gamma}^r, \phi_{\gamma}^r) \mathbf{a}_T^H(\theta_{\gamma}^t, \phi_{\gamma}^t) \quad (4)$$

where ρ represents the average path loss, α_{γ} is the complex gain for the γ -th path, $\{M_{tx}, M_{ty}\}$ are the number of antenna elements in each row and column of the transmitter array, $\{M_{rx}, M_{ry}\}$ are the elements in each row and column of the receiver array respectively. The transmitter antenna array response $\mathbf{a}_T(\theta_{\gamma}^t, \phi_{\gamma}^t) = \mathbf{a}_T(\theta_{\gamma}^t) \otimes \mathbf{a}_T(\phi_{\gamma}^t)$ with azimuth and elevation responses given as [23]:

$$\begin{aligned} \mathbf{a}_T(\theta_{\gamma}^t) &= \frac{1}{\sqrt{M_{tx}}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\theta_{\gamma}^t)}, \dots, e^{-j2\pi \frac{d}{\lambda} \sin(\theta_{\gamma}^t)(M_{tx}-1)} \right]^T, \\ \mathbf{a}_T(\phi_{\gamma}^t) &= \frac{1}{\sqrt{M_{ty}}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_{\gamma}^t)}, \dots, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_{\gamma}^t)(M_{ty}-1)} \right]^T \end{aligned} \quad (5)$$

where d is the separation between antenna elements, λ is the wavelength, $\theta_{\gamma}^t, \phi_{\gamma}^t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ are the normalized spatial azimuth and elevation AoD respectively. Similarly, the receiver antenna array response $\mathbf{a}_R(\theta_{\gamma}^r, \phi_{\gamma}^r) = \mathbf{a}_R(\theta_{\gamma}^r) \otimes \mathbf{a}_R(\phi_{\gamma}^r)$ for normalized spatial AoA $\theta_{\gamma}^r, \phi_{\gamma}^r \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ where θ_{γ}^r represents azimuth AoA and ϕ_{γ}^r represents elevation AoA [22].

Summarily, transmitter antenna array response $\mathbf{A}_T \in \mathbb{C}^{N_T \times L}$ and receiver antenna array response $\mathbf{A}_R \in \mathbb{C}^{N_R \times L}$ for UPA with L paths can be given as:

$$\begin{aligned} \mathbf{A}_T &= [\mathbf{a}(\theta_1^t, \phi_1^t), \mathbf{a}(\theta_2^t, \phi_2^t), \dots, \mathbf{a}(\theta_L^t, \phi_L^t)], \\ \mathbf{A}_R &= [\mathbf{a}(\theta_1^r, \phi_1^r), \mathbf{a}(\theta_2^r, \phi_2^r), \dots, \mathbf{a}(\theta_L^r, \phi_L^r)]. \end{aligned} \quad (6)$$

In more compact form the channel matrix given in (4) can be represented as [22]:

$$\mathbf{H} = \mathbf{A}_R \text{diag}(\boldsymbol{\alpha}) \mathbf{A}_T^H \quad (7)$$

where $\boldsymbol{\alpha} = \sqrt{M_{tx} M_{ty} M_{rx} M_{ry} / \rho} [\alpha_1, \alpha_2, \dots, \alpha_L]^T$.

2.3 Orthogonal Antenna Array Response

In order to generalize the antenna array response for both azimuth and elevation angles, grid points $N_g \gg L$ are uniformly taken in [22], [24], i.e. $\theta_{\zeta} \in 2\pi\zeta/N_g, \zeta \in \{1, 2, \dots, N_g\}$ and $\phi_{\eta} \in 2\pi\eta/N_g, \eta \in \{1, 2, \dots, N_g\}$. As a result the channel model as per (7) modifies with $\boldsymbol{\alpha} = \sqrt{M_{tx} M_{ty} M_{rx} M_{ry} / \rho} [\alpha_1, \alpha_2, \dots, \alpha_{N_g}]^T$ having non-zero channel gains at the corresponding azimuth and elevation angles at both ends.

In [20], the grid angles are defined in a non-uniform manner where $\theta \in [0, \pi]$ for the antenna array response of ULA. The resulting antenna array response matrix is orthogonal for $N_g \geq N$, where N is the number of antennas at the transmitter/receiver.

In case of UPA for mm-wave massive MIMO with critically spaced elements i.e. $d = \frac{\lambda}{2}$, having M_x rows and M_y columns of antenna elements, the azimuth angle $\theta \in [-\pi/2, \pi/2]$ and elevation angle $\phi \in [-\pi/2, \pi/2]$ are shown in Fig. 2. The azimuth angle grid $\mathcal{G}_{\theta} = \theta_i$ (for $i = 1, 2, \dots, N_{g\theta}$ and $N_{g\theta} \geq M_x$) and elevation angle grid $\mathcal{G}_{\phi} = \phi_k$ (for $k = 1, 2, \dots, N_{g\phi}$ and $N_{g\phi} \geq M_y$) are defined in a non-uniform manner in $\{-\frac{\pi}{2}, \frac{\pi}{2}\}$, such that $\sin(\theta_i)$ and $\sin(\phi_k)$ are well defined for the entire grid and remain uniformly distributed in $[-1, 1]$. Specifically, θ_i and ϕ_k are determined to satisfy:

$$\sin(\theta_i) = \frac{2}{N_{g\theta}}(i-1) + 1, \quad \sin(\phi_k) = \frac{2}{N_{g\phi}}(k-1) + 1. \quad (8)$$

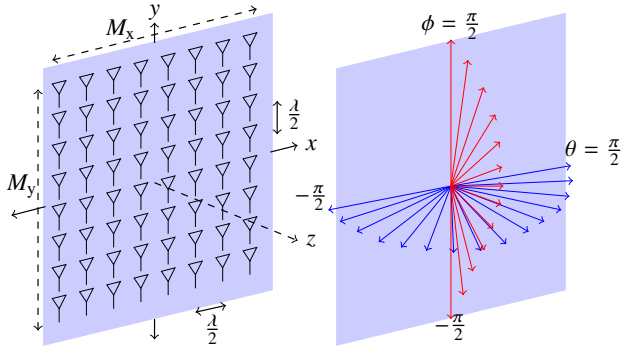


Fig. 2. Uniform planar array.

From (5), the antenna array response for critically spaced elements ($d = \frac{\lambda}{2}$) having azimuth angle θ_i and elevation angle ϕ_k can be written as:

$$\begin{aligned} \mathbf{a}(\theta_i) &= \frac{1}{\sqrt{M_x}} \left[1, e^{-j\pi \sin(\theta_i)}, e^{-j\pi \sin(\theta_i)2}, \dots, e^{-j\pi \sin(\theta_i)(M_x-1)} \right]^T, \\ \mathbf{a}(\phi_k) &= \frac{1}{\sqrt{M_y}} \left[1, e^{-j\pi \sin(\phi_k)}, e^{-j\pi \sin(\phi_k)2}, \dots, e^{-j\pi \sin(\phi_k)(M_y-1)} \right]^T. \end{aligned} \quad (9)$$

Collectively, the antenna array response for azimuth and elevation angles taken from their respective grids \mathcal{G}_θ and \mathcal{G}_ϕ are given as:

$$\begin{aligned} \mathbf{A}(\theta) &= [\mathbf{a}(\theta_1) \ \mathbf{a}(\theta_2) \ \dots \ \mathbf{a}(\theta_{N_{g\theta}})], \\ \mathbf{A}(\phi) &= [\mathbf{a}(\phi_1) \ \mathbf{a}(\phi_2) \ \dots \ \mathbf{a}(\phi_{N_{g\phi}})]. \end{aligned} \quad (10)$$

Correspondingly, the overall antenna array response is given as $\mathbf{A} = \mathbf{A}(\theta) \otimes \mathbf{A}(\phi)$. The Kronecker product $\mathbf{a}(\theta_i, \phi_k)$ is given as:

$$\begin{aligned} \mathbf{a}(\theta_i, \phi_k) &= \\ &= \frac{1}{\sqrt{M_x M_y}} \left[\left\{ 1, e^{-j\pi \sin(\phi_k)}, e^{-j\pi \sin(\phi_k)2}, \dots, e^{-j\pi \sin(\phi_k)(M_y-1)} \right\}^T, \right. \\ &e^{-j\pi \sin(\theta_i)} \left\{ 1, e^{-j\pi \sin(\phi_k)}, e^{-j\pi \sin(\phi_k)2}, \dots, e^{-j\pi \sin(\phi_k)(M_y-1)} \right\}^T, \\ &e^{-j\pi \sin(\theta_i)2} \left\{ 1, e^{-j\pi \sin(\phi_k)}, e^{-j\pi \sin(\phi_k)2}, \dots, e^{-j\pi \sin(\phi_k)(M_y-1)} \right\}^T, \\ &\vdots \\ &\left. e^{-j\pi \sin(\theta_i)(M_x-1)} \left\{ 1, e^{-j\pi \sin(\phi_k)}, e^{-j\pi \sin(\phi_k)2}, \dots, \right. \right. \\ &\left. \left. \dots, e^{-j\pi \sin(\phi_k)(M_y-1)} \right\}^T \right]. \end{aligned}$$

Accordingly, the transmitter antenna array response matrix \mathbf{A}_T and receiver antenna array response \mathbf{A}_R are given as:

$$\begin{aligned} \mathbf{A}_T &= \left[\mathbf{a}(\theta_1^t, \phi_1^t) \ \mathbf{a}(\theta_1^t, \phi_2^t) \ \dots \ \mathbf{a}(\theta_1^t, \phi_{N_{g\phi_1}^t}) \ \mathbf{a}(\theta_2^t, \phi_1^t) \ \dots \right. \\ &\quad \left. \dots \ \mathbf{a}(\theta_{N_{g\theta_1}^t}, \phi_{N_{g\phi_1}^t}^t) \right], \\ \mathbf{A}_R &= \left[\mathbf{a}(\theta_1^r, \phi_1^r) \ \mathbf{a}(\theta_1^r, \phi_2^r) \ \dots \ \mathbf{a}(\theta_1^r, \phi_{N_{g\phi_r}^r}) \ \mathbf{a}(\theta_2^r, \phi_1^r) \ \dots \right. \\ &\quad \left. \dots \ \mathbf{a}(\theta_{N_{g\theta_r}^r}, \phi_{N_{g\phi_r}^r}^r) \right] \end{aligned} \quad (11)$$

where (θ_i^t, ϕ_k^t) represent the transmitter (azimuth, elevation) angles (with grid resolutions $N_{g\theta_t}, N_{g\phi_t}$) and (θ_i^r, ϕ_k^r) represent the receiver (azimuth, elevation) angles (with grid resolutions $N_{g\theta_r}, N_{g\phi_r}$) respectively.

The resulting antenna array response matrices for transmitter and receiver have full rank which is proven in Appendix 1.

Lemma 1 Let the azimuth angle grid (\mathcal{G}_θ) and elevation angle grid (\mathcal{G}_ϕ) satisfy (8), the number of row and column elements at the transmitter and receiver be $M_{Tx}, M_{Ty}, M_{Rx}, M_{Ry}$ respectively, the grid resolution for azimuth and elevation angles at the transmitter be $N_{g\theta_t} \geq M_{Tx}$ and $N_{g\phi_t} \geq M_{Ty}$, the grid resolution for azimuth and elevation angles at the receiver be $N_{g\theta_r} \geq M_{Rx}$ and $N_{g\phi_r} \geq M_{Ry}$, while the elements of the antenna array are critically spaced i.e, $d = \frac{\lambda}{2}$, the antenna array responses fulfill,

$$\mathbf{A}_T \mathbf{A}_T^H = \frac{N_{g\theta_t} N_{g\phi_t}}{M_{Tx} M_{Ty}} \mathbf{I}_{N_T} \text{ and } \mathbf{A}_R \mathbf{A}_R^H = \frac{N_{g\theta_r} N_{g\phi_r}}{M_{Rx} M_{Ry}} \mathbf{I}_{N_R} \quad (12)$$

where \mathbf{I}_{N_T} and \mathbf{I}_{N_R} are identity matrices.

The antenna array responses for the transmitter \mathbf{A}_T and receiver \mathbf{A}_R form part of the dictionary for the system and their orthogonality allows minimized coherence of the sensing matrix for efficient compressed sensing (CS) based problem formulation.

3. Channel Formulation and Estimation

The estimation of geometric mm-wave channel model for UPA given in (4) calls for finding an estimate of the azimuth and elevation AoD from the transmitter and azimuth and elevation AoA at the receiver along with the complex gain vector for each of the paths. The limited scattering nature of mm-wave leads to fewer paths in the channel, which makes the complex gains vector sparse.

3.1 Sparse Channel Formulation

In order to formulate the sparse channel, the channel model (7) formed by orthogonal transmit and receive antenna arrays is substituted in the collective channel received vectors matrix form given in (3) as:

$$\mathbf{Y} = \sqrt{P} \mathbf{W}^H \mathbf{A}_R \text{diag}(\alpha) \mathbf{A}_T^H \mathbf{X} + \mathbf{W}^H \mathbf{N}. \quad (13)$$

To highlight the sparse nature of the channel [22], the matrix \mathbf{Y} is vectorized as:

$$\mathbf{y}_{\text{vec}} = \sqrt{P} \text{vec}(\mathbf{W}^H \mathbf{A}_R \text{diag}(\alpha) \mathbf{A}_T^H \mathbf{X}) + \text{vec}(\mathbf{W}^H \mathbf{N}),$$

$$\mathbf{y}_{\text{vec}} = \sqrt{P} \underbrace{(\mathbf{X}^T \mathbf{A}_T^* \otimes \mathbf{W}^H \mathbf{A}_R)}_{\mathbf{F}} \alpha + \text{vec}(\mathbf{W}^H \mathbf{N}). \quad (14)$$

In CS formulation, the mutual coherence of a matrix provides sufficient recovery guarantee of a sparse vector. The coherence of a sensing matrix \mathbf{F} which is based on the Kronecker product is given as [25]:

$$\mu(\mathbf{F}) = \max \{ \mu(\mathbf{X}^T \mathbf{A}_T^*), \mu(\mathbf{W}^H \mathbf{A}_R) \}. \quad (15)$$

This decoupling of the Kronecker product allows us to focus on individual coherences of the precoding matrix \mathbf{X}^T with orthogonal transmitter antenna array response \mathbf{A}_T^* as defined in Section 2.3 for a particular azimuth/elevation angular grids $N_{g\theta_t}/N_{g\phi_t}$, and similarly for combining matrix \mathbf{W}^H with orthogonal receiver antenna array response \mathbf{A}_R .

3.2 Channel Estimation

The reconstruction of the complex gains sparse vector α leads to the recovery of azimuth and elevation angles of departure/arrival respectively. The reconstruction of the sparse vector based on (14) is given as:

$$\min_{\hat{\alpha}, \hat{\theta}^t, \hat{\phi}^t, \hat{\theta}^r, \hat{\phi}^r} \|\alpha\|_0, \text{ such that } \|\mathbf{y}_{\text{vec}} - \mathbf{F}\alpha\|_F \quad (16)$$

where $\hat{\alpha}$ is the estimated complex gains vector, the estimated AoD azimuth angles vector $\hat{\theta}^t$, the estimated AoD elevation angles vector $\hat{\phi}^t$, the estimated AoA azimuth angles vector $\hat{\theta}^r$ and the estimated AoA elevation angles vector $\hat{\phi}^r$ respectively.

4. Pilot Beams Design

Mutual coherence of a sensing matrix is an important attribute and its minimization improves CS performance [26]. Measurement matrices are adapted or coherence optimized with respect to dictionaries in [27], [28].

The overall coherence of the channel model (15) depends on as low coherence as possible between precoder matrix \mathbf{X}^T and the orthogonal transmitter antenna array response matrix \mathbf{A}_T^* and similarly for combiner matrix \mathbf{W}^H and the orthogonal receiver antenna array response matrix \mathbf{A}_R . The antenna array matrices \mathbf{A}_T^* and \mathbf{A}_R are the transmitter and receiver dictionaries for the system. The coherence minimization problem with RF processors as unitary matrices ($\mathbf{X}_{\text{RF}}/\mathbf{W}_{\text{RF}}$) is difficult, therefore, proposed coherence minimization focuses on adaptation of baseband precoder (\mathbf{X}_{BB}) and combiner (\mathbf{W}_{BB}) with respect to their respective dictionaries. Grassmannian codebook generation algorithm given in [29] is adapted for dictionary adaptation of the baseband precoder/combiner matrix. The algorithm optimizes the coherence of initial random complex matrix $\mathbf{U} \in \mathbb{C}^{Z \times C}$ with $Z < C$. It performs Lagrange based optimization of the objective function,

$$g(\mathbf{U}, \lambda) = \sum_{b \neq s} \left(|\langle \mathbf{u}_b, \mathbf{u}_s \rangle|^2 - \beta^2 \mu_{\text{bd}}^2 \right)^p + \sum_{b=1}^C \lambda_b (\|\mathbf{u}_b\|^2 - 1) \quad (17)$$

where $\mathbf{u}_b, \mathbf{u}_s$ are column vectors of matrix \mathbf{U} , λ are the Lagrange multipliers with unit radius, $p \in \{2^1, 2^2, \dots, 2^9\}$ is the free parameter, $\beta = 0.5$ is the factor to improve stability during convergence of the algorithm and μ_{bd} is the lower bound on coherence of matrix \mathbf{U} [29], given as:

$$\mu_{\text{bd}}(\mathbf{U}) = \begin{cases} \sqrt{\frac{C-Z}{Z(C-1)}}, & \text{if } C \leq Z^2 \\ \max \left(\sqrt{\frac{1}{Z}}, \sqrt{\frac{2C-Z^2-Z}{(Z+1)(C-Z)}}, 1 - 2C \frac{1}{Z-1} \right), & \text{if } Z^2 < C \leq 2(Z^2 - 1) \\ \max \left(\sqrt{\frac{2C-Z^2-Z}{(Z+1)(C-Z)}}, 1 - 2C \frac{1}{Z-1} \right), & \text{if } 2(Z^2 - 1) < C. \end{cases} \quad (18)$$

The force vector to cover the cumulative effect of all other vectors on each vector is derived by finding the equilibrium conditions for (17) as:

$$\mathbf{f}_b^{(\delta)} = -2 \sum_{b \neq s} \left(|\langle \mathbf{u}_b^{(\delta)}, \mathbf{u}_s^{(\delta)} \rangle|^2 - \beta^2 \mu_{\text{bd}}^2 \right)^{p-1} \langle \mathbf{u}_s^{(\delta)}, \mathbf{u}_b^{(\delta)} \rangle \mathbf{u}_s^{(\delta)} \quad (19)$$

where underlining highlights normalization and $\mathbf{u}_b^{(\delta)}$ represents unit norm vector at index b in δ^{th} iteration.

After application of the forces on each vector the matrix is updated. Sequential optimization process moves to the next value of p causing coarse to fine adjustments of vectors to achieve better convergence, once condition $\|\mathbf{u}_b^{(\delta)} - \mathbf{u}_b^{(\delta-1)}\| < \epsilon$ is met.

In order to modify the algorithm for dictionary adaptation of the baseband precoding matrix $\mathbf{X}_{\text{BB}}^T \in \mathbb{C}^{N_x \times N_T}$ with respect to dictionary $\mathbf{A}_T^* \in \mathbb{C}^{N_T \times N_{g\theta} N_{g\phi}}$ while having fixed RF component \mathbf{X}_{RF}^T as unitary matrix, the rows of an initial random uniformly distributed phase matrix $\mathbf{X}_{\text{rd}}^T \in \mathbb{C}^{N_x \times N_T}$ should become as incoherent to the columns of dictionary \mathbf{A}_T^* as possible. To achieve this, a new input matrix $\mathbf{B} \in \mathbb{C}^{N_T \times \{N_x + N_{g\theta} N_{g\phi}\}}$ is formed by concatenating initial random matrix \mathbf{X}_{rd} with the dictionary matrix \mathbf{A}_T^* . The concatenated matrix \mathbf{B} is formed in a manner where once given as input to the modified algorithm provides an opportunity to make precoding vectors incoherent with themselves as well as the dictionary vectors.

In Algorithm 1, Line 8 [29], the conditions for progression of the iterative process to next value of free parameter p are set where either the number of iterations δ reaches maximum iterations i.e., δ_{max} or all column vectors of the prospective precoding matrix ($\mathbf{b}_i^{(\delta)}$ representing i^{th} column with $i \in \{1, 2, \dots, N_x\}$) of matrix \mathbf{B} in δ^{th} iteration) converge onto themselves and fulfill the condition $\|\mathbf{b}_i^{(\delta)} - \mathbf{b}_i^{(\delta-1)}\| \leq \epsilon$. This modified condition for the loop for each value of free parameter p allows convergence of only precoding matrix part of the concatenated matrix \mathbf{B} . In Algorithm 1, Line 9, the resultant force vector for each column vector ($1 : N_x$) involving contributions from all other vectors is calculated.

The interim matrix is partially updated after dampened application of force (Algorithm 1, Line 10). The update of the matrix in each iteration only alters the initial N_x columns of the matrix which correspond to the baseband precoding matrix and keeps the dictionary part unaltered. Even though the dictionary is orthogonal, the adapted algorithm has slower convergence due to updation of a limited part of the matrix. Correspondingly, the error threshold used to measure convergence of the precoding matrix part of the matrix \mathbf{B} is lowered from 10^{-10} in [29] to 10^{-12} and the maximum iterations δ_{\max} for each value of free parameter p is increased from 10^5 to 5×10^5 to avoid premature advancement of the algorithm and achieve better convergence. Dictionary adapted pilot vectors matrix \mathbf{X}_{BB} (baseband precoding matrix) is taken as output of the algorithm by extracting initial N_x columns of matrix \mathbf{B} after final convergence. The same algorithm can be used to adapt $\mathbf{W}_{\text{BB}}^{\text{H}}$ (baseband combining matrix) with respect to its dictionary \mathbf{A}_{R} .

Algorithm 1: Coherence optimized pilot beam patterns.

```

Let the input matrices  $\mathbf{X}_{\text{rd}}$  and  $\mathbf{A}_{\text{T}}^*$  be taken using (3) and (11)
Initialize:
1:  $\mathbf{B} = [\mathbf{X}_{\text{rd}} \ \mathbf{A}_{\text{T}}^*] \in \mathbb{C}^{N_{\text{T}} \times (N_x + N_{\text{g}\theta} N_{\text{g}\phi})}$            ▶ Input Matrix
2:  $\mu_{\text{bd}}$                                ▶ Coherence bound for precoding matrix (18)
3:  $\alpha_{\text{init}} \leftarrow 0.04, \alpha \leftarrow \alpha_{\text{init}}$            ▶ Dampening factor
4:  $\delta_{\text{max}} \leftarrow 5 \times 10^5$                                ▶ Max iterations
5:  $p_{\text{max}} \leftarrow 2^9, p \leftarrow 2,$                                ▶ Free parameter
6:  $\epsilon = 10^{-12}$                                            ▶ Error threshold
7:  $i \in \{1, 2, \dots, N_x\}$                                ▶ Column index of baseband precoding matrix
   while  $p \leq p_{\text{max}}$  do
      $\delta = 1,$                                            ▶ Initialize iteration index
     while  $\delta \leq \delta_{\text{max}}$  & any  $\|\mathbf{b}_i^{(\delta)} - \mathbf{b}_i^{(\delta-1)}\| \geq \epsilon$  do
       9: Calculate  $\mathbf{F}^{(\delta)} = [\mathbf{f}_1^{(\delta)} \ \mathbf{f}_2^{(\delta)} \ \dots \ \mathbf{f}_{N_x}^{(\delta)}]$  ▶ Force calculation (19)
       10:  $\mathbf{B}^{(\delta)}[:, 1 : N_x] = \mathbf{B}^{(\delta-1)}[:, 1 : N_x] + \alpha \mathbf{F}^{(\delta)}[:, 1 : N_x]$ 
           ▶ Update & normalize
        $\delta = \delta + 1$                                        ▶ Increase iteration
     end
     13:  $p \leftarrow 2p, \alpha \leftarrow \frac{\alpha_{\text{init}}}{(p-1)}$  ▶ Adjust free parameter and dampening factor
   end
Output:  $\mathbf{X}_{\text{BB}} = \mathbf{B}[:, 1 : N_x]$  ▶ Dictionary adapted pilot vectors

```

In Fig. 3a and 3b, the coherence profile of the off-diagonal values of the Gram matrices based on concatenation of RF-beamformer (DFT matrix), baseband precoding/combining matrices and antenna array response matrices is shown. The random precoding matrix \mathbf{X}_{rd} and combining matrix \mathbf{W}_{rd} are based on random uniformly distributed phases. \mathbf{X}_{BB} is the baseband precoding matrix adapted to its dictionary \mathbf{A}_{T}^* and \mathbf{W}_{BB} is the baseband combining matrix adapted to its dictionary \mathbf{A}_{R} . In the distribution, the efficacy of the algorithm is evident where the worst case coherence is considerably reduced and the spread of the off-diagonal values is also narrowed as compared to the random phases matrices.

In Fig. 3c, the distribution of off-diagonal values of the Gram matrices based on system transmitter side component ($\mathbf{X}_{\text{RF}}^{\text{T}} \mathbf{X}_{\text{BB}}^{\text{T}} \mathbf{A}_{\text{T}}^*$ and $\mathbf{X}_{\text{RF}}^{\text{T}} \mathbf{X}_{\text{rd}}^{\text{T}} \mathbf{A}_{\text{T}}^*$) is given. In Fig. 3d, distribution of off-diagonal values of the Gram matrices based on system

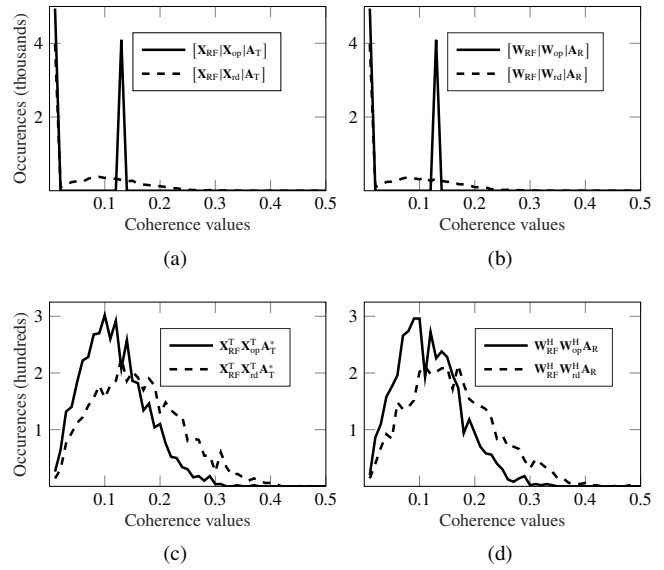


Fig. 3. Off-diagonal values distribution based on Gram matrices with $N_{\text{T}} = N_{\text{R}} = 8 \times 8$, $N_x = N_y = 32$, $N_{\text{g}\theta} = N_{\text{g}\phi} = 8$, $\{\mathbf{X}_{\text{rd}}, \mathbf{X}_{\text{op}}, \mathbf{W}_{\text{rd}}, \mathbf{W}_{\text{op}}\} \in \mathbb{C}^{64 \times 32}$, $\{\mathbf{X}_{\text{RF}}, \mathbf{W}_{\text{RF}}\} \in \mathbb{C}^{64 \times 64}$, $\mathbf{A}_{\text{T}}, \mathbf{A}_{\text{R}} \in \mathbb{C}^{64 \times 64}$.

receiver side component ($\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{W}_{\text{BB}}^{\text{H}} \mathbf{A}_{\text{R}}$ and $\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{W}_{\text{rd}}^{\text{H}} \mathbf{A}_{\text{R}}$) is shown. It is revealed that the overall coherence profile of the system for the dictionary adapted coherence optimized baseband precoding/combining matrices is lower as compared to the random phase baseband precoding/combining matrices which merits better sparse channel estimation probability. Moreover, if precoding/combining matrices $\mathbf{X}_{\text{BB}}/\mathbf{W}_{\text{BB}}$ are taken as Grassmannian codebooks with the appropriate dimensions which are optimized from initial random distributed phase matrices using algorithm in [29] without dictionary adaptation, the resulting coherence profile for the system would be worse than the one after coherence optimization involving adaptation to their respective dictionaries.

In Fig. 4, the coherence variation of the overall system with optimized baseband precoder/combiner matrices is shown while having varying grid resolutions for each fixed square ($M_{\text{tx}} = M_{\text{ty}} = M_{\text{rx}} = M_{\text{ry}}$) antenna configuration considered. Each plot shows the coherence variation for a particular antenna configuration, where the grid resolution starts with the resolution equivalent to the antenna configuration and then progressively finer grid resolutions. It is revealed that increase in grid resolution for any antenna configuration increases coherence. It is also worth mentioning that the coherence for antenna configurations with equivalent grid resolution steadily decreases as larger antenna configurations are considered. The line joining coherence for antenna array equivalent grid resolution follows an exponential-like distribution. Based on these observations, an optimum antenna configuration with suitable resolution can be selected for a particular scenario. Furthermore, the normalized mean squared error (NMSE) profile for such variation can help ascertain which configuration is best suited.

Algorithm 1 depends on the orientation and coherence of the initial random seed \mathbf{X}_{rd} which is given as input to the algorithm. For certain configurations, due to the geometry of the sensing matrix and the dictionary considered, the progressive algorithm takes longer to converge. At times the algorithm doesn't converge to a desirable mutual coherence level with respect to a given quantized antenna array response taken as dictionary.

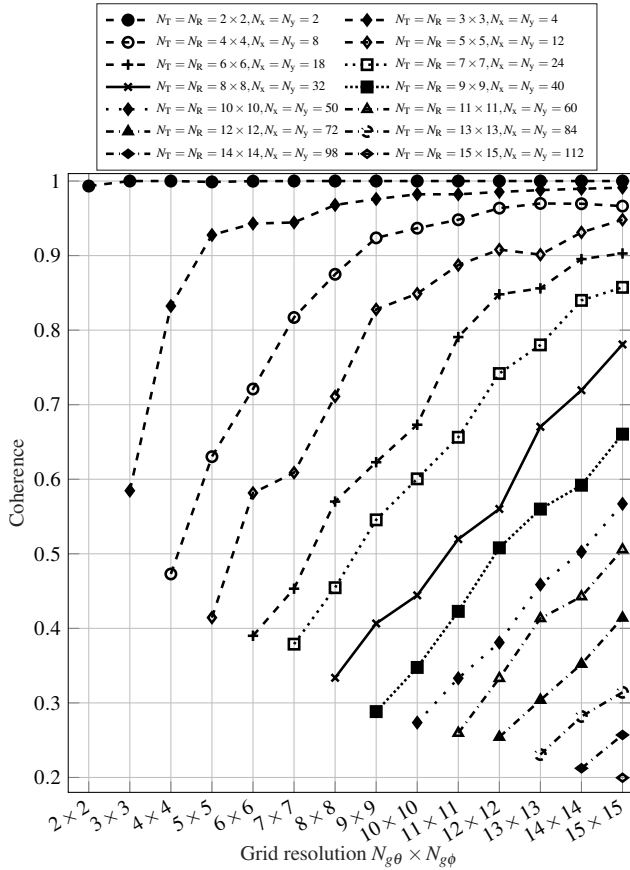


Fig. 4. Coherence grid comparison $\mu = \max\{\mu(\mathbf{X}^T \mathbf{A}_T^*), \mu(\mathbf{W}^H \mathbf{A}_R)\}$ with $\mathbf{X} \in \mathbb{C}^{(M_{Tx} \times M_{Ty}) \times N_x}$, $\mathbf{A}_T \in \mathbb{C}^{(M_{Tx} \times M_{Ty}) \times N_{g\theta} N_{g\phi}}$ and $\mathbf{W} \in \mathbb{C}^{N_y \times M_{Tx} \times M_{Ty}}$, $\mathbf{A}_R \in \mathbb{C}^{(M_{Tx} \times M_{Ty}) \times N_{g\theta} N_{g\phi}}$.

The algorithm is implemented in Matlab R2017a, and the machine used has Intel Core i7-6700 CPU @ 3.40GHz octa-core 64-bit processor with 32GB RAM. In order to highlight the time complexity of the algorithm, five test trials for each antenna configuration considered with varying antenna angular grid resolutions based dictionaries are performed. In Fig. 5, box plots for the computation times taken by the three UPA antenna configurations i.e., $\{\mathbf{X} \in \mathbb{C}^{(4 \times 4) \times 8}, \mathbf{A}_T \in \mathbb{C}^{(4 \times 4) \times N_{g\theta} N_{g\phi}}\}$, $\{\mathbf{X} \in \mathbb{C}^{(6 \times 6) \times 18}, \mathbf{A}_T \in \mathbb{C}^{(6 \times 6) \times N_{g\theta} N_{g\phi}}\}$ and $\{\mathbf{X} \in \mathbb{C}^{(8 \times 8) \times 32}, \mathbf{A}_T \in \mathbb{C}^{(8 \times 8) \times N_{g\theta} N_{g\phi}}\}$ are shown. It is observed that the computation time increases for each configuration once finer grid resolutions are considered. Moreover, the time complexity also increases once larger antenna arrays are considered.

Table 1 provides comparison of the proposed approach with the recent work related to TS designing in hybrid mm-wave massive MIMO.

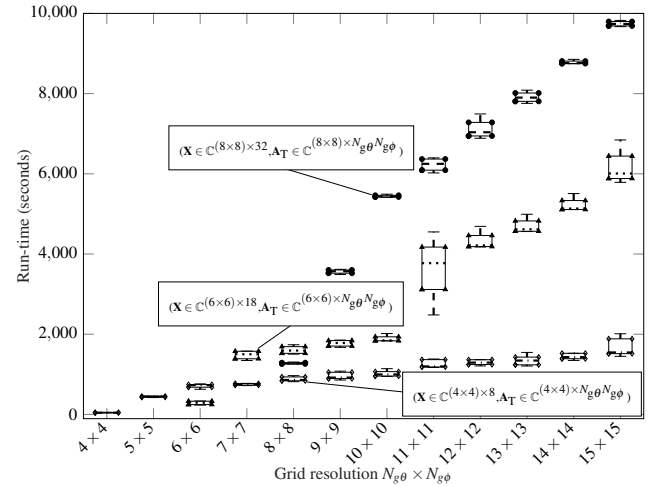


Fig. 5. Computation time variation comparison in 5 test trials.

Technique	AOA/AOD		Antenna	TS design	Complexity	Recovery algorithm	
	Distribution	Range					
		θ					ϕ
Ma et al. [17]	Continuous	$\{0, 2\pi\}$	–	ULA	Coherence minimized	Low	Block-OMP
Manoj and Kannu [18]	Uniform	$\{-\pi, \pi\}$	–	ULA	Random	Low	Generalized Block-OMP
Nguyen et al. [19]	Uniform	$\{0, 2\pi\}$	$\{-\frac{\pi}{2}, \frac{\pi}{2}\}$	UPA	MMSE via OMP	Medium	OMP
					Weighted-MMSE via OMP	Medium	
Lee et al. [20]	Uniform $\cos(\theta)$	Orthogonal Quantized $\{0, \pi\}$	–	ULA	Coherence minimized	Low	OMP
Proposed	Uniform $\sin(\theta)$ and $\sin(\phi)$	Orthogonal Quantized $\{-\frac{\pi}{2}, \frac{\pi}{2}\}$	Orthogonal Quantized $\{-\frac{\pi}{2}, \frac{\pi}{2}\}$	UPA	Coherence optimized	High	Iterative re-weight super resolution

Tab. 1. TS design comparison for mm-wave massive MIMO.

5. Simulation Results

In this section, the performance of the proposed dictionary adapted precoding/combining matrices with proposed orthogonal antenna array responses for UPA is evaluated.

5.1 Error Analysis

NMSE is defined as $10 \log_{10} [\mathbb{E}\{\|\mathbf{H} - \mathbf{H}_{\text{est}}\|^2 / \|\mathbf{H}\|^2\}]$ is compared for UPA in Fig. 6. The channel estimation is based on iterative re-weight super resolution algorithm [21]. In the simulation $\mathbf{X}_{\text{RF}}/\mathbf{W}_{\text{RF}}$ are assumed to be unitary matrices based on DFT matrix for all the variants considered. In baseband, the proposed dictionary adapted precoding \mathbf{X}_{BB} and combiner \mathbf{W}_{BB} matrices based system is compared with coherence minimized sensing matrix based system [20] and for hybrid precoding based on random uniformly distributed phase [21]. The results are observed for 1000 channel realizations for each level of SNR. In Fig. 6, the transmitter N_T and receiver N_R UPA have 8×8 elements, while having RF chains $N_T^{\text{RF}} = N_R^{\text{RF}} = 8$, transmission block $N_T^{\text{Block}} = \frac{N_T}{N_T^{\text{RF}}} = 8$, receiver block $N_R^{\text{Block}} = \frac{N_R}{N_R^{\text{RF}}} = 8$, transmission and reception beams $N_x = N_y = 32$ and number of paths $L = 5$ respectively. The signal to noise ratio is defined as $\frac{P\sigma_\alpha^2}{\sigma_n^2}$, where complex path gains α_i are assumed to be normally distributed with 0 mean, σ_α^2 variance and σ_n^2 is the noise variance. The proposed technique shows better NMSE performance owing to lower coherence of the overall system. The NMSE performance

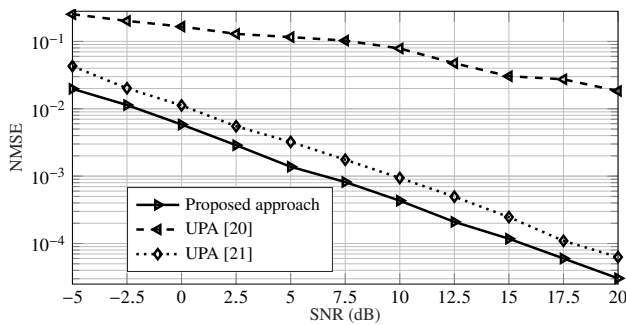


Fig. 6. NMSE comparison $\mathbf{X} \in \mathbb{C}^{(8 \times 8) \times 32}$, $\mathbf{A}_T \in \mathbb{C}^{(8 \times 8) \times 64}$ and $\mathbf{W} \in \mathbb{C}^{32 \times (8 \times 8)}$, $\mathbf{A}_R \in \mathbb{C}^{(8 \times 8) \times 64}$.

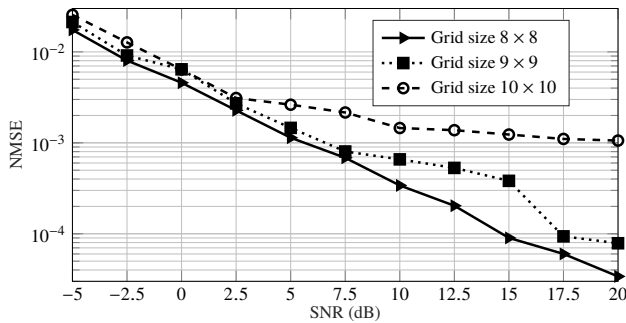


Fig. 8. NMSE comparison $L = 10$, $\mathbf{X} \in \mathbb{C}^{(8 \times 8) \times 32}$, $\mathbf{A}_T \in \mathbb{C}^{(8 \times 8) \times N_{g\theta} N_{g\phi}}$ and $\mathbf{W} \in \mathbb{C}^{32 \times (8 \times 8)}$, $\mathbf{A}_R \in \mathbb{C}^{(8 \times 8) \times N_{g\theta} N_{g\phi}}$.

of the proposed hybrid precoding system for UPA is similar to the results achieved for uniform linear array in [21].

In order to observe the variation in NMSE for increasing number of paths, in Fig. 7, the transmitter N_T and receiver N_R UPA have 14×14 elements, while having RF chains $N_T^{\text{RF}} = N_R^{\text{RF}} = 14$, transmission block $N_T^{\text{Block}} = \frac{N_T}{N_T^{\text{RF}}} = 14$, receiver block $N_R^{\text{Block}} = \frac{N_R}{N_R^{\text{RF}}} = 14$, transmission and reception beams $N_x = N_y = 98$ and number of paths $L = \{5, 10, 15\}$ respectively. It is observed that the NMSE performance of the proposed approach doesn't deteriorate much with increasing number of paths.

In order to observe the variation in NMSE for increasing grid size while keeping antenna size fixed, consider transmitter N_T and receiver N_R UPA having 8×8 elements, while having RF chains $N_T^{\text{RF}} = N_R^{\text{RF}} = 8$, transmission block $N_T^{\text{Block}} = \frac{N_T}{N_T^{\text{RF}}} = 8$, receiver block $N_R^{\text{Block}} = \frac{N_R}{N_R^{\text{RF}}} = 8$, transmission and reception beams $N_x = N_y = 32$, number of paths $L = 10$ and grid sizes $N_{g\theta} \times N_{g\phi} \in \{(8 \times 8), (9 \times 9), (10 \times 10)\}$ in Fig. 8 respectively. Furthermore, in Fig. 9, transmitter N_T and receiver N_R UPA having 12×12 antenna elements, while having RF chains $N_T^{\text{RF}} = N_R^{\text{RF}} = 12$, transmission block $N_T^{\text{Block}} = \frac{N_T}{N_T^{\text{RF}}} = 12$, receiver block $N_R^{\text{Block}} = \frac{N_R}{N_R^{\text{RF}}} = 12$, transmission and reception beams $N_x = N_y = 72$, number of paths $L = 10$ and grid sizes $N_{g\theta} \times N_{g\phi} \in \{(12 \times 12), (13 \times 13), (14 \times 14)\}$ respectively are considered. It is observed that the error profile worsens with increase in grid size for a particular antenna configuration.

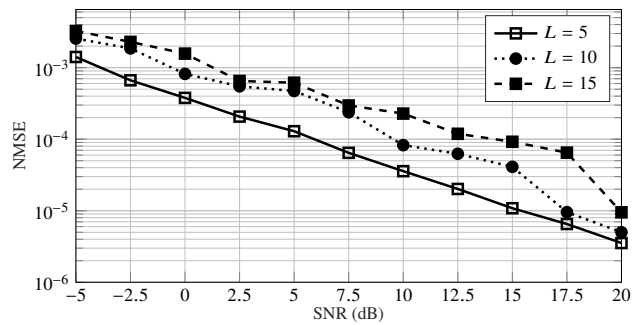


Fig. 7. NMSE comparison $L = \{5, 10, 15\}$, $\mathbf{X} \in \mathbb{C}^{(14 \times 14) \times 98}$, $\mathbf{A}_T \in \mathbb{C}^{(14 \times 14) \times 196}$ and $\mathbf{W} \in \mathbb{C}^{98 \times (14 \times 14)}$, $\mathbf{A}_R \in \mathbb{C}^{(14 \times 14) \times 196}$.

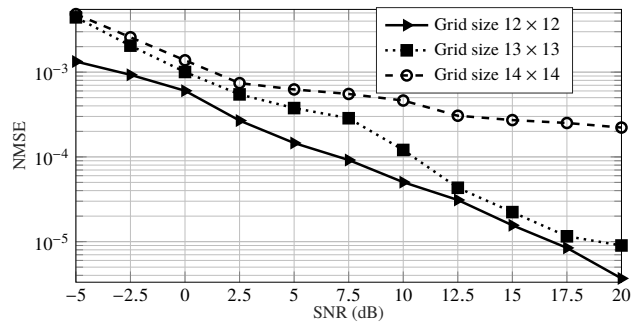


Fig. 9. NMSE comparison $L = 10$, $\mathbf{X} \in \mathbb{C}^{(12 \times 12) \times 72}$, $\mathbf{A}_T \in \mathbb{C}^{(12 \times 12) \times N_{g\theta} N_{g\phi}}$ and $\mathbf{W} \in \mathbb{C}^{72 \times (12 \times 12)}$, $\mathbf{A}_R \in \mathbb{C}^{(12 \times 12) \times N_{g\theta} N_{g\phi}}$.

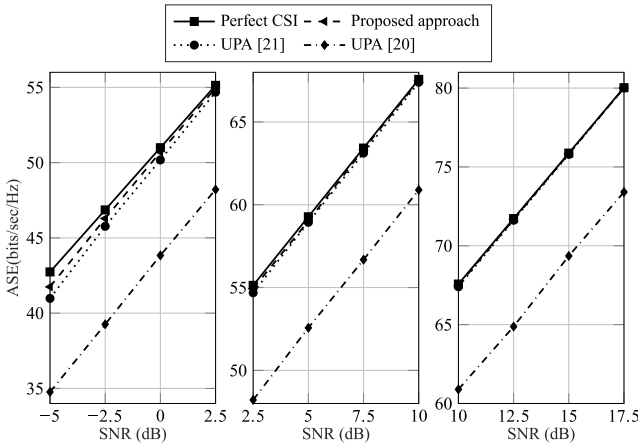


Fig. 10. ASE comparison $\mathbf{X} \in \mathbb{C}^{(8 \times 8) \times 32}$, $\mathbf{A}_T \in \mathbb{C}^{(8 \times 8) \times 64}$ and $\mathbf{W} \in \mathbb{C}^{32 \times (8 \times 8)}$, $\mathbf{A}_R \in \mathbb{C}^{(8 \times 8) \times 64}$, $L = 5$.

5.2 Spectral Efficiency

The achievable spectral efficiency (ASE) for random symbols transmitted is defined in [4] as:

$$R = \log_2 \left(\mathbf{I}_{N_x} + \frac{\rho}{N_x} \mathbf{R}_n^{-1} \mathbf{W}_{BB}^* \mathbf{W}_{RF}^* \mathbf{H} \mathbf{X}_{RF} \mathbf{X}_{BB} \times \mathbf{X}_{BB}^* \mathbf{X}_{RF}^* \mathbf{H} \mathbf{W}_{RF} \mathbf{W}_{BB} \right)$$

where $\mathbf{R}_n = \sigma_n^2 \mathbf{W}_{BB}^* \mathbf{W}_{RF}^* \mathbf{W}_{RF} \mathbf{W}_{BB}$ is the noise covariance matrix after combining. In Fig. 10, the ASE profile for transmitter N_T and receiver N_R UPA having 8×8 elements, while having RF chains $N_T^{\text{RF}} = N_R^{\text{RF}} = 8$, transmission block $N_T^{\text{Block}} = \frac{N_T}{N_T^{\text{RF}}} = 8$, receiver block $N_R^{\text{Block}} = \frac{N_R}{N_R^{\text{RF}}} = 8$, transmission and reception beams $N_x = N_y = 32$, number of paths $L = 5$ is considered. It is observed that the proposed approach reaches the perfect channel state information (CSI) case earlier than the existing techniques.

6. Conclusion

In this paper, two dimensional angular grid based (azimuth and elevation angle) antenna array representation for UPA is suggested. It is shown that the array representation remains full rank for grid resolution equivalent to the number of antennas in the array or for the case where finer grid resolutions are considered. Training vectors are designed to reduce the overall coherence of the mm-wave hybrid MIMO system. Lowered coherence of the equivalent sensing matrix for the system allows better channel estimation. Iterative re-weight super resolution algorithm is used for open-loop channel estimation and the NMSE profile shows considerable advantage over existing techniques. Similarly, achievable sum rate efficiency profile reveals that the proposed approach closely follows and reaches the perfect channel state information case.

References

- [1] RAPPAPORT, T. S., SUN, S., RIMMA, M., et al. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE Access*, 2013, vol. 1, p. 335–349. DOI: 10.1109/ACCESS.2013.2260813
- [2] RAPPAPORT, T. S., XING, Y., MACCARTNEY, G., et al. Overview of millimeter wave communications for fifth-generation (5G) wireless networks - With a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, 2017, vol. 65, no. 12, p. 6213–6230. DOI: 10.1109/TAP.2017.2734243
- [3] AKDENIZ, M. R., LIU, Y., SAMIMI, M. K., et al. Millimeter wave channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, 2014, vol. 32, no. 6, p. 1164–1179. DOI: 10.1109/JSAC.2014.2328154
- [4] AYACH, O. E., RAJAGOPAL, S., ABU-SURRA, S., et al. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 2014, vol. 13, no. 3, p. 1499–1513. DOI: 10.1109/TWC.2014.011714.130846
- [5] AHMED, I., KHAMMARI, H., SHAHID, A., et al. A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives. *IEEE Communications Surveys & Tutorials*, 2018, vol. 20, no. 4, p. 3060–3097. DOI: 10.1109/COMST.2018.2843719
- [6] LEE, J., KANG, M., OH, J., et al. Space-time alignment for channel estimation in millimeter wave communication with beam sweeping. In *IEEE Global Communications Conference (GLOBECOM)*. Singapore, 2017, p. 1–7. DOI:10.1109/GLOCOM.2017.8254894
- [7] QIN, Q., GUI, L., CHENG, P., et al. Time-varying channel estimation for millimeter wave multiuser MIMO systems. *IEEE Transactions on Vehicular Technology*, 2018, vol. 67, no. 10, p. 9435–9448. DOI: 10.1109/TVT.2018.2854735
- [8] SHAHMANSOORI, A., GARCIA, G. E., DESTINO, G., et al. Position and orientation estimation through millimeter-wave MIMO in 5G systems. *IEEE Transactions on Wireless Communications*, 2018, vol. 17, no. 3, p. 1822–1835. DOI: 10.1109/TWC.2017.2785788
- [9] YANG, J., WEI, Z., ZHANG, X., et al. Correlation based adaptive compressed sensing for millimeter wave channel estimation. In *IEEE Wireless Communications and Networking Conference (WCNC)*. San Francisco (CA, USA), 2017, p. 1–6. DOI: 10.1109/WCNC.2017.7925685
- [10] QI, B., WANG, W., WANG, B. Off-grid compressive channel estimation for mm-Wave massive MIMO with hybrid precoding. *IEEE Communications Letters*, 2019, vol. 23, no. 1, p. 108–111. DOI: 10.1109/LCOMM.2018.2878557
- [11] PARK, S., HEATH, R. W. Spatial channel covariance estimation for the hybrid MIMO architecture: A compressive sensing-based approach. *IEEE Transactions on Wireless Communications*, 2018, vol. 17, no. 12, p. 8047–8062. DOI: 10.1109/TWC.2018.2873592
- [12] HU, R., TONG, J., XI, J., et al. Matrix completion-based channel estimation for mmwave communication systems with array-inherent impairments. *IEEE Access*, 2018, vol. 6, p. 62915–62931. DOI: 10.1109/ACCESS.2018.2877432
- [13] ZHANG, D., WANG, Y., LI, X., et al. Hybridly connected structure for hybrid beamforming in mmwave massive MIMO systems. *IEEE Transactions on Communications*, 2018, vol. 66, no. 2, p. 662–674. DOI: 10.1109/TCOMM.2017.2756882
- [14] WANG, Y., LIU, A., XIA, X., et al. Exploiting the clustered sparsity for channel estimation in hybrid analog-digital massive MIMO systems. *IEEE Access*, 2019, vol. 7, p. 4989–5000. DOI: 10.1109/ACCESS.2018.2887297

- [15] SRIVASTAVA, S., MISHRA, A., RAJORIYA, A., et al. Quasi-static and time-selective channel Estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian learning (SBL) based approaches. *IEEE Transactions on Signal Processing*, 2019, vol. 67, no. 5, p. 1251–1266. DOI: 10.1109/TSP.2018.2890058
- [16] VIZZIELLO, A., SAVAZZI, P., CHOWDHURY, K. R. A Kalman based hybrid precoding for multi-user millimeter wave MIMO systems. *IEEE Access*, 2018, vol. 6, p. 55712–55722. DOI: 10.1109/ACCESS.2018.2872738
- [17] MA, X., YANG, F., LIU, S., et al. Design and optimization on training sequence for mmwave communications: A new approach for sparse channel estimation in massive MIMO. *IEEE Journal on Selected Areas in Communications*, 2017, vol. 35, no. 7, p. 1486–1497. DOI: 10.1109/JSAC.2017.2698978
- [18] MANOJ, A., KANNU, A. P. Channel estimation strategies for multi-user mm wave systems. *IEEE Transactions on Communications*, 2018, vol. 66, no. 11, p. 5678–5690. DOI: 10.1109/TCOMM.2018.2854188
- [19] NGUYEN, D. H. N., LE, L. B., LE-NGOC, T., et al. Hybrid MMSE precoding and combining designs for mmwave multiuser systems. *IEEE Access*, 2017, vol. 5, p. 19167–19181. DOI: 10.1109/ACCESS.2017.2754979
- [20] LEE, J., GIL, G., LEE, Y. H. Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications. *IEEE Transactions on Communications*, 2016, vol. 64, no. 6, p. 2370–2386. DOI: 10.1109/TCOMM.2016.2557791
- [21] HU, C., DAI, L., MIR, T., et al. Super-resolution channel estimation for mmwave massive MIMO with hybrid precoding. *IEEE Transactions on Vehicular Technology*, 2018, vol. 67, no. 9, p. 8954–8958. DOI: 10.1109/TVT.2018.2842724
- [22] ALKHATEEB, A., AYACH, O. E., LEUS, G., et al. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, 2014, vol. 8, no. 5, p. 831–846. DOI: 10.1109/JSTSP.2014.2334278
- [23] HEATH, R. W., GONZALEZ-PRELCIC, N., RANGAN, S., et al. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 2016, vol. 10, no. 3, p. 436–453. DOI: 10.1109/JSTSP.2016.2523924
- [24] MENDEZ-RIAL, R., RUSU, C., GONZALEZ-PRELCIC, N., et al. Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches? *IEEE Access*, 2016, vol. 4, p. 247–267. DOI: 10.1109/ACCESS.2015.2514261
- [25] JOKAR, S., MEHRMANN, V. Sparse solutions to underdetermined Kronecker product systems. *Linear Algebra and its Applications*, 2009, vol. 431, no. 12, p. 2437–2447. DOI: 10.1016/j.laa.2009.08.005
- [26] ZELNIK-MANOR, L., ROSENBLUM, K., ELDAR, Y. C. Sensing matrix optimization for block-sparse decoding. *IEEE Transactions on Signal Processing*, 2011, vol. 59, no. 9, p. 4300–4312. DOI: 10.1109/TSP.2011.2159211
- [27] ELAD, M. Optimized projections for compressed sensing. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 12, p. 5695–5702. DOI: 10.1109/TSP.2007.900760
- [28] ZORLEIN, H., AKRAM, F., BOSSERT, M. Dictionary adaptation in sparse recovery based on different types of coherence. In *2nd International Workshop on Compressed Sensing applied to Radar (CoSeRa)*. Bonn (Germany), 2013, p. 1–3.
- [29] LAUE, H. E. A., DU PLESSIS, W. P. A coherence-based algorithm for optimizing rank-1 Grassmannian codebooks. *IEEE Signal Processing Letters*, 2017, vol. 24, no. 6, p. 823–827. DOI: 10.1109/LSP.2017.2690466

About the Authors ...

Faisal AKRAM received his bachelors degree in Telecommunication Engineering from National University of Science and Technology (NUST), Pakistan, in 2005, the masters degree in Communication Technology from University of Ulm, Germany, in 2013. He is currently a Ph.D. candidate at the NUST, Pakistan. His research interests include compressed sensing, wireless communication, mm-wave hybrid MIMO systems and channel coding.

Imran RASHID received the B.E. degree in Electrical (Telecomm) Engineering from the NUST, Pakistan, in 1999, the M.Sc. degree in Telecomm Engineering (Optical Communication) from D.T.U. Denmark in 2004, and the Ph.D. degree in Mobile Communication from the University of Manchester, U.K., in 2011. He is currently the Chief Instructor at Engineering Wing, MCS, NUST, Pakistan. His research interests are in mobile and wireless communication, MIMO systems, compressed sensing for MIMO OFDM systems, massive MIMO systems, M2M for mobile systems, cognitive radio networks and cyber security.

Abdul GHAFOOR obtained his Ph.D. Control Systems in 2008 from University of Western Australia. Currently, he is Head of Information Security department at National University of Sciences and Technology, Pakistan. His research interests include model and controller reduction, image processing (enhancement, fusion, segmentation and watermarking), cognitive radios and ground penetration radar.

Adil Masood SIDDIQUI received his bachelor degree in Telecommunications Engineering from Military College of Signals, Rawalpindi Pakistan in 1994, masters degree in Electronics and Telecommunication and Ph.D. in Electrical from University of Engineering and Technology, Lahore Pakistan, in 2005 and 2009 respectively. He is on the faculty of Military College of Signals, NUST Since 2009. His research interest includes image registration, de-noising, image enhancement and defogging. He has number of research publication at his credit.

Appendix A: Proof of Lemma 1

Proof 1 The row of \mathbf{A}_T corresponding to particular row element x and column element y of the transmitter is given as:

$$\mathbf{A}_T(x, y) = \frac{1}{\sqrt{M_{tx}M_{ty}}} \left[e^{-j\pi(x \sin \theta_1^t + y \sin \phi_1^t)}, e^{-j\pi(x \sin \theta_1^t + y \sin \phi_2^t)}, \dots, e^{-j\pi(x \sin \theta_1^t + y \sin \phi_{N_g \phi_t}^t)}, e^{-j\pi(x \sin \theta_2^t + y \sin \phi_1^t)}, \dots, e^{-j\pi(x \sin \theta_{N_g \theta_t}^t + y \sin \phi_{N_g \phi_t}^t)} \right]$$

where $x \in \{1, 2, \dots, M_{tx}\}$ and $y \in \{1, 2, \dots, M_{ty}\}$.

Using (8),

$$e^{-j\pi(x \sin \theta_i^t + y \sin \phi_k^t)} = e^{-j\pi \left[x \left\{ \frac{2}{N_{g\theta_t}}(i-1)+1 \right\} + y \left\{ \frac{2}{N_{g\phi_t}}(k-1)+1 \right\} \right]}$$

$$= e^{-j\pi \left[\frac{2x}{N_{g\theta_t}}(i-1) + \frac{2y}{N_{g\phi_t}}(k-1) + x + y \right]}.$$

$$\mathbf{A}_T(x, y) = \frac{1}{\sqrt{M_{tx}M_{ty}}} \left[e^{-j\pi(x+y)}, e^{-j\pi \left(\frac{2}{N_{g\phi_t}}y + x + y \right)}, \dots, \right.$$

$$e^{-j\pi \left(\frac{2y}{N_{g\phi_t}}(N_{g\phi_t}-1) + x + y \right)}, e^{-j\pi \left(\frac{2}{N_{g\theta_t}}x + x + y \right)},$$

$$\left. \dots, e^{-j\pi \left\{ \frac{2x}{N_{g\theta_t}}(N_{g\theta_t}-1) + \frac{2y}{N_{g\phi_t}}(N_{g\phi_t}-1) + x + y \right\}} \right].$$

Now to check orthogonality, $\mathbf{A}_T(x_1, y_1)\mathbf{A}_T^H(x_2, y_2)$ is given as:

$$= \frac{1}{M_{tx}M_{ty}} \sum_{i=1}^{N_{g\theta_t}} \sum_{k=1}^{N_{g\phi_t}} e^{-j\pi \left[\frac{2}{N_{g\theta_t}} \{ i(x_1+x_2) - x_1 - x_2 \} \right]}$$

$$e^{-j\pi \left[\frac{2}{N_{g\phi_t}} \{ k(y_1+y_2) + y_1 - y_2 \} + x_1 + y_1 + x_2 + y_2 \right]}$$

where x_1, x_2 are the respective row antenna elements and y_1, y_2 are the respective column antenna elements. When $x_1 = x_2$ and $y_1 = y_2$,

$$\mathbf{A}_T(x, y)\mathbf{A}_T^H(x, y) = \frac{1}{M_{tx}M_{ty}} \sum_{i=1}^{N_{g\theta_t}} \sum_{k=1}^{N_{g\phi_t}} e^{-j\pi \left\{ \frac{2}{N_{g\theta_t}}x(2i-2) \right\}}$$

$$e^{-j\pi \left\{ \frac{2}{N_{g\phi_t}}y(2k-2) + 2x + 2y \right\}}$$

$$= \frac{N_{g\theta_t}N_{g\phi_t}}{M_{tx}M_{ty}}.$$

When $x_1 \neq x_2$ or $y_1 \neq y_2$, $\mathbf{A}_T(x_1, y_1)\mathbf{A}_T^H(x_2, y_2) = 0$. The above statements can also be verified by induction.

$$\mathbf{A}_T\mathbf{A}_T^H = \begin{cases} \frac{N_{g\theta_t}N_{g\phi_t}}{M_{tx}M_{ty}}, & \text{for } x_1 = x_2; y_1 = y_2, \\ 0, & \text{otherwise.} \end{cases}$$

$\mathbf{A}_R\mathbf{A}_R^H = \frac{N_{g\theta_t}N_{g\phi_t}}{M_{rx}M_{ry}}\mathbf{I}_{N_R}$ follows similarly.