

## Retraction

# Retracted: Channel Estimation for Switch-Based Millimeter-Wave Communications via Atomic Norm

### International Journal of Antennas and Propagation

Received 19 December 2023; Accepted 19 December 2023; Published 20 December 2023

Copyright © 2023 International Journal of Antennas and Propagation. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] T. Ma, X. Fan, and X. Wu, "Channel Estimation for Switch-Based Millimeter-Wave Communications via Atomic Norm," *International Journal of Antennas and Propagation*, vol. 2021, Article ID 7356963, 9 pages, 2021.

## Research Article

# Channel Estimation for Switch-Based Millimeter-Wave Communications via Atomic Norm

Tao Ma <sup>1,2</sup>, Xiangning Fan,<sup>1</sup> and Xiaohuan Wu<sup>3</sup>

<sup>1</sup>School of Information Science and Engineering, Southeast University, Nanjing 210096, China

<sup>2</sup>NARI Group Corporation, Nanjing, China

<sup>3</sup>School of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Correspondence should be addressed to Tao Ma; 230198982@seu.edu.cn

Received 15 July 2021; Revised 16 August 2021; Accepted 31 August 2021; Published 13 September 2021

Academic Editor: Fangqing Wen

Copyright © 2021 Tao Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Channel estimation is a challenging issue in millimeter-wave massive multiple-input-multiple-output (MIMO) communication systems due to the large number of antennas in the transceiver. Existing methods are usually based on phase shifters which may not be a simple circuit at mmWave band. In this paper, we construct a switch-based architecture for analog processors from the coarray point of view and then propose an atomic  $\ell_0$ -norm minimization problem. We then propose an efficient algorithm to solve this problem based on Wirtinger projection. Since the proposed method requires no angle discretization, it does not suffer from grid mismatch effect that greatly deteriorates the estimation performance of grid-based channel estimation methods. Compared to the atomic norm minimization (ANM) method, our method does not involve vectorization of the channel matrix and hence the dimensionality of the problem is much less than that of ANM. We show that our method is able to provide comparable estimation performance to ANM but with much less computational time. Extensive simulations are carried out to verify the effectiveness of our proposed method.

## 1. Introduction

Millimeter-wave (mmWave) communications is a key technology for the 5<sup>th</sup> generation (5G) mobile communication system. Compared to the sub-6 GHz communication system, the main differentiating factor for mmWave communication systems is the tenfold increase in carrier frequency. This difference can provide multigigabit services which are able to meet future traffic demand [1]. However, the mmWave signals in the high-frequency band suffer from large pathloss; thus, the power of the received signal to be detected by the receiver can be negligible and reliable communication cannot be achieved. To solve this problem, high-resolution beamforming by using massive multiple-input-multiple-output (MIMO) is essential in combating the large pathloss for mmWave communication systems. By using massive MIMO, the transmitter can concentrate the transmitted power on a specific direction to highly improve

the power of the received signal. On the other hand, the small wavelength of mmWave signal allows hundreds of antennas to be accommodated within a reasonable physical size, e.g., the  $8 \times 8$  antenna array in a hand-held unit, making high-resolution beamforming with large-scale antenna arrays possible.

Nevertheless, high-resolution beamforming requires accurate full channel state information (CSI) which is difficult to obtain due to the large number of antennas. Conventional channel estimation methods suffer from high training overhead and complexity [2, 3]. To address this issue, a codebook-based beam searching strategy is proposed to find the correct pair between the beamformer and combiner [4, 5]. Although the hierarchical search can be incorporated to reduce complexity to some extent, the performance heavily depends on the predefined training beam codebook [6]. Another approach is to exploit the sparse nature of the mmWave channel. Due to large

pathloss, there only exist a few ray components between the transmitter and receiver, i.e., the channel is sparse in space [7, 8]. With a high degree of freedom provided by a large number of antennas, we are able to find the angle of arrival (AoA) and angle of departure (AoD) as well as the complex gain of each ray [9, 10]. Then, the channel estimation problem can be formulated as an angle estimation one. Existing methods include compressive sensing- (CS-) based methods [11–13] and subspace-based methods [14–17]. CS-based methods formulate the channel estimation problem as a sparse signal recovery one and utilize the CS recovery methods such as orthogonal matching pursuit (OMP) [18] to retrieve the sparse signal where the indices of nonzero elements indicate the AoAs and AoDs. The subspace method in [14, 19] employs beamsearch 2-D MUSIC to estimate the channel. However, for CS-based methods, formulating the sparse model requires discretizing the angle space into a set of predefined angle grids and then assuming that the AoAs and AoDs exactly lie on the grids [20, 21]. Since the angle space is continuous rather than discrete, this discretizing procedure will bring in nonnegligible bias between the true angle and the closest grid. We call this the grid mismatch effect. Thus, a dense grid set is appealing since its bias can be small. However, since the dimensionality of the sparse model is proportional to the size of the grid set, a dense grid set may suffer from high computational cost. Moreover, a dense grid may conflict the restricted isometry property (RIP), and thus, it is not easy to find the balance between accuracy and efficiency. For 2-D MUSIC, finding the angles also require discretizing, and thus, it may also encounter computational issues when the grid set is large.

Recently, a gridless method which does not require angle discretization is proposed [22–26]. It employs the atomic norm minimization (ANM) concept into angle estimation and then reformulates the channel estimation problem as the semidefinite programming (SDP) which can be solved by CVX [27]. The ANM method does not suffer from the accuracy and efficiency issues caused by gridding. Theoretical analysis shows that the ANM method is an asymptotic maximum likelihood (ML) estimator [28] and its complexity is also immune to the size of the grid set. Although ANM shows excellent estimation performance in angle estimation [29] or channel estimation [24, 25], the main obstacle of the ANM estimator is the computational issue since solving SDP by CVX is time-consuming. Especially, in the full-dimensional MIMO case, ANM requires to solve an  $n^2$ -dimensional SDP, where  $n$  denotes the number of antennas on the transmitter or receiver which may be large [24]. Thus, it is urgent to derive a new algorithm for ANM methods to reduce computational complexity.

The ANM-based channel estimation method [24] only considers the phase shifters in hybrid architecture of the mmWave system. The hybrid architecture can achieve near-optimal performance compared to the fully digital transceivers [1]. However, the phase shifter-based network is not a simple circuit at the mmWave band [30]. Another type of architecture is to employ switch-based networks [31]. It is shown that the switch-based network is preferred in a range of operating conditions [30]. From the viewpoint of array

structures, antenna selection is relative to sparse arrays. Different selection strategies result in different sparse array architectures. From the coarray perspective [32], the antenna selection strategy having the longest uniform coarray part enjoys the best estimation performance. Several recently proposed sparse arrays such as coprime array [32], nested array [33, 34], and fractal array [35] have good coarray property and can be used in competitive antenna selection strategies for mmWave channel estimation. The coprime array has been incorporated into mmWave channel estimation in [36], and the nested array has been used in channel estimation and tracking in [37]. However, these papers only consider a single user with one antenna rather than multiuser or multiantenna.

In this paper, we consider the switch architecture for the analog processor in channel estimation of mmWave massive MIMO systems. We first exploit the antenna selection from the coarray point of view and then propose an atomic  $\ell_0$ -norm minimization problem. Compared to ANM, our method has much less dimensionality of the problem and hence is much more efficient than ANM. We also propose an efficient algorithm to solve this problem based on Wirtinger projection. Our method requires no angle discretization and hence is immune to the grid mismatch effect. We also carry out simulations to show the superiority of our method.

Notations:  $\mathbb{C}$  and  $\mathbb{Z}$  denote the sets of complex numbers and integers, respectively.  $\mathbf{A}^*$ ,  $\mathbf{A}^T$ , and  $\mathbf{A}^H$  denote the conjugate, transpose, and conjugate transpose of matrix  $\mathbf{A}$ , respectively.  $\text{vec}(\mathbf{A})$  denotes the vectorization operator that stacks matrix  $\mathbf{A}$  column by column.  $\mathbf{A} \odot \mathbf{B}$  and  $\mathbf{A} \otimes \mathbf{B}$  are the Khatri–Rao and Kronecker products of matrices  $\mathbf{A}$  and  $\mathbf{B}$ , respectively.  $\text{tr}(\bullet)$  and  $\text{rank}(\bullet)$  denote the trace and rank operators.  $\mathbf{I}_N$  denotes the identity matrix of size  $N \times N$ .  $\|\mathbf{A}\|_1$ ,  $\|\mathbf{A}\|_2$ , and  $\|\mathbf{A}\|_F$  denote the  $\ell_1$ -norm,  $\ell_2$ -norm, and Frobenius norm of  $\mathbf{A}$ , respectively.  $\mathbf{A} \geq 0$  means that matrix  $\mathbf{A}$  is positive semidefinite (PSD). For a vector  $\mathbf{x}$ ,  $\text{diag}(\mathbf{x})$  denotes a diagonal matrix with the diagonal elements being the elements of vector  $\mathbf{x}$  in turn.

The rest of this paper is organized as follows: Section 2 introduces the coarray concept and the system model. Section 3 provides our proposed method. Extensive simulations are provided in Section 4, and Section 5 concludes the whole paper.

## 2. Preliminary and System Model

**2.1. Coarray Concept.** In array signal processing, the aperture of an array is an important factor for angle estimation. A larger array aperture can bring in high estimation accuracy and super resolution. But, increasing the inter-element spacing is not a positive way to extend the aperture because a uniform linear array (ULA) with interelement spacing being greater than half-wavelength suffers from angle ambiguity. In this case, pseudo AoAs or AoDs will prevent us from correctly identifying the true positions. Fortunately, we can exploit the coarray concept to solve this problem. It is shown that we can construct a sparse linear array (SLA) with much larger aperture if its coarray has a long uniform part without holes [32]. Denote  $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_M\}$  as the antenna

indices where  $\Omega_1 < \Omega_2 < \dots < \Omega_M$  and each element is a positive integer. Then, the coarray is defined as

$$\mathcal{D} = \{m_1 - m_2 + 1 : m_1, m_2 \in \Omega, m_1 \geq m_2\}. \quad (1)$$

For instance, denote the array  $\Omega = \{1, 2, 5, 7\}$ ; then, its coarray  $\mathcal{D} = \{1, 2, 3, 4, 5, 6, 7\}$  which can be regarded as a 7-element ULA and has a large aperture without angle ambiguity. Thus, some special SLAs such as coprime array and nested array having a larger aperture can provide super resolution and satisfying performance.

**2.2. System Model.** Consider the mmWave massive MIMO communication system with a single user shown in Figure 1, where the transmitter is equipped with  $M_t$  RF chains and  $N_t > M_t$  antennas and the receiver is equipped with  $M_r$  RF chains and  $N_r > M_r$  antennas. The interelement spacing of each array is set to half-wavelength to avoid angle ambiguity. The analog architectures of both the transmitter and receiver are implemented by using switches. In particular, each switch is connected to a specific RF chain and can build connection between the RF chain and any antenna. The antennas selected by the switches are activated to transmit data, while other antennas remain idle. The selection strategy can be specified by a  $\{0, 1\}$  matrix denoted by  $\mathbf{F} \in \mathbb{C}^{N_t \times M_t}$  for the transmitter and  $\mathbf{W} \in \mathbb{C}^{N_r \times M_r}$  for the receiver. Let  $\Omega_F = \{\Omega_1, \Omega_2, \dots, \Omega_{M_t}\}$  and  $\Phi_W = \{\Phi_1, \Phi_2, \dots, \Phi_{M_r}\}$  denote the element indices of the transmit antennas and receiver antennas, respectively, where  $\Omega_1 < \Omega_2 < \dots < \Omega_{M_t} \leq N_t$ ,  $\Phi_1 < \Phi_2 < \dots < \Phi_{M_r} \leq N_r$ , and each element is a positive integer. In the following, we provide a simple example to better demonstrate the selection matrix. For simplicity, we only consider the transmitter. Let  $N_t = 7$  and  $M_t = 4$ ; if we select the antenna indexed by  $\Omega_F = \{1, 2, 5, 7\}$ , then the antenna array is an SLA and the selection matrix  $\mathbf{F}$  is as follows:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

From equation (2), it can be seen that the  $m$ -th column of matrix  $\mathbf{F}$  contains all zeros but a single one at the  $\Omega_m$ -th position. Note that if we let  $\Omega_F = \{1, 2, 3, 4\}$ , the transmitted antenna array is a short ULA having a shorter aperture than the SLA case. The discrete-time transmitted signal is therefore given by

$$\mathbf{x} = \mathbf{F}\mathbf{s}, \quad (3)$$

where  $\mathbf{s} \in \mathbb{C}^{M_t}$  denotes the signal after the digital processor. For simplicity, we consider a narrowband block-fading propagation channel which gives that

$$\mathbf{y}_s = \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{s} + \mathbf{W}^H \mathbf{n}, \quad (4)$$

where  $\mathbf{y}_s \in \mathbb{C}^{M_r}$  denotes the received signal at the RF chains,  $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$  denotes the channel matrix, and  $\mathbf{n} \in \mathbb{C}^{N_r}$  is the additive Gaussian noise with zero mean. During  $M_t$  successive time slots, the received signal at the receiver can be given as

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H} \mathbf{S} + \mathbf{W}^H \mathbf{N}, \quad (5)$$

where  $\mathbf{S}$  is the transmitted signal and  $\mathbf{N}$  is the noise matrix. For the training phase, we assume  $\mathbf{S} = \mathbf{I}_{M_t}$ , and therefore,

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H} \mathbf{F} + \mathbf{W}^H \mathbf{N}. \quad (6)$$

Our goal is to estimate  $\mathbf{H}$  given  $\mathbf{Y}$ .

**2.3. Channel Model.** In the mmWave massive MIMO system, the number of rays between the transmitter and receiver is limited. The channel  $\mathbf{H}$  can be expressed as

$$\mathbf{H} = \sum_{k=1}^K \alpha_k \mathbf{a}_r(\theta_k) \mathbf{a}_t^H(\phi_k), \quad (7)$$

where  $\alpha_k$ ,  $\theta_k$ , and  $\phi_k$  denote the complex gain, AoA, and AoD of the  $k$ -th ray, respectively, and  $\mathbf{a}_t(\phi_k)$  and  $\mathbf{a}_r(\theta_k)$  denote the steering vectors of the transmitter and receiver arrays with respect to the  $k$ -th ray, respectively, and can be given as

$$\begin{aligned} \mathbf{a}_t(\phi_k) &= \left[ 1, e^{j \cdot 1 \cdot ((2\pi d/\lambda) \sin(\phi_k))}, \dots, e^{j \cdot (N_t-1) \cdot ((2\pi d/\lambda) \sin(\phi_k))} \right]^T, \\ \mathbf{a}_r(\theta_k) &= \left[ 1, e^{j \cdot 1 \cdot ((2\pi d/\lambda) \sin(\theta_k))}, \dots, e^{j \cdot (N_r-1) \cdot ((2\pi d/\lambda) \sin(\theta_k))} \right]^T, \end{aligned} \quad (8)$$

where  $\lambda$  and  $d$  denote the wavelength and the spacing between adjacent antennas, respectively. Equation (7) can be compactly rewritten as

$$\mathbf{H} = \mathbf{A}_r \mathbf{\Sigma} \mathbf{A}_t^H, \quad (9)$$

where  $\mathbf{A}_t = [\mathbf{a}_t(\phi_1), \mathbf{a}_t(\phi_2), \dots, \mathbf{a}_t(\phi_K)]$ ,  $\mathbf{A}_r = [\mathbf{a}_r(\theta_1), \mathbf{a}_r(\theta_2), \dots, \mathbf{a}_r(\theta_K)]$ , and  $\mathbf{\Sigma} = \text{diag}([\alpha_1, \dots, \alpha_K])$ . Substituting equation (9) into model (6), we have

$$\mathbf{Y} = \mathbf{W}^H \mathbf{A}_r \mathbf{\Sigma} \mathbf{A}_t^H \mathbf{F} + \mathbf{W}^H \mathbf{N}. \quad (10)$$

Vectorizing  $\mathbf{Y}$  results in,

$$\begin{aligned} \mathbf{y} &= \text{vec}(\mathbf{Y}) \\ &= (\mathbf{F}^T \otimes \mathbf{W}^H) \mathbf{h} + \text{vec}(\mathbf{W}^H \mathbf{N}) \\ &= (\mathbf{F}^T \otimes \mathbf{W}^H) (\mathbf{A}_t^* \circ \mathbf{A}_r) \mathbf{z} + \text{vec}(\mathbf{W}^H \mathbf{N}), \end{aligned} \quad (11)$$

where  $\mathbf{h} = \text{vec}(\mathbf{H})$  and  $\mathbf{z} = [\alpha_1, \dots, \alpha_K]^T$ .

### 3. The Proposed Channel Estimation Method

**3.1. The Proposed Method.** Different from the ANM-based method in [23], we directly operate on model (10) rather than its vectorized version. First, we set up the following atom set:

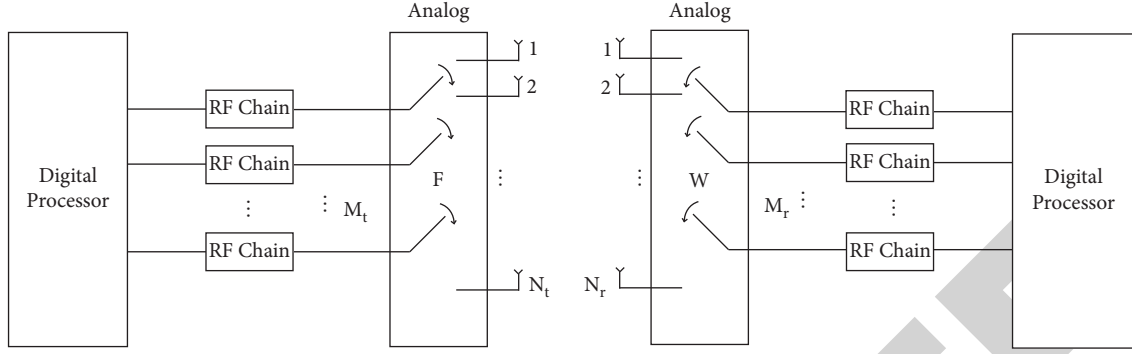


FIGURE 1: Block diagram of the mmWave massive MIMO communication system with an analog processor implemented using switches.

$$\tilde{\mathcal{A}} = \{ \mathbf{a}_r(\theta) \mathbf{a}_t^H(\phi) : \theta, \phi \in (-90^\circ, 90^\circ) \}, \quad (12)$$

based on which we can formulate the atomic  $\ell_0$ -norm of the channel matrix  $\mathbf{H}$  as

$$\|\mathbf{H}\|_{\tilde{\mathcal{A}},0} = \inf \left\{ K : \mathbf{H} = \sum_{k=1}^K \alpha_k \tilde{\mathbf{A}}, \tilde{\mathbf{A}} \in \tilde{\mathcal{A}}, \alpha_k \in \mathbb{C} \right\}. \quad (13)$$

We then propose the following optimization problem:

$$\min_{\mathbf{H}} \|\mathbf{H}\|_{\tilde{\mathcal{A}},0} \text{ s.t. } \|\mathbf{Y} - \mathbf{W}^H \mathbf{H} \mathbf{F}\|_F \leq \beta. \quad (14)$$

However, the above problem is a semi-infinite programming (SIP) which cannot be efficiently solved in polynomial time. To solve this problem, we have the following theorem.

**Theorem 1.** Assume that  $K < \min(N_t, N_r)$ , then  $\|\mathbf{H}\|_{\tilde{\mathcal{A}},0}$  equals the optimal value of the following rank minimization problem:

$$\begin{aligned} & \min_{\mathbf{u}, \mathbf{v}, \mathbf{W}} \text{rank}[\mathbf{W}], \\ & \text{s.t. } \mathbf{W} = \begin{bmatrix} \mathbf{T}(\mathbf{v}) & \mathbf{H}^H \\ \mathbf{H} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0, \end{aligned} \quad (15)$$

where  $\mathbf{T}(\mathbf{u})$  and  $\mathbf{T}(\mathbf{v})$  are Toeplitz matrices.

*Proof.* First, for arbitrary decomposition of  $\mathbf{H}$  as  $\mathbf{H} = \sum_{k=1}^K \alpha_k \mathbf{a}_r(\theta_k) \mathbf{a}_t^H(\phi_k)$ , we can construct matrices  $\mathbf{T}(\mathbf{u}) = \sum_{k=1}^K |\alpha_k|^2 \mathbf{a}_r(\theta_k) \mathbf{a}_r^H(\theta_k)$  and  $\mathbf{T}(\mathbf{v}) = \sum_{k=1}^K \alpha_k \mathbf{a}_t(\phi_k) \mathbf{a}_t^H(\phi_k)$  and then we have

$$\begin{aligned} \tilde{\mathbf{W}} &= \begin{bmatrix} \mathbf{T}(\mathbf{v}) & \mathbf{H}^H \\ \mathbf{H} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \\ &= \sum_{k=1}^K \begin{bmatrix} \mathbf{a}_t(\phi_k) \mathbf{a}_t^H(\phi_k) & \alpha_k \mathbf{a}_t(\phi_k) \mathbf{a}_r^H(\theta_k) \\ \alpha_k \mathbf{a}_r(\theta_k) \mathbf{a}_t^H(\phi_k) & |\alpha_k|^2 \mathbf{a}_r(\theta_k) \mathbf{a}_r^H(\theta_k) \end{bmatrix} \\ &= \sum_{k=1}^K \begin{bmatrix} \mathbf{a}_t(\phi_k) \\ \alpha_k \mathbf{a}_r(\theta_k) \end{bmatrix} \begin{bmatrix} \mathbf{a}_t(\phi_k) \\ \alpha_k \mathbf{a}_r(\theta_k) \end{bmatrix}^H \geq 0. \end{aligned} \quad (16)$$

It follows that the optimal solution of (15)  $K^* \leq \text{rank}(\tilde{\mathbf{W}}) = K = \|\mathbf{H}\|_{\tilde{\mathcal{A}},0}$ . On the other hand, if we find

the optimal solution of (15) as  $\{\mathbf{H}^*, \mathbf{u}^*, \mathbf{v}^*\}$ , then we have  $\text{rank}(\mathbf{W}^*) = K^* \leq K$ . It follows from the Vandermonde decomposition [38] that  $\mathbf{T}(\mathbf{u}^*) = \sum_{k=1}^{K^*} |\alpha_k^*|^2 \mathbf{a}_r(\theta_k^*) \mathbf{a}_r^H(\theta_k^*)$  and  $\mathbf{T}(\mathbf{v}^*) = \sum_{k=1}^{K^*} \alpha_k^* \mathbf{a}_t(\phi_k^*) \mathbf{a}_t^H(\phi_k^*)$ . Then, since  $\mathbf{H}^*$  lies in the column space of  $\mathbf{T}(\mathbf{u}^*)$  and row space of  $\mathbf{T}(\mathbf{v}^*)$ , there exists  $\tilde{\alpha}^*$  such that  $\mathbf{H}^* = \sum_{k=1}^{K^*} \tilde{\alpha}_k^* \mathbf{a}_r(\theta_k^*) \mathbf{a}_t^H(\phi_k^*)$ . Hence,  $\|\mathbf{H}^*\|_{\tilde{\mathcal{A}},0} \leq K^*$ . Therefore, it can be concluded that  $\|\mathbf{H}\|_{\tilde{\mathcal{A}},0} = K^*$ .  $\square$

According to Theorem 1, model (14) can be rewritten as

$$\begin{aligned} & \min_{\mathbf{H}, \mathbf{u}, \mathbf{v}, \mathbf{W}} \text{rank}[\mathbf{W}], \\ & \text{s.t. } \mathbf{W} = \begin{bmatrix} \mathbf{T}(\mathbf{v}) & \mathbf{H}^H \\ \mathbf{H} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0, \\ & \|\mathbf{Y} - \mathbf{W}^H \mathbf{H} \mathbf{F}\|_F \leq \beta. \end{aligned} \quad (17)$$

It can be seen that, compared to the ANM model [23], the proposed model in (17) has much smaller problem dimensionality.

However, directly solving this model is difficult due to the nonconvex rank operator. One possible approach is to relax the rank operator to the trace operator. The relaxed problem is convex and can be solved by CVX. Nevertheless, CVX is also an inefficient solver. In the following, we propose an efficient method to solve (17) based on Wirtinger projection.

First, we define two matrices sets as

$$\begin{aligned} \mathcal{M} &= \{ \mathbf{M} : \text{rank}(\mathbf{M}) \leq K \}, \\ \mathcal{N} &= \left\{ \mathbf{N} : \mathbf{N} = \begin{bmatrix} \mathbf{V} & \mathbf{H}^H \\ \mathbf{H} & \mathbf{U} \end{bmatrix} \geq 0, \|\mathbf{Y} - \mathbf{W}^H \mathbf{H} \mathbf{F}\|_F \leq \beta \right\}, \end{aligned} \quad (18)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  have Toeplitz structure. Solving model (17) is equivalent to finding a matrix that both in  $\mathcal{M}$  and  $\mathcal{N}$  or solving the following problem:

$$\min_{\mathbf{M} \in \mathcal{M}, \mathbf{N} \in \mathcal{N}} \|\mathbf{M} - \mathbf{N}\|_F^2. \quad (19)$$

One effective way is to alternatively update  $\mathbf{M}$  and  $\mathbf{N}$  until convergence. Based on the Wirtinger strategy [39], we formulate the following update rule to solve problem (19):

$$\begin{aligned}\mathbf{M}^{(t+1)} &= \mathcal{P}_{\mathcal{M}}(\mathbf{M}^{(t)} - \delta_1(\mathbf{M}^{(t)} - \mathbf{N}^{(t)})), \\ \mathbf{N}^{(t+1)} &= \mathcal{P}_{\mathcal{N}}(\mathbf{N}^{(t)} - \delta_2(\mathbf{N}^{(t)} - \mathbf{M}^{(t+1)})),\end{aligned}\quad (20)$$

where the superscript  $(t)$  denotes the  $t$ -th iteration,  $\mathcal{P}_{\mathcal{M}}(\bullet)$  and  $\mathcal{P}_{\mathcal{N}}(\bullet)$  denote the projection procedures onto matrices sets  $\mathcal{M}$  and  $\mathcal{N}$ , respectively, and  $\delta_1$  and  $\delta_2$  are two user-defined parameters. The next target is to find the two projection operators  $\mathcal{P}_{\mathcal{M}}(\bullet)$  and  $\mathcal{P}_{\mathcal{N}}(\bullet)$ .

Projection  $\mathcal{P}_{\mathcal{M}}(\mathbf{X})$  is to find the best rank- $K$  approximation that can be formulated as follows. We first apply singular value decomposition onto  $\mathbf{X}$  as  $\mathbf{X} = \mathbf{U}_X \Sigma_X \mathbf{V}_X^H$ ; then, projection  $\mathcal{P}_{\mathcal{M}}(\mathbf{X})$  is

$$M_{\mathcal{M}}(\mathbf{X}) = \mathbf{U}_K \Sigma_K \mathbf{V}_K^H, \quad (21)$$

where  $\mathbf{U}_K$  and  $\mathbf{V}_K$  are the first  $K$  columns of  $\mathbf{U}_X$  and  $\mathbf{V}_X$ , respectively, and  $\Sigma_K$  is the corresponding singular value matrix.

Denoting  $\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{\mathbf{V}} & \text{amp}; & \tilde{\mathbf{H}}^H \\ \tilde{\mathbf{H}} & \text{amp}; & \tilde{\mathbf{U}} \end{bmatrix}$ , projection  $M_{\mathcal{N}}(\tilde{\mathbf{X}})$  contains four sequential subprojections:  $\mathcal{P}_{\mathcal{N}_1} \rightarrow \mathcal{P}_{\mathcal{N}_2} \rightarrow \mathcal{P}_{\mathcal{N}_3} \rightarrow \mathcal{P}_{\mathcal{N}_4}$  where  $\mathcal{P}_{\mathcal{N}_1}$  and  $\mathcal{P}_{\mathcal{N}_2}$  are to project  $\tilde{\mathbf{U}}$  and  $\tilde{\mathbf{V}}$  onto the set of Toeplitz matrix, respectively,  $\mathcal{P}_{\mathcal{N}_3}$  is to project  $\tilde{\mathbf{H}}$  onto the ball defined by  $\{\mathbf{H}: \|\mathbf{Y} - \mathbf{W}^H \mathbf{H} \mathbf{F}\|_F \leq \beta\}$ , and  $\mathcal{P}_{\mathcal{N}_4}$  is to project  $\tilde{\mathbf{X}}$  onto a positive semidefinite matrix set. In particular,  $M_{\mathcal{N}_1}(\mathbf{U})$  is defined as

$$\mathcal{P}_{\mathcal{N}_1}(\mathbf{U}) = \mathbf{T}(\mathbf{u}), \quad (22)$$

where the  $i$ -th element of  $\mathbf{u}$  is  $\mathbf{u}_i = (1/(N_r - i + 1)) \sum_i \mathbf{U}_{m,n}$  where  $i = n - m + 1$ ,  $n \geq m$ , and  $m, n = 1, \dots, N_r$ . Similarly,  $\mathcal{P}_{\mathcal{N}_2}(\mathbf{V})$  is

$$\mathcal{P}_{\mathcal{N}_2}(\mathbf{V}) = \mathbf{T}(\mathbf{v}), \quad (23)$$

where the  $i$ -th element of  $\mathbf{v}$  is  $\mathbf{v}_i = (1/(N_t - i + 1)) \sum_i \mathbf{V}_{m,n}$  where  $i = n - m + 1$ ,  $n \geq m$ , and  $m, n = 1, \dots, N_t$ . Projection  $\mathcal{P}_{\mathcal{N}_3}$  is to project  $\mathbf{W}^H \tilde{\mathbf{H}} \mathbf{F}$  onto a ball with center  $\mathbf{Y}$  and radius  $\beta$ . To realize  $\mathcal{P}_{\mathcal{N}_3}$ , we first define

$$\tilde{\mathbf{H}}^{\text{proj}} = \mathbf{W} \left( \mathbf{Y} + \frac{\mathbf{W}^H \tilde{\mathbf{H}} \mathbf{F} - \mathbf{Y}}{\|\mathbf{W}^H \tilde{\mathbf{H}} \mathbf{F} - \mathbf{Y}\|_F} \beta \right) \mathbf{F}^H. \quad (24)$$

Then,  $M_{\mathcal{N}_3}$  can be given as

$$[\mathcal{P}_{\mathcal{N}_3}]_{m,n} = \begin{cases} \begin{bmatrix} \tilde{\mathbf{H}}^{\text{proj}} \end{bmatrix}_{m,n}, & \text{if } m \in \Phi_W, n \in \Omega_F, \\ \tilde{\mathbf{H}}_{m,n}, & \text{otherwise.} \end{cases} \quad (25)$$

For  $\mathcal{P}_{\mathcal{N}_4}$ , we first apply eigen-decomposition to  $\tilde{\mathbf{X}}$  and have  $\tilde{\mathbf{X}} = \mathbf{U}_{\tilde{\mathbf{X}}}^- \text{diag}(\sigma_{\tilde{\mathbf{X}}}^-) \mathbf{U}_{\tilde{\mathbf{X}}}^H$ . Then, projection  $\mathcal{P}_{\mathcal{N}_4}$  can be given as

$$\mathcal{P}_{\mathcal{N}_4}(\tilde{\mathbf{X}}) = \mathbf{U}_{\tilde{\mathbf{X}}}^+ \text{diag}(\sigma_{\tilde{\mathbf{X}}}^+) (\mathbf{U}_{\tilde{\mathbf{X}}}^+)^H, \quad (26)$$

where  $\sigma_{\tilde{\mathbf{X}}}^{\pm}$  denotes the positive part of  $\sigma_{\tilde{\mathbf{X}}}^-$  and  $\mathbf{U}_{\tilde{\mathbf{X}}}^{\pm}$  denotes the corresponding eigen-vectors.

The proposed method converges when  $(\|\mathbf{N}^{(t+1)} - \mathbf{N}^{(t)}\|_F / \|\mathbf{N}^{(t)}\|_F) \leq \varepsilon$ , where  $\varepsilon$  is a predefined threshold.

**3.2. The Case of  $N_t > \Omega_{M_t}$ .** It should be noted that, given the number of RF chains  $M_t$  at the transmitter (we take the transmitter as an example.), the array aperture is limited by its maximum tag  $\Omega_{M_t}$ , which may not equal  $N_t$ . If  $N_t > \Omega_{M_t}$ , finding the accurate channel matrix  $\mathbf{H}$  from model (17) is difficult since there may exist too many variables to be determined in (17). Thus, instead of directly obtaining  $\mathbf{H}$ , we alternatively first retrieve the angle information based on which we can find the channel gain. And then, the channel matrix  $\mathbf{H}$  can be well estimated. Specifically, we first let  $\Omega_1 = 1$  and replace  $\mathbf{F} \in \mathbb{Z}^{N_t \times M_t}$  by  $\bar{\mathbf{F}} \in \mathbb{Z}^{\Omega_{M_t} \times M_t}$ . Similarly, we obtain  $\bar{\mathbf{W}} \in \mathbb{Z}^{\Phi_{M_r} \times M_r}$ . Then, received signal model (10) can be rewritten as

$$\mathbf{Y} = \bar{\mathbf{W}}^H \bar{\mathbf{A}}_r \Sigma \bar{\mathbf{A}}_t^H \bar{\mathbf{F}} + \bar{\mathbf{W}}^H \bar{\mathbf{N}}, \quad (27)$$

where  $\bar{\mathbf{A}}_t = [\bar{\mathbf{a}}_t(\phi_1), \bar{\mathbf{a}}_t(\phi_2), \dots, \bar{\mathbf{a}}_t(\phi_K)]$  and  $\bar{\mathbf{A}}_r = [\bar{\mathbf{a}}_r(\theta_1), \bar{\mathbf{a}}_r(\theta_2), \dots, \bar{\mathbf{a}}_r(\theta_K)]$  with

$$\begin{aligned}\bar{\mathbf{a}}_t(\phi_k) &= \begin{bmatrix} 1, e^{j \cdot \frac{2\pi d}{\lambda} \sin(\phi_k)}, \dots, e^{j(\Omega_{M_t}-1) \cdot \frac{2\pi d}{\lambda} \sin(\phi_k)} \end{bmatrix}^T, \\ \bar{\mathbf{a}}_r(\theta_k) &= \begin{bmatrix} 1, e^{j \cdot \frac{2\pi d}{\lambda} \sin(\theta_k)}, \dots, e^{j(\Phi_{M_r}-1) \cdot \frac{2\pi d}{\lambda} \sin(\theta_k)} \end{bmatrix}^T.\end{aligned}\quad (28)$$

It is easy to see that  $\bar{\mathbf{a}}_t(\phi_k)$  and  $\bar{\mathbf{a}}_r(\theta_k)$  are subvectors of  $\mathbf{a}_t(\phi_k)$  and  $\mathbf{a}_r(\theta_k)$ , respectively. Based on truncated model (27), we propose the following truncated problem:

$$\begin{aligned}\min_{\bar{\mathbf{H}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}, \bar{\mathbf{W}}} & \text{rank}[\bar{\mathbf{W}}], \\ \text{s.t.} & \bar{\mathbf{W}} = \begin{bmatrix} \mathbf{T}(\bar{\mathbf{v}}) & \bar{\mathbf{H}}^H \\ \bar{\mathbf{H}} & \mathbf{T}(\bar{\mathbf{u}}) \end{bmatrix} \geq 0, \\ & \|\mathbf{Y} - \bar{\mathbf{W}}^H \bar{\mathbf{H}} \mathbf{F}\|_F \leq \beta.\end{aligned}\quad (29)$$

The optimal solution  $\bar{\mathbf{H}} \in \mathbb{C}^{\Phi_{M_r} \times \Omega_{M_t}}$  is only a submatrix of the full channel matrix  $\mathbf{H}$ . To estimate  $\mathbf{H}$ , we should first find the AoAs and AoDs from (29). From the proof of Theorem 1, we can see that  $\mathbf{T}(\bar{\mathbf{u}})$  and  $\mathbf{T}(\bar{\mathbf{v}})$  contain the AoAs and AoDs information, respectively, and the two Toeplitz matrices can be regarded as the noiseless covariance matrices of a ULA. Thus, the traditional ESPRIT method can be applied to find the angle estimates. Alternatively, we can also find the angle estimates from the Vandermonde decomposition theorem [40]. The channel matrix can then be constructed according to (7) after finding the channel gain by the LS method.

## 4. Simulation Results

In this section, we evaluate the channel estimation performance of our proposed method on the switch-based mmWave massive MIMO system. We also consider other

methods including MUSIC [41], OMP [18],  $L1$  minimization [42], ANM [22], and decoupled ANM (DANM) [43] for comparison. It should be pointed out that DANM is applied with the alternating direction method of multipliers (ADMM) [44] which is a fast solver. We assume that  $N_t = 16$ ,  $N_r = 32$ ,  $M_t = 6$ , and  $M_r = 10$  and the antennas are placed with half-wavelength spacing. The channel estimation performance is evaluated by

$$\text{NMSE} = \mathbb{E} \left[ \frac{\|\mathbf{H} - \mathbf{H}^\circ\|_F^2}{\|\mathbf{H}\|_F^2} \right], \quad (30)$$

based on 400 independent trials. The number of paths is set to  $K = 3$ . The AoAs and AoDs are randomly generated from  $[-90^\circ, 90^\circ]$  for each path. The SNR is defined as  $\text{SNR} = (P_t/\sigma^2)$ , where  $P_t$  and  $\sigma^2 = 1$  denote the transmitted power and noise power, respectively. For our method, we assume  $\delta_1 = \delta_2 = 0.8$  and  $\beta = \sqrt{M_t M_r + 2\sqrt{M_t M_r} \sigma}$  to the upper bound of noise energy.

**4.1. Convergence Performance.** We first evaluate the convergence of our method. Denote  $\Delta\mathbf{U}^{(t+1)} = (\|\mathbf{H}^{(t+1)} - \mathbf{H}^{(t)}\|_F^2 / \|\mathbf{H}^{(t)}\|_F^2)$  as the variation of the estimated channel matrices between the  $(t+1)$ -th and  $t$ -th iterations. The threshold  $\varepsilon$  is set to  $10^{-3}$ . We consider different SNR scenarios and show the relationship between  $\Delta\mathbf{U}$  and the iteration in Figure 2, from which we can observe that for different SNRs,  $\Delta\mathbf{U}$  decreases rapidly and our method converges after about 300 iterations.

**4.2. Different Antenna Selection Strategies.** In this section, we compare the performance of different antenna selection strategies. We select two representative SLA structures, nested and coprime arrays, and random selection strategy is also taken into consideration. In particular, we set  $\Omega_F^{\text{nested}} = \{1, 2, 3, 4, 8, 12\}$  and  $\Phi_W^{\text{nested}} = \{1, 2, 3, 4, 5, 10, 15, 20, 25, 30\}$  for the nested structure and  $\Omega_F^{\text{coprime}} = \{1, 3, 4, 5, 7, 10\}$  and  $\Phi_W^{\text{coprime}} = \{1, 4, 6, 7, 10, 11, 13, 16, 21, 26\}$  for the coprime structure. For random selection strategy, we always first select the first antenna and then randomly select  $M_t - 1$  antennas from  $\{2, \dots, 16\}$  for the transmitter and  $M_r - 1$  antennas from  $\{2, \dots, 32\}$  for the receiver. The NMSEs of our proposed method based on these three strategies are shown in Figure 3 with the SNR varying from  $-10$  dB to  $8$  dB. It can be seen that the nested array enjoys the best estimation performance. The coprime array shows inferior performance compared to nested array since it has shorter successive uniform part in its coarray. The random strategy shows the worst performance.

**4.3. Performance Comparison.** Next, we evaluate the channel estimation performance of our proposed method with other representative methods. For MUSIC, OMP, and  $L1$  which require discretizing the angle space, we consider two grid resolutions  $2^\circ$  and  $3^\circ$ . The NMSEs of these methods are shown in Figure 4(a). We can see that ANM enjoys the best estimation performance in the compared SNR region. The

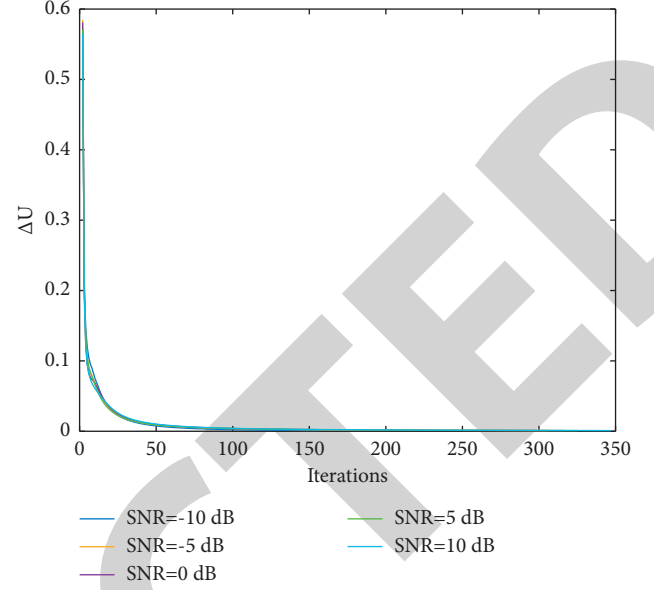


FIGURE 2: Convergence of the proposed method.

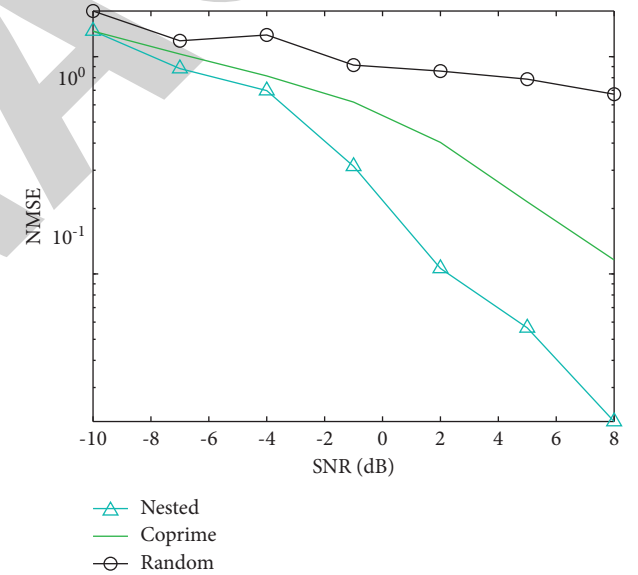


FIGURE 3: NMSE comparison of three different antenna selection strategies.

proposed method is superior to other methods except ANM in most cases. For the three grid-based methods with grid resolution  $3^\circ$ , they show unsatisfying accuracy in the low SNR region while suffer from grid effect when SNR becomes large. Thus, we can see that when SNR is larger than  $4$  dB, the gaps between these methods and our method become large. For grid resolution  $2^\circ$ , these methods show better performance than for the case of grid resolution  $3^\circ$ . This is because reducing the grid interval can relieve the grid mismatch effect, yielding higher estimation accuracy. We also show the running time of these methods in Figure 4(b). Since ANM requires to solve a large-dimensional SDP, it has the largest running time as compared to other methods. The  $L1$  method

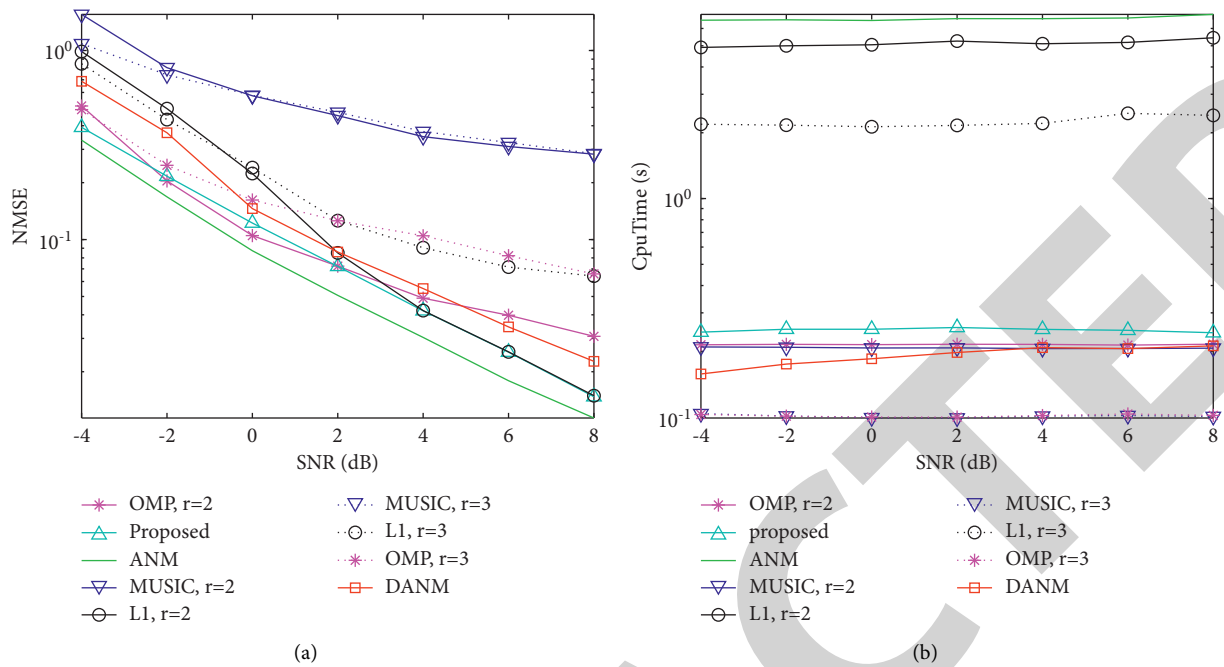


FIGURE 4: Channel estimation performance comparison with respect to the SNR. Two grid resolutions  $2^\circ$  and  $3^\circ$  are considered. (a) NMSE comparison. (b) CPU time comparison.

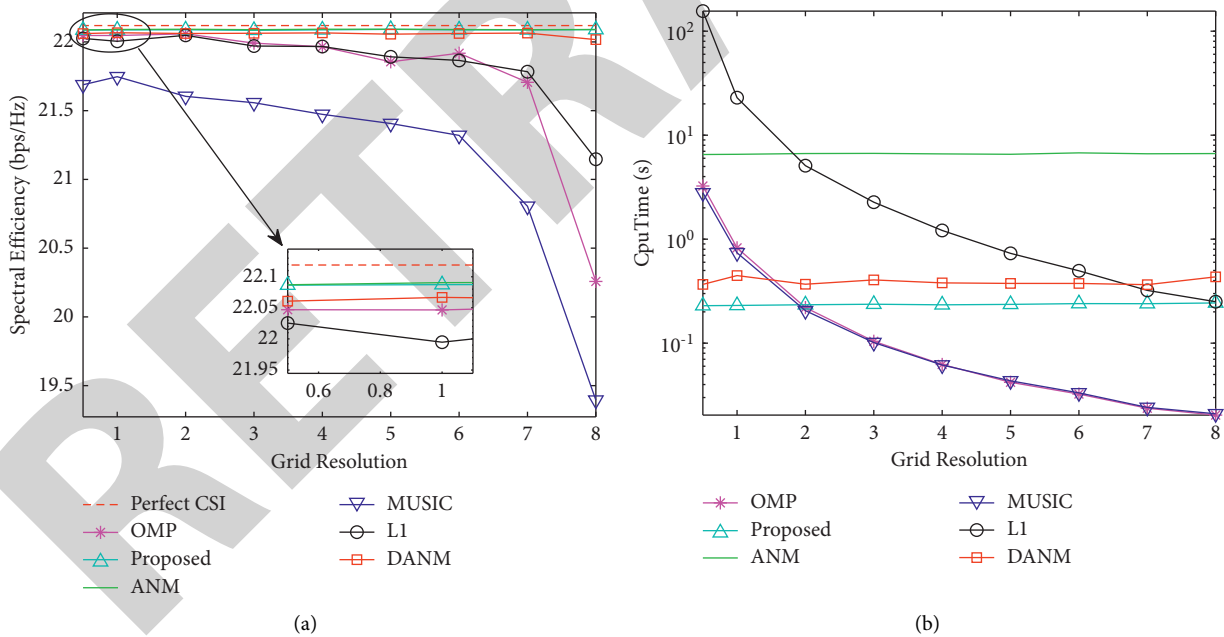


FIGURE 5: Spectral efficiency comparison with respect to the grid resolution. SNR = 5 dB. (a) Spectral efficiency comparison. (b) CPU time comparison.

requires solving a BPDN problem by CVX, and hence, it also suffers from high computational complexity, with respect to the two grid resolutions. DANM, MUSIC, and OMP are much faster than the ANM and L1 method, especially in the case of grid resolution  $2^\circ$ . Our method shows comparable computational efficiency to OMP and MUSIC with grid

resolution  $3^\circ$  but has better estimation performance as shown in Figure 4(a).

We also evaluate the spectral efficiency of these methods and show the simulation results in Figure 5 with different grid resolutions ranging from  $0.5^\circ$  to  $8^\circ$ . The spectral efficiency with perfect CSI is also considered as the upper



bound. From Figure 5(a), it can be seen that since ANM, DANM, and our method are immune to the angle discretization, they are not affected by the grid resolution and can coincide with the perfect CSI. For OMP, MUSIC, and  $L1$  methods, their performance deteriorates as the grid becomes sparser. Although the spectral efficiency of the  $L1$  method and OMP can approach our method, DANM and ANM in dense grid cases, from Figure 5(b), it can be seen that their computational times increase exponentially as the grid resolution decreases and are several times slower than our proposed method if the grid resolution is less than  $1^\circ$ .

## 5. Conclusion

In this paper, we proposed an atomic  $\ell_0$ -norm-based channel estimation method for switch-based mmWave massive MIMO communication systems. The proposed method exploits the coarray property of sparse arrays to select antennas and then formulates an atomic  $\ell_0$ -norm minimization problem which is efficiently solved based on Wirtinger projection. The proposed method is shown to have higher computational efficiency than ANM at comparable estimation performance. Compared to grid-based methods such as OMP, MUSIC, and  $L1$ , our method does not require angle discretization and hence is immune to the grid mismatch effect, leading to higher estimation accuracy.

## Data Availability

All data, models, or codes used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Science and Technology Project of SGCC (State Grid Corporation of China): "Development and Application of Power 5G Terminal based on Modular Design" (SGZJXT00JSJS2000454).

## References

- [1] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.
- [2] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity mimo-fbmc sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.
- [3] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse Bayesian learning based channel estimation in fbmc/oqam industrial iot networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.
- [4] J. Zhang, Y. Huang, Q. Shi, J. Wang, and L. Yang, "Codebook design for beam alignment in millimeter wave communication systems," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4980–4995, 2017.
- [5] Z. Xiao, T. He, P. Xia, and X. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3380–3392, 2016.
- [6] M. Xiao, S. Mumtaz, Y. Huang et al., "Millimeter wave communications for future mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.
- [7] M. K. Samimi and T. S. Rappaport, "3-d millimeter-wave statistical channel model for 5G wireless system design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2207–2225, 2016.
- [8] T. A. Thomas, H. C. Nguyen, G. R. MacCartney, and T. S. Rappaport, "3d mmwave channel model proposal," in *Proceedings of the 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, pp. 1–6, Vancouver, Canada, September 2014.
- [9] X. Wang, M. Huang, and L. Wan, "Joint 2d-dod and 2d-doa estimation for coprime emvscmimo radar," *Circuits, Systems, and Signal Processing*, vol. 40, no. 6, pp. 1–17, 2021.
- [10] F. Wen, J. Shi, and Z. Zhang, "Closed-form estimation algorithm for emvs-mimo radar with arbitrary sensor geometry," *Signal Processing*, vol. 186, no. 8, Article ID 108117, 2021.
- [11] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.
- [12] A. Alkhateeb, G. Leus, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: how many measurements are needed?" in *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2909–2913, South Brisbane, Australia, April 2015.
- [13] S. Sun and T. S. Rappaport, "Millimeter wave mimo channel estimation based on adaptive compressed sensing," in *Proceedings of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 47–53, Paris, France, May 2017.
- [14] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-d beamspace music method," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5384–5394, 2017.
- [15] C. Tsai, Y. Liu, and A. Wu, "Efficient compressive channel estimation for millimeter-wave large-scale antenna systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 9, pp. 2414–2428, 2018.
- [16] F. Dong, W. Wang, Z. Huang, and P. Huang, "High-resolution angle-of-arrival and channel estimation for mmwave massive mimo systems with lens antenna array," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 12963–12973, 2020.
- [17] L. Wan, K. Liu, Y.-C. Liang, and T. Zhu, "Doa and polarization estimation for non-circular signals in 3-d millimeter wave polarized massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3152–3167, 2021.
- [18] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [19] J. Shi, F. Wen, and T. Liu, "Nested mimo radar: coarrays, tensor modeling, and angle estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 1, pp. 573–585, 2021.

- [20] X. Wu, W.-P. Zhu, and J. Yan, "Direction of arrival estimation for off-grid signals based on sparse bayesian learning," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 2004–2016, 2016.
- [21] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2013.
- [22] G. Tang, B. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.
- [23] Z. Yang, L. Xie, and P. Stoica, "Vandermonde decomposition of multilevel Toeplitz matrices with application to multidimensional super-resolution," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3685–3701, 2016.
- [24] Y. Tsai, L. Zheng, and X. Wang, "Millimeter-wave beamformed full-dimensional mimo channel estimation based on atomic norm minimization," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6150–6163, 2018.
- [25] H. Chu, L. Zheng, and X. Wang, "Super-resolution mmwave channel estimation for generalized spatial modulation systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 6, pp. 1336–1347, 2019.
- [26] Z. Tian, Z. Zhang, and Y. Wang, "Low-complexity optimization for two-dimensional direction-of-arrival estimation via decoupled atomic norm minimization," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3071–3075, New Orleans, LA, USA, March 2017.
- [27] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," *Global Optimization*, pp. 155–210, 2008.
- [28] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing*, vol. 7, pp. 509–581, Elsevier, Amsterdam, Netherlands, 2018.
- [29] X. Wu, "Localization of far-field and near-field signals with mixed sparse approach: a generalized symmetric arrays perspective," *Signal Processing*, vol. 175, Article ID 107665, 2020.
- [30] R. Mèndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid mimo architectures for millimeter wave communications: phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.
- [31] A. Alkhateeb, Y. Nam, J. Zhang, and R. W. Heath, "Massive mimo combining with switches," *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 232–235, 2016.
- [32] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.
- [33] P. Pal and P. P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.
- [34] J. He, L. Li, and T. Shu, "Sparse nested arrays with spatially spread orthogonal dipoles: high accuracy passive direction finding with less mutual coupling," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, p. 1, 2021.
- [35] R. Cohen and Y. C. Eldar, "Sparse fractal array design with increased degrees of freedom," in *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4195–4199, Brighton, UK, May 2019.
- [36] S. Haghghatshoar and G. Caire, "Massive mimo channel subspace estimation from low-dimensional projections," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 303–318, 2017.
- [37] X. Wu, W.-P. Zhu, and J. Yan, "Channel estimation and tracking with nested sampling for fast-moving users in millimeter-wave communication," *Digital Signal Processing*, vol. 94, pp. 29–37, 2019.
- [38] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3139–3153, 2015.
- [39] M. Cho, J. Cai, S. Liu, Y. C. Eldar, and W. Xu, "Fast alternating projected gradient descent algorithms for recovering spectrally sparse signals," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4638–4642, Shanghai, China, March 2016.
- [40] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4959–4973, 2014.
- [41] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [42] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.
- [43] J. F. Cai, W. Xu, and Y. Yang, "Large scale 2D spectral compressed sensing in continuous domain," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5905–5909, New Orleans, LA, USA, March 2017.
- [44] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.