RECEIVER DESIGN FOR MASSIVE MIMO WIRELESS SYSTEMS

by

Ping Xu
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Master of Science
Electrical and Computer Engineering

Committee:

____________________	Dr. Zhi Tian, Thesis Director
____________________	Dr. Yariv Ephraim, Committee Member
____________________	Dr. Brian L. Mark, Committee Member
____________________	Dr. Monson H. Hayes, Department Chair
____________________	Dr. Kenneth S. Ball, Dean, Volgenau School of Engineering

Date: ____________________

Spring Semester 2018
George Mason University
Fairfax, VA

Receiver Design for Massive MIMO Wireless Systems

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

By

Ping Xu
Bachelor of Science
Northwestern Polytechnical University, 2015

Director: Dr. Zhi Tian, Professor
Department of Electrical and Computer Engineering

Spring Semester 2018
George Mason University
Fairfax, VA

Dedication

I dedicate this thesis to my parents, my sister and my brother.

Acknowledgments

My first and earnest gratitude goes to my supervisor, Prof. Zhi Tian. Her ideas and motivation are invaluable throughout my research work. She deserves many thanks not only for helping me complete this thesis but also for providing me continuous encouragement and invaluable advice during my graduate study at George Mason University. I feel so fortunate to be under her supervision. I would like to sincerely thank Prof. Yariv Ephraim and Prof. Brian L. Mark for sparing their valuable time to read this manuscript and being my committee members. I would like also to express my thanks to my labmates and friends. Special thanks go to Dr. Yue Wang and Dr. Zhe Zhang for their inspiring discussions and help. Lastly, but most importantly, I thank my parents for their tremendous support and endless self-sacrifices.

Table of Contents

List of Tables

List of Figures

Abstract

RECEIVER DESIGN FOR MASSIVE MIMO WIRELESS SYSTEMS

Ping Xu

George Mason University, 2018

Thesis Director: Dr. Zhi Tian

Massive multiple-input multiple-output (MIMO) systems that employ a large number of antennas at both receivers and transmitters have been widely considered for adoption in next generation (5G) wireless networks. The deployment of massive MIMO promises to enhance the received signal power for communications over millimeter-wave (mmWave) spectrum, which in turn increases the throughput and system efficiency. Notwithstanding the advantages of massive MIMO, several major technical challenges arise, which include the difficulty and complexity in hardware implementation, precoder design and channel estimation. In this thesis, we mainly focus on strategies that address the training overhead issue for mmWave massive MIMO channel estimation. By utilizing the sparsity feature in the angular domain of mmWave channels, we propose a gridless compressive sensing (CS) technique based on atomic norm minimization (ANM). Particularly for massive MIMO systems involving two-dimensional angle estimation, we develop a decoupled ANM (D-ANM) approach that offers high-accuracy channel estimation at low complexity and little training overhead. The proposed D-ANM approach is applied to mmWave massive MIMO systems with uniform rectangular array employed at base station and extended to the multi-user case. Investigation on the use of D-ANM for channel estimation in wideband mmWave

SIMO-OFDM systems is also carried out to cope with frequency-selective channel fading.

Chapter 1: Introduction

1.1 Background and Motivation

Massive multiple-input multiple-output (MIMO) systems that employ a large number of antennas at both receivers and transmitters have been widely considered for adoption in next generation (5G) wireless networks [1–5]. The deployment of massive MIMO promises to enhance the received signal power for communications over millimeter-wave (mmWave) spectrum, which in turn increases the throughput and system efficiency [4–8]. Moreover, the short wavelength associated with mmWave frequency facilitates implementation of massive antennas packed in small physical areas [3, 9]. In addition, moving to the mmWave regime means that large portions of unused spectrum that support orders of magnitude larger bandwidths (10s of GHz) comparing with existing systems can be used, which tackles the spectrum crisis at the current wireless frequencies [8]. Therefore, massive MIMO is usually coupled with mmWave communications to provide wireless services with good coverage, low latency, and high data rate for 5G communications [1, 3, 4].

Notwithstanding the advantages of massive MIMO, several major technical challenges arise in the efficient realization of mmWave and massive MIMO gains in practice, which include the difficulty and complexity in hardware implementation, precoder design and channel estimation [9]. The implicit assumption that a complete radio frequency (RF) chain is dedicated for each antenna in traditional MIMO systems does not work any longer for massive MIMO systems due to the high cost and power consumption of mixed signal components. While the analog-only beamforming/combining architectures that do all the required processing in the analog domain fail to perform sophisticated multi-stream or multi-user processing [10, 11], a widely adopted scheme that balances the system performance and hardware limitations is the hybrid analog/digital architecture proposed in [12, 13]. Hybrid architecture reduces the number of RF chains and partly alleviates the hardware constraints while effectively collecting the antenna gains through precoding. However, it introduces new problems on channel estimation and precoder design [13]. To achieve optimal design of precoding matrices, it is necessary to obtain complete knowledge of the channel state information (CSI), which is also critical for efficient transmission in wireless communication systems. However, the acquisition of CSI in mmWave massive MIMO systems using traditional channel estimation approaches is infeasible considering the high computational complexity and heavy training overhead induced by the large number of antennas [9].

It has been demonstrated that a mmWave channel usually exhibits a sparse multi-path structure due to limited scattering, suggesting that only a small number of significant paths dominate the propagation [14–16]. Therefore, compressive sensing (CS) based techniques are proposed to utilize the sparsity feature of mmWave channels to reduce training overhead in coherent channel estimation [13, 17–23]. However, CS-based approaches may suffer from considerable performance degradation in practice because they rely on an idealized on-grid assumption. For a MIMO channel, such an assumption means that the values of the angle of arrival (AoA) and angle of departure (AoD) of the dominant paths have to lie on some known grids [23], while in practice, the values of AoA and AoD are continuous and off-the-grid. Therefore, CS techniques experience a power leakage effect due to basis mismatch of the on-grid assumption, which has been analyzed in [24].

In order to estimate the MIMO channel characterized by continuous angular information, we resort to continuous spatial frequency or angle estimation techniques, for which considerable efforts have been put to circumvent the basis mismatch problem [25–29]. Among them, atomic norm minimization (ANM) has been proposed to deal with the continuous parameters directly, bypassing the on-grid assumption [26–29]. ANM is developed as a gridless CS approach, which retains the benefits of CS in terms of light-weight training overhead. In particular, by exploiting the Vandermonde structure of a signal, ANM attains super-resolution estimation, therefore outperforms the CS-based methods [26–28]. Moreover, different from the traditional super-resolution subspace-based methods that need to collect measurements from multiple snapshots to acquire signals' statistical information, ANM can work with only one or a few snapshots, which leads to short sensing time.

ANM can be further extended for two-dimensional (2D) harmonic retrieval through vectorization [28], which motivates us to apply the 2D vectorized ANM (V-ANM) on estimating narrowband mmWave massive MIMO channels based on the knowledge that both the transmit and receive uniform linear arrays (ULAs) exhibit Vandermonde structures. However, the computational complexity of V-ANM becomes unaffordable as the antenna size increases in massive MIMO scenarios, even in a truncated version [30]. The high complexity comes from the optimization formulation of V-ANM in the form of semi-definite programming (SDP), whose computational complexity is mainly contributed by the size of the semi-definite matrix in the optimization constraint that involves a large-size two-level Toeplitz matrix constructed from the vectorized channel matrix [28].

The goal of this thesis is to develop a low-complexity high-accuracy channel estimation approach for mmWave massive MIMO systems. While ANM-based approach has shown to attain high estimation accuracy at little training overhead, the remaining key challenge is to reduce the computational complexity, which is the focus of this thesis research.

1.2 Contributions of Thesis

To harness the potential of mmWave massive MIMO systems, it is important to obtain accurate CSI through channel estimation. With this motivation and the aforementioned challenges, the problem tackled in this thesis is to develop low-complexity high-accuracy channel estimation solutions. The primary contributions of this thesis can be summarized as follows:

1. We propose a low-complexity high-accuracy channel estimation algorithm for mmWave massive MIMO systems. Leveraging the sparsity feature of mmWave channels, we develop a sparse formulation of the mmWave channel estimation problem based on the decoupled atomic norm minimization (D-ANM). The optimization problem is then

solved by semi-definite programming (SDP). The approach is motivated by our work originally designed for 2D harmonic retrieval [31]. Applying D-ANM in mmWave massive MIMO channel estimation achieves super-resolution performance at reduced complexity. Unlike V-ANM that couples 2D angular information to form one large-size two-level Toeplitz matrix for SDP via vectorization, D-ANM does not involve vectorization. Instead, D-ANM decouples the large-size two-level Toeplitz matrix into two small-size one-level Toeplitz matrices, which leads to a reduced-size SDP for efficient computation. The decoupling is done by introducing a new atom set based on the channel structure in the D-ANM framework without losing optimality [31]. As the size of the semi-definite matrix is reduced dramatically, the computational load is decreased by several orders [31]. Since the number of antennas in massive MIMO systems is usually on the order of hundreds, the computational complexity reduction is nontrivial and benefits practical applications.

2. We apply the low-complexity high-accuracy D-ANM based channel estimation approach to the single-input multiple-output (SIMO) scenario where a 2D uniform rectangular array (URA) is deployed at the base station (BS). Simulation shows that channel estimation for this scenario can be done even with only one transmit symbol. We also apply the proposed algorithm to a downlink multi-user massive MIMO system with some modifications. The proposed algorithm utilizes the sparsity feature of mmWave channels and achieves good performance for multi-user systems.

3. We propose to estimate wideband mmWave channels using the low-complexity high-accuracy D-ANM based algorithm. To alleviate the frequency-selectivity of wideband channels, we resort to the orthogonal frequency division multiplexing (OFDM) technique to convert a frequency-selective fading channel into a group of flat fading channels. The antenna setting in this scenario is SIMO with an ULA deployed at the BS and a single antenna for the mobile station (MS). The 2D Vandermonde structures are

incorporated in spatial dimension (resulted from ULA) and frequency dimension (resulted from OFDM). D-ANM can therefore be applied to estimate the OFDM-SIMO channel.

1.3 Notations and Abbreviations

We use the following notations throughout this thesis: a is a scalar, $\mathbf{a}$ is a vector, $\mathbf{A}$ is a matrix, and $\mathcal{A}$ represents a set. $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, conjugate, and conjugate transpose of a matrix or vector, respectively. $\|\mathbf{a}\|_1$ and $\|\mathbf{a}\|_2$ are the ℓ_1 norm and ℓ_2 norm of $\mathbf{a}$, respectively. $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix with the diagonal elements constructed from $\mathbf{a}$. $\|\mathbf{A}\|_F$, is the Frobenius norm of $\mathbf{A}$. $\mathrm{tr}(\mathbf{A})$ is the trace of $\mathbf{A}$. The operation $\mathrm{vec}(\cdot)$ stacks all the columns of a matrix into a vector. $\otimes$ is the Kronecker product of matrices or vectors and $*$ denotes the convolution operation.

The abbreviations used in this dissertation are summarized in Table 1.1.

1.4 Organization of Thesis

The rest of this paper is organized as follows. Chapter 2 presents some preliminaries on wireless communication systems and atomic norm minimization. Chapter 3 presents the channel model and formulates the problem for single-user massive MIMO system. The low-complexity high-accuracy channel estimation technique based on D-ANM is also proposed. Chapter 4 applies the proposed D-ANM to the implementation of URA at the BS and multi-user scenarios. Chapter 5 extends the D-ANM based channel estimation algorithm to a SIMO-OFDM system. Conclusions and future works are finally presented in Chapter 6.

Table 1.1: Summary of abbreviations

5G	the fifth generation
AoA	angle of arrival
AoD	angle of departure
ANM	atomic norm minimization
AWGN	additive white gaussian noise
BS	base station
CIR	channel impulse response
CFR	channel frequency response
CP	cyclic prefix
CS	compressive sensing
CSI	channel state information
D-ANM	decoupled atomic norm minimization
DFT	discrete fourier transform
IDFT	inverse discrete fourier transform
ISI	intersymbol interference
mmWave	millimeter-wave
MIMO	multiple-input multiple-output
MS	mobile station
MSE	mean-squared-error
MMSE	minimum mean square error
OFDM	orthogonal frequency division multiplexing
RF	radio frequency
SDP	semi-definite programming
SIMO	single-input multiple-output
SNR	signal-to-noise ratio
UE	user equipment
ULA	uniform linear array
URA	uniform rectangular array
V-ANM	vectorized atomic norm minimization

Chapter 2: Preliminaries

2.1 Wireless Channel Models

Consider a typical scattering environment where N_r receive antennas receive a supposition of multiple attenuated, delayed, and phase/frequencey shifted copies of the original signal from N_t transmit antennas. Those copies caused by reflection, diffraction and scattering from the surrounding objects are called multipath signal components. Without loss of generality, consider uniform linear arrays (ULAs) with half-wavelength spacing, i.e. the antenna spacing d and signal wavelength λ are related by $d = \frac{\lambda}{2}$. Then the underlying time-varying frequency response matrix of the channel in terms of the underlying physical paths can be expressed as [9, 17]

$$\mathbf{H}(t, f) = \sqrt{\frac{N_r N_t}{\rho}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_r(\theta_{r,l}) \mathbf{a}_t^H(\theta_{t,l}) e^{-j2\pi\tau_l f} e^{j2\pi\nu_l t}, \tag{2.1}$$

where ρ is the average path loss, L represents the total number of paths. In mmWave environment, the value of L is usually small, meaning that the channel is dominated by only L strong paths, therefore referred as 'sparse' channel. The parameters α_l, $\theta_{r,l}$, $\theta_{t,l}$, τ_l and ν_l represent the complex path gain, angle of arrival (AoA) at the receiver, angle of departure (AoD) from the transmitter, relative time delay and the Doppler shift associated with the lth path, respectively. For the time delay and Doppler shift, we have $\tau_l \in [0, \tau_{max}]$ and $\nu_l \in [-\nu_{max}/2, \nu_{max}/2]$, where τ_{max} and ν_{max} are the delay spread and (two-sided) Doppler spread of the channel, respectively. The two steering vectors $\mathbf{a}_r$ and $\mathbf{a}_t$ associated

Table 2.1: Classification of wireless channels on the basis of channel and signaling parameters

Channel Classification	$W\tau_{max}$	$T\nu_{max}$
Nonselective Channels	$\ll 1$	$\ll 1$
Frequency-selective Channels	≥ 1	$\ll 1$
Time-selective Channels	$\ll 1$	≥ 1
Doubly-selective Channels	≥ 1	≥ 1

with receive and transmit antenna arrays are given by

$$\begin{aligned} \mathbf{a}_r(\theta_{r,l}) &= \frac{1}{\sqrt{N_r}}[1, e^{j2\pi\frac{d}{\lambda}\sin\theta_{r,l}}, \ldots, e^{j2\pi\frac{(N_r-1)d}{\lambda}\sin\theta_{r,l}}]^T, \\ \mathbf{a}_t(\theta_{t,l}) &= \frac{1}{\sqrt{N_t}}[1, e^{j2\pi\frac{d}{\lambda}\sin\theta_{t,l}}, \ldots, e^{j2\pi\frac{(N_t-1)d}{\lambda}\sin\theta_{t,l}}]^T. \end{aligned} \tag{2.2}$$

Suppose that the channel varies sufficiently slow over the signal duration T, that is, the Doppler shift is small for all paths, or equivalently, $T\nu_{max} \ll 1$, then (2.1) can be expressed as

$$\mathbf{H}(f) = \sqrt{\frac{N_r N_t}{\rho}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_r(\theta_{r,l}) \mathbf{a}_t^H(\theta_{t,l}) e^{-j2\pi\tau_l f}. \tag{2.3}$$

If in addition, the bandwidth W of the channel is sufficiently small so that $W\tau_{max} \ll 1$, then a narrowband spatial model for the channel matrix is obtained:

$$\mathbf{H} = \sqrt{\frac{N_r N_t}{\rho}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_r(\theta_{r,l}) \mathbf{a}_t^H(\theta_{t,l}). \tag{2.4}$$

It is obvious that the wireless channels can be broadly classified as nonselective, frequency selective, time selective, or doubly selective using the signal and channel parameters, see Table 2.1 for exact definitions [17]. In this thesis, we mainly study the SIMO frequency-selective channels with ULAs employed at the receivers, SIMO nonselective channels with uniform rectangular arrays (URAs) employed at the receivers and MIMO nonselective channels with ULAs employed at both receivers and transmitters. More details are given in the corresponding chapters.

2.2 Atomic Norm Minimization (ANM)

Throughout science and engineering, one of the fundamental but challenging tasks is to deduce the state or structure of a system from its partial, noisy measurements. The difficulty mainly comes from the acquisition of enough measurements relative to the ambient dimension of the signal of interest. In practice, however, the structures of those models of interesting signals are usually constrained so that they only have a few degrees of freedom relative to their large dimensions. For instance, the signature of a disease is usually constituted by a small number of genes, a molecular configuration may be completely specified by a sparse collection of geometric constraints, etc. These kind of low-dimensional structures play an important role in converting the ill-posed inverse problems to well-posed ones and can be unified and formulated as convex optimization problems [32].

In the framework of applying convex optimization to solve the linear under-determined inverse problems, the class of simple models considered are formed as linear combinations of a few atoms from some (possibly infinite) elementary atom set. Some well-studied examples include ℓ_1-minimization for sparse recovery where the atoms are unit-norm one-sparse vectors, nuclear norm minimization for low-rank matrix completion where the atoms are unit-norm rank-one matrices, and so on. The optimization problem is therefore to minimize the norms induced by the convex hull of the atom set and is referred to as atomic norm minimization (ANM). To illustrate how ANM works, we apply it on a simple sparse signal recovery problem.

Suppose a signal $\mathbf{x} \in \mathbb{C}^{M \times 1}$ can be denoted as a combination of some complex exponentials, which in matrix-vector form is

$$\mathbf{x} = \sum_{l=1}^{K} s_l \mathbf{a}(f_l) \tag{2.5}$$

where $s_l \in \mathbb{C},\ f_l \in [0, 1),\ [\mathbf{a}(f_l)]_i = e^{j2\pi(i-1)f_l},\ i \in \{1, 2, ..., M\}$, K is the unknown model order with $K < M$.

The atom set of a signal is defined as its simplest building blocks, the same way as the unit-norm one-sparse vectors for sparse signals, unit-norm rank-one matrices for low-rank matrices and so on. From the denotation of $\mathbf{x}$ in (2.5), its atom set is therefore a collection of $\mathbf{a}(f_l)$, denoted as $\mathcal{A} = \{\mathbf{a}(f) : f \in [0, 1)\}$.

The atomic norm of $\mathbf{x}$ induced by its atom set is defined accordingly as

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf \left\{ \sum_l |s_l| \,\middle|\, \mathbf{x} = \sum_l s_l \mathbf{a}(f_l), \mathbf{a}(f_l) \in \mathcal{A} \right\}. \tag{2.6}$$

Equation (2.6) is obtained by convexifying the representation of $\mathbf{x}$ using the smallest number of complex exponential components, that is

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf \left\{ K \,\middle|\, \mathbf{x} = \sum_{l=1}^{K} s_l \mathbf{a}(f_l), \mathbf{a}(f_l) \in \mathcal{A} \right\},$$

which, however, is an NP-hard problem and can not be solved by convex optimization. Basically, the definition of atomic norm enforces sparsity in the atom set $\mathcal{A}$. Interested readers are referred to [32] for a detailed discussion about atomic norm.

Given the atom set defined above, the attempt on recovering $\mathbf{x}$ can be done via minimizing its atomic norm. In noise free case, the measurements obtained is denoted as $\mathbf{x}^\star$

and the optimization goes as follow:

$$\begin{aligned} \min_{\mathbf{x}} \quad & ||\mathbf{x}||_{\mathcal{A}} \\ \text{s.t.} \quad & \mathbf{x} = \mathbf{x}^{\star}. \end{aligned} \tag{2.7}$$

From the definition of atom set and atomic norm, we can see that ANM works directly on continuous parameter space and utilizes the sparse feature of the signal of interest, therefore, it avoids the basis mismatch problem encountered in CS techniques and reduces the number of measurements needed for recovery. However, for the sparse signal to be recovered, it may have infinite many possible combinations, therefore to search all the possible combinations so that the minimum value of $\sum_l |s_l|$ can be obtained makes ANM untractable. One way to turn the untractable ANM problem into a tractable one is to explore the structure of the signal of interest. Researchers in [27] utilize the Vandermonde structure of the signal and exploit the semidefinite characterization of ANM, obtaining the following proposition.

Proposition 2.2.1. *For* $\mathbf{x} \in \mathbb{C}^M$ *that can be expressed as a linear combination of complex exponentials,*

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf \left\{ \frac{1}{2M}\text{trace}(\text{Toep}(\mathbf{u})) + \frac{1}{2}\upsilon : \begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & \upsilon \end{bmatrix} \succeq \mathbf{0} \right\}, \tag{2.8}$$

where $\text{Toep}(\mathbf{u})$ *denotes the Toeplitz matrix whose first column is equal to* $\mathbf{u}$ *and is constructed exactly from the atoms that constitute signal* $\mathbf{x}$ *when its atomic norm is minimized.*

Proof: *See Appendix A.*

Therefore, to recover $\mathbf{x}$ with its noise-free measurements $\mathbf{x}^{\star}$ using ANM in its original form (2.7) is equivalent to solve the following semidefinite programming (SDP):

$$\begin{aligned}
&\min_{\mathbf{u},\mathbf{x},\upsilon} \frac{1}{2M}\text{trace}(\text{Toep}(\mathbf{u})) + \frac{1}{2}\upsilon \\
&\text{s.t.} \begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & \upsilon \end{bmatrix} \succeq \mathbf{0} \\
&\mathbf{x} = \mathbf{x}^\star.
\end{aligned} \tag{2.9}$$

For the noisy case where $\mathbf{y} = \mathbf{x} + \mathbf{w}$, with $\mathbf{y} \in \mathbb{C}^M$ and $\mathbf{w} \in \mathbb{C}^M$ be the noisy measurements and additive white gaussian noise (AWGN), respectively, the SDP for recovering signal $\mathbf{x}$ is formulated as

$$\begin{aligned}
&\min_{\mathbf{u},\mathbf{x},\upsilon} \frac{\mu}{2}(\frac{1}{M}\text{trace}(\text{Toep}(\mathbf{u})) + \upsilon) + \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 \\
&\text{s.t.} \begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & \upsilon \end{bmatrix} \succeq \mathbf{0},
\end{aligned} \tag{2.10}$$

where μ is the regularization parameter. The SDP can be solved using the CVX toolbox [33].

So far we have introduced how the ANM problem arises and be solved. However, the estimation parameters involved here is only one-dimensional (1D), to solve the two-dimensional (2D) estimation problems such as 2D harmonic retrieval, approaches that include converting the signal to fit the ANM model has been proposed [28]. In [28], the signal of interest is vectorized and an enlarged ANM problem is formed and optimized by the corresponding SDP. Notice that the narrowband mmWave massive MIMO channel exhibits the same structure as the 2D signal in [28], we propose a channel estimation method based on the 2D ANM. More details are given in Section 3.3.

Chapter 3: Receiver Design for Millimeter-wave Massive MIMO Systems

3.1 Channel Model and Problem Formulation

Consider a point-to-point millimeter-wave (mmWave) massive MIMO wireless system with ULAs of N_t transmit and N_r receive antennas[1]. The array steering vectors of transmit and receive antenna arrays are denoted as $\mathbf{a}_r(\theta_r)$ and $\mathbf{a}_t(\theta_t)$, where θ_r and θ_t are angular directions of arriving and departing plane waves [9]. Receiver design in this thesis mainly focuses on developing estimator to obtain accurate channel state information (CSI). CSI is usually obtained by performing channel estimation within a period shorter than the coherence time such that the channel is invariant during a training block of time length T. Suppose that there is no Doppler shift and the channel experiences flat fading. As introduced in Chapter 1, at mmWave frequency, a MIMO channel experiences limited scattering, leading to a sparse multi-path propagation. For simplicity, assume that only L significant paths exist in the mmWave massive MIMO channel, with $L \ll \min\{N_r, N_t\}$. Then the spatial model for the channel in the matrix form $\mathbf{H}$, where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, can be expressed as a linear combination of L dominant paths as [9]

$$\begin{aligned}\mathbf{H} &= \sqrt{\frac{N_r N_t}{\rho}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_r(\theta_{r,l}) \mathbf{a}_t^{\mathrm{H}}(\theta_{t,l}) \\ &= \sum_{l=1}^{L} s_l \mathbf{a}_r(\theta_{r,l}) \mathbf{a}_t^{\mathrm{H}}(\theta_{t,l}) \\ &= \mathbf{A}_r(\boldsymbol{\theta}_r) \mathbf{S} \mathbf{A}_t^{\mathrm{H}}(\boldsymbol{\theta}_t),\end{aligned} \tag{3.1}$$

[1] Generalization to a multi-user scenario and other types of uniform arrays will be discussed in Chapter 4.

where ρ is the average path loss, α_l represents the fading coefficient of l-th propagation path, $\{\theta_{r,l}\}$ and $\{\theta_{t,l}\}$ are the AoA to the transmitter and AoD from the receiver of l-th propagation path, respectively, ranging from 0 to 2π in radian. The path gains $s_l = \sqrt{\frac{N_r N_t}{\rho}}\alpha_l$ are assumed to be uncorrelated and $\mathbf{S} = \text{diag}([s_1, s_2, \ldots, s_l])$. The normalized ULA steering vectors of transmit and receive antennas are given as (2.2), which are

$$\begin{aligned} \mathbf{a}_r(\theta_{r,l}) &= \frac{1}{\sqrt{N_r}}[1, e^{j2\pi\frac{d}{\lambda}\sin\theta_{r,l}}, \ldots, e^{j2\pi\frac{(N_r-1)d}{\lambda}\sin\theta_{r,l}}]^T, \\ \mathbf{a}_t(\theta_{t,l}) &= \frac{1}{\sqrt{N_t}}[1, e^{j2\pi\frac{d}{\lambda}\sin\theta_{t,l}}, \ldots, e^{j2\pi\frac{(N_t-1)d}{\lambda}\sin\theta_{t,l}}]^T, \end{aligned} \tag{3.2}$$

where the antenna spacing d and signal wavelength λ are related by $d = \frac{\lambda}{2}$. It is clear that the steering matrices $\mathbf{A}_r(\boldsymbol{\theta}_r) = [\mathbf{a}_r(\theta_{r,1}), \ldots, \mathbf{a}_r(\theta_{r,L})]$ and $\mathbf{A}_t(\boldsymbol{\theta}_t) = [\mathbf{a}_t(\theta_{t,1}), \ldots, \mathbf{a}_t(\theta_{t,L})]$ incorporate Vandermonde structures, which are useful properties and will be utilized later.

For coherent channel estimation, the training symbol vector $\mathbf{x}_i$ is transmitted over N_t antennas at each time instant i ($i = 1, \cdots, T$) and passes through the channel $\mathbf{H}$. The received measurements $\mathbf{y}_i$ are given by

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \mathbf{w}_i, \qquad i = 1, \cdots, T, \tag{3.3}$$

where $\mathbf{w}_i \in \mathbb{C}^{N_r}$ is the AWGN at the receiver whose distribution is independent of $\mathbf{H}$ and $\mathbf{x}_i$. The signal model can be concisely written in matrix form as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \tag{3.4}$$

where $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T] \in \mathbb{C}^{N_r \times T}$, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T] \in \mathbb{C}^{N_t \times T}$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_T] \in \mathbb{C}^{N_r \times T}$ are the received data, transmitted data and the noise term over a time length T, respectively.

The task of receiver design in terms of channel estimation boils down to estimating the

CSI in terms of $\mathbf{H}$ from $\{\mathbf{Y}, \mathbf{X}\}$, where $\mathbf{Y}$ is corrupted by AWGN noise $\mathbf{W}$. CSI acquisition is required for the design of transmitter (precoder) and receiver (combiner) for reliable data transmission. However, conventional channel estimation methods do not work well in massive MIMO systems due to the large number of antennas, resulting in tremendous training overhead and long sensing time. Therefore, this thesis tries to develop new techniques that can not only reduce training overhead but also achieve high estimation accuracy. Before introducing the proposed low-complexity high-accuracy algorithm for channel estimation, two related works are summarized first.

3.2 Related Work I: Compressive Sensing (CS) Techniques

When the antenna size is large, as in massive MIMO systems, an ideal on-grid assumption can be adopted, that is the angles in (3.2) fall onto L of the N_r or N_t directions on the equal-spaced grids, or to say $\theta_{r,l} \in \{0, \frac{2\pi}{N_r}, \dots, \frac{(N_r-1)2\pi}{N_r}\}$ and $\theta_{t,l} \in \{0, \frac{2\pi}{N_t}, \dots, \frac{(N_t-1)2\pi}{N_t}\}$, even though the L locations are unknown [19]. By this assumption, a virtual channel representation is applied to characterize the sparse MIMO channel along fixed virtual receive and transmit directions [9, 17], that is,

$$\mathbf{H} = \mathbf{F}_r \check{\mathbf{H}} \mathbf{F}_t^H, \tag{3.5}$$

where $\mathbf{F}_r \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{F}_t \in \mathbb{C}^{N_t \times N_t}$ are unitary discrete Fourier transform (DFT) matrices. In this sense, the virtual channel matrix $\check{\mathbf{H}}$ and the original channel matrix $\mathbf{H}$ are unitarily equivalent. The virtual channel matrix $\check{\mathbf{H}}$ is sparse and contains only L nonzero entries corresponding to the L paths, with AoA information reflected by the nonzero rows of $\check{\mathbf{H}}$ and AoD information reflected by the nonzero columns of $\check{\mathbf{H}}$.

The sparsity feature of mmWave channel motivates the utilization of compressive sensing (CS) techniques [34], which states that the reconstruction of a sparse signal can be done using a small number of compressive samples collected from its linear projections. To utilize

CS, the sparsity of $\check{\mathbf{H}}$ is illustrated in $\check{\mathbf{h}} := \text{vec}(\check{\mathbf{H}})$, and the system model is reformulated as [17]:

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = ((\mathbf{F}_t^H \mathbf{X})^T \otimes \mathbf{F}_r)\check{\mathbf{h}} + \text{vec}(\mathbf{W}), \tag{3.6}$$

where $\otimes$ represents Kronecker product. CS suggests that $\check{\mathbf{h}}$ can be recovered by ℓ_1 minimization as:

$$\min_{\check{\mathbf{h}}} \quad \mu\|\check{\mathbf{h}}\|_1 + \|\mathbf{y} - ((\mathbf{F}_t^H \mathbf{X})^T \otimes \mathbf{F}_r)\check{\mathbf{h}}\|_2^2 \tag{3.7}$$

where μ is the regularization parameter.

After recovering $\check{\mathbf{h}}$, the estimated channel $\hat{\mathbf{H}}$ can be obtained by doing inverse spatial Fourier transform on $\check{\mathbf{H}}$. However, the estimation performance of CS techniques suffers when the on-grid assumption is not guaranteed, that is, either AoA or AoD are continuously-valued off-the-grid. As a result, the CS-based channel estimation methods may not work effectively in practice.

3.3 Related Work II: Vectorized ANM (V-ANM)

For recovery of sparse continuously-valued signals, gridless CS in the form of atomic norm minimization (ANM) has been advocated for frequency/angle estimation. While Section 2.2 gives some fundamental knowledge on how 1D ANM works for a sparse signal reconstruction with continuous parameters, in this section, we show an direct application of 2D ANM in mmWave massive MIMO channel estimation, which stems from the 2D harmonic retrieval work in [28]. To utilize the 2D ANM approach for channel estimation, we vectorize both sides of (3.4) to yield

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = (\mathbf{X}^T \otimes \mathbf{I})\mathbf{h} + \text{vec}(\mathbf{W}), \tag{3.8}$$

where $\mathbf{I}$ is an identity matrix of size $N_r \times N_r$. Due to the vectorization operation in (3.8), we call this technique vectorized ANM (V-ANM).

The vectorization operation on the channel matrix $\mathbf{H}$ in (3.1) shows that the vectorized channel $\mathbf{h}$ is of the form

$$\mathbf{h} = \text{vec}(\mathbf{H}) = \sum_{l=1}^{L} s_l \mathbf{a}_t^*(\theta_{t,l}) \otimes \mathbf{a}_r(\theta_{r,l}) = \sum_{l=1}^{L} s_l \mathbf{a}_{2D}(\boldsymbol{\theta}_l), \tag{3.9}$$

where $\boldsymbol{\theta}_l = (\theta_{r,l}, \theta_{t,l})$, and $\mathbf{a}_{2D}(\boldsymbol{\theta}_l) = \mathbf{a}_t^*(\theta_{t,l}) \otimes \mathbf{a}_r(\theta_{r,l})$ is an extended 2D array response vector of length $N_r N_t$.

In (3.9), the vectorized channel $\mathbf{h}$ can be seen as a linear combination of L atoms $\mathbf{a}_{2D}(\boldsymbol{\theta}_l)$, and those atoms belong to the atom set $\mathcal{A}_V$, which is defined as

$$\mathcal{A}_V \triangleq \{\mathbf{a}_{2D}(\boldsymbol{\theta}), \boldsymbol{\theta} \in [0, 2\pi) \times [0, 2\pi)\}. \tag{3.10}$$

The atomic norm of $\mathbf{h}$ over its atom set $\mathcal{A}_V$ is then given by

$$\|\mathbf{h}\|_{\mathcal{A}_V} \triangleq \inf \left\{ \sum_l |s_l| \,\middle|\, \mathbf{h} = \sum_l s_l \mathbf{a}_{2D}(\boldsymbol{\theta}_l)\ , \mathbf{a}_{2D}(\boldsymbol{\theta}_l) \in \mathcal{A}_V \right\}. \tag{3.11}$$

Similar as the 1D ANM framework given in Section 2.2, the vector $\mathbf{h}$ is recovered when (3.11) is minimized, which involves searching all possible combinations over the continuously-valued angle space of infinite size. Hence, minimizing the atomic norm of $\mathbf{h}$ in the form of (3.11) is not tractable. To reach a computational feasible solution, the Vandermonde structure of atoms can be utilized to convert (3.11) into an equivalent semi-definite programming (SDP) as follows [28]:

$$\|\mathbf{h}\|_{\mathcal{A}_V} \triangleq \inf \left\{ \frac{1}{2}\Big(\upsilon + \text{trace}(\mathbf{T}_{2D}(\mathbf{u})\Big) \,\middle|\, \begin{bmatrix} \mathbf{T}_{2D}(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^H & \upsilon \end{bmatrix} \succeq \mathbf{0} \right\}. \tag{3.12}$$

Here both υ and $\mathbf{u}$ are optimization variables, $\mathbf{T}_{2D}(\mathbf{u})$ is a two-level Toeplitz matrix of size $N_r N_t \times N_r N_t$ constructed from $\mathbf{a}_{2D}(\boldsymbol{\theta})$, vector $\mathbf{u} \in \mathbb{C}^{N_r N_t \times 1}$ is of size $N_r N_t$ related with

$\mathbf{a}_{2D}(\boldsymbol{\theta})$ [28]. Specifically, $\mathbf{T}_{2D}(\mathbf{u})$ is defined by $\mathbf{u}$ as follows:

$$\mathbf{T}_{2D}(\mathbf{u}) = \begin{bmatrix} \mathbf{T}_0 & \mathbf{T}_1 & \cdots & \mathbf{T}_{N_t-1} \\ \mathbf{T}_{-1} & \mathbf{T}_0 & \cdots & \mathbf{T}_{N_t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}_{-(N_t-1)} & \mathbf{T}_{-(N_t-2)} & \cdots & \mathbf{T}_0 \end{bmatrix}, \tag{3.13}$$

where each block $\mathbf{T}_b$ $(-(N_t-1) \leq b \leq N_t - 1)$ is an $N_r \times N_r$ Toeplitz matrix of the form

$$\mathbf{T}_b = \begin{bmatrix} u_{b,0} & u_{b,1} & \cdots & u_{b,(N_r-1)} \\ u_{b,-1} & u_{b,0} & \cdots & u_{b,(N_r-2)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{b,-(N_r-1)} & u_{b,-(N_r-2)} & \cdots & u_{b,0} \end{bmatrix}. \tag{3.14}$$

For illustration, a pseudocolor plot of the elements of a two-level Toeplitz matrix with $N_t = 8$ and $N_r = 6$ is drawn in Figure 3.1. It is clear that the Toeplitz matrix has two levels where the 8×8 matrices of size 6×6 in the first level are symmetric and the matrices of size 6×6 in the second level are Toeplitz.

Based on the two-level Toeplitz structure of $\mathbf{T}_{2D}(\mathbf{u})$, the convex optimization of V-ANM in the form of SDP can be formulated as

$$\begin{aligned} \min_{\upsilon,\mathbf{u},\mathbf{h}} \quad & \frac{\mu}{2}\big(\upsilon + \text{trace}\big(\mathbf{T}_{2D}(\mathbf{u})\big)\big) + \|\mathbf{y} - (\mathbf{X}^T \otimes \mathbf{I})\mathbf{h}\|_2^2 \\ \text{s.t.} \quad & \begin{pmatrix} \upsilon & \mathbf{h}^{\mathrm{H}} \\ \mathbf{h} & \mathbf{T}_{2D}(\mathbf{u}) \end{pmatrix} \succeq \mathbf{0}. \end{aligned} \tag{3.15}$$

In (3.15), the first term of the objective function, along with the constraint, is the ANM component for 2D channel reconstruction, the second term $\|\cdot\|_2^2$ in the objective function is the least squares fit of the measurement model (3.8) in the presence of noise, and μ is the

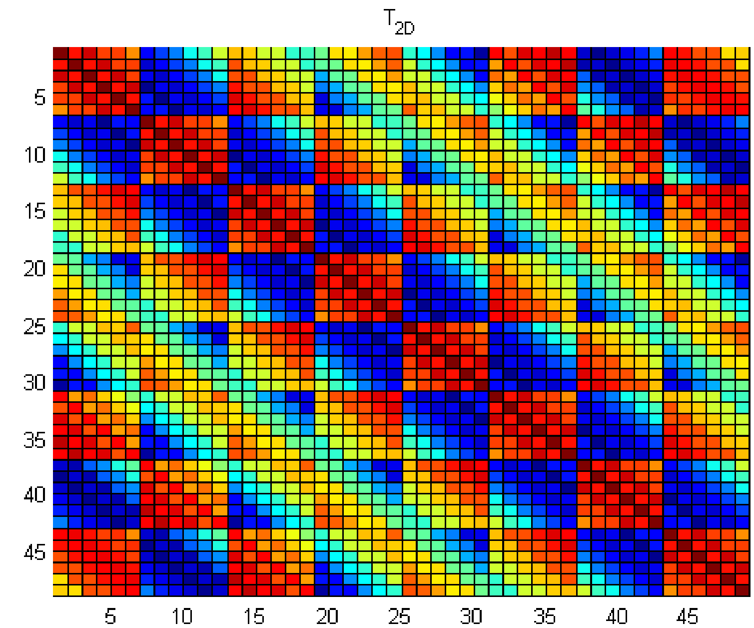

Figure 3.1: Two-level Toeplitz matrix with $N_t = 8, N_r = 6$.

regularization parameter controlling the trade-off between the Toeplitz structure and the noise tolerance to the observation [26].

As V-ANM deals with continuous angular values, it does not have the basis mismatch problem as in CS-based methods and thus obtains better estimation performance. However, V-ANM leads to much higher computational load than CS-based methods because of the much enlarged size $(N_r N_t + 1) \times (N_r N_t + 1)$ of the semi-definite matrix in the constraint of (3.15). The computational complexity goes extremely high when both N_r and N_t become large for massive MIMO systems. In training stage, a low-complexity algorithm is always desired so that channel estimation does not consume too much resources and more resources can be used for subsequent transmission. This motivates us to develop super-resolution channel estimation algorithm by utilizing both the sparsity feature and the Vandermonde structure of mmWave MIMO channel to achieve high performance at reduced computational complexity and shorter training time.

3.4 Proposed Work: Decoupled ANM (D-ANM)

To retain the benefits of V-ANM at remarkably reduced computational complexity, an efficient algorithm for 2D harmonic retrieval based on decoupled atomic norm minimization (D-ANM) is developed in [31]. In this section, we show how it is adopted for channel estimation in mmWave massive MIMO setting.

Instead of vectorizing the channel matrix $\mathbf{H}$, we directly express it as

$$\mathbf{H} = \sum_{l=1}^{L} s_l \mathbf{A}_D(\boldsymbol{\theta}_l), \tag{3.16}$$

where $\mathbf{A}_D(\boldsymbol{\theta}_l) = \mathbf{a}_r(\theta_{r,l})\mathbf{a}_t^{\mathrm{H}}(\theta_{t,l})$ can be viewed as the atoms forming $\mathbf{H}$. Here, the 2D angular information in both AoA and AoD is well incorporated in $\mathbf{A}_D(\boldsymbol{\theta})$, in contrast to the augmented vectors $\mathbf{A}_{2D}(\boldsymbol{\theta})$ in (3.10).

Accordingly, we define the new atom set as follows [31]:

$$\begin{aligned} \mathcal{A}_D &= \{\mathbf{A}_D(\boldsymbol{\theta}), \quad \boldsymbol{\theta} \in [0, 2\pi) \times [0, 2\pi)\} \\ &= \left\{\mathbf{a}_r(\theta_r)\mathbf{a}_t^{\mathrm{H}}(\theta_t), \quad \theta_r \in [0, 2\pi), \theta_t \in [0, 2\pi)\right\}. \end{aligned} \tag{3.17}$$

Then, the atomic norm of channel matrix $\mathbf{H}$ over its atom set $\mathcal{A}_D$ is given by

$$\|\mathbf{H}\|_{\mathcal{A}_D} \triangleq \inf \left\{ \sum_l |s_l| \;\middle|\; \mathbf{H} = \sum_l s_l \mathbf{A}_D(\boldsymbol{\theta}_l),\ \mathbf{A}_D(\boldsymbol{\theta}_l) \in \mathcal{A}_D \right\}. \tag{3.18}$$

The atomic norm of $\mathbf{H}$ is minimal only when those atoms corresponding to the true $\theta_{r,l}$ and $\theta_{t,l}$ are selected to linearly describe $\mathbf{H}$. For theoretical proof, one can refer to [31]. Notice that the vectorization operation from $\mathbf{H}$ to $\mathbf{h}$ is a one-to-one mapping, thus D-ANM shares the same optimality as V-ANM.

The equivalent SDP formulation of (3.18) is given by [31]:

$$\|\mathbf{H}\|_{\mathcal{A}_D} \triangleq \inf \left\{ \frac{1}{2}\Big(\mathrm{trace}(\mathbf{T}(\mathbf{u}_r)) + \mathrm{trace}(\mathbf{T}(\mathbf{u}_t))\Big) \middle| \begin{bmatrix} \mathbf{T}(\mathbf{u}_t) & \mathbf{H}^H \\ \mathbf{H} & \mathbf{T}(\mathbf{u}_r) \end{bmatrix} \succeq \mathbf{0} \right\}. \tag{3.19}$$

Here the two Toeplitz matrices $\mathbf{T}(\mathbf{u}_r)$ and $\mathbf{T}(\mathbf{u}_t)$ are constructed from $\mathbf{a}_r(\theta_r)$ and $\mathbf{a}_t(\theta_t)$, respectively. Parameterized by $\mathbf{u}_r \in \mathbb{C}^{N_r}$ and $\mathbf{u}_t \in \mathbb{C}^{N_t}$, $\mathbf{T}(\mathbf{u}_r)$ and $\mathbf{T}(\mathbf{u}_t)$ are of the forms

$$\begin{aligned} \mathbf{T}(\mathbf{u}_r) &= \begin{bmatrix} u_{r0} & u_{r1} & \cdots & u_{r(N_r-1)} \\ u_{r(-1)} & u_{r0} & \cdots & u_{r(N_r-2)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{r(1-N_r)} & u_{r(2-N_r)} & \cdots & u_{r0} \end{bmatrix}, \\ \mathbf{T}(\mathbf{u}_t) &= \begin{bmatrix} u_{t0} & u_{t1} & \cdots & u_{t(N_t-1)} \\ u_{t(-1)} & u_{t0} & \cdots & u_{t(N_t-2)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t(1-N_t)} & u_{t(2-N_t)} & \cdots & u_{t0} \end{bmatrix}. \end{aligned} \tag{3.20}$$

It is obvious that the above Toeplitz matrices have a much simpler structure and a considerably smaller size than the 2D Toeplitz matrix $\mathbf{T}_{2D}(\mathbf{u})$ in (3.13), which would make the associated computation much easier.

Therefore, estimating a narrow-band mmWave massive MIMO channel can be done by recovering channel matrix $\mathbf{H}$ through D-ANM, which is formulated as

$$\begin{aligned} \min_{\mathbf{u}_r, \mathbf{u}_t, \mathbf{H}} \quad & \frac{\mu}{2}\Big(\mathrm{trace}(\mathbf{T}(\mathbf{u}_r)) + \mathrm{trace}(\mathbf{T}(\mathbf{u}_t))\Big) \\ & + \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & \begin{pmatrix} \mathbf{T}(\mathbf{u}_t) & \mathbf{H}^{\mathrm{H}} \\ \mathbf{H} & \mathbf{T}(\mathbf{u}_r) \end{pmatrix} \succeq \mathbf{0}. \end{aligned} \tag{3.21}$$

The channel in terms of matrix $\mathbf{H}$ can be directly recovered by solving the SDP without any additional operations. Furthermore, as the two-level Toeplitz matrix $\mathbf{T}_{2D}(\mathbf{u})$ in V-ANM is decoupled into two one-level Toeplitz matrices $\mathbf{T}(\mathbf{u}_r)$ and $\mathbf{T}(\mathbf{u}_t)$ in D-ANM, the constraint size in SDP formulation is reduced from $(N_r N_t + 1) \times (N_r N_t + 1)$ in (3.15) to $(N_r + N_t) \times (N_r + N_t)$ in (3.21). As the computational complexity of ANM based methods mainly depends on the size of the SDP constraint, D-ANM has much lower computational complexity than V-ANM, which is attractive to massive MIMO systems where the antenna size is huge.

3.5 Simulation Results

In this section, we evaluate the performance of the proposed D-ANM based approach for a point-to-point MIMO channel over mmWave, and compare its performance with V-ANM and CS-based method in different simulation setups. The estimation performance is first testified in terms of mean-squared-error (MSE) of the estimated channel $\hat{\mathbf{H}}$ averaged over 1000 Monte Carlo trials. In each trial, the AoA/AoD and path gains of the MIMO channel are all randomly generated. The training data $\mathbf{X}$ are binary symbols randomly generated according to the Bernoulli distribution with equal probability. Due to the very high computational load of V-ANM, we set $N_r = N_t = 8,\ L = 2$ for illustration, though the proposed D-ANM approach can be applied to large-scale cases. We first test the performance of these methods in terms of average MSE versus different signal-to-noise ratio (SNR). All the three methods are implemented using the CVX toolbox [33]. The MSE performance is given in Figure 3.2. From Figure 3.2, we can see that both V-ANM and D-ANM outperform the CS-based method in terms of lower MSE with the same SNR value. Moreover, when increasing SNR, the MSE values for ANM-based methods keep decreasing while the MSE curve for CS-based method flattens out, which corroborates that the basis mismatch problem is inevitable in CS techniques and that ANM-based methods can achieve super-resolution in angle/frequency estimation.

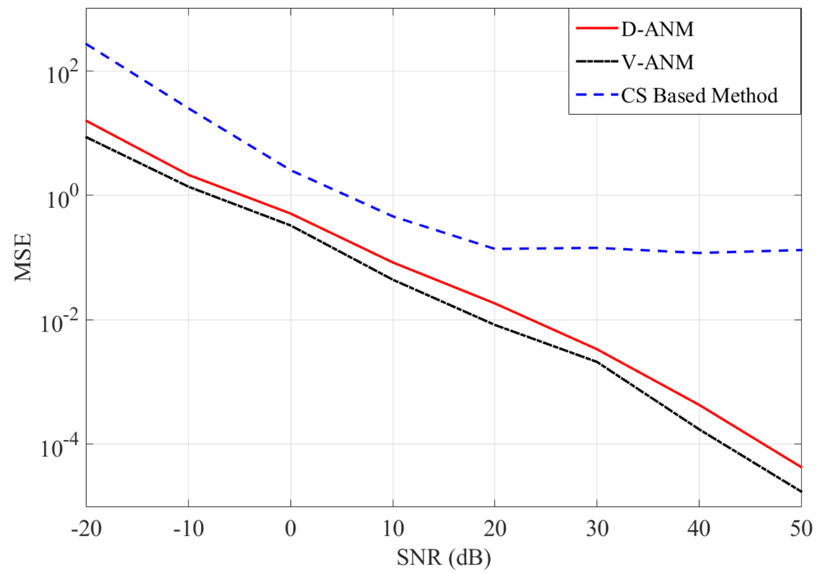

Figure 3.2: MSE versus SNR ($N_r = N_t = 8,\ L = 2,\ T = 6$).

Figure 3.3 shows how the MSE varies with sensing time T. It is shown that the two ANM-based methods consume less sensing time than the CS-based method for the same estimation accuracy. Further, the MSE curves of both ANM-based methods flatten out when T reaches a moderate value while the CS-based method needs larger T, which indicates that ANM can afford to use less training resources for the same estimation accuracy than the CS-based method. Such saving in training time may improve the throughput of data transmission as more time resources can be allocated for data transmission.

For the two ANM-based methods, both Figure 3.2 and Figure 3.3 indicate that D-ANM always exhibits a trivial performance gap compared with V-ANM, which results from the decoupling step. However, for the very large antenna size of massive MIMO systems, the small performance gap is worthy to exchange for the significant reduction in implementation complexity, which will be corroborated by the upcoming simulation results.

We also test the spectral efficiency and bit error rate (BER) performance to evaluate the impact of channel estimation on the achievable system. For simplicity, binary phase shift keying (BPSK) is adopted. Figure 3.4 and Figure 3.5 show the spectral efficiency

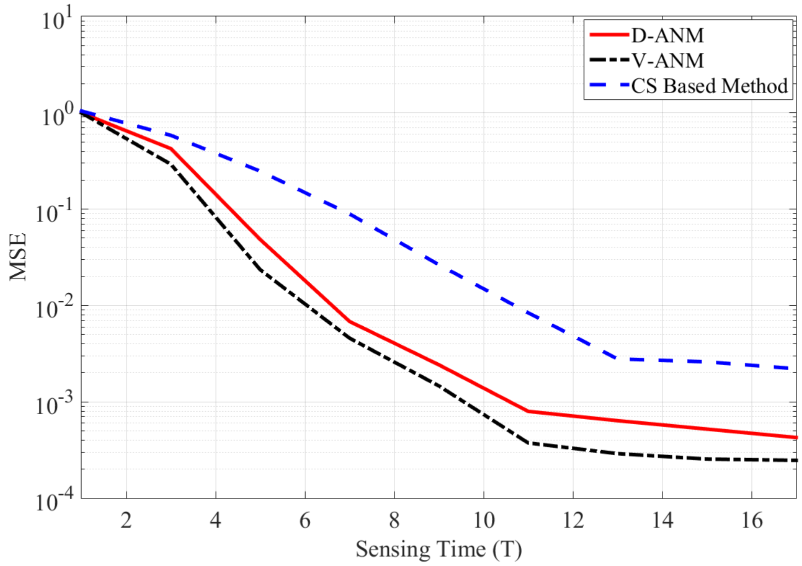

Figure 3.3: MSE versus sensing time T $(N_r = N_t = 8,\ L = 2,\ \text{SNR} = 15\ dB)$.

versus the number of antennas and the sensing time T, respectively. Both figures indicate that the spectral efficiency provided by V-ANM and D-ANM are much better than that of CS-based method. Even though the resolution of CS-based method increases as the antenna size increases, ANM-based methods still outperform the CS ones because the former can achieve super-resolution performance. Moreover, the gap on spectral efficiency between D-ANM and V-ANM are negligible when T becomes large. The spectral efficiency versus SNR is shown in Figure 3.6 and the bit error rate (BER) performance versus SNR is shown in Figure 3.7. Both figures show that D-ANM and V-ANM offer nearly the same spectral efficiency and BER performance, and outperform the CS-based method.

Figure 3.8 compares the computational cost of the three methods, measured by the average running time of one trial for the setting $N_r = N_t = 8$, $L = 2$, $T = 6$ and SNR $= 15$ dB. It can be seen that V-ANM has much higher computational complexity than D-ANM. According to [31], the computational complexities of ANM-based methods depend on the size of the Toeplitz matrices in their semi-definite constraints. For V-ANM, the matrix's size $(N_rN_t + 1) \times (N_rN_t + 1)$ results in a time complexity of $\mathcal{O}(N_r^{3.5}N_t^{3.5}\log(1/\epsilon))$, where ϵ

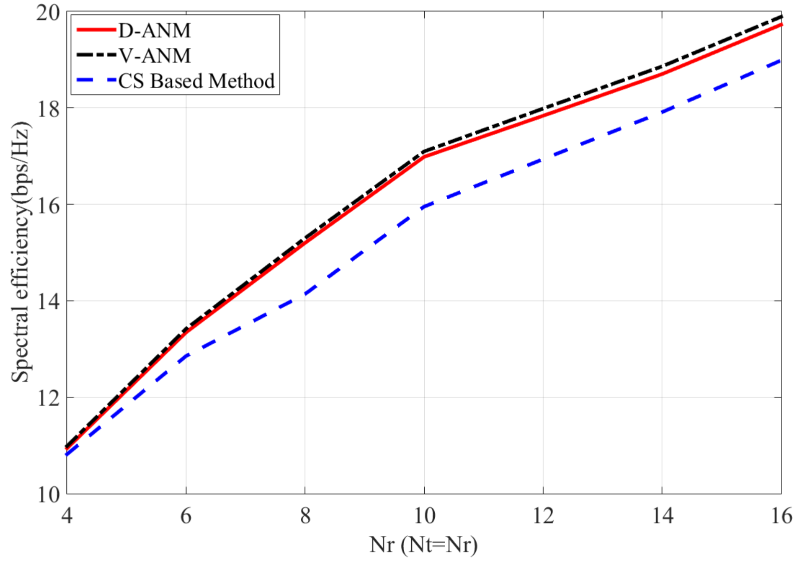

Figure 3.4: Spectral efficiency versus N_r ($N_r = N_t,\ L = 2,\ T = 6,\ \text{SNR} = 15\ dB$).

is the desired recovery precision [31]. In contrast, for D-ANM, the matrix's size shrinks to $(N_r+N_t)\times(N_r+N_t)$ only, resulting in a reduced time complexity as $\mathcal{O}((N_r+N_t)^{3.5}\log(1/\epsilon))$. When $N_r = N_t = N$, the computational load is reduced on the order of $\mathcal{O}(N^{3.5})$ from V-ANM to D-ANM. Thus, D-ANM enjoys huge complexity reduction over V-ANM, which is attractive for massive MIMO scenarios, especially when the performance gaps in estimation stage between D-ANM and V-ANM can be neglected in transmission stage in terms of spectral efficiency and BER performance. To summarize, the proposed D-ANM based channel estimation technique achieves better estimation performance at low computational complexity.

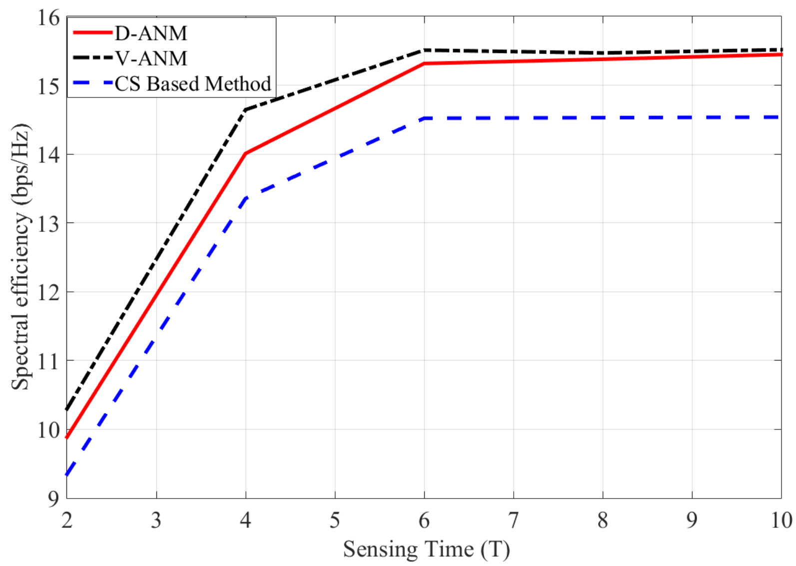

Figure 3.5: Spectral efficiency versus sensing time T ($N_r = N_t = 8,\ L = 2,\ \mathrm{SNR} = 15\ dB$).

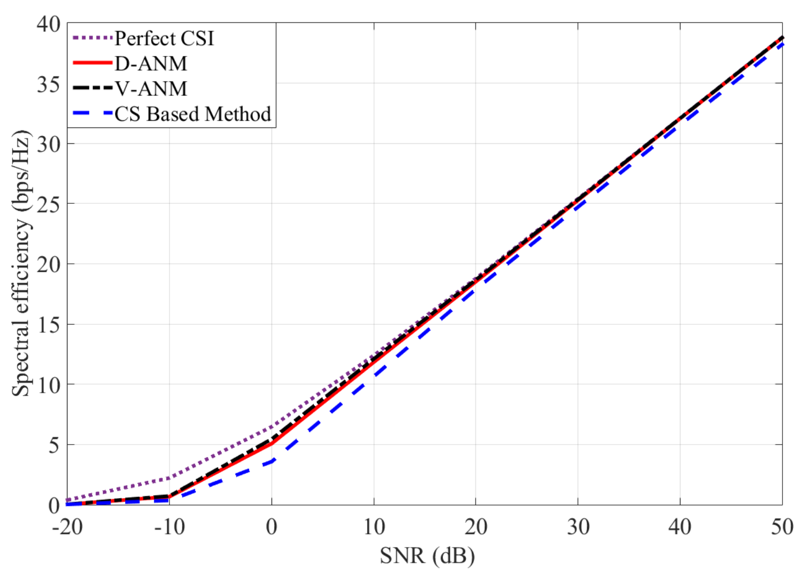

Figure 3.6: Spectral efficiency versus SNR ($N_r = N_t = 8,\ L = 2,\ T = 6$).

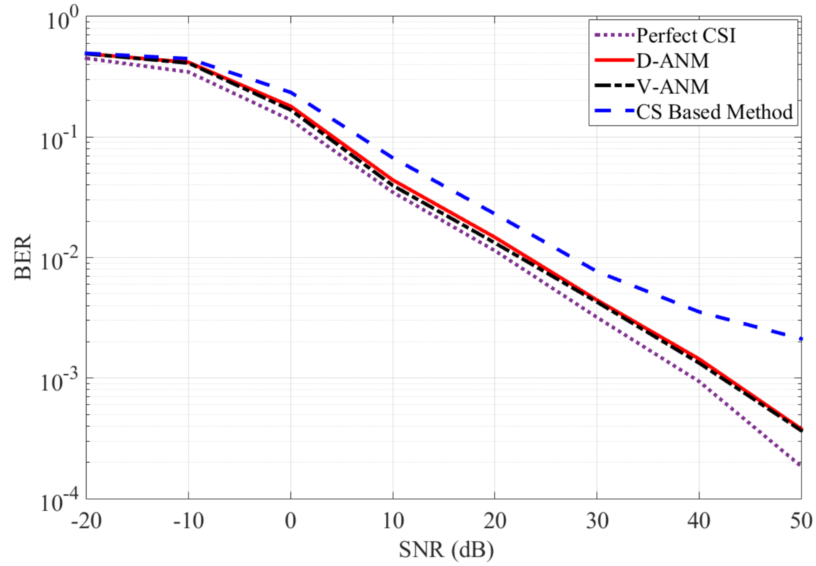

Figure 3.7: BER versus SNR $(N_r = N_t = 8,\ L = 2,\ T = 6)$.

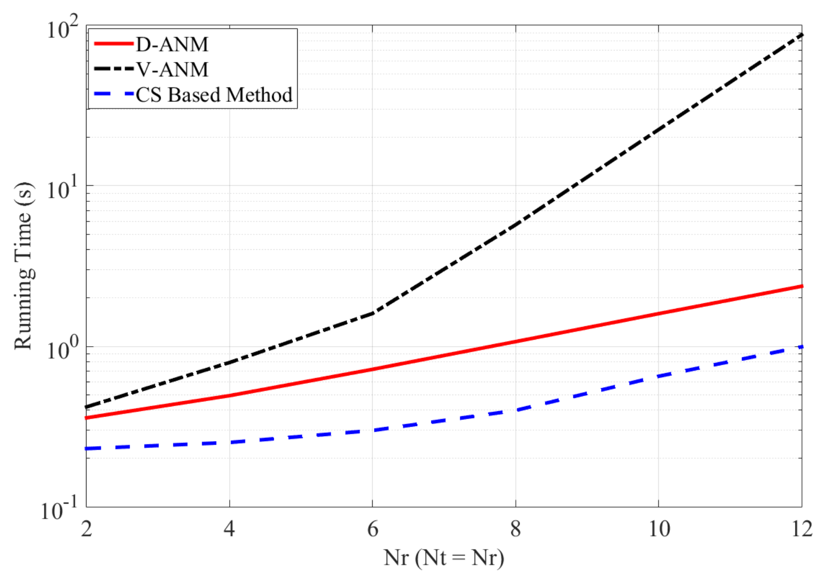

Figure 3.8: Computational complexity: Running time versus N_r $(N_r = N_t,\ L = 2,\ T = 6,\ \mathrm{SNR} = 15dB)$.

Chapter 4: Extensions

In this chapter, we apply D-ANM to two practical cases where either a uniform rectangular array (URA) is employed at the base station (BS) or multi-user interference is present in a mmWave massive MIMO system.

4.1 Uniform Rectangular Arrays

Consider a single-input multiple-output (SIMO) setting in which the BS employs an $M_r \times M_c$ ($M_r, M_c \gg 1$) URA. The uplink channel estimation is done by letting the BS collect and process the training data transmitted from a single-antenna mobile station (MS). Again, assume there exists only L dominant paths. Following [35], the channel model for a SIMO URA setting is formed as

$$\mathbf{H}_R = \sum_{l=1}^{L} s_l \mathbf{A}_R(\mathbf{\Phi}_l), \tag{4.1}$$

where the array steering matrix $\mathbf{A}_R(\mathbf{\Phi}_l)$ for the l-th path can be decoupled into column and row vectors as

$$\begin{aligned}
\mathbf{a}_c(\varphi_{c,l}) &= \frac{1}{\sqrt{M_c}}[1, e^{j2\pi \frac{d_c}{\lambda} \sin\varphi_{c,l}}, \ldots, e^{j2\pi \frac{(M_c-1)d_c}{\lambda} \sin\varphi_{c,l}}]^T, \\
\mathbf{a}_r(\varphi_{r,l}) &= \frac{1}{\sqrt{M_r}}[1, e^{j2\pi \frac{d_r}{\lambda} \sin\varphi_{r,l}}, \ldots, e^{j2\pi \frac{(M_r-1)d_r}{\lambda} \sin\varphi_{r,l}}]^T,
\end{aligned} \tag{4.2}$$

with $\sin\varphi_{c,l} = \sin\alpha_l$, $\sin\varphi_{r,l} = \sin\beta_l \cos\alpha_l$, α_l, β_l being the elevation and azimuth AoA, respectively, and d_c and d_r being the column and row antenna spacings, respectively.

Accordingly, we can define the atom set for the SIMO URA channel matrix $\mathbf{H}_R$ as

$$\begin{aligned}\mathcal{A}_R &= \{\mathbf{A}_R(\mathbf{\Phi}), \quad \mathbf{\Phi} \in [0, 2\pi) \times [0, 2\pi)\} \\ &= \left\{\mathbf{a}_c(\varphi_{c,l})\mathbf{a}_r^{\mathrm{H}}(\varphi_{r,l}), \quad \varphi_{c,l} \in [0, 2\pi), \varphi_{r,l} \in [0, 2\pi)\right\}.\end{aligned} \tag{4.3}$$

The system model for this scenario therefore is formulated as

$$\mathbf{Y}_i = \mathbf{H}_R x_i + \mathbf{W}_i, \quad i = 1, \dots, T, \tag{4.4}$$

where x_i is the pilot symbol sent from MS and $\mathbf{W}_i$ is the AWGN collected at the BS at time instant i. Correspondingly, the estimation of $\mathbf{H}_R$ by D-ANM can be formulated as

$$\begin{aligned}\min_{\mathbf{v}_c, \mathbf{v}_r, \mathbf{H}_R} \quad & \frac{\mu}{2}\Big(\mathrm{trace}\big(\mathbf{T}(\mathbf{v}_c)\big) + \mathrm{trace}\big(\mathbf{T}(\mathbf{v}_r)\big)\Big) \\ & + \sum_{i=1}^{T} \|\mathbf{Y}_i - \mathbf{H}_R x_i\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & \begin{pmatrix} \mathbf{T}(\mathbf{v}_r) & \mathbf{H}_R^{\mathrm{H}} \\ \mathbf{H}_R & \mathbf{T}(\mathbf{v}_c) \end{pmatrix} \succeq \mathbf{0},\end{aligned} \tag{4.5}$$

where $\mathbf{v}_c \in \mathbb{C}^{M_c}$ and $\mathbf{v}_r \in \mathbb{C}^{M_r}$, $\mathbf{T}(\mathbf{v}_c)$ and $\mathbf{T}(\mathbf{v}_r)$ are of the same structures as $\mathbf{T}(\mathbf{u}_r)$ and $\mathbf{T}(\mathbf{u}_t)$ defined in Section 3.4. In (4.5), the matrix-form array geometry is reflected through two small-size vectors $\mathbf{v}_c$ and $\mathbf{v}_r$ to reduce computational complexity, without loss of optimality. This is a key advantage of the proposed D-ANM method for channel estimation. In contrast, when angular information is concerned, traditional array processing techniques for URA arrays require the vectorization of the array manifold matrix while the D-ANM avoids vectorization, which not only drastically reduces the computational cost but also shortens the sensing time.

The pilot symbols x_i sent from transmitter is only a scalar at each time instant i for

$i = 1, \ldots, T$, however, (4.5) works well even for $T = 1$, meaning that the sensing time can be reduced to $T = 1$ with only one pilot symbol sent. This is because (4.5) effectively utilizes the structure of $\mathbf{H}_R$, which is a large channel matrix constructed from only a few angular parameters. Therefore, D-ANM provides efficient channel estimation technique for SIMO systems with URAs employed at the BS, with only one training symbol needed for estimation.

4.2 Multi-user Massive MIMO

Now consider an uplink channel estimation where a BS serves J ($J \geq 2$) MS's, each of which is equipped with a ULA. The system model extends that in (3.4) to

$$\mathbf{Y} = \sum_{j=1}^{J} \mathbf{H}_j \mathbf{X}_j + \mathbf{W}, \tag{4.6}$$

where $\mathbf{Y} \in \mathbb{C}^{N_r \times T}$ is the received training data collected at the BS during T sensing time slots, $\mathbf{H}_j \in \mathbb{C}^{N_{rj} \times N_{tj}}$ is the channel between j-th user and the BS which is also dominated by L_j paths, $\mathbf{X}_j \in \mathbb{C}^{N_{tj} \times T}$ is the training symbol sent from j-th user and $\mathbf{W} \in \mathbb{C}^{N_r \times T}$ is the AWGN.

As the channel in an mmWave massive MIMO system is very sparse, it enables a chance to estimate all channels $\mathbf{H}_j$ simultaneously with high accuracy. Given all pilot symbols $\mathbf{X}_j$ transmitted from J transmitters and the measurements $\mathbf{Y}$ collected from the receiver, the

BS estimates all $\mathbf{H}_j$ from $\mathbf{Y}$ as follows:

$$
\begin{aligned}
\min_{\{\mathbf{u}_{rj},\mathbf{u}_{tj},\mathbf{H}_j\}_{j=1}^{J}} \quad & \sum_{j=1}^{J} \frac{\mu}{2}\Big(\mathrm{trace}\big(\mathbf{T}(\mathbf{u}_{rj})\big) + \mathrm{trace}\big(\mathbf{T}(\mathbf{u}_{tj})\big)\Big) \\
& + \|\mathbf{Y} - \sum_{j=1}^{J} \mathbf{H}_j \mathbf{X}_j\|_{\mathrm{F}}^2 \\
\text{s.t.} \quad & \begin{pmatrix} \mathbf{T}(\mathbf{u}_{tj}) & \mathbf{H}_j^{\mathrm{H}} \\ \mathbf{H}_j & \mathbf{T}(\mathbf{u}_{rj}) \end{pmatrix} \succeq \mathbf{0}, \qquad j = 1, \ldots, J.
\end{aligned} \tag{4.7}
$$

Each channel $\mathbf{H}_j$ holds the same structure as $\mathbf{H}$ defined in Section 3.1 for $j = 1, \ldots, J$, therefore (4.7) adequately captures the structured information for channel estimation. From the formulation we can see that the estimation accuracy is affected by the number of users in the system, the inter-user interference and the length of training block T.

4.3 Simulation Results

For the mmWave wireless system that employs an URA at the BS, we simulate its spectral efficiency using the channel estimated by D-ANM, and compare it with the ideal case where the perfect CSI is known. The URA is composed of $M_c = M_r = 8$ antennas and the number of dominant paths in the channel is set to be $L = 2$. The number of training symbols sent at the receiver side at each time instant is only one. Still, D-ANM achieves good estimation performance, as depicted by Figure 4.1.

For the multi-user system with J users, we set $N_{rj} = N_r$, $N_{tj} = N_t$, $L_j = L$, $\forall j$, and $T = 6$. We test the cases for $J = 2, 3, 4$, and measure their spectral efficiencies and BER performance, shown in Figure 4.2 and Figure 4.3, respectively. The gap on spectral efficiencies between D-ANM and the perfect CSI increases when the number of users goes large. This is because the number of unknowns to be estimated in D-ANM increases as the number of accessing users increases, resulting in a decrease in estimation accuracy, which

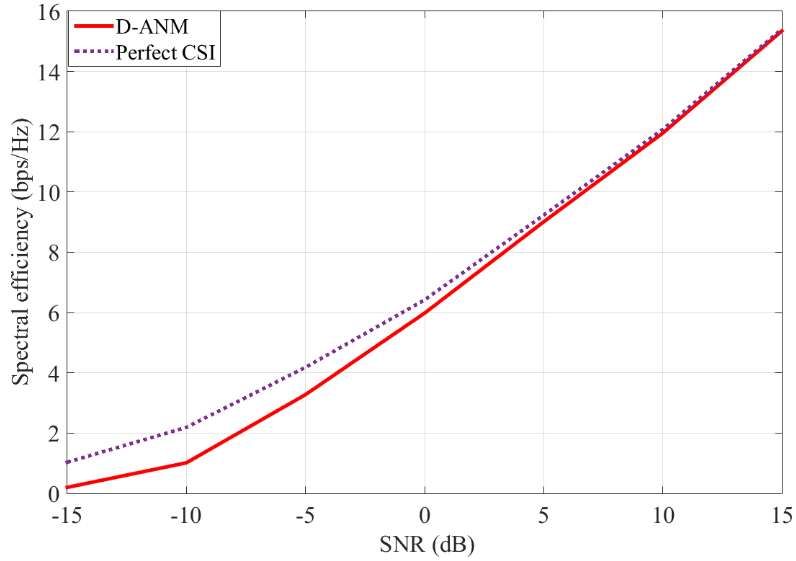

Figure 4.1: Spectral efficiency versus SNR for URA $(Mc = Mr = 8,\ L = 2)$.

is also corroborated by BER performance. The estimation accuracy can be improved by increasing either the antenna size or sensing time. Overall, the modified D-ANM approach is appealing for mmWave massive MIMO multi-user systems.

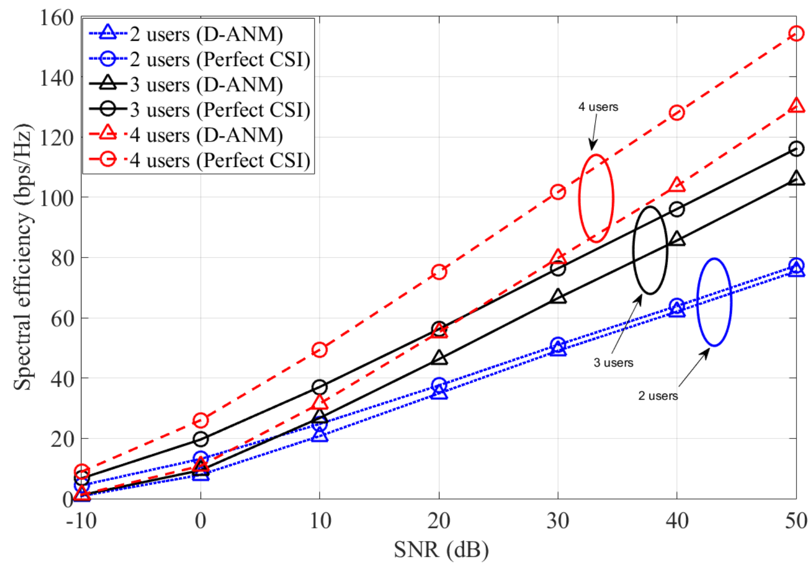

Figure 4.2: Spectral efficiency versus SNR for multi-user $(Nr = Nt = 8,\ L = 2,\ T = 6,\ \text{SNR} = 15dB)$.

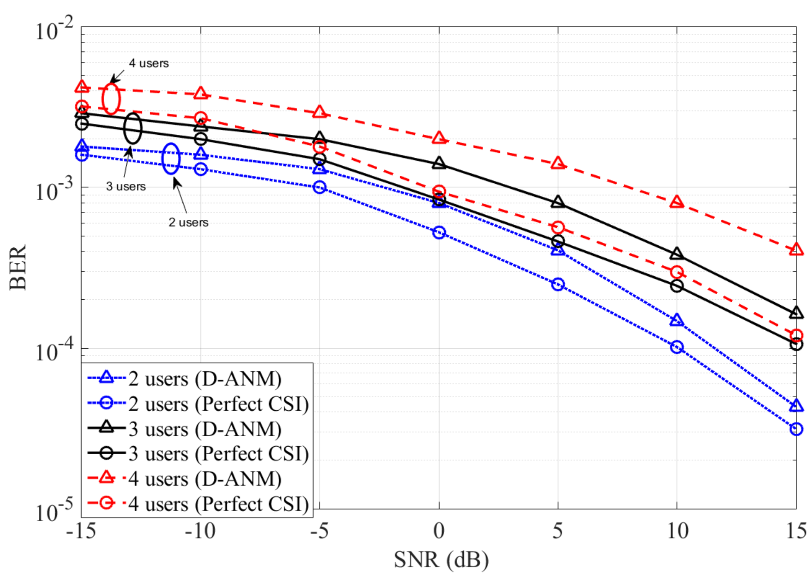

Figure 4.3: BER versus SNR for multi-user $(Nr = Nt = 8,\ L = 2,\ T = 6,\ \text{SNR} = 15dB)$.

Chapter 5: D-ANM for Frequency-selective SIMO-OFDM Systems

5.1 Background and Motivation

Previous chapters discuss the channel estimation for narrowband mmWave massive MIMO/SIMO systems, however in reality, mmWave channels are wideband and frequency-selective with severe path-loss [9]. To combat the severe path-loss and achieve reliable communications, beamforming with large-scale antenna arrays is inevitable, which requires the knowledge of AoA or AoD [36]. As implementing beamforming with large-scale antenna arrays in BS's is much easier than that in user equipments (UEs) and most often users may not want to carry so many antennas, therefore, in this work, we consider a wide-band SIMO system where the large-scale antenna array in the form of ULA is only employed at the BS. While the channel between the BS and UE is time-dispersive, resulting in frequency-selective fading, we adopt the orthogonal frequency division multiplexing (OFDM) technique to convert the frequency-selective fading into flat fading.

Again, to utilize the potential performance gain of mmWave large-scale antenna systems, CSI has to be obtained accurately by doing channel estimation. However, traditional channel estimation approaches can not be directly applied to SIMO-OFDM systems due to the heavy training overhead and high computational complexity caused by the large number of antennas [37]. By adopting the usual assumption in most literatures that each scatterer in an mmWave channel contributes to only one single path which is parameterized by its delay and AoA, CS-based methods can be applied for SIMO-OFDM channel estimation. While CS-based methods succeed in estimating the channel with much less pilot overhead or obtaining higher accuracy with a constant number of pilots compared with traditional

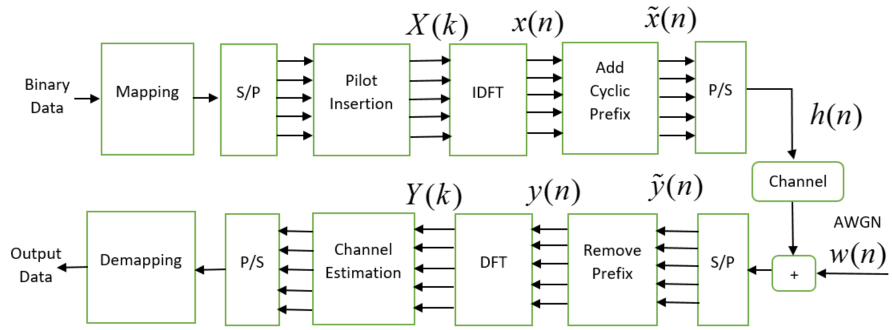

Figure 5.1: Baseband OFDM.

methods [36, 38, 39], they suffer from considerable performance degradation in practice due to their on-grid assumptions, which in the SIMO-OFDM setup means that the values of delays and AoAs have to lie on some predefined grids [23].

Motivated by the framework of D-ANM for 2D mmWave channel estimation with high accuracy at reduced complexity and little training overhead, we propose a D-ANM based approach for SIMO-OFDM systems after observing that Vandermonde structures are incorporated in both spatial dimension (resulted from ULA) and frequency dimension (resulted from OFDM). Before introducing the estimation scheme, fundamentals of OFDM systems are first described.

5.2 Description of OFDM Systems

Orthogonal frequency division multiplexing (OFDM) technique is usually adopted to convert a frequency-selective channel into a parallel collection of frequency-flat channels by dividing the entire channel into a group of narrow subchannels [40]. Moreover, when utilizing cyclic prefix (CP), which is done by extending an OFDM symbol with some portion of its head or tail, intersymbol interference (ISI) is avoided. A block diagram of a baseband OFDM system is shown in Figure 5.1.

Suppose that the OFDM system divides the channel into N parallel subchannels. The binary information-bearing data are coded and modulated, after which an inverse discrete Fourier transform (IDFT) is done to obtain the time domain OFDM symbol $\{x(n)\}$ of length N, where

$$\begin{aligned} x(n) &= IDFT\{X(k)\} \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi kn/N}, \ n = 0, \ldots, N-1. \end{aligned} \tag{5.1}$$

The CP of length N_{cp} is added to $x(n)$ to eliminate ISI, resulting in the time sample $\tilde{x}(n)$, where

$$\tilde{x}(n) = \begin{cases} x(N+n), & n = -N_{cp}, -N_{cp}+1, \ldots, -1 \\ x(n), & n = 0, 1, \ldots, N-1 \end{cases}. \tag{5.2}$$

The transmitted signal $\tilde{x}(n)$ then passes through the frequency-selective channel with additive noise, resulting in measurements $\tilde{y}(n)$:

$$\tilde{y}(n) = \tilde{x}(n) * h(n) + w(n), \ -N_{cp} \leq n \leq N-1, \tag{5.3}$$

where $w(n)$ and $h(n)$ are the AWGN and channel impulse response (CIR) between the transmit antenna and receive antenna, respectively. The CIR is modeled as

$$h(n) = \sum_{l=1}^{L} \alpha_l \delta(n - \tau_l), \ 0 \leq n \leq N-1, \tag{5.4}$$

where L is the number of paths, α_l and τ_l are the complex path gain and normalized propagation delay of l-th path, respectively.

At the receiver, the prefix of $\tilde{y}$ consisting of N_{cp} samples is removed, resulting in a received data sequence of length N:

$$y(n) = \tilde{y}(n + N_{cp}),\ n = 0, 1 \ldots, N - 1. \tag{5.5}$$

Afterwards, $y(n)$ is sent to DFT block to obtain the frequency domain data:

$$\begin{aligned} Y(k) &= DFT\{y(n)\} \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y(n) e^{-j2\pi kn/N}, \\ k &= 0, \ldots, N-1. \end{aligned} \tag{5.6}$$

Theoretically, ISI is totally eliminated. With $H(k) = DFT\{h(n)\}$ be the N-point DFT channel frequency response (CFR) of $h(n)$, the relationship between frequency domain data $Y(k)$ and CFR $H(k)$ in the absence of noise is given in [41], which is

$$Y(k) = H(k)X(k),\ k = 0, \ldots, N - 1. \tag{5.7}$$

With AWGN $w(n)$, and $W(k) = DFT\{w(n)\}$, the received signal in frequency domain is of the form

$$Y(k) = H(k)X(k) + W(k),\ k = 0, \ldots, N - 1. \tag{5.8}$$

In compact vector-matrix form, the received frequency domain sample of all N subcarriers is

$$\mathbf{Y} = \text{diag}(\mathbf{X})\mathbf{H} + \mathbf{W}, \tag{5.9}$$

where $\mathbf{X} = [X(0), \ldots, X(N-1)]$, $\mathbf{H} = [H(0), \ldots, H(N-1)]^T$ and $\mathbf{W} = [W(0), \ldots, W(N-1)]^T$ are row vectors. Note that both $\mathbf{X}$ and $\mathbf{W}$ are frequency domain data. Traditional channel estimation approaches such as least square (LS) estimator, minimum mean square error (MMSE) estimator are usually done in frequency domain with pilot aided.

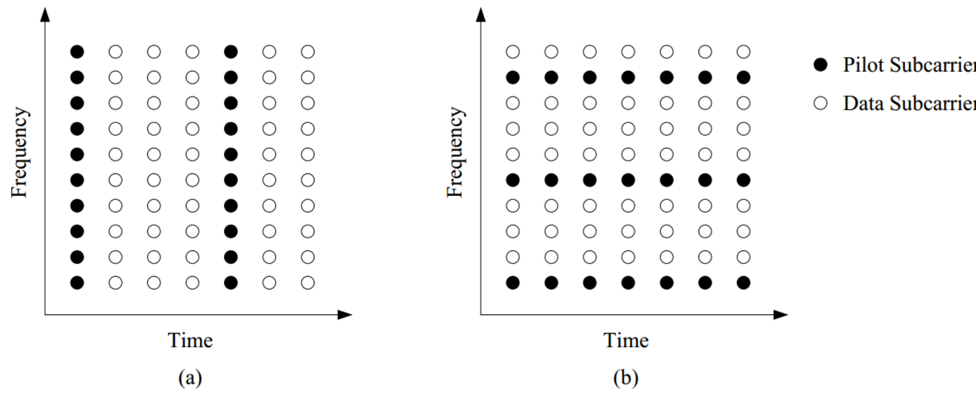

Figure 5.2: Pilot arrangements: (a) block-type pilot arrangement and (b) comb-type pilot arrangement.

To insert pilot symbols, there are two different arrangements, the block-type pilot arrangement and the comb-type pilot arrangement, as shown in Figure 5.2. Block-type channel estimation usually assumes slow-fading, meaning that the channel is constant over one or more OFDM symbol periods, which is not true in practice for mmWave channels. Therefore in this work, to meet the need for channel equalization or tracking in fast fading scenarios, we adopt the comb-type pilot arrangement to do channel estimation, in which the pilot symbols are multiplexed with the data within an OFDM symbol. The main idea in comb-type channel estimation is to estimate the channel conditions at the pilot subcarriers first and then estimate the overall channel by means of interpolation.

As the estimation performance using comb-type pilot arrangement is directly affected by the number and/or locations of pilot subcarriers used for the initial estimation [42], therefore, to obtain an accurate estimation, the number of pilot subcarriers needs to be high, making traditional LS estimator and MMSE estimator suffer from high computational complexity.

5.3 Channel Model and Problem Formulation

Assume that the time dispersive channel to be estimated is composed of L scatters and each scatter contributes to one single path parameterized by its delay and AoA. The time-domain CIR for the receive antennas, denoted by $\mathbf{h}(n)$ is given by [41]:

$$\mathbf{h}(n) = \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l)\delta(n - \tau_l),\ 0 \le n \le N-1 \tag{5.10}$$

where $\mathbf{h}(n) = [h_1(n), \ldots, h_{N_r}(n)]^T$, α_l and τ_l are the complex path gain and normalized propagation delay of l-th path, respectively. The receive array response vector of l-th path associated with corresponding AoA θ_l is by $\mathbf{a}(\theta_l) = [1, e^{j2\pi \sin\theta_l d/\lambda}, \ldots, e^{j2\pi \sin\theta_l (N_r-1)d/\lambda}]^T$, where d and λ are the antenna spacing and the wavelength, respectively.

The CFR at k-th subcarrier is given by

$$\begin{aligned} \mathbf{H}(k) &= \mathrm{DFT}\{\mathbf{h}(n)\} \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l)\delta(n-\tau_l) e^{-j2\pi kn/N} \\ &= \frac{1}{\sqrt{N}} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l) e^{-j2\pi k\tau_l/N} \\ k &= 0, \ldots, N-1. \end{aligned} \tag{5.11}$$

Stack the CFR of all N subcarriers into a matrix gives the expression of $\mathbf{H} \in \mathbb{C}^{N_r \times N}$ in

(5.11) as

$$
\begin{aligned}
\mathbf{H} &= [\mathbf{H}(0), \ldots, \mathbf{H}(N-1)] \\
&= \frac{1}{\sqrt{N}} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l) \mathbf{e}^T(\tau_l) \\
&= \mathbf{A}(\boldsymbol{\theta}) \mathbf{S} \mathbf{E}^T(\boldsymbol{\tau}) \\
&= \sum_{l=1}^{L} s_l \mathbf{A}_{\mathbf{E}}(\boldsymbol{f}_l)
\end{aligned}
, \tag{5.12}
$$

where $\mathbf{e}(\tau_l) = [1, e^{j2\pi\tau_l/N}, \ldots, e^{j2\pi(N-1)\tau_l/N}]^T$, $s_l = \frac{1}{\sqrt{N}}\alpha_l$, $\mathbf{E}(\boldsymbol{\tau}) = [\mathbf{e}(\tau_1), \ldots, \mathbf{e}(\tau_L)]$, $\mathbf{S} = \mathrm{diag}[s_1, \ldots, s_L]$, $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_L)]$, $\mathbf{A}_{\mathbf{E}}(\boldsymbol{f}_l) = \mathbf{a}(\theta_l)\mathbf{e}^T(\tau_l)$.

Comb-type pilots are used for channel estimation, but we only consider transmitting one OFDM symbol at each time instant. The OFDM symbol before CP addition is composed of N_p pilot symbols and $N - N_p$ information bearing symbols. The N_p pilot symbols are transmitted over N_p uniformly separated subcarriers whose indices are in the set $\mathcal{P} := \{p_1, p_2, \ldots, p_{N_p}\}$ with $p_i = (i-1)M + 1$, where $M = N/N_p$. The transmitted pilot vector and received pilot sample vector are $\mathbf{X}_{\mathcal{P}} = [X(p_1), X(p_2), \ldots, X(p_{N_p})]^T$ and $\mathbf{Y}_{\mathcal{P}} = [\mathbf{Y}(p_1), \mathbf{Y}(p_2), \ldots, \mathbf{Y}(p_{N_p})]$, respectively and they are related by

$$
\mathbf{Y}_{\mathcal{P}} = \mathbf{H}_{\mathcal{P}} \mathrm{diag}(\mathbf{X}_{\mathcal{P}}) + \mathbf{W}_{\mathcal{P}}, \tag{5.13}
$$

where $\mathbf{W}_{\mathcal{P}} = [\mathbf{W}(p_1), \mathbf{W}(p_2), \ldots, \mathbf{W}(p_{N_p})] \in \mathbb{C}^{N_r \times N_p}$ is the AWGN and the CFR $\mathbf{H}_{\mathcal{P}} =$

$[\mathbf{H}(p_1), \mathbf{H}(p_2), \ldots, \mathbf{H}(p_{N_p})] \in \mathbb{C}^{N_r \times N_p}$ of N_p subcarriers is of the form

$$\begin{aligned}\mathbf{H}_{\mathcal{P}} &= \frac{1}{\sqrt{N}} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l) \mathbf{e}_{\mathcal{P}}^T(\tau_l) \\ &= \mathbf{A}(\boldsymbol{\theta}) \mathbf{S} \mathbf{E}_{\mathcal{P}}^T(\boldsymbol{\tau}) \qquad\qquad , \\ &= \sum_{l=1}^{L} s_l \mathbf{A}_{\mathcal{P}}(\boldsymbol{f}_l)\end{aligned} \tag{5.14}$$

where $\mathbf{e}_{\mathcal{P}}(\tau_l) = [e^{j2\pi(p_1-1)\tau_l/N}, \ldots, e^{j2\pi(p_{N_p}-1)\tau_l/N}]^T$ and $\mathbf{E}_{\mathcal{P}}(\boldsymbol{\tau}) = [\mathbf{e}_{\mathcal{P}}(\tau_1), \ldots, \mathbf{e}_{\mathcal{P}}(\tau_L)]$.

To estimate the CFR in the form of matrix $\mathbf{H}$, the CFR at subcarriers $\mathcal{P}$ in the form of $\mathbf{H}_{\mathcal{P}}$ needs to be estimated from $\{\mathbf{Y}_{\mathcal{P}}, \mathbf{X}_{\mathcal{P}}\}$ first, where $\mathbf{Y}_{\mathcal{P}}$ is corrupted by noise $\mathbf{W}_{\mathcal{P}}$, then $\mathbf{H}$ can be obtained through interpolation.

5.4 D-ANM based Channel Estimation

Following the frame work of D-ANM, we define the atom set for the SIMO-OFDM pilot symbol incorporated channel $\mathbf{H}_{\mathcal{P}}$ as

$$\begin{aligned}\mathcal{A}_{\mathcal{P}} &= \{\mathbf{A}_{\mathcal{P}}(\boldsymbol{f}), \quad \boldsymbol{f} \in [0, 2\pi) \times [0, \tau_{max})\} \\ &= \left\{\mathbf{a}(\theta_l)\mathbf{e}_{\mathcal{P}}^T(\tau_l), \quad \theta_l \in [0, 2\pi), \tau_l \in [0, \tau_{max})\right\}.\end{aligned} \tag{5.15}$$

Note that the channel in equation (5.12) and (5.14) both have the 2-D Vandermonde structures, therefore, we can adopt (3.21) to estimate the channel through pilots:

$$\begin{aligned}\min_{\mathbf{u}_r,\mathbf{u}_\tau,\mathbf{H}_\mathcal{P}} \quad & \frac{\mu}{2}\Big(\mathrm{trace}(\mathbf{T}(\mathbf{u}_r)) + \mathrm{trace}(\mathbf{T}(\mathbf{u}_\tau))\Big) \\ & + \|\mathbf{Y}_\mathcal{P} - \mathbf{H}_\mathcal{P}\mathrm{diag}(\mathbf{X}_\mathcal{P})\|_\mathrm{F}^2 \\ \text{s.t.} \quad & \begin{pmatrix} \mathbf{T}(\mathbf{u}_\tau) & \mathbf{H}_\mathcal{P}^\mathrm{H} \\ \mathbf{H}_\mathcal{P} & \mathbf{T}(\mathbf{u}_r) \end{pmatrix} \succeq \mathbf{0}. \end{aligned} \tag{5.16}$$

The estimation of $\mathbf{H}_\mathcal{P}$ is done after collecting all frequency domain data, that is, after an DFT operation on the time domain measurements $y(n)$. Once $\mathbf{H}_\mathcal{P}$ is estimated, $\mathbf{H}$ can be obtained through the following operation:

$$\mathbf{H} = DFT_N\{IDFT_{N_p}\{\mathbf{H}_\mathcal{P}\}\}. \tag{5.17}$$

By applying D-ANM to SIMO-OFDM systems, basis mismatch problem that exists in CS techniques can be avoided. The continuously-valued AoA and time delay can be obtained easily using matrix pencil and pairing (MaPP) [31]. The angular information can be utilized for beamforming in large-scale antenna arrays.

5.5 Simulation Results

In the SIMO-OFDM system, we consider deploying $N_r = 8$ antennas in the form of ULA at the BS and dividing the wideband carrier into $N = 64$ narrowband subcarriers using OFDM. The number of pilots used for estimation is $N_{cp} = 8$, transmitted on the subcarriers of indices {1, 9, 17, 25, 33, 41, 49, 57}. We adopt a BPSK modulation and randomly generate AoA and time delays in the range of $[0, 2\pi]$ and $[0, \tau_{max})$. The overall channel $\mathbf{H}$ is obtained using (5.17) after obtaining the estimates of $\mathbf{H}_\mathcal{P}$. To evaluate the performance of D-ANM, we test the MSE and spectral efficiency versus SNR, shown in Figure 5.3 and Figure 5.4, respectively.

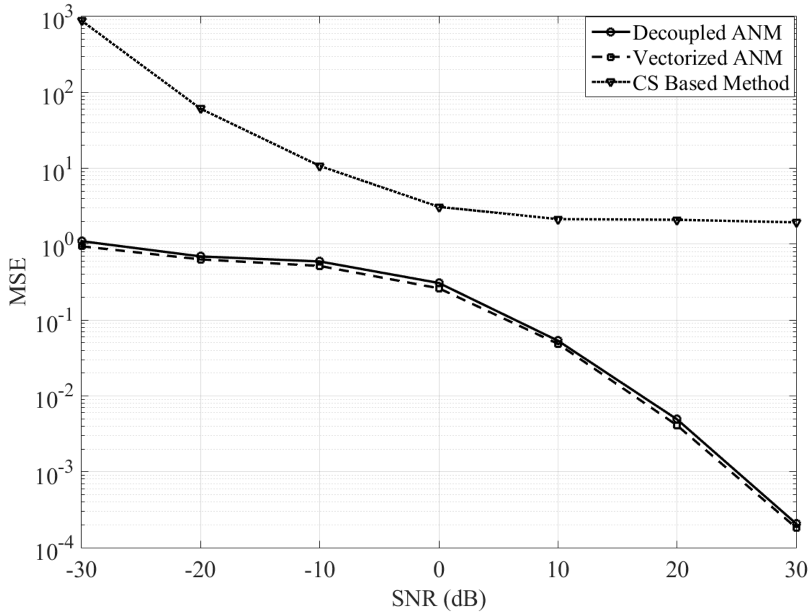

Figure 5.3: MSE versus SNR ($N_r = 8,\ L = 2,\ N = 64,\ N_{cp} = 8$).

From the simulations, we can see that D-ANM achieves better performance than CS-based techniques, as the parameters that characterize the channel are continuously-valued and off-the-grid. Further work will be done to exploit the effects of pilot locations on estimation performance.

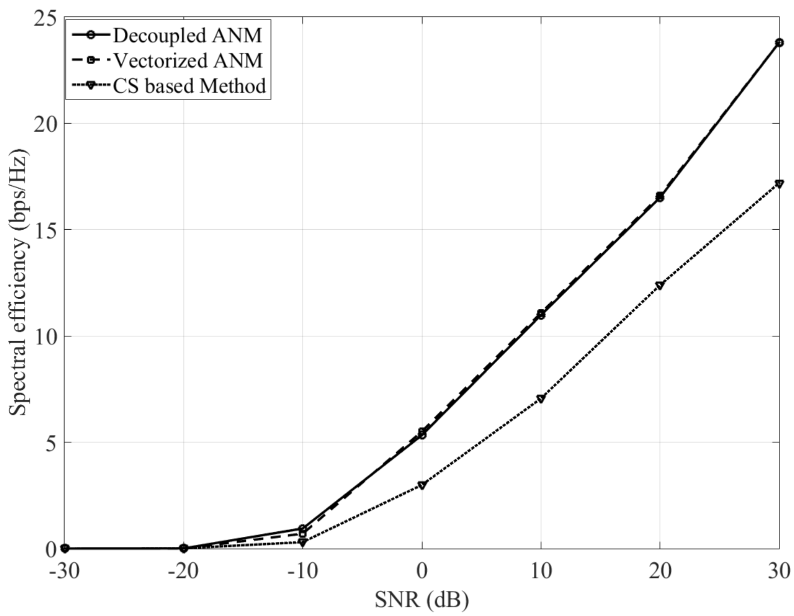

Figure 5.4: Spectral efficiency versus SNR ($N_r = 8,\ L = 2,\ N = 64,\ N_{cp} = 8$).

Chapter 6: Conclusions and Future Works

6.1 Conclusions

In this thesis, by leveraging the framework of D-ANM for 2D harmonic retrieval and utilizing the sparsity feature of mmWave channels, we develop a low-complexity high-accuracy channel estimation approach for mmWave large-scale antenna array systems. The proposed method retains the super-resolution property provided by ANM, and outperforms the state-of-the-art CS-based methods in both estimation performance and training resource saving. Meanwhile, through decoupling, D-ANM achieves good estimation performance at much lower computational complexity. As a result, it is very attractive in offering a desired trade-off between estimation performance and computational complexity for practical mmWave massive MIMO systems. We also apply the proposed algorithm to two practical scenarios, including the URA antenna implementation and the multi-user case. Simulations show that D-ANM has good performance for both cases. Finally, we consider frequency-selective channels and develop a pilot-aided channel estimation approach based on D-ANM for the SIMO-OFDM systems.

6.2 Future Works

There are several possible directions for future research:

1. In Chapter 3, we only estimated a 2D mmWave channel, with the assumption that the channel experiences frequency flat fading and there is no multipath delay in the system. However, in practice, the wireless environment is more complex, therefore, it would be both meaningful and interesting if we can extend D-ANM to higher dimensions and take time delays and Doppler spread into considerations.

2. Throughout the thesis, we focused on estimating digital channels at low complexities, without considering the hardware constraints. However, it is impossible to realize the digital baseband precoding and channel estimation in massive MIMO systems due to the high power consumption caused by the large antenna size. Therefore, we need to consider hybrid analog/digital architectures for implementation and develop more practical channel estimation algorithms.

3. The channel estimation step occupies resources for transmission, therefore, it may be of significance if non-coherent detection can be done with desired performance. The non-coherent detection may be realized by adopting multiple symbol differential detection (MSDD) scheme.

Appendix A: Proof of Proposition 2.2.1

Denote the value of the right hand side of equation (2.8) as SDP($\mathbf{x}$). Let $\mathbf{x} = \sum_l s_l \mathbf{a}(f_l)$, where $s_l = |s_l|e^{j\theta_l}$. Define $\mathbf{u} = \sum_l |s_l|\mathbf{a}(f_l)$, $\upsilon = \sum_l |s_l|$, and

$$\text{Toep}(\mathbf{u}) = \sum_l |s_l|\mathbf{a}(f_l)\mathbf{a}^H(f_l).$$

Therefore,

$$\begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & \upsilon \end{bmatrix} = \sum_l |s_l| \begin{bmatrix} \mathbf{a}(f_l) \\ e^{-j\theta_l} \end{bmatrix} \begin{bmatrix} \mathbf{a}(f_l) \\ e^{-j\theta_l} \end{bmatrix}^H \succeq \mathbf{0}, \tag{A.1}$$

indicating SDP($\mathbf{x}$) $\leq \frac{1}{2M}\text{trace}(\text{Toep}(\mathbf{u})) + \frac{1}{2}\upsilon = \sum_l |s_l| = ||\mathbf{x}||_{\mathcal{A}}$.

Conversely, suppose for some $\mathbf{u}$ and $\mathbf{x}$ satisfying

$$\begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^H & \upsilon \end{bmatrix} \succeq \mathbf{0}, \tag{A.2}$$

then we have $\text{Toep}(\mathbf{u}) \succeq \mathbf{0}$ and $\text{Toep}(\mathbf{u}) \succeq \frac{1}{\upsilon}\mathbf{x}\mathbf{x}^H$ by Schur complement condition. Form a Vandermonde decomposition

$$\text{Toep}(\mathbf{u}) = \mathbf{V}\mathbf{D}\mathbf{V}^H,$$

where $\mathbf{V} = [\mathbf{a}(f_1), ..., \mathbf{a}(f_r)], \mathbf{D} = \text{diag}([d_1, ..., d_r])$ with d_l's being real and positive values and $r = \text{rank}(\text{Toep}(\mathbf{u}))$. Since $\mathbf{V}\mathbf{D}\mathbf{V}^H = \sum_l d_l \mathbf{a}(f_l)\mathbf{a}^H(f_l)$ and $\|\mathbf{a}(f_l)\| = \sqrt{M}$, we have $\frac{1}{M}\text{trace}(\text{Toep}(\mathbf{u})) = \text{trace}(\mathbf{D})$.

Using this Vandermonde decomposition and the matrix inequality in equation(A.2), then it follows that $\mathbf{x}$ falls within the column space of Toep($\mathbf{u}$), or equivalently, $\mathbf{x} = \sum_l w_l \mathbf{a}(f_l) = \mathbf{V}\mathbf{w}$ for some vector $\mathbf{w} = [\cdots w_l \cdots]^T$. Let $\mathbf{q}$ be any vector such that $\mathbf{V}^H\mathbf{q} = \text{sign}(\mathbf{w})$, where

$\mathrm{sign}(\mathbf{w})$ is the sign vector of $\mathbf{w}$, then $\mathrm{trace}(\mathbf{D}) = \mathbf{q}^H\mathbf{V}\mathbf{D}\mathbf{V}^H\mathbf{q} \geq \frac{1}{\upsilon}\mathbf{V}\mathbf{w}\mathbf{w}^H\mathbf{V} = \frac{1}{\upsilon}(\sum_l |w_l|)^2$, which implies that

$$\begin{aligned}\frac{1}{2M}\mathrm{trace}(\mathrm{Toep}(\mathbf{u})) + \frac{1}{2}\upsilon &= \frac{1}{2}\mathrm{trace}(\mathbf{D}) + \frac{1}{2}\upsilon \\ &\geq \sqrt{\mathrm{trace}(\mathbf{D})\upsilon} \geq \sum_l |w_l| = ||\mathbf{x}||_{\mathcal{A}},\end{aligned}$$

equivalently, $\mathrm{SDP}(\mathbf{x}) \geq ||\mathbf{x}||_{\mathcal{A}}$. Therefore we conclude $\mathrm{SDP}(\mathbf{x}) = ||\mathbf{x}||_{\mathcal{A}}$.

Bibliography

Bibliography

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[2] K. T. Truong and R. W. Heath, "Effects of Channel Aging in Massive MIMO Systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, 2013.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[4] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and Capacity of Millimeter-wave Cellular Networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.

[5] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five Disruptive Technology Directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[6] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications.* Prentice Hall, 2015.

[7] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave Cellular Wireless Networks: Potentials and Challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.

[8] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[9] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[10] J. Wang, "Beam Codebook based Beamforming Protocol for Multi-Gbps Millimeter-wave WPAN Systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, 2009.

[11] Y. M. Tsang, A. S. Poon, and S. Addepalli, "Coding the Beams: Improving Beamforming Training in Mmwave Communication System," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE.* IEEE, 2011, pp. 1–6.

[12] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[13] A. Alkhateeb, O. El 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.

[14] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental Study of MIMO Channel Statistics and Capacity via the Virtual Channel Representation," *IEEE Transactions on Wireless Communications*, 2007.

[15] H. Zhang, S. Venkateswaran, and U. Madhow, "Channel Modeling and MIMO Capacity for Outdoor Millimeter Wave Links," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*. IEEE, 2010, pp. 1–6.

[16] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband Millimeter-wave Propagation Measurements and Models using Adaptive-beam Antennas for Outdoor Urban Cellular Communications," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1850–1859, 2013.

[17] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[18] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured Compressive Sensing-based Spatio-temporal Joint Channel Estimation for FDD Massive MIMO," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 601–617, Feb. 2016.

[19] P. Schniter and A. Sayeed, "Channel Estimation and Precoder Design for Millimeter-wave Communications: The Sparse Way," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 273–277.

[20] X. Rao and V. K. N. Lau, "Distributed Compressive CSIT Estimation and Feedback for FDD Multi-user Massive MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.

[21] Z. Zhou and J. Fang, "Channel Estimation for Millimeter Wave MIMO-OFDM Systems via Low-Rank Tensor Decomposition," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7501–7516, Nov. 2016.

[22] T. Kim and D. J. Love, "Virtual AoA and AoD Estimation for Sparse Millimeter-wave MIMO Channels," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2015, pp. 146–150.

[23] Y. Wang, Z. Tian, S. Feng, and P. Zhang, "A Fast Channel Estimation Approach for Millimeter-wave Massive MIMO Systems," in *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*. IEEE, 2016, pp. 1413–1417.

[24] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to Basis Mismatch in Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[25] Z. Tan, P. Yang, and A. Nehorai, "Joint Sparse Recovery Method for Compressed Sensing with Structured Dictionary Mismatches," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4997–5008, Oct. 2014.

[26] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic Norm Denoising with Applications to Line Spectral Estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, 2013.

[27] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed Sensing Off the Grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[28] Y. Chi and Y. Chen, "Compressive Two-Dimensional Harmonic Retrieval via Atomic Norm Minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.

[29] 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.

[30] Y. Wang, P. Xu, and Z. Tian, "Efficient Channel Estimation for Massive MIMO Systems via Truncated Two-dimensional Atomic Norm Minimization," in *Communications (ICC), 2017 IEEE International Conference on.* IEEE, 2017, pp. 1–6.

[31] Z. Tian, Z. Zhang, and Y. Wang, "Low-complexity Optimization for Two-dimensional Direction-of-arrival Estimation via Decoupled Atomic Norm Minimization," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on.* IEEE, 2017, pp. 3071–3075.

[32] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The Convex Geometry of Linear Inverse Problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, Oct. 2012.

[33] "CVX: Matlab Software for Disciplined Convex Programming | CVX Research, Inc." [Online]. Available: http://cvxr.com/cvx/

[34] D. L. Donoho, "Compressed Sensing," *IEEE Transcations on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[35] G. Sommerkorn, D. Hampicke, R. Klukas, A. Richter, A. Schneider, and R. Thomä, "Uniform Rectangular Antenna Array Design and Calibration Issues for 2-D ESPRIT Application," in *Proceedings of the 4th European Personal Mobile Communications Conference*, 2001.

[36] C.-Y. Chen, W.-R. Wu, and C.-S. Gau, "Joint AoA and Channel Estimation for SIMO-OFDM Systems: A Compressive-sensing Approach," *Journal of Signal Processing Systems*, vol. 83, no. 2, pp. 191–205, 2016.

[37] J.-C. Shen, J. Zhang, K.-C. Chen, and K. B. Letaief, "High-dimensional CSI Acquisition in Massive MIMO: Sparsity-inspired Approaches," *IEEE Systems Journal*, vol. 11, no. 1, pp. 32–40, 2017.

[38] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, “Application of Compressive Sensing to Sparse Channel Estimation,” *IEEE Communications Magazine*, vol. 48, no. 11, 2010.

[39] P. Cheng, Z. Chen, Y. Rui, Y. J. Guo, L. Gui, M. Tao, and Q. Zhang, “Channel Estimation for OFDM Systems over Doubly Selective Channels: A Distributed Compressive Sensing based Approach,” *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4173–4185, 2013.

[40] Y. Shen and E. Martinez, “Channel Estimation in OFDM Systems,” *Freescale Semiconductor*, vol. 2, no. 6, 2006.

[41] A. Goldsmith, *Wireless Communications.* Cambridge university press, 2005.

[42] M. K. Ozdemir and H. Arslan, “Channel Estimation for Wireless OFDM Systems,” *IEEE Communications Surveys and Tutorials*, vol. 9, no. 2, pp. 18–48, 2007.

Curriculum Vitae

Ping Xu received her B.S. (with highest honors) from Northwestern Polytechnical University, China, in 2015. She has been working on the project of designing receivers for next generation wireless communications from Summer 2016 to Fall 2017. She is currently working towards the Ph.D. degree at George Mason University under the supervision of Dr. Zhi Tian.