Downlink Achievable Rate Analysis for FDD Massive MIMO Systems

By

Hayder Almosa

Submitted to the Department of Electrical Engineering and Computer Science and the
Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Prof. Erik Perrins, Chairperson

Prof. Lingjia Liu, Co-Chairperson

Committee members

Prof. Shannon Blunt

Prof. Rongqing Hui

Prof. Hongyi Cai

Date defended:

The Dissertation Committee for Hayder Almosa certifies
that this is the approved version of the following dissertation :

Downlink Achievable Rate Analysis for FDD Massive MIMO Systems

Prof. Erik Perrins, Chairperson

Prof. Lingjia Liu, Co-Chairperson

Date approved:

Abstract

Multiple-Input Multiple-Output (MIMO) systems with large-scale transmit antenna arrays, often called massive MIMO, are a very promising direction for 5G due to their ability to increase capacity and enhance both spectrum and energy efficiency. To get the benefit of massive MIMO systems, accurate downlink channel state information at the transmitter (CSIT) is essential for downlink beamforming and resource allocation. Conventional approaches to obtain CSIT for FDD massive MIMO systems require downlink training and CSI feedback. However, such training will cause a large overhead for massive MIMO systems because of the large dimensionality of the channel matrix. In this dissertation, we improve the performance of FDD massive MIMO networks in terms of downlink training overhead reduction, by designing an efficient downlink beamforming method and developing a new algorithm to estimate the channel state information based on compressive sensing techniques. First, we design an efficient downlink beamforming method based on partial CSI. By exploiting the relationship between uplink direction of arrivals (DoAs) and downlink direction of departures (DoDs), we derive an expression for estimated downlink DoDs, which will be used for downlink beamforming. Second, By exploiting the sparsity structure of downlink channel matrix, we develop an algorithm that selects the best features from the measurement matrix to obtain efficient CSIT acquisition that can reduce the

downlink training overhead compared with conventional LS/MMSE estimators. In both cases, we compare the performance of our proposed beamforming method with traditional methods in terms of downlink achievable rate and simulation results show that our proposed method outperform the traditional beamforming methods.

Acknowledgements

First and foremost, I thank Allah for his guidance and for blessing me with the special people who helped and supported me so much and special thanks to my loving family for their help and encouragement. Most importantly, I wish to thank my loving and supportive wife, Rajaa, and my four wonderful children, Fatimah, Malak, Rand, and Mohammed, who provide unending inspiration. Without their unlimited patience, I would have not been able to complete my doctoral degree.

I would like to express my deep appreciation, sincere thanks and gratitude to my advisors Dr. Erik Perrins and Dr. Lingjia Liu, for their continuing guidance, assistance and support throughout the progress of this research. I also extend my thanks to the members of my dissertation advisory committee Dr. Shannon Blunt, Dr. Rongqing Hui and Dr. Hongyi Cai for their helpful discussions.

Lastly, I would like to extend my sincere thanks to all my colleagues in the office, with whom I have shared moments of friendship and life stories. I thank them for their support throughout my journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Mobile access technology has been undergoing various revolutionary changes over the years. The rapid changes are due to capacity demands that result from growth of massive amounts of data over the past decade. In wireless communication systems, there is a huge demand for reliability, high data rates, energy efficiency, high mobility, and spectral efficiency. It is expected that mobile and wireless traffic volume will increase a thousand-fold over the next decade, which will be driven by the expected 50 billion devices connected to the cloud by 2020, all of which need to access and share data, anywhere and anytime [1, 2, 3]. Higher data throughput is required in the next generation of wireless cellular networks since the worldwide interest for data traffic is consistently developing. This objective can be accomplished without the requirement for more bandwidth or extra base stations if spectral efficiency is enhanced [4, 5]. Given the way that spectrum is turning into a scarce resource in cellular networks, orthogonal frequency division multiple access (OFDMA) has been proposed as a cutting edge air interface innovation to empower

spectrum efficiency and successfully battle frequency selective-fading. Because of its promising highlights, OFDMA is adopted in several wireless networks like Long Term Evolution (LTE) and IEEE 802.16m [6, 7]. To keep up with the fast wireless data traffic development, a vital objective of the next generation of wireless cellular networks is to enhance the data throughput by orders of magnitude; 100X and even 1000X higher throughput are often mentioned as 5G design goals [8, 4]. 4G networks, on the other hand, have only reached theoretical limits on data with the current technologies and, thus, are insufficient for accommodating these challenges. This means that groundbreaking wireless technologies are needed for solving the problems that are caused by trillions of wireless devices [9, 10, 11]. Researchers already have begun investigating beyond 4G wireless techniques. Additionally, statistics demonstrate that relying upon time and position, the use of both licensed and unlicensed spectrum could be low. In this manner, effective strategies required for spectrum sharing among various systems are of interest in a dynamic cellular network. The idea of software defined radios and cognitive radios was introduced to improve the effectiveness of frequency spectrum utilization [12]. The vision of next generation 5G cellular networks lies in providing higher capacity, higher data rates, less latency, massive device connectivity, better quality of service (QoS), and reduction in power consumption to improve battery life, compared to current 4G cellular network. Therefore, next generation wireless networks ought to provide economical power conservation mechanisms to realize longer battery life while at the same time providing increased user expertise and quality of service (QoS) [13]. In any case, the developing 5G arrangements will require somewhere around 100 MHz bandwidth, hundreds of antennas, and ultra-densely deployed base stations (BSs) to bolster an enormous number of clients as shown in Figure 1 [14]. These individual changes demonstrate that applying Nyquist's sampling theorem to 5G systems reminiscent of the prior 2G/3G/4G arrangements may result in additional difficul-

ties: restrictively huge overhead, excessively high complexity, and unexpected expense as well as power consumption because of the vast number of samples required. Moreover, compressive sensing (CS) offers a sub-Nyquist sampling method to deal with the under-determined linear system in a computationally effective manner, accordingly offering us chances to handle the previously mentioned difficulties.

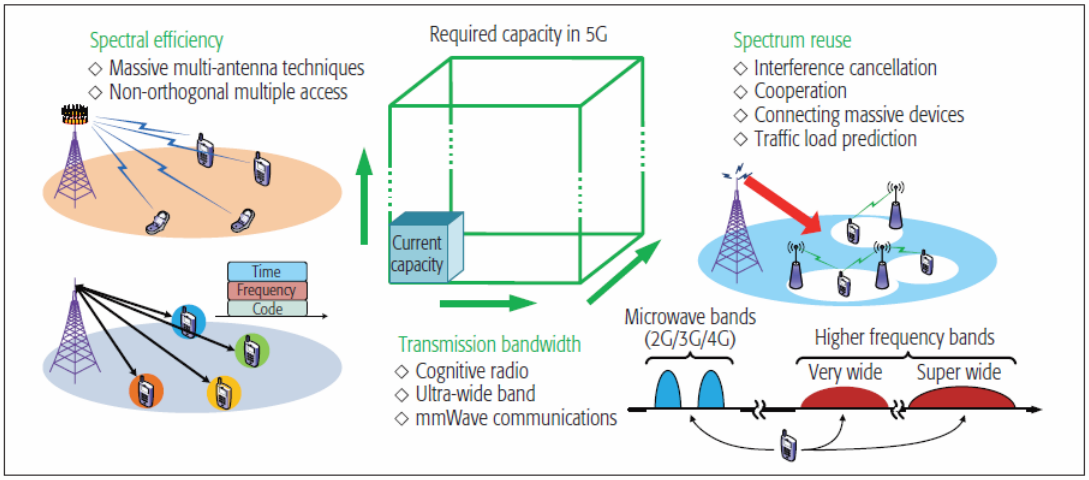

Figure 1.1: Three promising technical directions for 5G.

Full dimension MIMO, or FD-MIMO, has pulled in significant attention in wireless cellular networks in the past few years as a competitor innovation for next generation evolution. FD-MIMO uses a large number of antenna elements placed in a 2D plane to improve the spatial multiplexing and give more degrees of freedom in supporting users [15, 16]. 5G mobile communication technology beyond 2020 has the potential to provide access to information and sharing data anytime and anywhere for anything and anyone [17, 18, 19]. Recently, only a few main technical parameters for efficient performance of 5G have been approved, such as spectrum allocations. Bands with multiple frequencies can be introduced in 5G. Currently, bands at 6 GHz and lower have limited bandwidth and are fully occupied. In order to meet the growing expectation, millimeter wave

(mmWave) spectrum ranging from 24–100 GHz also got approval from the World Radiocommunication Conference 2015 (WRC-15). Many bands (in GHz: 24.25–27.5, 31.8–43.5, 45.5–50.2, 50.4–52.6, 66–76, and 81–86) were approved providing larger bandwidth [19, 20, 21]. However, WRC-19 will approve the final list. Bands of 60 GHz, not yet licensed, may also be used in the spectrum. mmWave spectrum will play a vital role in cutting edge 5G mmWave small cell systems. Unlike popular 6 GHz spectrum, mmWave spectrum has a lot of accessible bands, yet these experience the ill effects of high path loss [22, 23]. However, mmWave has different benefits, including limited inter-cell interference, low transmission latency, and improved security [24]. The 60 GHz spectrum demonstrates 20–40 dB expanded free space path loss and experiences 15–30 dB/km atmospheric absorption depending on climatic conditions. Luckily, as shown in Figure 2 [25], plenty of available spectrum around 60 GHz can be used to deliver these high rates.

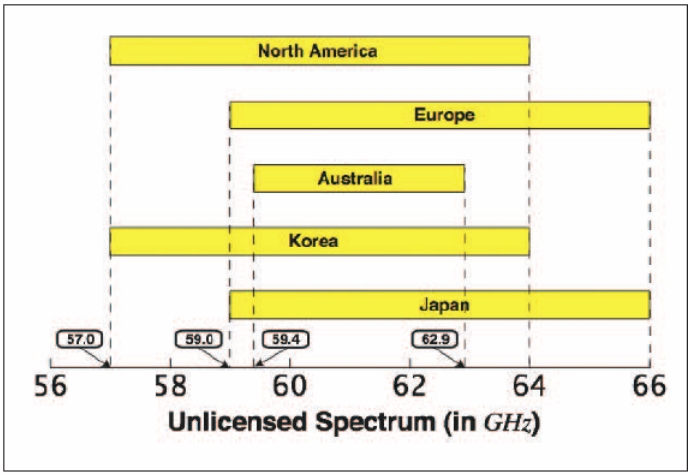

Figure 1.2: International unlicensed spectrum around 60 GHz.

This collection of frequencies in the spectrum will give better results for 5G. For coverage over wide areas, lower frequencies are included for personal and local communication; mmWave links with high rates can be included. Moreover, the most valuable frequency bands are below 6 GHz because these frequencies can provide good network coverage and service quality, while higher

bands might only work well under short-range line-of-sight conditions. The spectral efficiency (SE) has not seen any major improvements in previous network generations. Hence, it might be a factor that can be greatly improved in the future and possibly become the primary way to achieve high area throughput in 5G networks. Small cell (SC) networks are considered to be a key solution for the next generation of wireless cellular networks. However, synchronization between densely deployed small cells is one of the main issues in these networks. Even though SCs are right now connected chiefly for traffic offloading and indoor coverage, they can provide high capacity at a cost, in both indoor and outdoor situations. In principle, the system capacity scales directly with the cell intensity, so decreasing the cell size could viably enhance system capacity. Then again, since a shorter separation results in less path loss, the transmitting power could be decreased [26, 27, 28]. One of the key technologies for 5G wireless communication is massive MIMO. Massive MIMO significantly enhances spectral and energy efficiency [29, 30]. In massive MIMO systems, the effect of uncorrelated noise and small scale fading are eliminated, the number of users per cell are independent of the size of the cell and simple linear signal processing methods such as matched filters (MFs) can be used [31, 32, 33]. The leveraging of time-division duplexing (TDD) helps massive MIMO to scale to any desired degree with respect to number of service antennas [34, 35]. The addition of more antennas can be useful to increase throughput, reduce radiated power, uniformity in services everywhere in the cell and great simplicity in the signal processing in 5G technologies [36]. The TDD operation mode permits massive MIMO to be scalable in number of service antennas to any extent required. However, pilot contamination is the main factor that limits the performance of TDD massive MIMO. Pilot contamination occurs when users in adjacent cells share the same pilot sequences. The effect of pilot contamination is usually quantified as SINR (signal-to-interference-plus-noise ratio) saturation due to intercell interference

[37, 38]. If frequency-division duplexing (FDD) is utilized, the downlink and uplink use various frequency bands, then the CSI corresponding to the downlink and uplink is also different. The channel estimation for uplink through the BS is done by allowing the users to send different pilot sequences. The time needed to uplink the pilot transmission does not depend on the number of antennas at the BS. Moreover, for getting CSI for the downlink channels in the FDD systems, a 2-stage procedure is needed. The BS transmits the pilot symbols first to the users, and then the user feeds back the estimated CSI, complete or partial, for downlink channels to the BS. Channel sounding for FDD systems requires a large pilot overhead, which usually scales proportionally to the number of transmit antennas, so exploring efficient channel estimation techniques is an important issue for FDD massive MIMO systems [39, 32, 40, 41].

1.2 Proposed Research

Multiple-Input Multiple-Output (MIMO) systems with large-scale transmit antenna arrays, often called massive MIMO, are a very promising direction for 5G due to their ability to increase capacity and enhance both spectrum and energy efficiency. To get the benefit of massive MIMO, knowledge of CSIT is essential [42, 43, 44, 45]. In a TDD system, the BS obtains the CSI via uplink channel reciprocity, where the uplink training symbols scale linearly with the number of users, which is much less than the number of BS antennas. However, in a TDD massive MIMO system, the CSI estimation may be inaccurate because of calibration errors of RF chains and growing of reuse uplink training pilots when the number of users increases, which leads to inter-user interference and limits the system performance [46, 47]. In a FDD system, the uplink and downlink operations occur at different frequencies. The frequency separation for cellular systems is typically on the

order of 100 MHz and this is much larger than the coherence bandwidth of the channel (typically about 1MHz). Therefore, the small-scale fading of the channel impulse responses in the uplink and downlink are completely uncorrelated. In other words, the small-scale fading, and so, the instantaneous impulse response of the channel, are different for uplink and downlink [48]. In FDD systems, uplink and downlink channels fade independently if the frequency separation between them is large enough compared to the coherence bandwidth of the channel. As a result, the uplink channel information may not be applied directly for downlink beamforming. However, there are some properties that can be exploited for downlink beamforming from uplink channel information.

The primary downlink direction of arrival (DoA) can be modeled from the uplink by the equation below:

$$\theta_{DL} = \theta_{UL} + \Delta\theta \tag{1.1}$$

where θ_{DL} and θ_{UL} are the primary downlink and uplink DoAs (elevation or azimuth) respectively, and $\Delta\theta$ is the angular perturbation modeled as a Laplacian random process [49, 50]. In terms of spatial channel measurements, Hugl et al. [51] investigated the congruence of the directional properties of uplink and downlink as seen from the BS. They showed that the directional properties of the uplink and downlink radio channel are strongly correlated. According to measurement data evaluation, it is shown that the dominant DoA in uplink and downlink has only a minor deviation. Therefore, it is reasonable to utilize the dominant DoA which is estimated during uplink reception for downlink beamforming purposes. As a consequence, adaptive antenna downlink beamforming without terminal feedback is possible. The authors in [52] assumed that the dominant uplink DoA and downlink direction of departure (DoD) are the same. Using this information, the authors first

construct a precoding matrix composed of the downlink steering vectors. Then they constructed a second precoder to cancel the multi user interference. The second precoder is formed based on the first precoder, and constructed it in a zero-forcing (pseudo-inverse) type format. Finally, they computed the achievable rates based on these two beamformers. In [53], the authors assumed that the amplitude of a downlink path is equivalent to that of uplink one but differs in phase. Therefore, the user equipment (UE) only needs to feed the information of the phase back to the BS. The BS then reconstructs the downlink CSIT and performs beamforming.

It is clear that using the estimated uplink DoA directly for downlink beamforming is not optimum. Therefore, in Chapter 2, we design an efficient downlink beamforming method based on partial CSI. The partial CSI is the uplink dominant DoA which is obtained during uplink training phase. By exploiting the relationship between uplink (UL) direction of arrivals (DoAs) and downlink (DL) direction of departures (DoDs), we derive an expression for estimated downlink DoDs, which will be used for downlink beamforming to compare the performance with traditional methods in terms of downlink achievable rate that we derived. Simulation results also verify that, in terms of achievable rate, our proposed method outperform the traditional beamforming method.

In FDD sytems, the BS acquire CSI by broadcasting a sequence of downlink training pilots, then users feed back the estimated CSI to the BS via uplink signaling channels. Traditional downlink channel estimation techniques such as least square (LS) and minimum mean square error (MMSE) estimators require that the downlink training length scales linearly with the number of BS antennas, which becomes prohibitively large in massive MIMO systems. In addition, the CSI feedback becomes challenging due to large number of BS antennas [54, 42]. Based on experimental results of massive MIMO channels, it turns out that the user channel matrices exhibit sparsity due to limited local scatterers at the BS. In [42], the authors exploit the hidden joint sparsity struc-

ture in the user channel matrices to reduce the downlink training overhead and develop a joint orthogonal matching pursuit (OMP) algorithm to recover CSIT at the BS jointly. By exploiting the spatially common sparsity of massive MIMO channels and the temporal channel correlation, the author in [46] proposed an adaptive channel estimation scheme using compressive sensing (CS) to acquire the CSI. In [54], a closed-loop CSIT estimation scheme with 1 bit feedback was proposed to learn the minimum required pilot training overhead that leads to accurate CSIT acquisition. By exploiting the block sparse nature of the angular domain channel representation, the author in [55] proposed a weighted block l_1 minimization based downlink CSI recovery scheme for FDD massive MIMO systems with two stages. In the first stage, the common/individual supports of the multiuser channel matrix are extracted, while the second stage uses the obtained support information for CSI recovery by conducting weighted block l_1 minimization. Furthermore, FDD system dominates current wireless cellular systems and exploring efficient downlink channel acquisition schemes is important. In Chapter 3, we investigate the channel estimation problem in FDD multi-user massive MIMO systems with spatially correlated channels and develop an efficient channel estimation algorithm that exploits the sparsity structure of the downlink channel matrix. The proposed algorithm selects the best features from the measurement matrix to obtain efficient CSI acquisition that can reduce the downlink training overhead compared with the conventional LS/MMSE channel estimators. We compare the performance of our proposed channel estimation method with traditional ones in terms of normalized mean square error (MSE). Simulation results verify that the proposed algorithm can significantly reduce the pilot overhead and has better performance compared with the traditional channel estimation methods. In Chapter 4, we propose a feedback scheme that exploits the sparsity structure of the downlink channel matrix and design a low-complexity beam-forming scheme for FDD multi-user massive MIMO systems with spatially

correlated channels. The proposed schemes consider codebook-based feedback, which is adopted in the LTE/LTE-advanced systems. Specifically, after estimating the downlink channel for each user during the downlink training phase, we design a low-complexity beam-forming scheme for downlink transmission based on an estimated channel matrix. We next update the channel matrix using limited feedback and calculate the downlink achievable rate as a performance metric to compare the results. Simulation results verify that the proposed scheme can significantly reduce the pilot overhead and has better performance compared with the traditional methods. Systems with a large number of antennas at the BS are costly due to the large number of radio frequency (RF) chains—one for each antenna—which in turn results in large power consumption. As the number of transmit antennas increases at the BS, the circuit power consumption will gradually increase by the number of antennas. Therefore, in Chapter 5, we investigate the problem of transmit antenna selection for FDD multi-user massive MIMO networks and design an efficient antenna selection algorithm that can reduce the hardware complexity and improve the system performance in terms of achievable sum-rate.

1.3 Research Impact and Contribution

In this dissertation, first, we aim to design an efficient downlink beamforming approach for FDD massive MIMO systems based on partial channel state information and considering the mismatch error. Second, we estimate the downlink channel using compressive sensing techniques to reduce the downlink training overhead and derive an efficient downlink beamforming based on CSI estimates. Figure 1.3 illustrates the summary of our contribution in each chapter. By completing these research tasks, the proposed research is of value for the design and analysis of the current

and emerging cellular networks.

- The novelty of this research lies in the ability to develop different solutions to help support FDD massive MIMO systems and meet 5G technology requirements. We expect that this research will provide great benefits to next generation networks by solving some of their challenges, from training overhead to channel estimation to spectral efficiency.

- Downlink training overhead is the main challenge for FDD massive MIMO systems to acquire channel state information at the BS, which is necessary for downlink beamforming. Most of the past research exploits the reciprocity between the uplink and downlink bands in terms of uplink dominant angles of arrival (AoAs), but they ignore the mismatch error between them. The literature shows that the angular perturbation between uplink dominant DoA and DoD can be modeled as a Laplacian random process. The proposed beamforming based on partial channel state information along with the consideration of the mismatch error can be considered as a promising solution for FDD massive MIMO systems to achieve better performance.

- Accurate downlink channel state information is crucial for downlink beamforming and resource allocation to get the benefits of massive MIMO systems. Due to the large number of antennas at the BS, conventional approaches for channel estimation are not efficient for FDD massive MIMO systems because it requires the downlink training overhead to be scaled with the number of BS antennas and hence large percentage of resources will be dedicated to channel estimation and limit the system performance. The proposed channel estimation algorithm improves the system performance in terms of downlink training overhead reduction.

- The proposed limited feedback scheme along with channel estimation algorithm enhances the overall system performance of FDD massive MIMO systems in terms of overhead reduction and spectral efficiency.

- The addition of more antennas at the BS can be useful to increase throughput, reduce radiated transmit power, achieve uniformity in services everywhere in the cell, and simplify the signal processing. Systems with a large number of antennas at the BS lead to a large number of radio frequency (RF) chains, which increases power consumption and degrades the system performance in terms of energy efficiency. Hence, we develop a low complexity iterative algorithm to solve the antenna selection problem and enhance the system performance in terms of downlink achievable sum-rate.

- System level simulation using MATLAB software is carried out to verify the analysis for realistic network scenarios.

1.4 Notation

Throughout this dissertation, normal letters are used for scalars. Boldface capital and lower case letters denote matrices and vectors, respectively. The transposition, the Hermitian transposition, and the determinant of a complex matrix $\mathbf{A}$ are denoted by $\mathbf{A}^T$, $\mathbf{A}^H$ and $|\mathbf{A}|$, respectively. A $N \times K$ matrix, with ones on its main diagonal and zeros on its off-diagonal entries, is denoted by $\mathbf{I}_{N\times K}$, while the identity matrix of size N is simply denoted by $\mathbf{I}_N$. An $N \times K$ all-zeros matrix is denoted by $\mathbf{0}_{N\times K}$. The sets of complex and real numbers are denoted by $\mathbb{C}$ and $\mathbb{R}$, respectively. A circularly symmetric complex Gaussian random variable (r.v.) is represented by $Z = X + jY \sim \mathscr{CN}(0, \sigma^2)$,

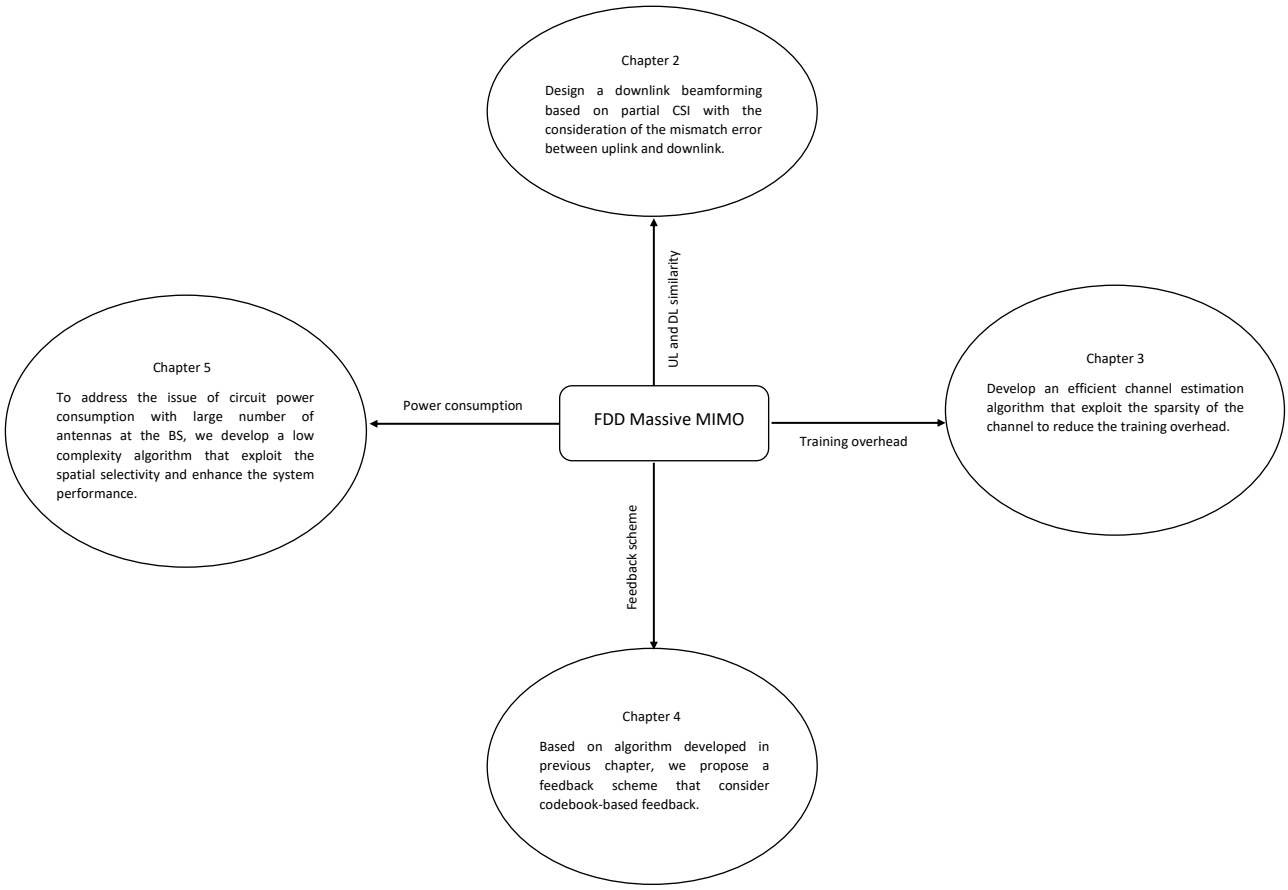

Figure 1.3: Summary of our contribution in each chapter.

where X and Y are independent and identically distributed (i.i.d.) normal r.v.'s from $\mathcal{N}(0, \frac{\sigma^2}{2})$. $\mathbb{E}[\cdot]$ represents the expectation operator. The trace of a square matrix $\mathbf{A} = [a_{ij}]_{n\times n}$ is defined as $\text{Tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. The vec$(\cdot)$ operator aligns all the elements of a matrix into a column vector by stacking the column vectors of the matrix, i.e., for $\mathbf{A} \in \mathbb{C}^{M\times N}$ then vec$(\mathbf{A}) \in \mathbb{C}^{MN\times 1}$. Moreover, the Kronecker and Hadamard product between two matrices $\mathbf{A}$ and $\mathbf{B}$ are symbolized by $\mathbf{A} \otimes \mathbf{B}$ and $\mathbf{A} \odot \mathbf{B}$, respectively. $\|\mathbf{a}\|_2$ and $\|\mathbf{A}\|_F$ denote the two-norm of the vector $\mathbf{a}$ and the Frobenius norm of the matrix $\mathbf{A}$, respectively. The inner product and angle are symbolized by $< \cdot, \cdot >$, and $\angle$, respectively.

Chapter 2

Downlink Achievable Rate Analysis based on partial CSI

2.1 Introduction

In this chapter, the downlink achievable rate for FDD massive MIMO systems based on partial channel state information will be discussed. This chapter is organized as follows. We begin with direction of arrival estimation in Section 2.2, after which the system model is discussed in Section 2.3. The uplink DoA estimation is discussed in detail in Section 2.4. The downlink framework is discussed in Section 2.5. In Section 2.5.1, we derive an expression for the downlink DoD by exploiting the relationship between uplink DoAs and downlink DoDs. The downlink achievable rate is discussed in Section 2.5.2. The performance of the proposed downlink beamforming has been discussed in Section 2.6.

2.2 Direction of Arrival Estimation

DoA estimation of narrow-band signals is a fundamental problem in many sensor array systems such as, direction finders (DF), radar, sonar, electronic surveillance, and mobile communications. Over the years, numerous methods have been proposed, for example, the MUltiple SIgnal Classification (MUSIC) technique and the Maximum-Likelihood (ML) technique [56]. Estimating the DoA of narrowband wave-fronts impinging on an array of sensors has long been of great research interest. Early methods used beam-forming techniques to electronically steer the array in one direction at a time and measure the corresponding output power. The major drawback of this approach is its low resolution, that is, the ability to resolve two closely spaced sources. The resolution was greatly improved with the introduction of the MUSIC technique [57], which used a subspace-based approach to find the DoA angles [58]. The goal of DoA estimation is to use the data received on the downlink at the base-station sensor array to estimate the directions of the signals from the desired sources as well as the directions of interference signals. The results of DoA estimation are then used to adjust the weights of the antenna array so that the radiated power is maximized towards the sources, and radiation nulls are placed in the directions of interference signals [59].

2.3 System Model

In this chapter, we consider a FDD Massive MIMO system with N_r antennas at the BS, and N_t antennas at the UE. In the uplink, the $N_r \times 1$ received signal, $\mathbf{y}^{ul}$, can be expressed as

$$\mathbf{y}^{ul} = \mathbf{H}^{ul}\mathbf{x}^{ul} + \mathbf{n}^{ul}, \tag{2.1}$$

where $\mathbf{H}^{ul}$ is the $N_r \times N_t$ uplink channel matrix, $\mathbf{x}^{ul}$ is the $N_t \times 1$ uplink transmit signal from the UE, and $\mathbf{n}^{ul}$ is the corresponding $N_r \times 1$ noise vector. In this chapter, we consider a clustered channel model, where each scattering cluster is assumed to contribute a single propagation path. Considering a finite number of resolvable paths between the transmitter and the receiver, the uplink channel matrix, $\mathbf{H}^{ul}$, can be represented as

$$\mathbf{H}^{ul} = \sum_{\ell=0}^{L-1} \alpha_{ul}(\ell)\mathbf{e}_{r,ul}(\ell)\mathbf{e}_{t,ul}^{H}(\ell), \tag{2.2}$$

where $\alpha_{ul}(\ell)$, $\mathbf{e}_{r,ul}(\ell)$, and $\mathbf{e}_{t,ul}^{H}(\ell)$ are respectively, the channel gain, $N_r \times 1$ receive antenna array response, and $N_t \times 1$ transmit antenna array response for the ℓ-th path of the channel between the BS and the UE. Moreover, L is the total number of paths for the channel, and $(\cdot)^H$ denotes the Hermitian. It is obvious that the transmit and receive antenna array responses depend on DoD of the transmit signal and DoA of the received signal, respectively. For the transmitter equipped with a uniform linear array (ULA), the transmit antenna array response can be described using the Vandermonde structure: $\mathbf{e}_{t,ul}^{H}(\ell) = \begin{bmatrix} 1 & e^{j\omega_\ell^{ul}} & \dots & e^{j(N_t-1)\omega_\ell^{ul}} \end{bmatrix}^T$, where $\omega_\ell^{ul} = (2\pi\Delta_t/\lambda^{ul})\cos\Omega_\ell^{ul}$, Δ_t is the spacing between the adjacent transmit antenna elements, Ω_ℓ^{ul} is the transmit angle (DoD) for the ℓ-th path of the uplink channel, and λ^{ul} is the uplink carrier wavelength. The antenna array at the BS is a planar array placed in the X–Z plane, with M_1 and M_2 antenna elements in vertical and horizontal directions, respectively. Accordingly, the number of receive antenna elements at the BS is $N_r = M_1 M_2$.

Since the antenna elements at the BS are placed in a 2D plane, for each resolvable path, there will be an azimuth DoA and an elevation DoA. Therefore, the receive antenna array response for the ℓ-th path can be expressed as $\mathbf{e}_{r,ul}(\ell) = \mathbf{a}(v_\ell^{ul}) \otimes \mathbf{a}(u_\ell^{ul})$, where $\otimes$ represents the Kronecker

product, and $\mathbf{a}(u_\ell^{ul}) = \begin{bmatrix} 1 & e^{ju_\ell^{ul}} & \dots & e^{j(M_1-1)u_\ell^{ul}} \end{bmatrix}^T$ and $\mathbf{a}(v_\ell^{ul}) = \begin{bmatrix} 1 & e^{jv_\ell^{ul}} & \dots & e^{j(M_2-1)v_\ell^{ul}} \end{bmatrix}^T$ can be viewed as the receive steering vectors of the elevation and azimuth angles, respectively. Here, $u_\ell^{ul} = \frac{2\pi\Delta_r}{\lambda^{ul}} \cos\theta_\ell^{ul}$ and $v_\ell = \frac{2\pi\Delta_r}{\lambda^{ul}} \sin\theta_\ell^{ul} \cos\phi_\ell^{ul}$ are the two receive spatial frequencies at the BS, Δ_r is the spacing between adjacent antenna elements in the receive antenna array, and θ_ℓ^{ul} and ϕ_ℓ^{ul} are the elevation and azimuth DoAs for the ℓ-th path of the channel, respectively.

2.4 Uplink DoA Estimation

In this section, we introduce a low complexity DoA-estimation algorithm based on Unitary ESPRIT in order to jointly estimate the elevation and azimuth angles. The uplink channel matrix in (2.2) can be written as

$$\mathbf{H}^{ul} = \mathbf{A}^{ul}\mathbf{D}^{ul}\mathbf{B}^{ul^H}, \tag{2.3}$$

where $\mathbf{A}^{ul} = \begin{bmatrix} \mathbf{e}_{r,ul}(0) & \mathbf{e}_{r,ul}(1) & \dots & \mathbf{e}_{r,ul}(L-1) \end{bmatrix}$, $\mathbf{D} = \text{diag}\begin{bmatrix} \alpha_{ul}(0) & \alpha_{ul}(1) & \dots & \alpha_{ul}(L-1) \end{bmatrix}$, and $\mathbf{B}(k) = \begin{bmatrix} \mathbf{e}_{t,ul}(0) & \mathbf{e}_{t,ul}(1) & \dots & \mathbf{e}_{t,ul}(L-1) \end{bmatrix}$. Now, collecting the received signal for V snapshots, we can extend (2.1) as follows

$$\mathbf{Y}^{ul} = \mathbf{H}^{ul}\mathbf{X}_{ul} + \mathbf{N}^{ul}, \tag{2.4}$$

where $\mathbf{Y}^{ul}$ is the $N_r \times V$ received signal matrix, $\mathbf{X}^{ul}$ is the $N_t \times V$ transmit signal matrix, and $\mathbf{N}^{ul}$ is

the corresponding $N_r \times V$ noise matrix. Using (2.3), we can write (2.4) in a compact form:

$$\mathbf{Y}^{ul} = \mathbf{A}^{ul}\mathbf{S}^{ul} + \mathbf{N}^{ul}, \tag{2.5}$$

where $\mathbf{S}^{ul} = \mathbf{D}^{ul}\mathbf{B}^{ul^H}\mathbf{X}_{ul}$ can be regarded as equivalent uplink transmit signal.

In order to perform unitary ESPRIT, we need to use forward–backward averaging on the received signal:

$$\begin{aligned}\mathbf{Y}^{ul^{fba}} &= \begin{bmatrix}\mathbf{Y}^{ul} & \mathbf{\Pi}_{N_r}\mathbf{Y}^{ul^*}\mathbf{\Pi}_{N_t}\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{A}^{ul}\mathbf{S}^{ul} & \mathbf{\Pi}_{N_r}\mathbf{A}^{ul^*}\mathbf{S}^{ul^*}\mathbf{\Pi}_{N_t}\end{bmatrix} \\ &+ \begin{bmatrix}\mathbf{N}^{ul} & \mathbf{\Pi}_{N_r}\mathbf{N}^{ul^*}\mathbf{\Pi}_{N_t}\end{bmatrix},\end{aligned} \tag{2.6}$$

where $\mathbf{A}^*$ represents the complex conjugate of $\mathbf{A}$, and $\mathbf{\Pi}_p$ denotes the $p \times p$ exchange matrix with ones on its antidiagonal and zeros elsewhere. The subspace decomposition of the signal space of the received signal through singular value decomposition then can be written as:

$$\begin{bmatrix}\mathbf{A}^{ul}\mathbf{S}^{ul} & \mathbf{\Pi}_{N_r}\mathbf{A}^{ul^*}\mathbf{S}^{ul^*}\mathbf{\Pi}_{N_t}\end{bmatrix} = \begin{bmatrix}\mathbf{U}_s & \mathbf{U}_n\end{bmatrix} \begin{bmatrix}\mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{V}_s^H \\ \mathbf{V}_n^H\end{bmatrix} \tag{2.7}$$

The array manifold matrix of an $M_1 \times M_2$ rectangular antenna array can be expressed as:

$$\mathbf{A}^{ul}(u_\ell, v_\ell) = \mathbf{a}(u_\ell^{ul})\mathbf{a}^T(v_\ell^{ul}),$$

that is, the 2D steering matrix can be decomposed into product of two 1D steering vectors [60]. We

can choose the two subarrays of the steering vector, $\mathbf{a}(u_\ell^{ul})$, with maximum overlap, that is, each having $M_1 - 1$ antenna elements. Because of the fixed displacement between the first and second subarrays, if the first $M_1 - 1$ elements are multiplied by $e^{ju_\ell^{ul}}$, the resulting vector will be equal to the vector containing the last $M_1 - 1$ components. This can be expressed as:

$$e^{ju_\ell^{ul}} \mathbf{J_1} \mathbf{a}(u_\ell^{ul}) = \mathbf{J_2} \mathbf{a}(u_\ell^{ul}), \tag{2.8}$$

where, $\mathbf{J_1}$ is an $(M_1 - 1) \times M_1$ selection matrix constructed by taking the first $(M_1 - 1)$ rows of $\mathbf{I}_{M_1}$ ($M_1 \times M_1$ Identity Matrix), and $\mathbf{J_2}$ is an $(M_1 - 1) \times M_1$ selection matrix constructed by taking the last $(M_1 - 1)$ rows of $\mathbf{I}_{M_1}$. A unitary and left-$\mathbf{\Pi}$ real matrix, $\mathbf{Q}_{M_1}$ can be constructed to change $\mathbf{a}(u_\ell^{ul})$ to a real-valued transformed steering vector, $\mathbf{a}^R(u_\ell^{ul})$,

$$\mathbf{a}^R(u_\ell^{ul}) = \mathbf{Q}_{M_1}^H \mathbf{a}(u_\ell^{ul}). \tag{2.9}$$

Assuming $M_1 = 2K$, $\mathbf{Q}_{M_1}$ can be constructed as

$$\mathbf{Q}_{2K} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K \\ \mathbf{\Pi}_K & j\mathbf{\Pi}_K \end{bmatrix}, \tag{2.10}$$

where, $\mathbf{I}_K$ is a $K \times K$ identity matrix, and $\mathbf{\Pi}_K$ is a $K \times K$ exchange matrix.

We have the following relation [61, 62]:

$$\tan(\frac{u_\ell^{ul}}{2}) \mathbf{K}_1 \mathbf{a}^R(u_\ell^{ul}) = \mathbf{K}_2 \mathbf{a}^R(u_\ell^{ul}).$$

Here, $\mathbf{K}_1 = \mathrm{Re}\{\mathbf{Q}_{M_1-1}^H \mathbf{J}_2 \mathbf{Q}_{M_1}\}$, $\mathbf{K}_2 = \mathrm{Im}\{\mathbf{Q}_{M_1-1}^H \mathbf{J}_2 \mathbf{Q}_{M_1}\}$. Extending the relation to a 2D antenna array:

$$\tan\left(\frac{u_\ell^{ul}}{2}\right)\mathbf{K}_1\mathbf{A}^R(u_\ell^{ul}, v_\ell^{ul}) = \mathbf{K}_2\mathbf{A}^R(u_\ell^{ul}, v_\ell^{ul}), \tag{2.11}$$

where,

$$\mathbf{A}^R(u_\ell^{ul}, v_\ell^{ul}) = \mathbf{Q}_{M_1}^H \mathbf{a}(u_\ell^{ul})\mathbf{a}^T(v_\ell^{ul})\mathbf{Q}_{M_2}^* = \mathbf{a}^R(u_\ell^{ul})(\mathbf{a}^R(v_\ell^{ul}))^T.$$

Furthermore, we can rewrite the formulation in (2.11) as:

$$\tan\left(\frac{u_\ell^{ul}}{2}\right)\mathbf{K}_{x1}\mathrm{vec}\left\{\mathbf{A}^R(u_\ell^{ul}, v_\ell^{ul})\right\} = \mathbf{K}_{x2}\mathrm{vec}\left\{\mathbf{A}^R(u_\ell^{ul}, v_\ell^{ul})\right\}$$

where $\mathbf{K}_{x1} \triangleq \mathbf{I}_{M_2} \otimes \mathbf{K}_1$, $\mathbf{K}_{x2} \triangleq \mathbf{I}_{M_2} \otimes \mathbf{K}_2$, and vec$\{.\}$ is the vectorization operation. Accordingly, we can specify an $M_1M_2 \times N_c$ real-valued array manifold matrix:

$$\mathbf{A}^R \triangleq \left[\mathrm{vec}\left\{\mathbf{a}^R(u_0^{ul}, v_0^{ul})\right\}, \ldots, \mathrm{vec}\left\{\mathbf{a}^R(u_{L-1}^{ul}, v_{L-1}^{ul})\right\}\right]$$

Then, we have the shift-invariance equation:

$$\mathbf{K}_{x1}\mathbf{A}^R\mathbf{\Omega}_x = \mathbf{K}_{x2}\mathbf{A}^R \tag{2.12}$$

where,

$$\mathbf{\Omega}_x \triangleq \mathrm{diag}\left\{\tan\left(\frac{u_0^{ul}}{2}\right), \ldots, \tan\left(\frac{u_{L-1}^{ul}}{2}\right)\right\} \tag{2.13}$$

It is important to note that after the unitary transformation, the matrices all become real-valued

matrices. This will significantly reduce the computational complexity.

Similarly, for $\mathbf{a}(v_\ell^{ul})$, we can conduct the same process. Let $\mathbf{K}_3 = \mathrm{Re}\left\{\mathbf{Q}_{M_2-1}^H \mathbf{J}_2' \mathbf{Q}_{M_2}\right\}$, and $\mathbf{K}_4 = \mathrm{Im}\left\{\mathbf{Q}_{M_2-1}^H \mathbf{J}_2' \mathbf{Q}_{M_2}\right\}$, where $\mathbf{J}_2'$ is the $(M_2-1)\times M_2$ matrix constructed by taking the last (M_2-1) rows of $\mathbf{I}_{M_2}$. Accordingly, we have

$$\mathbf{K}_{y1}\mathbf{A}^R\mathbf{\Omega}_y = \mathbf{K}_{y2}\mathbf{A}^R \tag{2.14}$$

where $\mathbf{K}_{y1} \triangleq \mathbf{K}_3 \otimes \mathbf{I}_{M_1}$, $\mathbf{K}_{y2} \triangleq \mathbf{K}_4 \otimes \mathbf{I}_{M_1}$, and

$$\mathbf{\Omega}_y \triangleq \mathrm{diag}\left\{\tan\left(\frac{v_0^{ul}}{2}\right),\ldots,\tan\left(\frac{v_{L-1}^{ul}}{2}\right)\right\} \tag{2.15}$$

Let $\mathbf{U}_s$ be the signal subspace and $\mathbf{T}$ be the nonsingular transformation matrix, we have $\mathbf{U}_s = \mathbf{A}^R\mathbf{T}$ since the array steering matrix $\mathbf{A}^R$ and the matrix $\mathbf{U}_s$ span the same column space in the absence of noise or with an infinite number of measurements. Under the noisy case or with a finite number of measurements, this expression holds approximately [63]. Substituting this relation into (2.12), we have

$$\mathbf{K}_{x1}\mathbf{U}_s\mathbf{\Psi}_x = \mathbf{K}_{x2}\mathbf{U}_s \tag{2.16}$$

where $\mathbf{\Psi}_x \triangleq \mathbf{T}^{-1}\mathbf{\Omega}_x\mathbf{T}$. Similarly, we also have

$$\mathbf{K}_{y1}\mathbf{U}_s\mathbf{\Psi}_y = \mathbf{K}_{y2}\mathbf{U}_s \tag{2.17}$$

where $\mathbf{\Psi}_y \triangleq \mathbf{T}^{-1}\mathbf{\Omega}_y\mathbf{T}$. From (2.16) and (2.17), we can solve for $\hat{\mathbf{\Psi}}_x$ and $\hat{\mathbf{\Psi}}_y$ based on the estimated signal subspace using LS-type methods. Let the eigenvalues of the $N_c \times N_c$ complex matrix $\hat{\mathbf{\Psi}}_x +$

$j\hat{\mathbf{\Psi}}_y$ be $\hat{\lambda}_\ell, \ell = 1,2,\ldots,N_c$. u_ℓ^{ul} and v_ℓ^{ul} can be estimated from:

$$\hat{u}_\ell^{ul} = 2\tan^{-1}\left\{\mathrm{Re}\left(\hat{\lambda}_\ell^{ul}\right)\right\} \quad \hat{v}_\ell^{ul} = 2\tan^{-1}\left\{\mathrm{Im}\left(\hat{\lambda}_\ell^{ul}\right)\right\}$$

Accordingly, 2D DoAs of interest are obtained through simple parameter transformation.

2.5 Downlink Framework

Accurate downlink channel state information at the transmitter (CSIT) is essential to utilize the benefit of massive MIMO. Conventional approaches to obtain CSIT for FDD massive MIMO systems require downlink training and CSI feedback. However, such training results in a large overhead cost for massive MIMO systems because of the large dimensionality of the channel matrix. In this section, we design an efficient downlink beamforming method based on partial CSI. By exploiting the relationship between uplink (UL) DoAs and downlink (DL) DoDs, we derive an expression for estimated downlink DoDs in Section (2.5.1), which will be used for downlink beamforming to compare the performance with traditional method in terms of downlink achievable rate that we derived in Section (2.5.2).

2.5.1 Downlink DoD Estimation

In this section, we will derive the expressions for downlink beamforming directions for both elevation ($\tilde{\theta}_{dl}$) and azimuth ($\tilde{\phi}_{dl}$) domains by using MMSE estimator. The joint estimation of the

dominant elevation and azimuth angles can be expressed as:

$$\left(\tilde{\theta}_{dl}, \tilde{\phi}_{dl}\right) = \mathbb{E}\left((\theta_{dl}, \phi_{dl}) | (\tilde{\theta}_{ul}, \tilde{\phi}_{ul})\right) \tag{2.18}$$

where $\tilde{\theta}_{ul} = \theta_{ul} + e_1$ and $\tilde{\phi}_{ul} = \phi_{ul} + e_2$ are the uplink estimated elevation and azimuth angles, respectively; θ_{ul} and ϕ_{ul} are the true uplink angles, and e_1 and e_2 are the corresponding estimation errors with zero mean. However, in this work, we assume that the uplink DoAs are perfectly estimated, and hence $e_1 = 0$ and $e_2 = 0$. Incorporation of uplink DoA estimation error [45] into the analysis is left for a future extension of the current work. Now, the relationship between uplink dominant DoA and downlink DoD can be defined as [49]:

$$\theta_{dl} = \tilde{\theta}_{ul} + \varepsilon_1 \tag{2.19}$$

$$\phi_{dl} = \tilde{\phi}_{ul} + \varepsilon_2 \tag{2.20}$$

where ε_1 and ε_2 are the perturbations modelled as a Laplacian random process and we assume these have zero mean. We can assume that the DoA-DoD mismatch errors of the elevation and azimuth domains are independent of each other. In other words, the mismatch error variables perturbs the elevation and azimuth angles independently, and therefore, we can carry out the analysis for the elevation and azimuth angles estimators separately. Now, the estimated DoD of the elevation angle can be expressed as:

$$\begin{aligned} \tilde{\theta}_{dl} = \mathbb{E}\{\theta_{dl}; \tilde{\theta}_{ul}\} &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \theta_{dl} \frac{\beta_\theta}{\sqrt{2}\sigma_{\theta,dl}} e^{-\frac{\sqrt{2}\ |\theta_{dl} - \tilde{\theta}_{ul}|}{\sigma_{\theta,dl}}} d\theta_{dl} \\ &= \frac{\beta_\theta}{\sqrt{2}\sigma_{\theta,dl}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \theta_{dl} e^{-\frac{\sqrt{2}\ |\theta_{dl} - \tilde{\theta}_{ul}|}{\sigma_{\theta,dl}}} d\theta_{dl}, \end{aligned} \tag{2.21}$$

where $\beta_\theta = \frac{1}{1-e^{-\frac{\pi}{\sqrt{2}\sigma_{\theta,dl}}}}$ is a normalization constant required for the function to integrate to one. We can write the expression as:

$$\tilde{\theta}_{dl} = \frac{\beta_\theta}{\sqrt{2}\,\sigma_{\theta,dl}} \left[\int_{-\frac{\pi}{2}}^{\tilde{\theta}_{ul}} \theta_{dl} e^{\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} d\theta_{dl} + \int_{\tilde{\theta}_{ul}}^{\frac{\pi}{2}} \theta_{dl} e^{-\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} d\theta_{dl} \right] \tag{2.22}$$

Now, using integration by parts, we can calculate the integrals as follows:

$$\begin{aligned} &\int_{-\frac{\pi}{2}}^{\tilde{\theta}_{ul}} \theta_{dl} e^{\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} d\theta_{dl} \\ &= \theta_{dl} \frac{\sigma_{\theta,dl}}{\sqrt{2}} e^{\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \Bigg]_{-\frac{\pi}{2}}^{\tilde{\theta}_{ul}} - \int_{-\frac{\pi}{2}}^{\tilde{\theta}_{ul}} \frac{\sigma_{\theta,dl}}{\sqrt{2}} e^{\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} d\theta_{dl} \\ &= \frac{\tilde{\theta}_{ul}\,\sigma_{\theta,dl}}{\sqrt{2}} + \frac{\pi\,\sigma_{\theta,dl}}{2\sqrt{2}} e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} - \left[\frac{\sigma_{\theta,dl}^2}{2} e^{\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \right]_{-\frac{\pi}{2}}^{\tilde{\theta}_{ul}} \\ &= \frac{\tilde{\theta}_{ul}\,\sigma_{\theta,dl}}{\sqrt{2}} + \frac{\pi\,\sigma_{\theta,dl}}{2\sqrt{2}} e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} - \frac{\sigma_{\theta,dl}^2}{2} + \frac{\sigma_{\theta,dl}^2}{2} e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \end{aligned} \tag{2.23}$$

Similarly, we calculate the second integral:

$$\begin{aligned} &\int_{\tilde{\theta}_{ul}}^{\frac{\pi}{2}} \theta_{dl} e^{-\frac{\sqrt{2}(\theta_{dl}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} d\theta_{dl} \\ &= \frac{\tilde{\theta}_{ul}\,\sigma_{\theta,dl}}{\sqrt{2}} - \frac{\pi\,\sigma_{\theta,dl}}{2\sqrt{2}} e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} + \frac{\sigma_{\theta,dl}^2}{2} - \frac{\sigma_{\theta,dl}^2}{2} e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \end{aligned} \tag{2.24}$$

By adding these two integrals, we get the final expression of $\tilde{\theta}_{dl}$ as:

$$\tilde{\theta}_{dl} = \frac{\beta_\theta}{\sqrt{2}\,\sigma_{\theta,dl}} \left[\frac{2\tilde{\theta}_{ul}\,\sigma_{\theta,dl}}{\sqrt{2}} + \frac{\pi\,\sigma_{\theta,dl}}{2\sqrt{2}} \left(e^{\frac{\sqrt{2}-\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} - e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \right) + \frac{\sigma^2_{\theta,dl}}{2} \left(e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} - e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\theta}_{ul})}{\sigma_{\theta,dl}}} \right) \right] \tag{2.25}$$

Similarly, we can obtain the DoD estimation for azimuth angle as:

$$\tilde{\phi}_{dl} = \frac{\beta_\phi}{\sqrt{2}\,\sigma_{\phi,dl}} \left[\frac{2\tilde{\phi}_{ul}\,\sigma_{\phi,dl}}{\sqrt{2}} + \frac{\pi\,\sigma_{\phi,dl}}{2\sqrt{2}} \left(e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\phi}_{ul})}{\sigma_{\phi,dl}}} - e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\phi}_{ul})}{\sigma_{\phi,dl}}} \right) + \frac{\sigma^2_{\phi,dl}}{2} \left(e^{\frac{\sqrt{2}(-\frac{\pi}{2}-\tilde{\phi}_{ul})}{\sigma_{\phi,dl}}} - e^{\frac{-\sqrt{2}(\frac{\pi}{2}-\tilde{\phi}_{ul})}{\sigma_{\phi,dl}}} \right) \right], \tag{2.26}$$

where $\beta_\phi = \frac{1}{1-e^{-\frac{\pi}{\sqrt{2}\sigma_{\phi,dl}}}}$.

2.5.2 Downlink Achievable Rate Analysis

In this section, we will derive the downlink achievable rate for our model. The downlink achievable rate can be written as:

$$R = \log_2 \left[det \left(\mathbf{I} + \frac{\mathbf{H}^{dl}\mathbf{Q}^{dl}\mathbf{H}^{dl^H}}{\sigma_n^2} \right) \right] \tag{2.27}$$

where σ_n^2 is the noise variance and $\mathbf{H}^{dl}$ is the downlink channel defined as:

$$\mathbf{H}^{dl} = \mathbf{B}^{dl}\mathbf{D}^{dl}\mathbf{A}^{dl^H} \tag{2.28}$$

where $\mathbf{D} = \text{diag}\begin{bmatrix}\alpha_{dl}(0) & \alpha_{dl}(1) & \dots & \alpha_{dl}(L-1)\end{bmatrix}$, L is the total number of paths, $\alpha_{dl}(l^{th})$ is the complex channel gain for l^{th} path, $\mathbf{B}^{dl} = \begin{bmatrix}\mathbf{e}_{r,dl}(0) & \mathbf{e}_{r,dl}(1) & \dots & \mathbf{e}_{r,dl}(L-1)\end{bmatrix}$,$\mathbf{e}_{r,dl}(\ell) = \begin{bmatrix}1 & e^{j\omega_\ell^{dl}} & \dots & e^{j(N_t-1)\omega_\ell^{dl}}\end{bmatrix}^T$, where $\omega_\ell^{dl} = (2\pi\Delta_t/\lambda^{dl})\cos\Omega_\ell^{dl}$, Δ_t is the spacing between the adjacent receive antenna elements, Ω_ℓ^{dl} is the receive angle (DoA) for ℓ-th path of the downlink channel, and λ^{dl} is the downlink carrier wavelength. Moreover, $\mathbf{A}^{dl} = \begin{bmatrix}\mathbf{e}_{t,dl}(0) & \mathbf{e}_{t,dl}(1) & \dots & \mathbf{e}_{t,dl}(L-1)\end{bmatrix}$, $\mathbf{e}_{t,dl}(\ell) = \mathbf{a}(v_\ell^{dl}) \otimes \mathbf{a}(u_\ell^{dl})$ where $\otimes$ represents the Kronecker product, and $\mathbf{a}(u_\ell^{dl}) = \begin{bmatrix}1 & e^{ju_\ell^{dl}} & \dots & e^{j(M_1-1)u_\ell^{dl}}\end{bmatrix}^T$ and $\mathbf{a}(v_\ell^{dl}) = \begin{bmatrix}1 & e^{jv_\ell^{dl}} & \dots & e^{j(M_2-1)v_\ell^{dl}}\end{bmatrix}^T$ can be viewed as the transmit steering vectors of the elevation and azimuth angles, respectively. Here, $u_\ell^{dl} = \frac{2\pi\Delta_r}{\lambda^{dl}}\cos\theta_\ell^{dl}$ and $v_\ell = \frac{2\pi\Delta_r}{\lambda^{dl}}\sin\theta_\ell^{dl}\cos\phi_\ell^{dl}$ are the two transmit spatial frequencies at the BS, Δ_r is the spacing between adjacent antenna elements in the transmit antenna array, and θ_ℓ^{dl} and ϕ_ℓ^{dl} are the elevation and azimuth DoDs for the ℓ-th path of the channel, respectively. Finally, $\mathbf{Q}^{dl}$ is the covariance matrix of the transmit signal, which can be defined as:

$$\begin{aligned}\mathbf{Q}^{dl} &= \mathbb{E}\left[\mathbf{w}x^{dl}{x^{dl}}^H\mathbf{w}^H\right] \\ &= \mathbf{w}\,\mathbb{E}\left[x^{dl}{x^{dl}}^H\right]\mathbf{w}^H = \mathbf{w}\mathbf{w}^H\end{aligned} \tag{2.29}$$

where $\mathbf{w}$ is the beamformer in the dominant direction, and the second equality holds due to the fact that the transmit signal x is a Gaussian random variable with zero mean and unit variance. Then the achievable rate can be expressed as

$$\begin{aligned}R &= \log_2\left[det\left(\mathbf{I} + \frac{\mathbf{B}^{dl}\mathbf{D}^{dl}{\mathbf{A}^{dl}}^H\mathbf{w}\mathbf{w}^H\mathbf{A}^{dl}\mathbf{D}^{dl*}{\mathbf{B}^{dl}}^H}{\sigma_n^2}\right)\right] \\ &= \log_2\left[det\left(\mathbf{I} + \frac{{\mathbf{A}^{dl}}^H\mathbf{w}\mathbf{w}^H\mathbf{A}^{dl}\mathbf{D}^{dl*}{\mathbf{B}^{dl}}^H\mathbf{B}^{dl}\mathbf{D}^{dl}}{\sigma_n^2}\right)\right]\end{aligned} \tag{2.30}$$

For the massive MIMO systems, $\frac{\mathbf{B}^{dl^H}\mathbf{B}^{dl}}{N_t} \approx \mathbf{I}$ so that the achievable rate can be simplified as

$$R = \log_2 \left[det \left(\mathbf{I} + \frac{\mathbf{A}^{dl^H}\mathbf{w}\mathbf{w}^H\mathbf{A}^{dl}\mathbf{D}^{dl^*}\mathbf{D}^{dl}N_t}{\sigma_n^2} \right) \right] \tag{2.31}$$

Now, the beamforming vector $\mathbf{w}$ can be defined as:

$$\mathbf{w} = \frac{1}{N_r}\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl}), \tag{2.32}$$

where $\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl})$ is steering vector corresponding estimated DoD. The term $\mathbf{A}^{dl^H}\mathbf{w}$ can now be written as:

$$\mathbf{A}^{dl^H}\mathbf{w} = \frac{1}{N_r}\begin{bmatrix} \mathbf{e}_{t,dl}^H(0)\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl}) \\ \vdots \\ \mathbf{e}_{t,dl}^H(\ell)\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl}) \\ \vdots \\ \mathbf{e}_{t,dl}^H(L-1)\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl}) \end{bmatrix}, \tag{2.33}$$

where, for massive MIMO systems all elements, will be zero except for the one that correspond to the dominant DoA, $\hat{\boldsymbol{\varphi}}_{ul}$:

$$\mathbf{A}^{dl^H}\mathbf{w} = \frac{1}{N_r}\begin{bmatrix} 0 \\ \vdots \\ \mathbf{e}_{t,dl}^H(\hat{\boldsymbol{\varphi}}_{ul})\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl}) \\ \vdots \\ 0 \end{bmatrix}. \tag{2.34}$$

Hence, $\mathbf{A}^{dl^H}\mathbf{w}\mathbf{w}^H\mathbf{A}^{dl}$ will be a diagonal matrix and the final expression for the achievable rate can be expressed as

$$R = \log_2\left(1 + \frac{|\mathbf{e}_{t,dl}^H(\hat{\boldsymbol{\varphi}}_{ul})\tilde{\mathbf{e}}_{t,dl}(\hat{\boldsymbol{\varphi}}_{dl})|^2\ |\alpha_{dl}(\hat{\boldsymbol{\varphi}}_{ul})|^2 N_t}{N_r^2\ \sigma_n^2}\right), \tag{2.35}$$

where $\alpha_{dl}(\hat{\boldsymbol{\varphi}}_{ul})$ is the channel gain corresponding to the dominant path.

2.6 Performance Evaluation of Downlink Beamforming

In this section, we analyze DoD estimation for FDD massive MIMO systems. As a performance metric, we use the downlink achievable rate to compare the results with the traditional method. We first evaluate ESPRIT-based uplink DoA estimation for FDD massive MIMO systems. As a performance metric, we use the root mean square error (RMSE) of the estimated elevation and azimuth angles, and the RMSE of any angle (either azimuth or elevation) of path ℓ is defined as $\sqrt{\mathbb{E}\{(\theta_\ell - \hat{\theta}_\ell)^2\}}$, where θ_ℓ and $\hat{\theta}_\ell$ denote the true and estimated angles, respectively, and $\mathbb{E}$

represents expectation over different channel realizations.

To evaluate the performance of DoA estimation, we assume that there are 4 dominant clusters, and each cluster contributes one resolvable path, which is typical for outdoor millimeter-wave communication systems at both 28 GHz and 73 GHz. The antenna spacing for both the received and transmit antennas, at the BS and the UE, is assumed to be 0.5λ. The number of receive antennas at the UE is set as 8. The number of transmit antennas at BS is set as 8×8. The azimuth and elevation DoAs are chosen randomly from the uniform distribution over support $U[-180^\circ, 180^\circ]$ and $U[50^\circ, 80^\circ]$, respectively, In this work, we invoke the far field assumption, and the wavefront impinging on the antenna array is assumed to be planar. Finally, the total available transmit power at each MS is assumed to be unity, and the SNR is defined as the ratio of the received signal power to the noise power.

The performance of the estimation of elevation and azimuth angles for 8×8 and 16×16 antenna arrays is shown in Figure 2.1 and Figure 2.2, respectively, where the RMSE of DoA estimation has been used as a performance metric.

It can be observed that as SNR increases, the estimation performance improves. Moreover, in the high SNR regime, an increase in the total number of BS antennas also has a positive impact on the estimation of both elevation and azimuth angles. In this work, we perform simulation for the downlink achievable rate based on four cases. First, we assume that the downlink beamformer is constructed by simply taking the uplink angles as the beamforming directions. This is the traditional way of beamforming. Second, the beamformer utilizes (2.25) just for elevation angles, and the azimuth angle is chosen as the same as the uplink azimuth angle. The third case is just the vice-versa of the second case—the BS uses (2.26) for azimuth and elevation angle is chosen as that of uplink. Finally, in the proposed method, we assume for beamforming in both elevation and

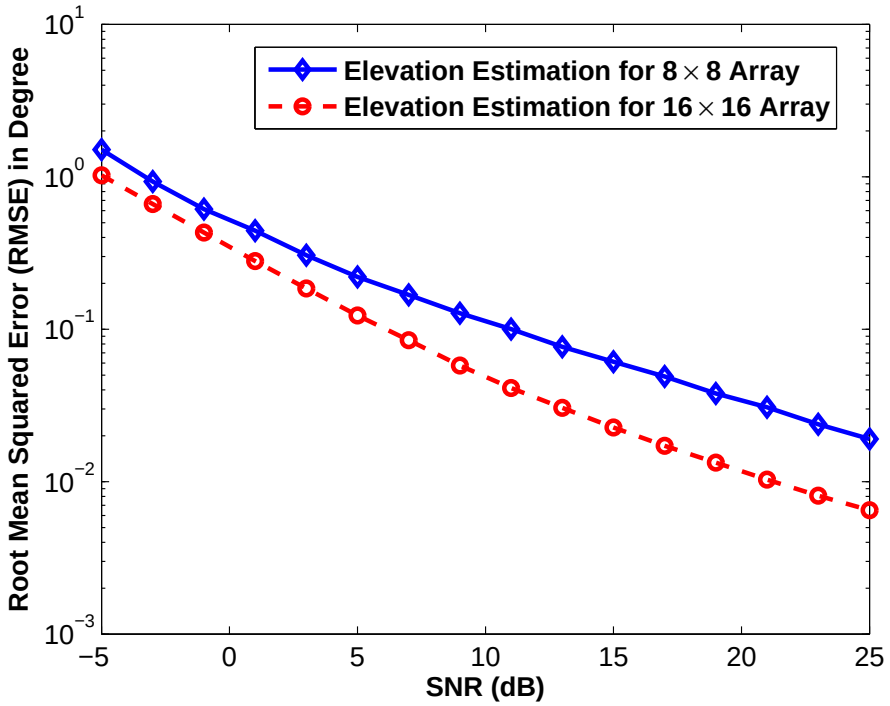

Figure 2.1: Elevation angle estimation for 8×8 and 16×16 antenna array for Gaussian and uniform DoA distribution.

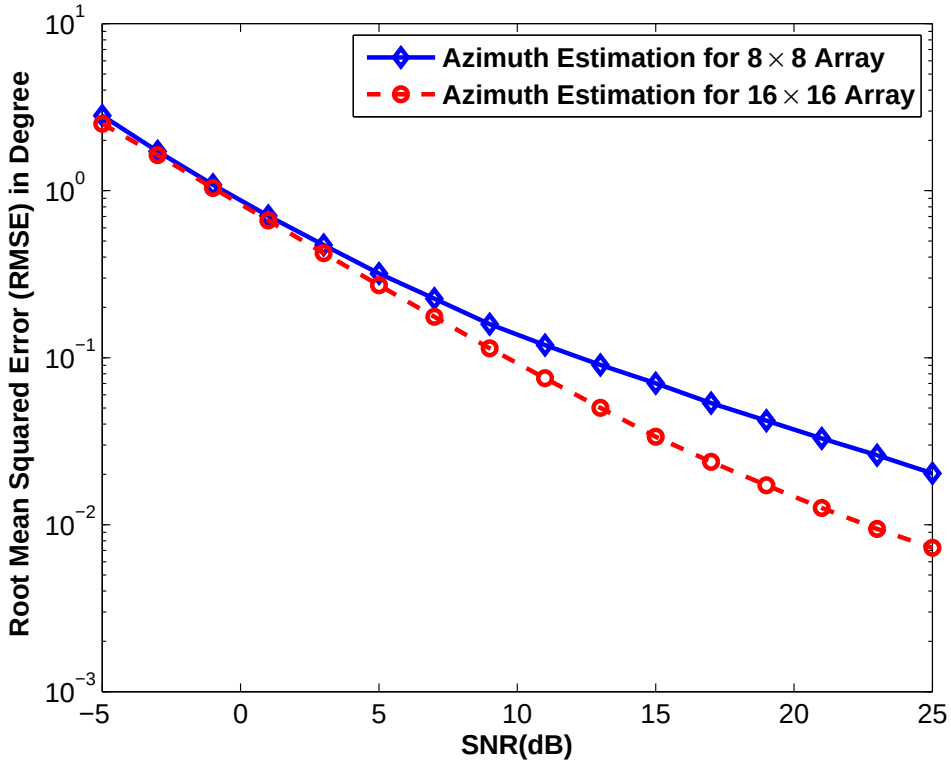

Figure 2.2: Azimuth estimation for 8×8 and 16×16 antenna array for Gaussian and uniform DoA distribution.

azimuth domain, the BS utilizes both (2.25) and (2.26). The results are shown in Figure 2.1, from which it is evident that the proposed beamforming method gives higher achievable rate than any of the other three cases.

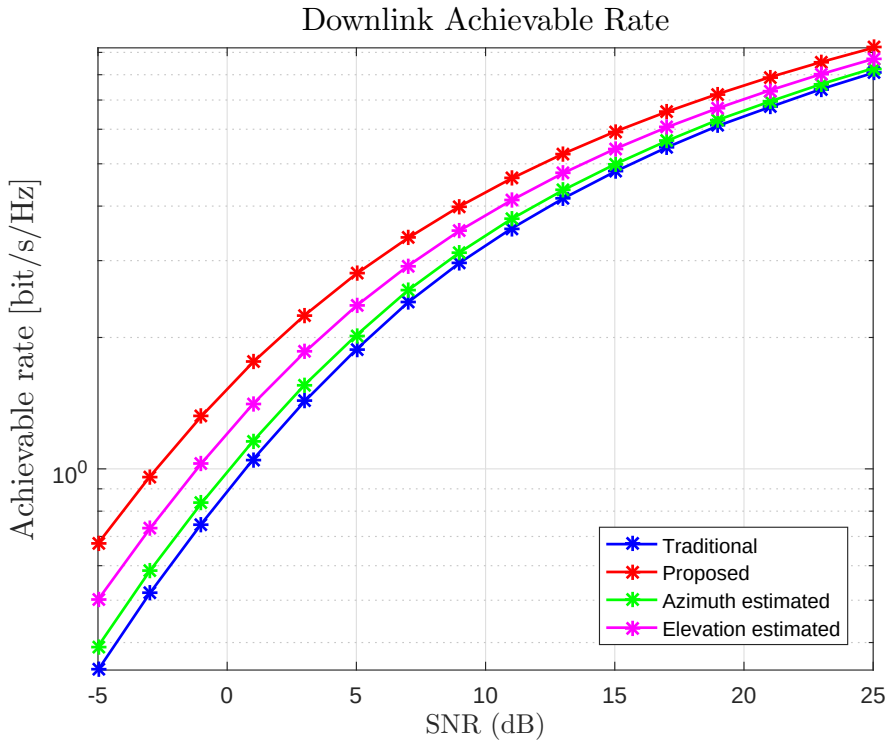

Figure 2.3: Downlink achievable rate vs. SNR for 8×8 URA at the BS.

Chapter 3

Channel Estimation for FDD Multi-User Massive MIMO systems

3.1 Introduction

In this chapter, channel estimation techniques for FDD multi-user massive MIMO systems will be discussed. This chapter is organized as follows. We begin with compressive sensing in Section 3.2, after which the system model is discussed in Section 3.3. The downlink framework is discussed in detail in Section 3.4. In Section 3.4.1 channel estimation techniques are discussed. In Section 3.5, Uplink CSIT feedback is discussed. The performance of channel estimation methods is discussed in Section 3.6.

3.2 Compressive Sensing

3.2.1 Motivation

The traditional method to acquire and reconstruct a continous-time bandlimited signal from a set of uniformly spaced samples is based on the Shannon–Nyquist sampling theorem, which states that the sampling rate (Nyquist rate) should be at least twice the highest frequency present in the signal of interest. Unfortunately, in many important and emerging applications including digital image and video cameras, the Nyquist rate is high, which results in many samples and makes compression a necessity prior to storage or transmission. One of the most popular techniques for signal compression is known as transform coding, which is based on finding a basis that provides a sparse or compressible representation for signals in a class of interest. A sparse representation of a signal of length N means that we can represent it with K nonzero coefficients where $K << N$. A compressible representation of a signal of length N means that it is well-approximated by a signal with only K nonzero coefficients. Both sparse and compressible signals can be represented with high accuracy by preserving only the values and locations of the largest coefficients of the signal. This process is called sparse approximation, and forms the foundation of transform coding schemes that exploit signal sparsity and compressibility, including the JPEG, JPEG2000, MPEG, and MP3 standards. In practice, signals tend to be compressible, rather than sparse. Mathematically, a compressible signal has a representation whose entries decay rapidly when sorted in order of decreasing magnitude. Leveraging the concept of transform coding, compressive sensing (CS) has emerged as a new framework for signal acquisition and sensor design that enables a potentially large reduction in the sampling and computation costs for sensing signals that have a sparse or compressible representation [64, 65, 66, 67, 68]

3.2.2 Fundamentals of Compressive Sensing

Compressive sensing (CS) is a simple and efficient signal acquisition technique that collects a few measurements about the signal of interest and later uses optimization techniques for reconstructing the original signal from what appears to be an incomplete set of measurements. There are two important aspects in CS: information preservation and recovery information. In CS, we obtain a unique solution for underdetermined linear system of equations $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{y} \in R^M$, $\mathbf{A} \in R^{M \times N}$ and $\mathbf{x} \in R^N$, exploiting the fact the signal is sparse (which means it has at most k elements that are non-zero, where $k \ll N$). The matrix ($\mathbf{A}$) is called the sensing matrix, measurement matrix, or dictionary. If the signal has a sparse representation then we can transform it into a certain subspace as a linear combination of some basis. Because the CS model is nonlinear, it presents a difficulty in the algorithms that perform the sparse recovery. In addition, it turns out to find the unique solution, the sensing matrix must satisfy certain properties:

- Restricted Isometry Property (RIP):

  Any k columns in $\mathbf{A}$ must be linearly independent. This property is difficult to calculate and takes time if N is large. However, it turns out that there is class of matrices that satisfy this property:

  - Gaussian Random Matrices.
  - Bernoulli Random Matrices.

  All the matrix that obey the RIP property provided that:

$$M \geq c\, k\, \log\left(\frac{N}{k}\right) \tag{3.1}$$

where c is small constant [66, 69, 70].

- Null Space Property (NSP):

  If we wish to be able to recover all sparse signals $\mathbf{x}$ from the measurements $\mathbf{Ax}$, then it is immediately clear that for any pair of distinct vectors $\mathbf{x}, \grave{\mathbf{x}} \in \sum_k$ we must have $\mathbf{Ax} \neq \mathbf{A}\grave{\mathbf{x}}$, since otherwise it would be impossible to distinguish $\mathbf{x}$ from $\grave{\mathbf{x}}$ based solely on the measurements $\mathbf{y}$. More formally, by observing that if $\mathbf{Ax} = \mathbf{A}\grave{\mathbf{x}}$ then $\mathbf{A}(\mathbf{x} - \grave{\mathbf{x}}) = 0$ with $(\mathbf{x} - \grave{\mathbf{x}}) \in \sum_{2k}$, we see that $\mathbf{A}$ uniquely represents all $\mathbf{x} \in \sum_k$ if and only if $\mathbb{N}(\mathbf{A})$ contains no vectors in $\sum_{2k}$. While there are many equivalent ways of characterizing this property, one of the most common is known as the spark. Therefore, the smallest number of columns in $\mathbf{A}$ that are linearly dependent, which is called spark ($\mathbf{A}$), should be greater than $2k$ [67].

- Mutual coherence:

  The mutual coherence of a matrix $\mathbf{A}$, $\mu(\mathbf{A})$ is the largest absolute normalized inner product between different columns of A.

$$\mu(\mathbf{A}) = \max_{1 \leq i < j \leq N} \frac{|\langle a_i, a_j \rangle|}{||a_i||_2 ||a_j||_2} \tag{3.2}$$

  where a_i is the i^{th} column of $\mathbf{A}$. The coherence of a matrix is always in the range $\mu(\mathbf{A}) \in \left[\sqrt{\frac{N-M}{M(N-1)}}, 1\right]$, where the lower bound is known as the Welch bound. The lower bound is approximately $\mu(\mathbf{A}) \geq \frac{1}{\sqrt{M}}$ [67, 65]. The definition of incoherence usually states that distances between sparse signals are approximately conserved as distances between their respective measurements generated by the sampling process. The following condition on $\mathbf{A}$

are required to guarantee uniqueness.

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{A})}\right) \tag{3.3}$$

Compressive sampling is mainly concerned with low coherence. Random matrices are largely incoherent with any fixed basis matrix [66, 70].

There are two methods to recover the sparse signal. The first are l_1 minimization methods, such as Basis Pursuit (BP), which use a linear programming optimization to recover the signal. This method provides strong guarantees and stability, but it requires large computational cost, in particular for large-scale applications. For example, a solver based on the interior point method has an associated computational complexity order of $O(m^2n^3)$. For many real-time applications, including wireless communication applications, the computational cost and time complexity of the l_1-norm minimization solver might be burdensome [69]. The second method, which uses greedy methods, calculate the support of the signal iteratively, such as Matching Pursuit (MP) and Orthogonal Matching Pursuit (OMP). In the OMP algorithm, columns of the matrix $\mathbf{A}$ are chosen one at a time using a greedy strategy. Specifically, in each iteration, a column maximally correlated with the modified observation is chosen. This is not necessarily optimal since the choice does not guarantee to pick the column associated with the nonzero element of $\mathbf{x}$. These methods are usually much faster than BP, but they cannot provide the same guarantees as the BP method [71, 67].

3.3 System Model

In this chapter, we consider a flat and block-fading FDD multi-user massive MIMO system with one BS equipped with N_t antennas that serves K non-cooperating users. Each user has N_r antennas. For the downlink, the $N_r \times 1$ received signal for k-th user, $\mathbf{y}_k^{\mathrm{d}l}$, can be expressed as

$$\mathbf{y}_k^{\mathrm{d}l} = \mathbf{H}_k^{\mathrm{d}l}\mathbf{x}^{\mathrm{d}l} + \mathbf{n}_k^{\mathrm{d}l}, \tag{3.4}$$

where $\mathbf{H}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times N_t}$ is the downlink channel matrix for the k-th user with zero mean and spatial CCM $\left(\mathbf{R}_k \in \mathbb{C}^{N_r N_t \times N_r N_t}\right)$, which is defined on the column stacking of the downlink channel matrix. Thus, $\mathrm{vec}(\mathbf{H}_k^{\mathrm{d}l}) \in \mathscr{CN}(0, \mathbf{R}_k)$. We assume that users have different spatial CCM. We compute the spatial CCM for each user based on different models include the one-ring model, the Laplacian model, and the exponential model, which are described in Section 3.6. $\mathbf{x}^{\mathrm{d}l} \in \mathbb{C}^{N_t \times 1}$ is the transmitted signal and $\mathbf{n}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times 1}$ is a complex Gaussian noise vector, where $\mathbf{n}_k^{\mathrm{d}l} \sim \mathscr{CN}(0, \mathbf{I}_{N_r})$. In this chapter, we consider the angular domain representation of the downlink channel matrix [72, 42], where the downlink channel matrix for the k-th user $\mathbf{H}_k^{\mathrm{d}l}$ can be represented as:

$$\mathbf{H}_k^{\mathrm{d}l} = \mathbf{A}_R \mathbf{H}_k^w \mathbf{A}_T^H \tag{3.5}$$

where $\mathbf{H}_k^w \in \mathbb{C}^{N_r \times N_t}$ is the angular domain channel matrix for the k-th user and $\mathbf{A}_R \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{A}_T \in \mathbb{C}^{N_t \times N_t}$ are defined as unitary matrices for the angular domain transformation at the UE and BS respectively. Based on [42, 47, 54], the angular domain channel matrix $\mathbf{H}_k^w$ is sparse due to limited local scattering effects at the BS as shown in Figure 3.1. The (i, j)-th nonzero entry of $\mathbf{H}_k^w$ indicates there is a spatial path from the j-th transmit direction of the BS to the i-th receive

direction of UE. We define the i-th row vector of $\mathbf{H}_k^w$ as $\mathbf{h}_{ki}$ and $\text{supp}(\mathbf{h}_{ki})$ denotes the index set of nonzero elements in $\mathbf{h}_{ki}$. Due to the fact that the row vectors of $\mathbf{H}_k^w$ usually have the same sparsity, the following assumption for $\mathbf{H}_k^w$ is considered as in [42]:

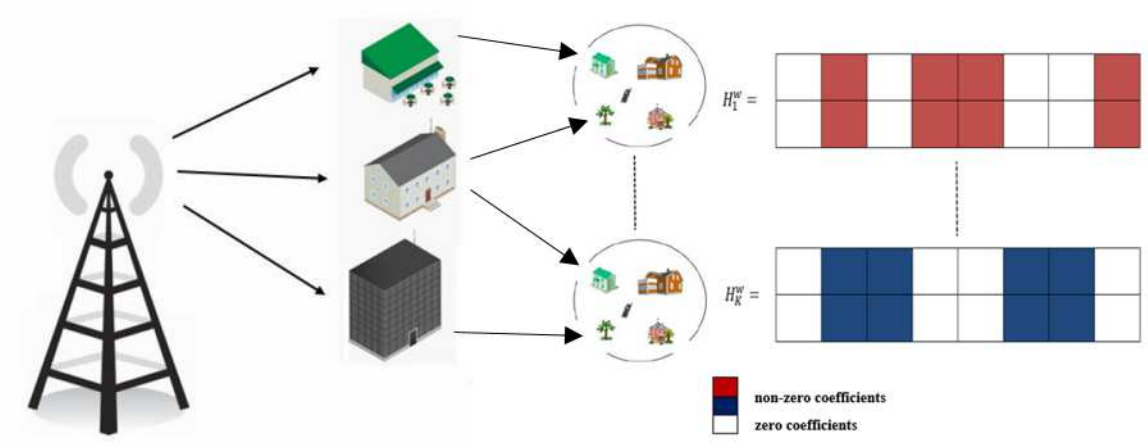

Figure 3.1: Sparsity structure of downlink channel matrix due to limited local scattering at the BS.

Assumption 1. *Define* $\mathbf{h}_{ki}$ *as the i-th row vector of* $\mathbf{H}_k^w$*, then there exists an index set* $\Omega_k, 0 < |\Omega_k| \ll N_t, \forall k$*, such that*

$$\text{supp}(\mathbf{h}_{k1}) = \text{supp}(\mathbf{h}_{k2}) = \cdots = \text{supp}(\mathbf{h}_{kN_r}) \stackrel{\Delta}{=} \Omega_k \tag{3.6}$$

3.4 Downlink Framework

Accurate downlink channel state information at the transmitter (CSIT) is essential to utilize the benefit of massive MIMO. Conventional approaches to obtain CSIT for frequency-division duplex (FDD) multi-user massive MIMO systems require downlink training and uplink CSI feedback. However, such training results in large overhead for massive MIMO systems because of the large dimensionality of the channel matrix. In this section, we investigate the channel estimation prob-

lem in FDD multi-user massive MIMO systems with spatially correlated channels and develop an efficient channel estimation algorithm that exploits the sparsity structure of the downlink channel matrix. The proposed algorithm selects the best features from the measurement matrix to obtain efficient CSI acquisition that can reduce the downlink training overhead compared with the conventional LS/MMSE channel estimators.

3.4.1 Downlink Channel Estimation

In this section, CSI estimation techniques using traditional methods such as least squares (LS), minimum mean squared error (MMSE), compressive sensing (CS), and backward feature selection (BFS) are considered. During the downlink training phase, the BS broadcasts a sequence of $T < \tau$ pilot vectors $[\mathbf{x}(1)\dots\mathbf{x}(t)\dots\mathbf{x}(T)]$, where τ is the channel coherence interval and data will be transmitted for a fraction $\left(\frac{\tau-T}{\tau}\right)$. Moreover, $\mathbb{E}||\mathbf{x}(t)||^2 = P$, $t = 1,\dots,T$, where P is the average transmit power at the BS per training time slot. The received signal vector for k-th user in the t-th time slot, i.e. $\mathbf{y}_k(t) \in \mathbb{C}^{N_r \times 1}$ can be written as:

$$\mathbf{y}_k(t) = \mathbf{H}_k^{\mathrm{d}l}\mathbf{x}(t) + \mathbf{n}_k(t), \;\; t = 1,\dots,T \tag{3.7}$$

where $\mathbf{H}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times N_t}$ is the quasi-static block fading downlink channel matrix, and $\mathbf{n}_k(t) \in \mathbb{C}^{N_r \times 1}$ is a complex Gaussian noise vector with $i.i.d.$ entries having zero mean and unit variance. Let $\mathbf{X} = [\mathbf{x}(1)\dots\mathbf{x}(T)] \in \mathbb{C}^{N_t \times T}$, $\mathbf{Y}_k = [\mathbf{y}_k(1)\dots\mathbf{y}_k(T)] \in \mathbb{C}^{N_r \times T}$ and $\mathbf{N}_k = [\mathbf{n}_k(1)\dots\mathbf{n}_k(T)] \in \mathbb{C}^{N_r \times T}$, respectively, be the pilot matrix with total power constraint $\mathrm{tr}(\mathbf{X}\mathbf{X}^H) = PT$, received signal at k-th user during downlink training phase, and the noise matrix. The received signal in 3.7 can be

equivalently represented as:

$$\mathbf{Y}_k = \mathbf{H}_k^{\mathrm{d}l}\mathbf{X} + \mathbf{N}_k, \;\; k = 1, \ldots, K \tag{3.8}$$

The CSI estimation based LS approach can be expressed as [73]:

$$\tilde{\mathbf{H}}_{k,LS}^{\mathrm{d}l} = \mathbf{Y}_k \mathbf{X}^{\dagger} \tag{3.9}$$

where $\mathbf{X}^{\dagger} = \mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1}$ is the pseudoinverse of $\mathbf{X}$ and $(\cdot)^H$ denotes the Hermitian transpose. By vectorizing the received signal in (3.8) and applying $\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\mathrm{vec}(\mathbf{B})$, where $\mathrm{vec}(\mathbf{X})$ is the column vector obtained by stacking the columns of $\mathbf{X}$, the received training signal at k-th user can be expressed as:

$$\mathrm{vec}(\mathbf{Y}_k) = \tilde{\mathbf{X}}\mathrm{vec}(\mathbf{H}_k^{\mathrm{d}l}) + \mathrm{vec}(\mathbf{N}_k) \tag{3.10}$$

where $\tilde{\mathbf{X}} \triangleq (\mathbf{X}^T \otimes \mathbf{I}_{N_t})$ and $\mathbf{I}_n$ denotes the n-dimension identity matrix. The MMSE estimator $\tilde{\mathbf{H}}_{k,MMSE}^{\mathrm{d}l}$ of the downlink channel matrix for k-th user can be written as [74]:

$$\mathrm{vec}(\tilde{\mathbf{H}}_{k,MMSE}^{\mathrm{d}l}) = \mathbf{R}_k\tilde{\mathbf{X}}^H(\tilde{\mathbf{X}}\mathbf{R}_k\tilde{\mathbf{X}}^H + \mathbf{I}_{TN_t})^{-1}\mathrm{vec}(\mathbf{Y}_k) \tag{3.11}$$

The traditional channel estimation methods such as LS and MMSE require that the optimal training length be at least the same as the number of transmit antennas [75], $T \geq N_t$, which leads to large overhead in downlink training and CSI feedback for FDD multi-user massive MIMO systems due to large number of BS antennas. Therefore, these methods are not efficient for FDD multi-user

massive MIMO systems. Since the downlink channel matrix exhibits sparsity structure in the angular domain, channel estimation based on the CS approach provides efficient CSI acquisition by utilizing prior knowledge of the downlink channel sparsity structure so that the overhead reduction can be achieved [46, 54]. In the CS method, the BS sends compressive training symbols $\mathbf{X} \in \mathbb{C}^{N_t \times T}$ with $T \ll N_t$ during the downlink training phase, then users estimate their respective downlink CSI and feed it back to the BS during uplink feedback. By using Equation 3.5, the received signal for the k-th user during the downlink training phase in (3.8) can be rewritten into the noisy CS recovery problem as:

$$\tilde{\mathbf{Y}}_k = \tilde{\mathbf{X}}\tilde{\mathbf{H}}_k^{\mathrm{d}l} + \tilde{\mathbf{N}}_k \tag{3.12}$$

where,

$$\begin{aligned}
\tilde{\mathbf{Y}}_k &= \sqrt{\frac{N_t}{PT}}\mathbf{Y}_k^H \mathbf{A}_R \quad \in \mathbb{C}^{T \times N_r} \\
\tilde{\mathbf{X}} &= \sqrt{\frac{N_t}{PT}}\mathbf{X}^H \mathbf{A}_T \quad \in \mathbb{C}^{T \times N_t} \\
\tilde{\mathbf{H}}_k^{\mathrm{d}l} &= (\mathbf{H}_k^w)^H \quad \in \mathbb{C}^{N_t \times N_r} \\
\tilde{\mathbf{N}}_k &= \sqrt{\frac{N_t}{PT}}\mathbf{N}_k^H \mathbf{A}_R \quad \in \mathbb{C}^{T \times N_r}
\end{aligned} \tag{3.13}$$

In the CS approach, the measurement matrix $\tilde{\mathbf{X}}$ must satisfy the restricted isometry property (RIP) in order to obtain efficient CSI recovery. The RIP is defined as follows: The measurement matrix $\tilde{\mathbf{X}} \in \mathbb{C}^{T \times N_t}$ satisfies the RIP of order k with isometry constant δ_k if $0 < \delta_k < 1$ and δ_k is the smallest

number such that:

$$(1-\delta_k)\,||\mathbf{h}||^2 \leq ||\tilde{\mathbf{X}}\mathbf{h}||^2 \leq (1+\delta_k)\,||\mathbf{h}||^2 \tag{3.14}$$

holds for all k sparse $\mathbf{h}$ [67]. The following theorem shows that the measurement matrix $\tilde{\mathbf{X}}$ satisfies the RIP.

Theorem 1. *The measurements matrix* $\tilde{\mathbf{X}} = \sqrt{\frac{N_t}{PT}}\mathbf{X}^H\mathbf{A}_T$ *with pilot matrix* $\mathbf{X} \in \mathbb{C}^{N_t \times T}$ *drawn from Rademacher distribution satisfies the RIP property.*

Proof. We have

$$[\mathbf{X}]_{i,j} = \begin{cases} \sqrt{\frac{P}{N_t}} & \text{with probability} = 0.5 \\ -\sqrt{\frac{P}{N_t}} & \text{with probability} = 0.5 \end{cases} \tag{3.15}$$

where $i = 1,\ldots,N_t,\ j = 1,\ldots,T$. Observe that the rows of matrix $\mathbf{X}$ have l_2 norm equal to $\sqrt{\frac{PT}{N_t}}$. We have A_T as a unitary matrix and $\|\mathbf{A}_T\|_2 = 1$. Let $\mathbf{B} = \sqrt{\frac{N_t}{PT}}\mathbf{X}^H$, then we observe that the columns $\mathbf{b}_i$ of matrix $\mathbf{B}$ hav ℓ_2 norm equal to one. Let $\mathbf{h} \in \mathbb{C}^{N_t}$ be a sparse vector with $\|\mathbf{h}\|_0 = s > 1$

and $\mathbf{z} = \mathbf{A}_T\mathbf{h}$. Thus, we have:

$$
\begin{aligned}
|\|\mathbf{B}\mathbf{A}_T\mathbf{h}\|_2^2 - \|\mathbf{h}\|_2^2| &= |\|\mathbf{B}\mathbf{A}_T\mathbf{h}\|_2^2 - \|\mathbf{A}_T\mathbf{h}\|_2^2| \\
&= |\|\mathbf{B}\mathbf{z}\|_2^2 - \|\mathbf{z}\|_2^2| \\
&= |(z_1 b_{11} + \ldots + z_{N_t} b_{1N_t})^2 + \ldots + \\
&\quad (z_1 b_{T1} + \ldots + z_{N_t} b_{TN_t})^2 - \sum_{i=1}^{N_t} z_i^2| \\
&= |2 \sum_{i<j} z_i z_j \langle b_i, b_j \rangle| \\
&\leq 2\mu \sum_{i<j} |z_i z_j| \\
&\leq \mu \left(\left(\sum_i |z_i| \right)^2 - \|\mathbf{z}\|_2^2 \right) \\
&\leq \mu \left(\|\mathbf{z}\|_1^2 - \|\mathbf{z}\|_2^2 \right) \\
&\leq \mu \left(N_t \|\mathbf{z}\|_2^2 - \|\mathbf{z}\|_2^2 \right) \\
&\leq \mu \left(N_t - 1 \right) \|\mathbf{z}\|_2^2 \\
&= \mu \left(N_t - 1 \right) \|\mathbf{h}\|_2^2
\end{aligned}
$$

Where $\|\mathbf{z}\|_1^2 \leq N_t \|\mathbf{z}\|_2^2$ and $\mu = \max_{i \neq j} |\langle b_i, b_j \rangle|$ is the mutual coherence of matrix $\mathbf{B}$. If we let $\delta = \mu(N_t - 1)$, then the matrix $\tilde{\mathbf{X}} = \mathbf{B}\mathbf{A}_T$ satisfies the RIP property:

$$
(1 - \delta)\|\mathbf{h}\|_2^2 \leq \|\mathbf{B}\mathbf{A}_T\mathbf{h}\|_2^2 \leq (1 + \delta)\|\mathbf{h}\|_2^2
$$

□

The OMP algorithm is an iterative greedy algorithm that selects a column from $\tilde{\mathbf{X}}$ that is most

Algorithm 1: OMP algorithm for CSI recovery

Input : $\tilde{\mathbf{Y}}_k, \tilde{\mathbf{X}}, \mathbf{A}_T, \mathbf{A}_R$ and sparsity level (s)
Output : $\hat{\mathbf{H}}_k^{dl}$

- **Initialization** : $r_k^0 = \tilde{\mathbf{y}}_{kj}, \tilde{\mathbf{h}}_{kj}^{dl} = 0$ where $\tilde{\mathbf{y}}_{kj}$ and $\tilde{\mathbf{h}}_{kj}^{dl}$ is the j-th column of $\tilde{\mathbf{Y}}_k$, $\tilde{\mathbf{H}}_k^{dl}$ respectively, $j = 1 : N_r$, $\text{supp}_k^0 = \emptyset$
- **while** $i \leq s$ **do**
- $l_k^i = \arg\max |\langle r_k^{(i-1)}, \tilde{\mathbf{X}} \rangle|$
- $\text{supp}_k^i = \text{supp}_k^{i-1} \cup l_k^i$
- $(\tilde{\mathbf{h}}_{kj}^{dl})^i = (\tilde{\mathbf{X}}_{\text{supp}_k^i})^\dagger \tilde{\mathbf{y}}_{kj}$
- $r_k^i = \tilde{\mathbf{y}}_{kj} - \tilde{\mathbf{X}}_{\text{supp}_k^i} (\tilde{\mathbf{h}}_{kj}^{dl})^i$
- **end** (**while**)
- $\hat{\mathbf{H}}_k^{dl} = \mathbf{A}_R (\tilde{\mathbf{H}}_k^{dl})^H \mathbf{A}_T^H$

correlated with the current residuals. Then, this column is added to the support set and the algorithm updates the approximation for the sparse vector by projecting $\tilde{\mathbf{y}}_{kj}$ orthogonally onto the columns of $\tilde{\mathbf{X}}$ associated with the current support set. The proposed OMP algorithm is summarized in Algorithm 1. The performance of the OMP algorithm depends on the selection of the best column's index from the measurement matrix $\tilde{\mathbf{X}}$. According to the CS literature, there is a high probability that OMP algorithm fails to select a correct element at the first iteration for certain random matrices [67]. To address this issue, in this chapter, we develop a BFS algorithm to acquire CSI. The BFS algorithm depends on the cost function. In our scenario, since we are interested in CSI estimation, we choose least squares as the cost function. The BFS algorithm selects the best features from the measurement matrix $\tilde{\mathbf{X}}$ based on ℓ_1 norm minimization and updates the feature set at each iteration. The proposed BFS algorithm is summarized in Algorithm 2.

Algorithm 2: BFS algorithm for CSI recovery

Input : $\tilde{\mathbf{Y}}_k, \tilde{\mathbf{X}}, \mathbf{A}_T, \mathbf{A}_R$ and sparsity level (s)
Output : $\hat{\mathbf{H}}_k^{dl}$

- **Initialization** : initialize feature set $F = \{1, \ldots, N_t\}$.
- **for** $i = N_t - s,$ **do**
- compute $(\tilde{\mathbf{H}}_k^{dl})_F = \tilde{\mathbf{X}}_F^{\dagger} \tilde{\mathbf{Y}}_k$
- compute $||\mathbf{h}_i||_1 = \sum_{l=1}^{N_r} |\mathbf{h}_{li}|$, where $\mathbf{h}_i$ is the i-th column of $(\tilde{\mathbf{H}}_k^{dl})_F^H$.
- update $F = F - \{i : \min ||\mathbf{h}_i||_1\}$
- **end** (**for**)
- $\hat{\mathbf{H}}_k^{dl} = \mathbf{A}_R (\tilde{\mathbf{H}}_k^{dl})^H \mathbf{A}_T^H$, where $(\tilde{\mathbf{H}}_k^{dl})^F = \tilde{\mathbf{X}}_F^{\dagger} \tilde{\mathbf{Y}}_k$, $(\tilde{\mathbf{H}}_k^{dl})^{[N_t] \backslash F} = 0$

3.5 Uplink CSIT Feedback

As we discussed earlier, in FDD multi-user massive MIMO systems, each user measures the downlink MIMO channel through the reference signals/pilots, and then feeds back the CSI to the BS using codebook-based channel feedback [76]. In this work, we consider the limited feedback scheme used in FDD LTE/LTE-Advanced standards. The CSI feedback in LTE/LTE-Advanced standards usually contains two kinds of information: the channel direction information (CDI), which is related to the eigen-directions of the underlying MIMO channel, and the channel quality indicator (CQI), which is related to the strength of the corresponding spatial directions.

In limited feedback systems such as multi-user massive MIMO systems, each user feeds back its CQI and CDI to the BS. CQI feedback is relatively straightforward where user k feeds back the quantized version of the singular values using scalar or vector quantization methods. CDI feedback is generally more involved relying on codebook-based feedback. After obtaining each

of the eigen-directions of the MIMO channel, say $\tilde{\mathbf{h}}_k^{(\ell)} = \mathbf{h}_k^{(\ell)}/\|\mathbf{h}_k^{(\ell)}\|$ for the ℓ-th eigen-direction, $\forall \ell \in \{1,\ldots,N_r\}$, user k quantizes it to $\hat{\mathbf{h}}_k^{(\ell)}$ using a random vector quantization codebook $\mathscr{C}$, which is known to both the BS and the user. In general, there will be multiple CDIs of the underlying MIMO channel; the exact number of CDIs to be fed back will impact the system performance and the feedback overhead. The indices of the quantized CDIs will be sent to the BS through a feedback link. Accordingly, the BS obtains $\hat{\mathbf{h}}_k^{(\ell)}$ and then uses these for downlink MIMO precoding. A quantization codebook $\mathscr{C}$ consisting of 2^β N_t-dimensional unit norm vectors is given by $\mathscr{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{2^\beta}\}$, where 2^β is the codebook size, β is the number of feedback bits per user, and $\mathbf{c}_m \in \mathbb{C}^{N_t \times 1}$ is a unit norm codeword, i.e., $\|\mathbf{c}_m\|^2 = 1$.

Each user chooses the CDI in the codebook that is closet to its eigen-direction, where closeness is usually measured in terms of the angle between the eigen-direction and the codeword in the codebook or equivalently the inner product. Hence, user k computes quantization index $q_k^{(\ell)}$ according to

$$q_k^{(\ell)} = \underset{m=1,\ldots,2^\beta}{\arg\max} |\langle \tilde{\mathbf{h}}_k^{(\ell)}, \mathbf{c}_m \rangle| = \underset{m=1,\ldots,2^\beta}{\arg\min} \sin^2(\angle(\tilde{\mathbf{h}}_k^{(\ell)}, \mathbf{c}_m)),$$

and feeds the index back to the BS. Upon reception of the index $q_k^{(\ell)}$, the BS can recover the CDI by searching for the corresponding entry in the codebook. In this way, the feedback overhead is significantly reduced with a penalty of quantization error in the finite rate feedback systems [77].

3.6 Performance Evaluation of Channel Estimation Techniques

In this section, we investigate the performance of channel estimation techniques that we discussed in Section 3.4.1 and compare the results of the proposed BFS algorithm with the CS based OMP

algorithm and traditional LS/MMSE estimators through Monte Carlo simulations. As a performance metric, we use normalized mean square error (NMSE) to evaluate the performance of each estimator. The NMSE for each user is defined as $\mathbb{E}\left(\frac{||\mathbf{H}_k^{dl}-\hat{\mathbf{H}}_k^{dl}||_F^2}{||\mathbf{H}_k^{dl}||_F^2}\right)$. The antenna spacing for both the transmit and receive antennas, at the BS and the user, is assumed to be 0.5λ. The number of receive antennas at the user is set to be 8. The number of transmit antennas at the BS is set to be 64. The average transmit SNR is 30 dB. For simplicity, we assume that all users have the same channel sparsity level, which is $s = 8$. First, we evaluate three methods to generate the CCM for an arbitrary user. The CCM for k-th user ($\mathbf{R}_k$) can be modeled by:

- One-ring model (OR): The (i, j)-th element of $\mathbf{R}_k$ for a uniform linear array (ULA) is given by [78, 79]

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{2\Delta}\int_{-\Delta+\theta_k}^{\Delta+\theta_k} e^{-j2\pi\frac{d}{\lambda}(i-j)\sin(x)}dx \tag{3.16}$$

  Where $\Delta, \alpha_k, \theta_k, d,$ and λ refer to the angular spread, average large-scale fading, mean angle of arrival (AoA) of the k-th user in the azimuth direction, antenna spacing, and carrier wavelength, respectively.

- Laplacian model: The (i, j)-th element of $\mathbf{R}_k$ for a uniform linear array (ULA) is given by [80, 81]

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{\sqrt{2}\Delta}\int_{-\pi+\theta_k}^{\pi+\theta_k} e^{-\frac{\sqrt{2}}{\Delta}|x-\theta_k|-j2\pi\frac{d}{\lambda}(i-j)\sin(x)}dx \tag{3.17}$$

  where $\Delta, \alpha_k, \theta_k, d,$ and λ are defined in the OR model.

- Exponential model: The (i,j)-th element of $\mathbf{R}_k$ is given by [82, 83]

$$[\mathbf{R}_k(\rho,\theta_k)]_{i,j} = \begin{cases} \left(\rho e^{j\theta_k}\right)^{j-i} & i \leq j \\ \left(\rho e^{-j\theta_k}\right)^{i-j} & i > j \end{cases} \tag{3.18}$$

  where ρ is the correlation factor, $0 \leq \rho \leq 1$. For simplicity, ρ is the same for all users while the AoA (θ_k) is different and uniformly distributed in $[-\pi,\pi)$.

Figure 3.2 shows the cumulative distribution function (CDF) for the three models discussed above under different parameters. We use the empirical CDF, $\hat{F}(x) = \frac{\sum_i \lambda_i \leq x}{n}$, where $\lambda_i, i = 1,\ldots,n$ are the singular values of the CCM $\mathbf{R}_k \in \mathbb{C}^{n\times n}$. The figure shows that a large number of singular values are close to zero, while a few are large for both OR and Laplacian models. The exponential model provides a slight variation in its singular values. Figure 3.3 shows the distribution of singular values for the three models discussed above. From this figure, we observe that both OR and Laplacian model curves are L-shaped, where many singular values are close to zero. This suggests their respective CCMs are not full rank and the effective rank (ER), (number of singular values that contribute to 99 percent of the total singular values), for $\Delta = \pi/12, d = 0.5\lambda,$ and $\rho = 0.6$, varies between 1 and 50, 1 and 106, and 1 and 491 for the OR, Laplacian and exponential models, respectively (observe the corner of the L-curve). On the other hand, the exponential model provides full-rank CCM because it has a slow variation in its singular values. In what follows, we consider the exponential model with correlation factor $\rho = 0.6$.

Figure 3.4 shows the average NMSE of each estimator vs. the training overhead (T). We observe that LS/MMSE estimators have poor performance due to the large number of BS antennas. The OMP algorithm shows better performance than traditional LS/MMSE estimators because it

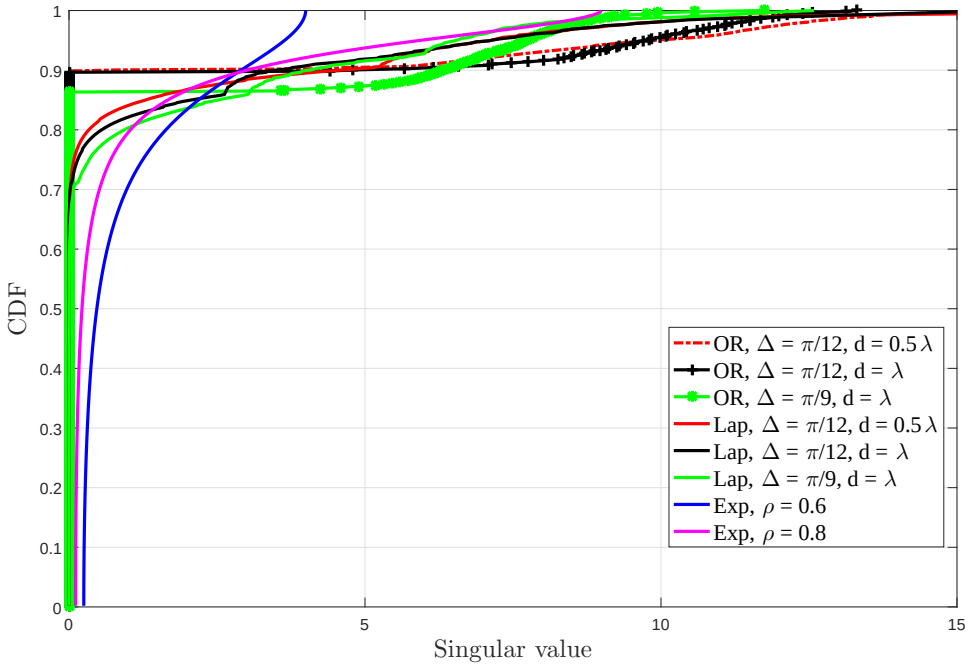

Figure 3.2: The CDF of the singular values of CCM for an arbitrary user with $N_t = 64, N_r = 8, \alpha = 1$ and $\theta = \pi/3$.

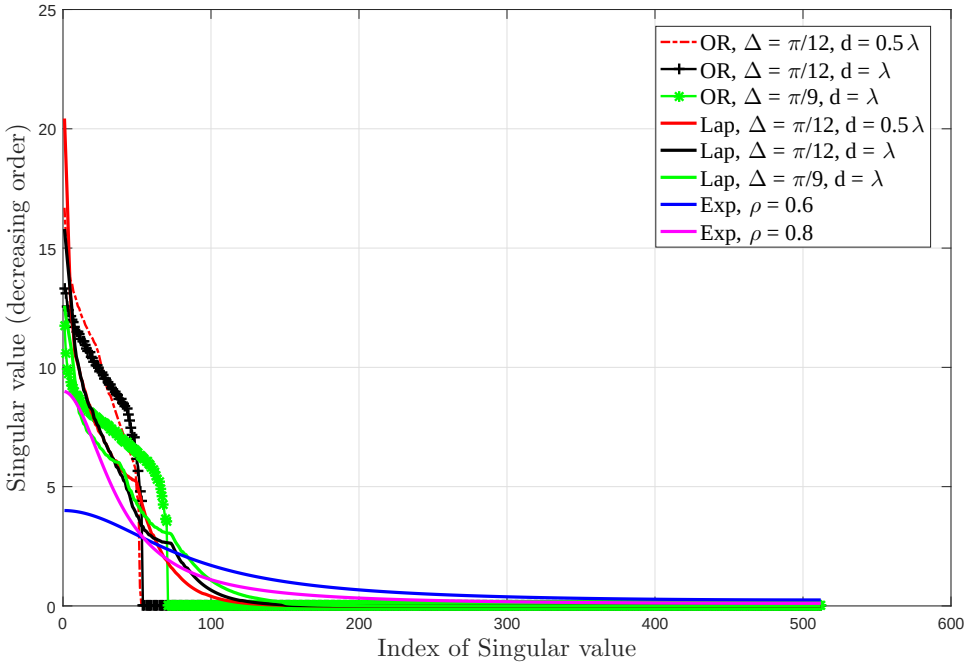

Figure 3.3: The distribution of the singular values of CCM for an arbitrary user with $N_t = 64, N_r = 8, \alpha = 1$ and $\theta = \pi/3$.

exploits the sparsity structure of the downlink channel matrix, but it has poor performance at the first iteration, which is a well-known property of the CS based greedy algorithm. Moreover, we observe that the BFS algorithm has better performance than the OMP algorithm especially at the first iteration because of its ability to select the best features from the measurement matrix based on ℓ_1 norm minimization.

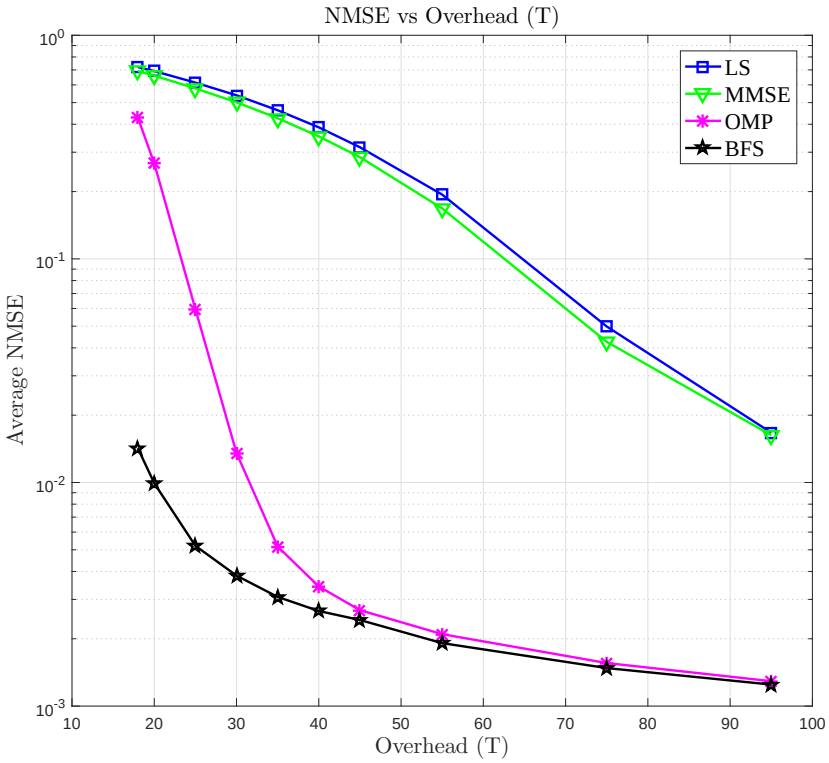

Figure 3.4: The normalized MSE performance of CSI vs. downlink training overhead (T) for $N_t = 64$ antennas at the BS.

In Figure 3.5, we compare the average NMSE of each estimator versus the transmit SNR. In this figure, we set the sparsity level to $s = 8$ and $T = 30$. The figure shows the proposed algorithm outperforms the traditional estimation methods and the CS based OMP algorithm under the same training length. Moreover, it shows the proposed algorithm achieves a better performance gain in

the high SNR region.

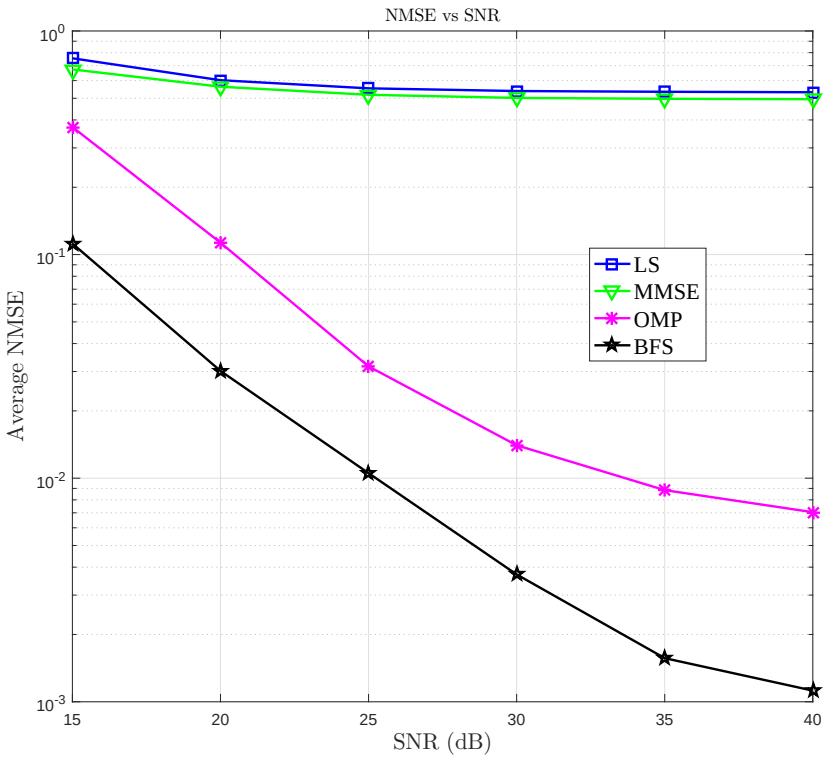

Figure 3.5: The normalized MSE performance of CSI vs. transmit SNR for $T = 30, s = 8, N_t = 64, N_r = 8$ and $K = 40$.

The average NMSE performance of each estimator versus the sparsity level is depicted in Figure 3.6. In this figure, we set the training length to $T = 30$ and the transmit SNR $= 30$dB. The figure shows the performance of traditional LS/MMSE estimators is not affected by the sparsity level, while channel estimation based on the CS approach depends on the sparsity level. We observe that as the sparsity level increases, the average NMSE increases and the proposed algorithm outperforms the traditional CS based OMP algorithm.

Figure 3.7 shows the average NMSE of each estimator vs. the number of transmit antennas at the BS. From this figure, we observe that as the number of transmit antennas at the BS increases,

the average NMSE increases and the proposed algorithm outperforms the traditional CS based OMP algorithm.

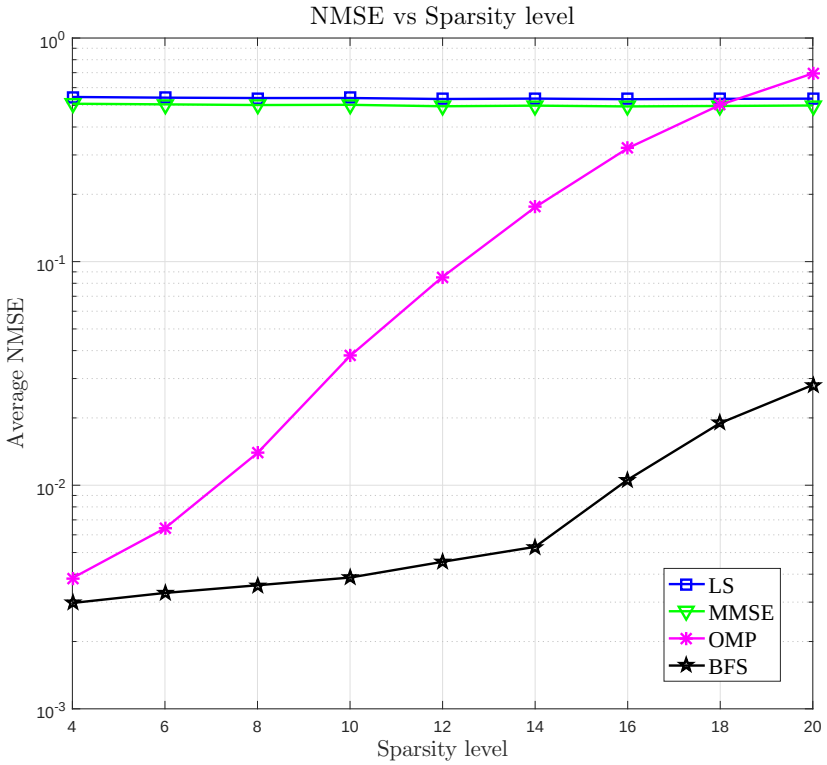

Figure 3.6: The normalized MSE performance of CSI vs. sparsity level for $T = 30, N_t = 64, N_r = 8$ and $K = 40$.

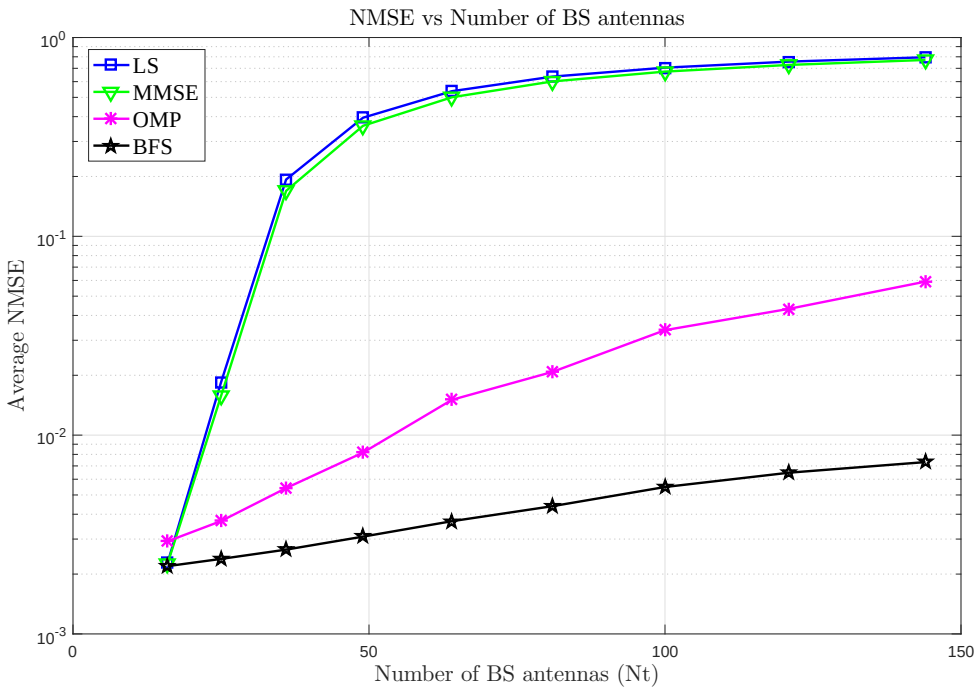

Figure 3.7: The normalized MSE performance of CSI vs. number of transmit antennas at the BS for $T = 30, N_r = 8$ and $K = 40$.

Chapter 4

Downlink Achievable Rate Analysis for FDD massive MIMO systems under Limited Feedback

4.1 Introduction

Massive multiple input multiple output (MIMO) (also called Large-Scale Antenna) systems are considered to be a promising technique for 5G cellular networks. In massive MIMO systems, a base station (BS) equipped with hundreds or thousands of active antenna elements (M) communicate with a limited number of scheduled users (K) such that $M/K > 10$ [30, 84, 85]. Due to the large number of antennas used at the BS, massive MIMO can significantly increased the system capacity by orders of magnitude compared to current 4G wireless systems and improve both energy consumption and spectrum efficiency. Another benefit of massive MIMO systems is alleviating inter-user interference with a simple linear precoder and receiver combiner such as

zero-forcing (ZF), matched filter (MF), and minimum mean square error(MMSE). Massive MIMO systems depends on spatial multiplexing, which requires the knowledge of the channel transfer function on both the uplink and downlink [86, 32, 87]. To take advantage of massive MIMO systems, knowledge of channel state information at the transmitter (CSIT) is essential in order to perform downlink beamforming and resource allocation [88, 42, 89]. In time-division duplexing (TDD) systems, the base station (BS) acquires CSI via uplink channel reciprocity, where uplink training symbols scale linearly with the number of users, which is much less than the number of the BS antennas. However, in TDD massive MIMO systems, the CSI estimation may be inaccurate because of calibration errors in the radio frequency (RF) chains [46, 47, 90]. In frequency-division duplex (FDD) systems, uplink and downlink channels fade independently if the frequency separation between them is large enough compared to the coherence bandwidth of the channel. As a result, the uplink channel information may not be applied directly for downlink beamforming. However, there are some properties that can be exploited for downlink beamforming from uplink channel information [88, 91]. In FDD massive MIMO systems, the BS obtains CSIT via uplink feedback. The length of the downlink pilot training is proportional to the number of transmit antennas, which is challenging in FDD massive MIMO systems due to the large number of transmit antennas at the BS. The experimental results for massive MIMO systems show that the downlink channel matrix exhibits sparsity in the angular domain representation due to limited local scatterers at the BS [88, 81]. In order to reduce the downlink training overhead, the authors in [42] exploited the hidden joint sparsity structure of the user channel matrices and developed a joint orthogonal matching pursuit (OMP) algorithm to recover CSIT at the BS jointly. The authors in [92] developed an optimal block OMP algorithm to estimate CSI at the BS based on user grouping in order to reduce the pilot overhead. They assumed that users in each group have the same channel cor-

relation matrix (CCM), which results in beam block sparsity for the channel of each group. The authors in [93] proposed a two stage algorithm to obtain downlink CSI for FDD massive MIMO systems by exploiting the common and individual sparsity structure of the angular domain representation of the downlink channel matrix. In the first stage, the algorithm extracts the common and individual channel support of each user channel matrix. The second stage obtains the support information for CSI acquisition based on weighted block ℓ_1-minimization. Most of the research that adopts the angular domain representation of the downlink channel matrix in FDD multi-user massive MIMO systems obtains the CSI jointly for all users at the BS based on their respective training measurement feedback. However, such feedback is not applicable in practice.

The performance of MIMO systems with limited feedback has been studied extensively in the literature. In [94], the author used a zero-forcing precoder to perform downlink transmission based on quantized downlink channels from each receiver, where each receiver sends only $\mathscr{B}$ feedback bits to the base station. The downlink achievable rate is analyzed for the system by using random vector quantization (RVQ) as a codebook. The authors in [95] assume that the BS is equipped with a uniform rectangle antenna array. They proposed a new codebook design by exploiting the array correlation structure. A low complexity precoding scheme was proposed in [77] for multi-cell LTE-advanced systems. They designed a downlink beam-forming scheme based on limited feedback to mitigate inter-cell interference. The authors in [96] proposed a strategy for a limited feedback scheme for massive MIMO systems called hybrid limited feedback with selective eigenvalue information (HLFSEI). They fed back only quantized channel information of selective eigenvalues to the BS and designed an optimal feedback allocation that maximizes the system capacity. The authors in [97] proposed an algorithm that minimized the upper bound of rate loss in MIMO uplink networks with a finite rate feedback channel. A new channel quantization scheme

was proposed in [98] for MIMO systems to improve the performance of MIMO channels in terms of interference alignment with limited feedback. They designed additional filters for each user that minimize the chordal distance to decrease the channel quantization error. However, all previous works consider a TDD system and the receiver has perfect knowledge of CSI. In FDD systems, the BS acquires the CSI via the feedback link. However, the knowledge of downlink CSI is much more complex in FDD massive MIMO systems due to the large dimensionality of the channel matrix, which leads to large overhead and degrades system performance [99, 100].

In this chapter, we investigate the problem of limited feedback for a downlink multi-user massive MIMO network and design an efficient limited feedback scheme. By exploiting our previous work for channel estimation of FDD multi-user massive MIMO systems [101], we are able to analyze the system performance in terms of downlink achievable rate.

This research differs from the previously mentioned studies particularly in its aim to reduce the training and feedback overhead further by exploiting the sparsity structure of the massive MIMO downlink channel matrix. Furthermore, this work also takes into account the effects of the finite-rate codebook based feedback adopted in the frequency division duplex (FDD) LTE-Advanced system. The main contributions of this research can be summarized as follows.

- We consider FDD multi-user massive MIMO systems with a correlated channel model and imperfect CSI. Most of the existing work considers TDD systems with perfect CSI available at the receiver. Channel estimation during downlink training phase is done based on our previous work.

- We propose a simple downlink transmission scheme based on limited feedback. To be specific, each user estimates its downlink channel during the downlink training phase. Next,

each user quantizes its channel estimate to $\mathscr{B}$ bits and feeds back a limited CSI to the BS. The BS receives $\mathscr{B}$ feedback bits from each user and uses our proposed scheme to perform downlink precoding based on the channel quantizations. The proposed scheme takes into account spatial channel correlation, the effects of finite-rate codebook based limited feedback adopted in the LTE-Advanced system, channel quantization errors, and CSI errors.

- We generate numerical results to evaluate the system performance. We consider the one-ring channel model, the Laplacian channel model, and the exponential channel model. We use the downlink achievable rate as a performance metric to compare our proposed limited feedback scheme in FDD multi-user massive MIMO systems in terms of imperfect CSI under various system parameters.

4.2 System Model

Consider a flat and block-fading FDD multi-user massive MIMO system with one BS equipped with N_t antennas that serves K non-cooperating users. Each user has N_r antennas. It is assumed that each co-scheduled user operating in the FDD MU-MIMO mode receives only one spatial stream (rank 1 transmission) as specified by the Rel-10 LTE-Advanced standard [102, Chapter 11]. A spatially correlated flat Rayleigh fading channel is assumed in this chapter where the elements of $\mathbf{H}_k^{dl}$ are modeled as i.i.d. complex Gaussian variables with zero-mean and spatial CCM. It is worth mentioning that in this chapter we assume users have different spatial CCM. Therefore, we compute the spatial CCM for each user based on different models including the one-ring model, the Laplacian model, and the exponential model (more details are discussed in Section 5.4). Moreover, we consider the angular domain representation of the downlink channel matrix [72, 42], where the

downlink channel matrix for the k-th user $\mathbf{H}_k^{\mathrm{d}l}$ can be represented as:

$$\mathbf{H}_k^{\mathrm{d}l} = \mathbf{A}_R \mathbf{H}_k^w \mathbf{A}_T^H \tag{4.1}$$

where $\mathbf{H}_k^w \in \mathbb{C}^{N_r \times N_t}$ is the angular domain channel matrix for the k-th user and $\mathbf{A}_R \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{A}_T \in \mathbb{C}^{N_t \times N_t}$ are defined as unitary matrices for the angular domain transformation at the UE and BS respectively. Based on [42, 47, 54], the angular domain channel matrix $\mathbf{H}_k^w$ is sparse due to limited local scattering effects at the BS as shown in Fig. 4.1. The (i, j)-th nonzero entry of $\mathbf{H}_k^w$ indicates there is a spatial path from the j-th transmit direction of BS to the i-th receive direction of the UE. We define the i-th row vector of $\mathbf{H}_k^w$ as $\mathbf{h}_{ki}$ and $\mathrm{supp}(\mathbf{h}_{ki})$ denotes the index set of nonzero elements in $\mathbf{h}_{ki}$. Due to the fact that the row vectors of $\mathbf{H}_k^w$ usually have the same sparsity, the following assumption for $\mathbf{H}_k^w$ is considered as in [42]:

Assumption 2. *Define $\mathbf{h}_{ki}$ as the i-th row vector of $\mathbf{H}_k^w$, then there exists an index set $\Omega_k, 0 < |\Omega_k| \ll N_t, \forall k$, such that*

$$\mathrm{supp}(\mathbf{h}_{k1}) = \mathrm{supp}(\mathbf{h}_{k2}) = \cdots = \mathrm{supp}(\mathbf{h}_{kN_r}) \stackrel{\Delta}{=} \Omega_k \tag{4.2}$$

The BS plans to communicate a symbol vector $\mathbf{s} = [s_1, \ldots, s_K]^T \in \mathbb{C}^{K \times 1}$ to its associated receivers, where s_k is the transmit symbol from the BS to the k receiver with unit power of $\mathbb{E}\{|s_k|^2\} = 1$. Prior to transmitting, the BS linearly precodes its downlink data vector $\mathbf{x} = \sum_{k=1}^K \sqrt{\rho_k}\mathbf{w}_k s_k$ where ρ_k stands for the transmit power allocated to user k; $\mathbf{w}_k$ denotes the beam-former that the BS uses to transmit the signal s_k to user k; and $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_k, ..., \mathbf{w}_K] \in \mathbb{C}^{N_t \times K}$ indicates the transmit precoding matrix with unit-norm columns. Hence, the transmit power constraint at the BS can be

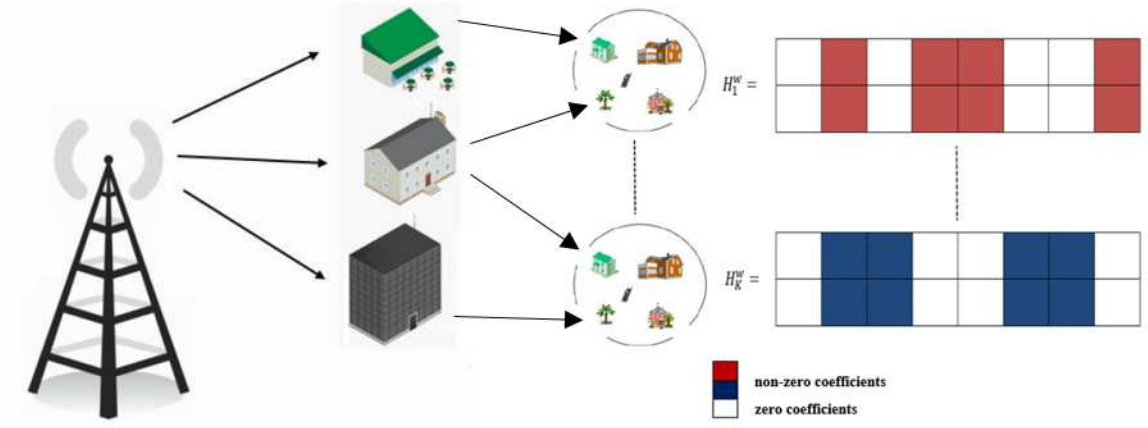

Figure 4.1: Illustration of downlink channel sparsity.

expressed as $p = \mathbb{E}\{\|\mathbf{x}\|^2\} = \sum_{k=1}^{K} \rho_k \mathbb{E}\{\|\mathbf{w}_k\|^2\} = \sum_{k=1}^{K} \rho_k$. Under our assumptions, the received signal vector $\mathbf{y}_k \in \mathbb{C}^{N_r \times 1}$ at user k can be written as:

$$\begin{aligned} \mathbf{y}_k^{\mathrm{d}l} &= \mathbf{H}_k^{\mathrm{d}l} \mathbf{x}^{\mathrm{d}l} + \mathbf{n}_k^{\mathrm{d}l} \\ &= \underbrace{\sqrt{\rho_k} \mathbf{H}_k^{\mathrm{d}l} \mathbf{w}_k s_k}_{\text{desired signal}} + \underbrace{\sum_{i=1, i \neq k}^{K} \sqrt{\rho_i} \mathbf{H}_k^{\mathrm{d}l} \mathbf{w}_i s_i}_{\text{intra-cell interference}} + \mathbf{n}_k^{\mathrm{d}l} \end{aligned} \tag{4.3}$$

where $\mathbf{n}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times 1}$ is a complex Gaussian noise vector, i.e., $\mathbf{n}^{\mathrm{d}l} \sim \mathscr{CN}(0, \mathbf{I}_{N_r})$, and $\mathbf{H}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times N_t}$ is the downlink channel matrix for the k-th user with zero mean and spatial CCM $\left(\mathbf{R}_k \in \mathbb{C}^{N_r N_t \times N_r N_t}\right)$, which is defined on the column stacking of the downlink channel matrix. Thus, $\mathrm{vec}(\mathbf{H}_k^{\mathrm{d}l}) \in \mathscr{CN}(0, \mathbf{R}_k)$. Accordingly, the signal-to-interference-plus-noise ratio (SINR) at the k^{th} user can

be expressed as:

$$\begin{aligned} \mathrm{SINR}_k &= \frac{\rho_k |\hat{\mathbf{H}}_k^{\mathrm{d}l} \mathbf{w}_k|^2}{\sigma_k^2 + \sum_{i=1;i\neq k}^{K} \rho_i |\hat{\mathbf{H}}_k^{\mathrm{d}l} \mathbf{w}_i|^2} \\ &= \frac{\rho_k \mathrm{Tr}\left(\mathbf{w}_k^H \hat{\mathbf{H}}_k^{\mathrm{d}l\,H} \hat{\mathbf{H}}_k^{\mathrm{d}l} \mathbf{w}_k \right)}{\sigma_k^2 + \sum_{i=1;i\neq k}^{K} \rho_i \mathrm{Tr}\left(\mathbf{w}_i^H \hat{\mathbf{H}}_k^{\mathrm{d}l\,H} \hat{\mathbf{H}}_k^{\mathrm{d}l} \mathbf{w}_i \right)} \end{aligned} \tag{4.4}$$

and the downlink achievable transmission rate for the k^{th} user can be expressed as $R_k = \log_2(1 + \mathrm{SINR}_k)$, Therefore, the sum rate of the whole network can be written as:

$$R_{\mathrm{sum}} = \sum_{k=1}^{K} \log 2 \left(1 + \frac{\rho_k \mathrm{Tr}\left(\mathbf{w}_k^H \mathbf{H}_k^{\mathrm{d}l\,H} \mathbf{H}_k^{\mathrm{d}l} \mathbf{w}_k \right)}{\sigma_k^2 + \mathbf{I}_k} \right) \tag{4.5}$$

where $\mathbf{I}_k = \sum_{i=1;i\neq k}^{K} \rho_i \mathrm{Tr}\left(\mathbf{w}_i^H \mathbf{H}_k^{\mathrm{d}l\,H} \mathbf{H}_k^{\mathrm{d}l} \mathbf{w}_i \right)$ denotes the multi-user interference caused by other users in the cell.

In order to calculate the throughput, first we estimate the channel and then based on the channel matrix we will find the optimum beamformers. To design the optimal linear precoding scheme, it is often desirable to maximize the output SINR for each user. However, this problem is known to be challenging due to its coupled nature and unavailability of closed-form solutions. A more tractable but suboptimal design is to enforce a zero-CCI requirement for each user, such as zeroforcing (ZF) beam-forming [103, 104]. However, the sum-rate of ZF beam-forming under codebook-based feedback usually suffers significantly from the feedback error [105]. In [106], the authors introduced the concept of signal-to-leakage-and-noise ratio (SLNR) as the optimization metric for a linear precoder design. This metric transforms a coupled optimization problem into a completely decoupled one, for which a closed-form solution is available. Therefore, in this chapter we will

maximize the SLNR with respect to the beamforming vectors. To further improve the system capacity, we adopt the coordinated beam-forming (CB) algorithm proposed in [107] that jointly optimizes the transmit and receive beam-forming via iteration. Also, we consider the existence of quantization errors in designing the beamformer vectors. In the next section we will estimate the downlink channel. Knowing the channel, the optimum beamformer, in the sense of maximizing SLNR, will be calculated.

4.3 Downlink Channel Estimation

In this section, CSI estimation techniques using traditional methods such as least squares (LS), minimum mean squared error (MMSE), compressive sensing (CS), and backward feature selection (BFS) are considered. During the downlink training phase, the BS broadcasts a sequence of $T < \tau$ pilot vectors $[\mathbf{x}(1)\dots\mathbf{x}(t)\dots\mathbf{x}(T)]$, where τ is the channel coherence interval and data will be transmitted for a fraction $\left(\frac{\tau-T}{\tau}\right)$. Moreover, $\mathbb{E}||\mathbf{x}(t)||^2 = P$, $t = 1,\dots,T$, where P is the average transmit power at the BS per training time slot. The received signal vector for k-th user in the t-th time slot, i.e. $\mathbf{y}_k(t) \in \mathbb{C}^{N_r \times 1}$ can be written as:

$$\mathbf{y}_k(t) = \mathbf{H}_k^{\mathrm{d}l}\mathbf{x}(t) + \mathbf{n}_k(t), \;\; t = 1,\dots,T \tag{4.6}$$

where $\mathbf{H}_k^{\mathrm{d}l} \in \mathbb{C}^{N_r \times N_t}$ is the quasi-static block fading downlink channel matrix, and $\mathbf{n}_k(t) \in \mathbb{C}^{N_r \times 1}$ is a complex Gaussian noise vector with $i.i.d.$ entries having zero mean and unit variance. Let $\mathbf{X} = [\mathbf{x}(1)\dots\mathbf{x}(T)] \in \mathbb{C}^{N_t \times T}$, $\mathbf{Y}_k = [\mathbf{y}_k(1)\dots\mathbf{y}_k(T)] \in \mathbb{C}^{N_r \times T}$ and $\mathbf{N}_k = [\mathbf{n}_k(1)\dots\mathbf{n}_k(T)] \in \mathbb{C}^{N_r \times T}$, respectively, be the pilot matrix with total power constraint $\mathrm{tr}(\mathbf{X}\mathbf{X}^H) = PT$, received signal at the

k-th user during the downlink training phase, and the noise matrix. The received signal in (4.6) can be equivalently represented as:

$$\mathbf{Y}_k = \mathbf{H}_k^{dl}\mathbf{X} + \mathbf{N}_k, \;\; k = 1, \ldots, K \tag{4.7}$$

The CSI estimation based LS approach can be expressed as [73]:

$$\tilde{\mathbf{H}}_{k,LS}^{dl} = \mathbf{Y}_k \mathbf{X}^{\dagger} \tag{4.8}$$

where $\mathbf{X}^{\dagger} = \mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1}$ is the pseudoinverse of $\mathbf{X}$ and $(\cdot)^H$ denotes the Hermitian transpose. By vectorizing the received signal in (4.7) and applying $\text{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\text{vec}(\mathbf{B})$, the received training signal at k-th user can be expressed as:

$$\text{vec}(\mathbf{Y}_k) = \tilde{\mathbf{X}}\text{vec}(\mathbf{H}_k^{dl}) + \text{vec}(\mathbf{N}_k) \tag{4.9}$$

where $\tilde{\mathbf{X}} \triangleq (\mathbf{X}^T \otimes \mathbf{I}_{N_t})$, the MMSE estimator $\tilde{\mathbf{H}}_{k,MMSE}^{dl}$ of the downlink channel matrix for k-th user can be written as [74]:

$$\text{vec}(\tilde{\mathbf{H}}_{k,MMSE}^{dl}) = \mathbf{R}_k\tilde{\mathbf{X}}^H(\tilde{\mathbf{X}}\mathbf{R}_k\tilde{\mathbf{X}}^H + \mathbf{I}_{TN_t})^{-1}\text{vec}(\mathbf{Y}_k) \tag{4.10}$$

The traditional channel estimation methods such as LS and MMSE require that the optimal training length be at least the same as the number of transmit antennas [75], $T \geq N_t$, which leads to large overhead in downlink training and CSI feedback for FDD multi-user massive MIMO systems due to large number of BS antennas. Therefore, these methods are not efficient for FDD multi-user

massive MIMO systems. Since the downlink channel matrix exhibits sparsity structure in the angular domain, channel estimation based on the CS approach provides efficient CSI acquisition by utilizing prior knowledge of the downlink channel sparsity structure so that the overhead reduction can be achieved [46, 54]. In the CS method, the BS sends compressive training symbols $\mathbf{X} \in \mathbb{C}^{N_t \times T}$ with $T \ll N_t$ during downlink training phase, then users estimate their respective downlink CSI and feed it back to the BS during uplink feedback. By using Equation (4.1), the received signal for the k-th user during the downlink training phase in (4.7) can be rewritten into the noisy CS recovery problem as:

$$\tilde{\mathbf{Y}}_k = \tilde{\mathbf{X}}\tilde{\mathbf{H}}_k^{\mathrm{d}l} + \tilde{\mathbf{N}}_k \tag{4.11}$$

where,

$$\begin{aligned}
\tilde{\mathbf{Y}}_k &= \sqrt{\frac{N_t}{PT}}\mathbf{Y}_k^H \mathbf{A}_R \quad \in \mathbb{C}^{T \times N_r} \\
\tilde{\mathbf{X}} &= \sqrt{\frac{N_t}{PT}}\mathbf{X}^H \mathbf{A}_T \quad \in \mathbb{C}^{T \times N_t} \\
\tilde{\mathbf{H}}_k^{\mathrm{d}l} &= (\mathbf{H}_k^w)^H \quad \in \mathbb{C}^{N_t \times N_r} \\
\tilde{\mathbf{N}}_k &= \sqrt{\frac{N_t}{PT}}\mathbf{N}_k^H \mathbf{A}_R \quad \in \mathbb{C}^{T \times N_r}
\end{aligned} \tag{4.12}$$

Assuming the channel state information is given by either Algorithm 3 or Algorithm 4, the problem of finding the optimum beamforming vectors can be formulated as follows.

Algorithm 3: OMP algorithm for CSI recovery

Input : $\tilde{\mathbf{Y}}_k, \tilde{\mathbf{X}}, \mathbf{A}_T, \mathbf{A}_R$ and sparsity level (s)
Output : $\hat{\mathbf{H}}_k^{dl}$

- **Initialization** : $r_k^0 = \tilde{\mathbf{y}}_{kj}, \tilde{\mathbf{h}}_{kj}^{dl} = 0$ where $\tilde{\mathbf{y}}_{kj}$ and $\tilde{\mathbf{h}}_{kj}^{dl}$ is the j-th column of $\tilde{\mathbf{Y}}_k$, $\tilde{\mathbf{H}}_k^{dl}$ respectively, $j = 1 : N_r$, $\text{supp}_k^0 = \emptyset$
- **while** $i \leq s$ **do**
- $l_k^i = \arg\max |\langle r_k^{(i-1)}, \tilde{\mathbf{X}} \rangle|$
- $\text{supp}_k^i = \text{supp}_k^{i-1} \cup l_k^i$
- $(\tilde{\mathbf{h}}_{kj}^{dl})^i = (\tilde{\mathbf{X}}_{\text{supp}_k^i})^\dagger \tilde{\mathbf{y}}_{kj}$
- $r_k^i = \tilde{\mathbf{y}}_{kj} - \tilde{\mathbf{X}}_{\text{supp}_k^i} (\tilde{\mathbf{h}}_{kj}^{dl})^i$
- **end** (**while**)
- $\hat{\mathbf{H}}_k^{dl} = \mathbf{A}_R (\tilde{\mathbf{H}}_k^{dl})^H \mathbf{A}_T^H$

$$\begin{aligned} \max_{\mathbf{w}_k} \quad & \sum_{k=1}^{K} \text{SLNR}_k \\ \text{s.t.} \quad & \sum_{k=1}^{K} \rho_k \leq P \end{aligned} \tag{4.13}$$

where SLNR_k can be written as

$$\text{SLNR}_k = \frac{\text{Tr}\big(\mathbf{w}_k^H \hat{\mathbf{H}}_k^{dl^H} \hat{\mathbf{H}}_k^{dl} \mathbf{w}_k\big)}{\frac{\sigma_k^2}{\rho_k} + \text{Tr}\big(\mathbf{w}_k^H (\sum_{i \neq k} \hat{\mathbf{H}}_i^{dl^H} \hat{\mathbf{H}}_i^{dl}) \mathbf{w}_k\big)} \tag{4.14}$$

The following result is adapted from [77] to express the optimal precoders in a closed form solution.

Proposition 1. *The optimal downlink beam-forming vector* $\mathbf{w}_k$ *that the base station uses to send*

Algorithm 4: BFS algorithm for CSI recovery

Input : $\tilde{\mathbf{Y}}_k, \tilde{\mathbf{X}}, \mathbf{A}_T, \mathbf{A}_R$ and sparsity level (s)
Output : $\hat{\mathbf{H}}_k^{\mathrm{d}l}$

- **Initialization** : initialize feature set $F = \{1, \ldots, N_t\}$.
- **for** $i = N_t - s,$ **do**
- compute $(\tilde{\mathbf{H}}_k^{\mathrm{d}l})_F = \tilde{\mathbf{X}}_F^{\dagger} \tilde{\mathbf{Y}}_k$
- compute $||\mathbf{h}_i||_1 = \sum_{l=1}^{N_r} |\mathbf{h}_{li}|$, where $\mathbf{h}_i$ is the i-th column of $(\tilde{\mathbf{H}}_k^{\mathrm{d}l})_F^H$.
- update $F = F - \{i : \min ||\mathbf{h}_i||_1\}$
- **end** (**for**)
- $\hat{\mathbf{H}}_k^{\mathrm{d}l} = \mathbf{A}_R (\tilde{\mathbf{H}}_k^{\mathrm{d}l})^H \mathbf{A}_T^H$, where $(\tilde{\mathbf{H}}_k^{\mathrm{d}l})^F = \tilde{\mathbf{X}}_F^{\dagger} \tilde{\mathbf{Y}}_k$, $(\tilde{\mathbf{H}}_k^{\mathrm{d}l})^{[N_t] \backslash F} = 0$

data to the k^{th} user in order to maximize the signal-to-leakage-plus noise ratio can be expressed as

$$\mathbf{w}_k^{opt} = \upsilon_{\max} \left\{ \left(\frac{\sigma_k^2}{\rho_k} \mathbf{I} + \bar{\mathbf{H}}_k^{dl^H} \bar{\mathbf{H}}_k^{dl} \right)^{-1} \hat{\mathbf{H}}_k^{dl^H} \hat{\mathbf{H}}_k^{dl} \right\} \tag{4.15}$$

where $\upsilon_{\max}(\mathbf{A})$ is the eigenvector corresponding to the largest eigenvalue of the matrix $\mathbf{A}$, and $\bar{\mathbf{H}}_k^{dl} = [\hat{\mathbf{H}}_1^{dl}, \ldots, \hat{\mathbf{H}}_{k-1}^{dl}, \hat{\mathbf{H}}_{k+1}^{dl}, \ldots, \hat{\mathbf{H}}_K^{dl}]^H$ represents the corresponding concatenated leakage channel.

4.4 Uplink CSIT Feedback

As we discussed earlier, in FDD multi-user massive MIMO systems, each user measures the downlink MIMO channel through the reference signals/pilots, and then feeds back the CSI to the BS using codebook-based channel feedback [76]. In this chapter, we consider the limited feedback

scheme used in FDD LTE/LTE-Advanced standards. The CSI feedback in LTE/LTE-Advanced standards usually contains two kinds of information: the channel direction information (CDI), which is related to the eigen-directions of the underlying massive MIMO channel, and the channel quality indicator (CQI), which is related to the strength of the corresponding spatial directions.

In limited feedback systems such as multi-user massive MIMO, each user feeds back its CQI and CDI to the BS. CQI feedback is relatively straightforward where user k feeds back the quantized version of the singular values using scalar or vector quantization methods. CDI feedback is generally more involved relying on codebook-based feedback. After obtaining each of the eigen-directions of the downlink massive MIMO channel, say $\tilde{\mathbf{h}}_k^{(\ell)} = \mathbf{h}_k^{(\ell)}/\|\mathbf{h}_k^{(\ell)}\|$ for the ℓ-th eigen-direction, $\forall \ell \in \{1,\ldots,N_r\}$, user k quantizes it to $\hat{\mathbf{h}}_k^{(\ell)}$ using a random vector quantization codebook $\mathscr{C}$, which is known to both the BS and the user. In general, there will be multiple CDIs of the underlying downlink massive MIMO channel; the exact number of CDIs to be fed back will impact the system performance and the feedback overhead. The indices of the quantized CDIs will be sent to the BS through a feedback link. Accordingly, the BS obtains $\hat{\mathbf{h}}_k^{(\ell)}$ and then uses these for downlink MIMO precoding. A quantization codebook $\mathscr{C}$ consisting of 2^β N_t-dimensional unit norm vectors is given by $\mathscr{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{2^\beta}\}$, where 2^β is the codebook size, β is the number of feedback bits per user, and $\mathbf{c}_m \in \mathbb{C}^{N_t \times 1}$ is a unit norm codeword, i.e., $\|\mathbf{c}_m\|^2 = 1$.

Each user chooses the CDI in the codebook that is closest to its eigen-direction, where closeness is usually measured in terms of the angle between the eigen-direction and the codeword in the codebook or equivalently the inner product. Hence, user k computes quantization index $q_k^{(\ell)}$ according to

$$q_k^{(\ell)} = \underset{m=1,\ldots,2^\beta}{\arg\max} |\langle \tilde{\mathbf{h}}_k^{(\ell)}, \mathbf{c}_m \rangle| = \underset{m=1,\ldots,2^\beta}{\arg\min} \sin^2(\angle(\tilde{\mathbf{h}}_k^{(\ell)}, \mathbf{c}_m)),$$

and feeds the index back to the BS. Upon reception of the index $q_k^{(\ell)}$, the BS can recover the CDI by searching for the corresponding entry in the codebook.

It is worth noting that feedback overhead of the system depends on 1) the number of CDIs to be fed back to each BS, and 2) the number of BSs in the cooperating set. Hence, in this way, the feedback overhead is significantly reduced with a penalty of quantization error in the finite rate feedback systems [77].

The relationship between the full CDI and the quantized CDI can be established by the following lemma [105]:

Lemma 1. *The relationship between the full CDI and the quantized CDI can be expressed by*

$$\tilde{\mathbf{h}}_k^{(\ell)} = \sqrt{1 - z_k^{(\ell)}}\hat{\mathbf{h}}_k^{(\ell)} + \sqrt{z_k^{(\ell)}}\mathbf{r}_k^{(\ell)}, \tag{4.16}$$

where $z_k^{(\ell)} = 1 - |\tilde{\mathbf{h}}_k^{(\ell)^H}\hat{\mathbf{h}}_k^{(\ell)}|^2$ *is distributed according to the quantization error distribution and is independent of* $\mathbf{r}_k^{(\ell)} \in \mathbb{C}^{N_t \times 1}$ *which represents a unit norm vector isotropically distributed in the null space of* $\hat{\mathbf{h}}_k^{(\ell)}$.

As we mentioned earlier, the downlink channel matrix for the k-th user can be represented as

$$\mathbf{H}_k^{\mathrm{d}l} = \mathbf{A}_R \mathbf{H}_k^w \mathbf{A}_T^H \tag{4.17}$$

where $\mathbf{H}_k^w \in \mathbb{C}^{N_r \times N_t}$ is the angular domain channel matrix for the k-th user and is sparse. Moreover, $\mathbf{A}_R \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{A}_T \in \mathbb{C}^{N_t \times N_t}$ are the unitary matrices for the angular domain transformation at the UE and BS, respectively. In the case of limited feedback, only a quantized version of the estimated channel matrices, namely $\hat{\mathbf{H}}_k^{\mathrm{d}l}$, are available at the BSs, where each user uses B bits to quantize

$\hat{\mathbf{H}}_k^{\mathrm{d}l}$ to $\hat{\mathbf{H}}_k^{\mathrm{dl},Q}$. We can then perform the beamformer design where $\hat{\mathbf{H}}_k^{\mathrm{dl},Q}$ are treated as the estimated channels while we maximize the sum SLNR. To distinguish these beam-forming matrices from those designed in the previous section, we denote these beam-forming matrices as $\hat{\mathbf{W}}_k$. To this end, let $\hat{\mathbf{H}}_k^{\mathrm{d}l} = \mathbf{U}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}} \mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}} \mathbf{V}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}^H$ be the singular value decomposition (SVD) of the $N_r \times N_t$ channel matrix $\hat{\mathbf{H}}_k^{\mathrm{d}l}$, with $\mathbf{U}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}$ and $\mathbf{V}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}} \triangleq [\mathbf{v}_k^{(1)}, \ldots, \mathbf{v}_k^{(N_t)}]$, two unitary matrices with dimension $N_r \times N_r$ and $N_t \times N_t$, respectively, and $\mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}$ an $N_r \times N_t$ matrix with main diagonal $(\lambda_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}^{(1)}, \ldots, \lambda_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}^{(d)})$ and zeros on its off-diagonal, which $d \triangleq \min(N_t, N_r)$.

It is clear that the BS only needs to know the quantized version of CDI, $\hat{\mathbf{V}}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}$, and CQI, $\hat{\mathbf{\Lambda}}_{\hat{\mathbf{H}}_k^{\mathrm{d}l}}$, to obtain the transmit beam-formers. To this end, user k feeds back the quantized version of the singular values using scalar or vector quantization methods while employing a codebook-based feedback to send the CDI to the BSs. In this regard, after obtaining each of the eigen-directions of the estimated downlink massive MIMO channel, say $\tilde{\mathbf{v}}_k^{(\ell)}$, user k quantizes it to $\mathbf{v}_k^{(\ell),Q}$ using a random vector quantization codebook $\mathscr{C}$, as discussed earlier in this section.

The following lemma shows that the subspace of the true channel matrix can be decomposed as the weighted sum of the quantized channel and an independent and isotropic quantization error term [77].

Lemma 2. *The quantization $\mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{dl}}^{Q} \mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{H,Q}$ of the estimated channel $\mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{H}$ follows the following decomposition:*

$$\tilde{\mathbf{\Lambda}}_{\hat{\mathbf{H}}_k^{dl}} \tilde{\mathbf{V}}_{\hat{\mathbf{H}}_k^{dl}}^{H} \triangleq \mathbf{A}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{Q} + \mathbf{B}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{R}_{\hat{\mathbf{H}}_k^{dl}} \tag{4.18}$$

where $\tilde{\mathbf{\Lambda}}_{\hat{\mathbf{H}}_k^{dl}} \tilde{\mathbf{V}}_{\hat{\mathbf{H}}_k^{dl}}^{H} \in \mathbb{C}^{N_r \times N_t}$ is an orthonormal basis for the subspace spanned by the columns of $\mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{H}$; $\mathbf{A}_{\hat{\mathbf{H}}_k^{dl}} \triangleq \mathbf{\Lambda}_{\mathbf{H}_k^{dl}}^{Q} (\mathbf{I}_{N_t} - \mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}})^{1/2}$; $\mathbf{B}_{\hat{\mathbf{H}}_k^{dl}} \triangleq \mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{dl}}^{Q} \mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}}^{1/2}$; $\mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}}$ is a $N_t \times N_t$ matrix with main diago-

nal $(z_k^{(1)},\ldots,z_k^{(N_t)})$ *and represents the quantization error and satisfies* $Tr\left\{\mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}}^H\mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}}\right\} \triangleq \mathrm{d}^2(\hat{\mathbf{H}}_{\mathrm{k}}^{\mathrm{dl}},\hat{\mathbf{H}}_{\mathrm{k}}^{\mathrm{dl}})$*;* *and* $\mathbf{R}_{\hat{\mathbf{H}}_k^{dl}} = [\mathbf{r}_k^{(1)},\ldots,\mathbf{r}_k^{(N_t)}]^H \in \mathbb{C}^{N_t\times N_t}$ *is an orthonormal basis for an isotropically distributed (complex)* N_t*-dimensional plane in the null space of* $\mathbf{\Lambda}_{\hat{\mathbf{H}}_k^{dl}}^{Q}\mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{H,Q}$*. The quantities* $\mathbf{A}_{\hat{\mathbf{H}}_k^{dl}}$*,* $\mathbf{B}_{\hat{\mathbf{H}}_k^{dl}}$ *and* $\mathbf{V}_{\hat{\mathbf{H}}_k^{dl}}^{Q}$ *are distributed independent of each other, as are the pair* $\mathbf{R}_{\hat{\mathbf{H}}_k^{dl}}$ *and* $\mathbf{Z}_{\hat{\mathbf{H}}_k^{dl}}$*.*

Taking Lemma 2 into account, the precoding vector $\hat{\mathbf{w}}_k$ can be designed by solving (4.13). However, due to the limited feedback framework, it is reasonable to consider the beam-forming design via maximizing an expected SLNR, averaging over all possible channel realizations. Hence, the problem of interest can be formulated as

$$\begin{aligned} \max_{\hat{\mathbf{w}}_k} \quad & \sum_{k=1}^{K}\mathbb{E}_{\mathbf{H}}\{\mathrm{SLNR}_k\} \\ \text{s.t.} \quad & \sum_{k=1}^{K}\rho_k \le P \end{aligned} \tag{4.19}$$

It is worth mentioning that, not only it is difficult to derive a closed-form expression of $\mathbb{E}\{\mathrm{SLNR}_k\}$, but it is also hard to obtain a low-complexity algorithm to obtain the beam-forming vectors. To tackle this issue, instead of maximizing the expected value of SLNR, we maximize the lower bound of $\mathbb{E}\{\mathrm{SLNR}_k\}$. The following result adapted from [77, Theorem 2] allows us to provide a closed form solution for each transmitter precoder.

Proposition 2. *The optimal precoders that are able to maximize the lower bound of* $\sum_{k=1}^{K}\mathbb{E}_{\mathbf{H}}\{\mathrm{SLNR}_k\}$ *can be obtained by extracting the leading K columns of matrix* $\mathbf{J}_k$*, where* $\mathbf{J}_k$ *is the generalized eigenmatrix of the pair* $\sum_{k=1}^{K}\mathbb{E}\{\hat{\mathbf{H}}_k^{dl^H}\hat{\mathbf{H}}_k^{dl}\}$ *and* $\frac{\sigma_k^2}{\rho_k}\mathbf{I}_k + \sum_{i\neq k}\hat{\mathbf{H}}_i^{dl^H}\hat{\mathbf{H}}_i^{dl}$ *as follows*

$$\hat{\mathbf{W}}_k = \mu\mathbf{J}_k[\mathbf{I}_k;\mathbf{0}] \tag{4.20}$$

where μ is a scaling factor so that $\sum_{k=1}^{K} \rho_k \leq P$.

Feedback strategies: In the following, we introduce two feedback strategies: "Rank 1 feedback" and "Rank 2 feedback." In a typical cellular network, usually the largest singular value is dominant over the other singular values. In order to reduce the feedback overhead, each user only needs to feed back the quantized version of the dominant singular value and the corresponding eigen-direction, which is usually termed as "Rank 1 feedback," Without loss of generality, we assume that the first singular value is the largest one. Then, the optimal transmit beam-forming vectors $\hat{\mathbf{w}}_k$ calculated by the BS can be expressed as (4.20), where the quantized channel matrix has the form $[\hat{\lambda}^{(1)}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{v}_k^{(1)^H}, \mathbf{0}_{(N_r-1)\times N_t}]^T$.

Alternatively, each user can feed back two dominant directions: both singular values and their corresponding eigen-directions are fed back to the BS. Without loss of generality, we assume that the first two singular values are the dominant ones. Similar to Rank 1 feedback, the optimal solutions for the transmit beam-forming vectors can be expressed as (4.20) where the quantized channel matrix is of the form $[\hat{\lambda}^{(1)}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{v}_k^{(1)^H}, \hat{\lambda}^{(2)}_{\hat{\mathbf{H}}_k^{dl}} \mathbf{v}_k^{(2)^H}, \mathbf{0}_{(N_r-2)\times N_t}]^T$. Note that Rank 2 feedback doubles the overhead as opposed to Rank 1 feedback.

Performance analysis: To characterize the performance of the introduced algorithm with limited feedback, we examine the rate loss incurred by the proposed beam-forming algorithm, where channel estimates are used to calculate the columns of the precoders $\mathbf{w}_k$. The mean loss in sum-rate is then defined as $\triangle R_{sum} = \mathbb{E}_{\mathbf{H}}\{R_{sum}\} - \mathbb{E}_{\mathbf{H}}\{\hat{R}_{sum}\}$ where $\mathbb{E}_{\mathbf{H}}\{R_{sum}\}$ is the average sum rate from the introduced algorithm with estimated CSI and $\mathbb{E}_{\mathbf{H}}\{\hat{R}_{sum}\}$ is the rate achieved with limited feedback of estimated CSI. Using the received signal, the upper bound on mean loss in sum-rate will be achieved similar to the one in [77, (37)].

4.5 Performance Evaluation

In this section, we investigate the performance of the introduced limited feedback scheme for FDD multi-user massive MIMO systems that we discussed in Section 4.4 and compare the results of Rank 1 and Rank 2 feedback via Monte Carlo simulations. As a performance metric, we use normalized mean squared error (NMSE) for the channel estimation to evaluate the performance of each estimator and downlink achievable rate for the limited feedback scheme to evaluate the performance of Rank 1 and Rank 2 feedback. The NMSE for each user is defined as $\mathbb{E}\left(\frac{||\mathbf{H}_k^{dl}-\hat{\mathbf{H}}_k^{dl}||_F^2}{||\mathbf{H}_k^{dl}||_F^2}\right)$. The antenna spacing for both the transmit and receive antennas, at the BS and the user, is assumed to be 0.5λ. The noise at each receiver is assumed to be Gaussian with zero mean and unit variance. The number of receive antennas at the user is selected as 8. The number of transmit antennas at the BS is selected as 64. The transmit power at the BS is fixed at 43 dBm. The system bandwidth is taken to be 10 MHz. Codebook based random vector quantization is used through the simulation. For simplicity, we assume that all users have the same channel sparsity level, which is $s = 8$. First, we evaluate three methods to generate the CCM for an arbitrary user. The CCM for the k-th user ($\mathbf{R}_k$) can be modeled by:

- One-ring model (OR): The (i, j)-th element of $\mathbf{R}_k$ for a uniform linear array (ULA) is given by [78, 79]

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{2\Delta}\int_{-\Delta+\theta_k}^{\Delta+\theta_k} e^{-j2\pi\frac{d}{\lambda}(i-j)\sin(x)}dx \tag{4.21}$$

  Where $\Delta, \alpha_k, \theta_k, d,$ and λ refer to the angular spread, average large-scale fading, mean angle of arrival (AoA) of the k-th user in the azimuth direction, antenna spacing, and carrier

wavelength, respectively.

- Laplacian model: The (i,j)-th element of $\mathbf{R}_k$ for a uniform linear array (ULA) is given by [80, 81]

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{\sqrt{2}\Delta} \int_{-\pi+\theta_k}^{\pi+\theta_k} e^{-\frac{\sqrt{2}}{\Delta}|x-\theta_k| - j2\pi\frac{d}{\lambda}(i-j)\sin(x)} dx \tag{4.22}$$

  where $\Delta, \alpha_k, \theta_k, d,$ and λ are defined in the OR model.

- Exponential model: The (i,j)-th element of $\mathbf{R}_k$ is given by [82, 83]

$$[\mathbf{R}_k(\rho,\theta_k)]_{i,j} = \begin{cases} \left(\rho e^{j\theta_k}\right)^{j-i} & i \leq j \\ \left(\rho e^{-j\theta_k}\right)^{i-j} & i > j \end{cases} \tag{4.23}$$

  where ρ is the correlation factor, $0 \leq \rho \leq 1$. For simplicity, ρ is the same for all users while the AoA (θ_k) is different and uniformly distributed in $[-\pi, \pi)$.

Figure 4.2 shows the cumulative distribution function (CDF) for the three models discussed above under different parameters. We use the empirical CDF, $\hat{F}(x) = \frac{\sum_i \lambda_i \leq x}{n}$, where $\lambda_i, i = 1, \ldots, n$ are the singular values of the CCM $\mathbf{R}_k \in \mathbb{C}^{n \times n}$. The figure shows that a large number of singular values are close to zero, while a few are large for both OR and Laplacian models. The exponential model provides a slight variation in its singular values. From here on, we consider the exponential model with correlation factor $\rho = 0.6$.

The average NMSE of each estimator versus the downlink training overhead (T) is shown in Figure 4.3. It is obvious that LS/MMSE estimators are not efficient for FDD multi-user massive MIMO systems compared with channel estimation based on the CS approach. The OMP algorithm

shows better performance than traditional LS/MMSE estimators because it exploits the sparsity structure of the downlink channel matrix. Moreover, we observe that the BFS algorithm has better performance than the OMP algorithm especially at the first iteration because of its ability to select the best features from the measurement matrix based on ℓ_1 norm minimization.

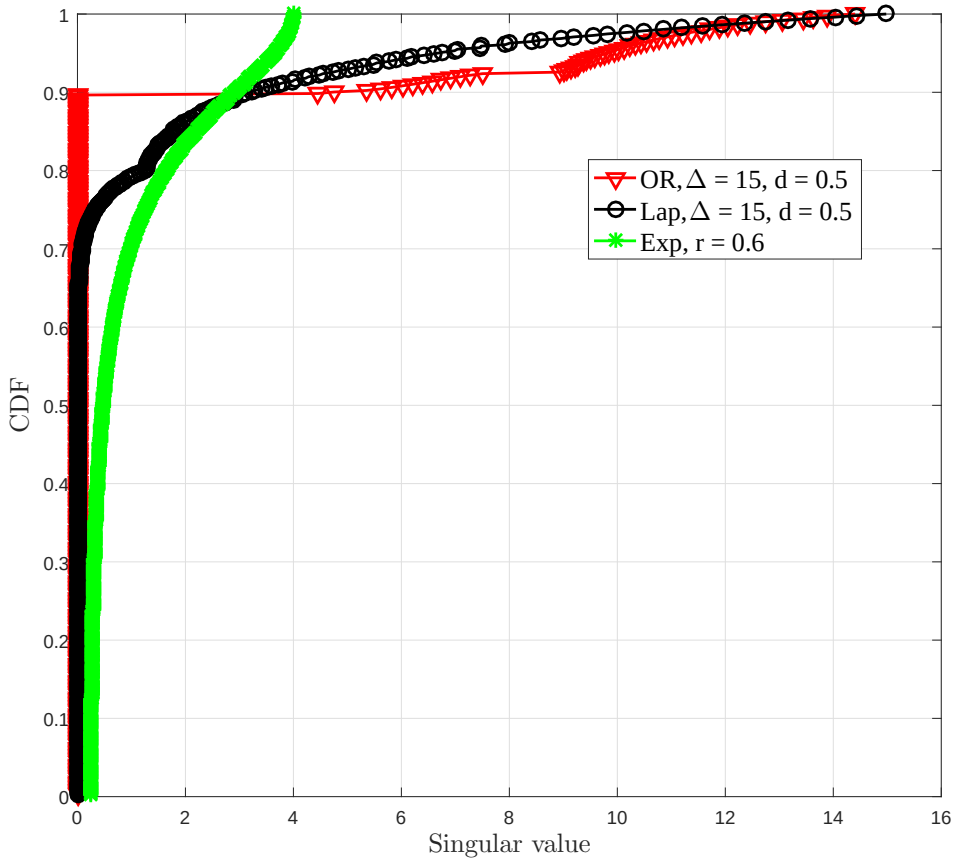

Figure 4.2: The CDF of the singular values of CCM for an arbitrary user with $N_t = 64, N_r = 8, \alpha = 1$ and $\theta = pi/4$.

In Figure 4.4, we plot the average sum-rate curve when $N_t = 64$ and $N_r = 8$ of our system versus the SNR under different feedback strategies: Rank 1 feedback and Rank 2 feedback with $\beta = 12$ with perfect CSI and imperfect CSI. We observe that Rank 2 feedback exhibits enhanced performance compared to Rank 1 feedback, as expected, because with Rank 2 feedback, each user

feeds back the two dominant eigenvectors and their corresponding eigenvalues to the BS. In this figure, we observe a performance gain of 22.9% for Rank 2 feedback at SNR $= 10$ dB under perfect CSI. The performance gain become more substantial in the high SNR regime because the second dominant eigenvalue cannot be ignored in this regime. However, Rank 2 feedback doubles the overhead as opposed to Rank 1 feedback. For the comparison with a different number of feedback bits, we depict the performance of feedback strategies when $\beta = 14$, $N_t = 64$ and $N_r = 8$ in Figure 4.5. In this figure, we observe that the performance gain exhibits better performance compared to the previous case because more feedback bits are required with large number of antennas in order to maintain the system performance. Moreover, we observe that the performance gap between Rank 1 feedback and Rank 2 feedback decreased as the number of feedback bits β increases due to reduction in channel quantization error.

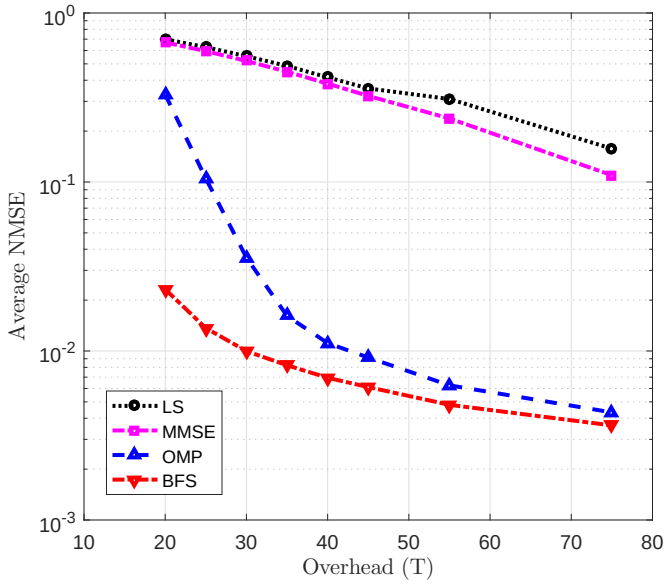

Figure 4.3: The average NMSE performance of CSI versus overhead (T) for $N_t = 64$ antennas at the BS.

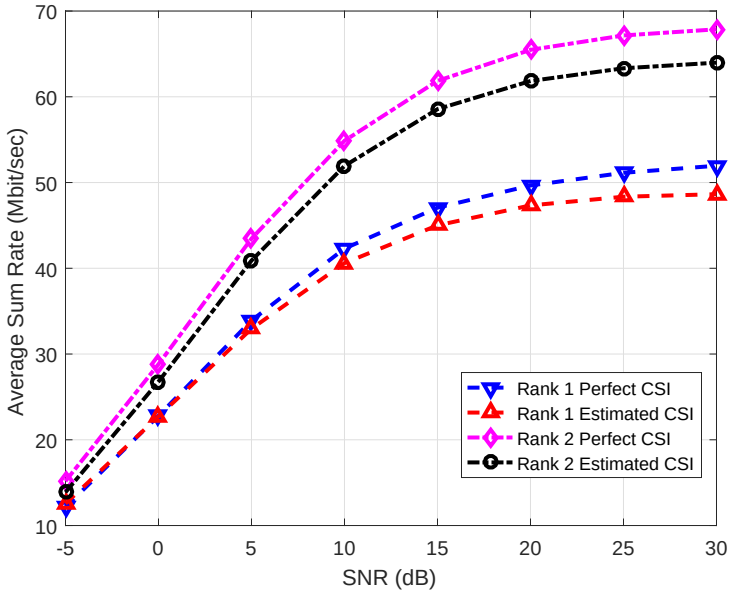

Figure 4.4: Average sum-rate versus SNR with $N_t = 64, N_r = 8,$, and $\beta = 12$.

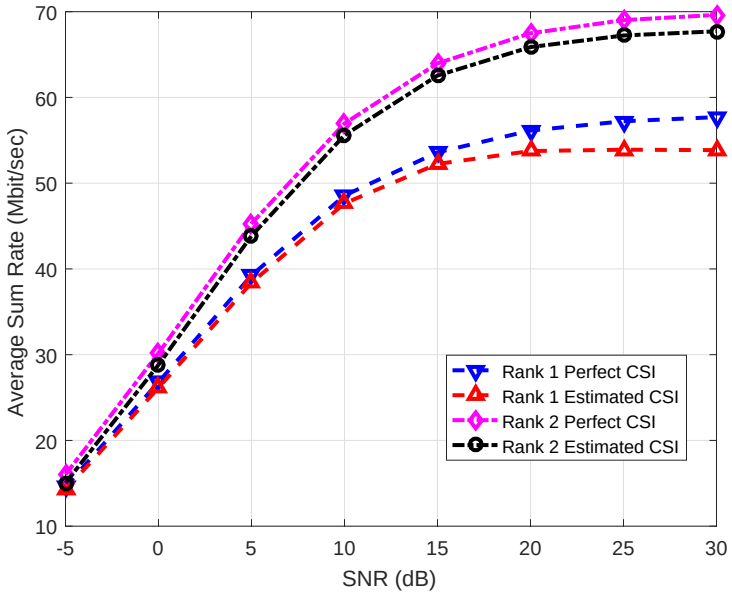

Figure 4.5: Average sum-rate versus SNR with $N_t = 64, N_r = 8,$, and $\beta = 14$.

Chapter 5

Antenna Selection for FDD Multi-User Massive MIMO systems

5.1 Introduction

Massive multiple-input multiple-output (MIMO) systems are a promising technique for next-generation wireless communication systems. MIMO has been widely studied recently to enhance the spectral and energy efficiency of cellular networks with simple linear signal processing methods or combining techniques, such as matched filter (MF) and zero-forcing (ZF) [108]. The addition of more antennas at the BS can be useful to increase throughput, reduce radiated transmit power, uniformity in services everywhere in the cell, and simplicity in the signal processing in 5G technologies. However, systems with a large number of antennas at the BS result a hardware cost due to the large number of radio frequency (RF) chains needed for each antenna, high computational complexity, and more power consumption [109, 110]. As the number of antennas increases at the BS, circuit power consumption will gradually increase by the number of antennas.

To address the issue of increasing circuit power consumption in massive MIMO systems, transmit antenna selection techniques can be used to limit the number of RF chains and enhance the system diversity gain by exploiting the spatial selectivity introduced by a large number of antennas. The optimal subset of antennas can be obtained by using an exhaustive searching method such as a brute force search. However, the computational complexity of an exhaustive search method grows exponentially with the number of antennas available at the BS [111]. Therefore, a low complexity suboptimal transmit antenna selection algorithm is required to improve system performance in terms of power consumption.

Antenna selection algorithms have been widely studied for conventional MIMO systems and a few works have been extended for massive MIMO systems. The authors in [111], propose a new algorithm for joint antenna selection and user scheduling that maximizes the achievable sum-rate of the system. The algorithm iteratively removes the transmit antennas and users that have minimum contribution to the system performance gain. In [112], the authors introduced the concept of sub-array antenna selection for massive MIMO systems. The idea of sub-array antenna selection is to divide the full-array antenna into sub-arrays and use only one antenna for each sub-array. The authors in [113] consider a point-to-point system and proposed a low complexity antenna selection algorithm based on the norm and correlation to enhance the energy efficiency in massive MIMO systems under a total power constraint. An iterative algorithm for transmit antenna selection based on $\ell_{1/2}$ regularity and group sparsity was proposed in [114]. They formulate an optimization problem that minimizes the transmission power at certain number of RF chains subject to SINR constraints at all users. The authors in [115] develop a transmit antenna selection strategy for massive MIMO systems that can reduce the CSI overhead and computational complexity and improve the system performance in terms of achievable sum-rate. They assume that each user can share infor-

mation about CSI feedback with other users. Later, each user will remove the channel coefficients corresponding to a certain subset of antennas and then send it back to the BS. The authors in [116], consider transmit antenna selection in the downlink of massive MIMO systems with measured massive MIMO channels. They select the set of antennas that maximizes the achievable sum-rate. The authors in [117], propose an antenna selection algorithm for massive MIMO systems based on rectangular maximum volume (RMV) approach. The proposed algorithm can choose a rectangular submatrix with maximum volume from the channel matrix. A low complexity antenna selection algorithm based on constructive interference performance metric was proposed in [118] to reduce the computational costs. The proposed algorithm exploits the antenna diversity by selecting the antenna set with optimal inter channel interference (ICI).

In this chapter, we investigate the problem of transmit antenna selection for multi-user massive MIMO networks and design an efficient antenna selection algorithm that can reduce the hardware complexity and improve the system performance in terms of achievable sum-rate. This chapter is organized as follows. We begin with the system model in Section 5.2, after which the problem formulation is discussed in Section 5.3. The performance evaluation of the proposed algorithm is discussed in Section 5.4.

5.2 System Model

In this chapter, we consider a flat and block-fading single-cell FDD multi-user massive MIMO system with one BS equipped with N_t antennas that serves K single-antenna non-cooperating scheduled users, $N_t \gg K$. The BS performs antenna selection by choosing K antennas among the N_t

antennas, and scheduling K users within the cell to be served simultaneously. The BS determines

$$\mathbb{A} : \text{the set of transmit antennas selected} \tag{5.1}$$

In the downlink, the $K \times 1$ received signal vector $\mathbf{y}$ at the K scheduled users antennas can be expressed as

$$\mathbf{y} = \mathbf{H}^H \mathbf{x} + \mathbf{n}, \tag{5.2}$$

where $(\cdot)^H$ denotes the Hermitian of the matrix $\mathbf{H}$. $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the signal vector transmitted by the BS, which satisfies the power constraint $\mathbb{E}\left\{\|\mathbf{x}\|^2\right\} = P$, where P is the average transmit power available at the BS. $\mathbf{n} \in \mathbb{C}^{K \times 1}$ is a complex circularly symmetric Gaussian noise vector, where $\mathbf{n} \sim \mathscr{CN}(0, \mathbf{I}_K)$. $\mathbf{H} \in \mathbb{C}^{N_t \times K}$ is the overall downlink channel matrix between the BS and the scheduled users, where $\mathbf{H} = [\mathbf{h}_1 \dots \mathbf{h}_k \dots \mathbf{h}_K]$. In this chapter, we consider Rayleigh correlated channel coefficients, where the channel vector $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ for user k has zero mean and spatial channel correlation matrix (CCM), $\mathbf{R}_k \in \mathbb{C}^{N_t \times N_t}$. Thus, $\mathbf{h}_k \in \mathscr{CN}(0, \mathbf{R}_k)$.

In this chapter, we adopt a one ring (OR) model, which is a appropriate model to capture the antenna correlation in massive MIMO systems. In the OR model, the BS is elevated with no significant local scatters and each user is located at azimuth θ and distance D from the BS, which is surrounded by a ring of local scatters at radius R. The angular spread (AS) is approximated as $\Delta \approx \arctan(\frac{R}{D})$ because D and R are typically large compared to the antenna spacing. Throughout the chapter, we assume that each user k is characterized by (θ_k, Δ) of its angle of arrival (AoA) and

AS. The correlation between the channel coefficients of antennas for k-th user is given by

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{2\Delta} \int_{-\Delta+\theta_k}^{\Delta+\theta_k} e^{-j2\pi \frac{d}{\lambda}(i-j)\sin(x)} dx \tag{5.3}$$

where α_k, θ_k, d, and λ refer to the average large-scale fading, mean angle of arrival (AoA) of the user k in the azimuth direction, antenna spacing, and carrier wavelength, respectively. The singular-value-decomposition (SVD) of CCM for user k can be written as

$$\mathbf{R}_k = \mathbf{U}_k \Lambda_k \mathbf{U}_k^H \tag{5.4}$$

where $\mathbf{U}_k \in \mathbb{C}^{N_t \times r_k}$ is a tall unitary matrix that contains the eigenvectors corresponding to the r_k eigenvalues of $\mathbf{R}_k$. Λ_k is $r_k \times r_k$ diagonal matrix whose elements are the eigenvalues of $\mathbf{R}_k$. In addition, the channel vector $\mathbf{h}_k$ for user k can be rewritten using Karhunen–Loeve representation as

$$\mathbf{h}_k = \mathbf{U}_k \Lambda_k^{\frac{1}{2}} \mathbf{v}_k \tag{5.5}$$

where $\mathbf{v}_k \in \mathbb{C}^{r_k \times 1}$ is an i.i.d. vector, $\mathbf{v}_k \sim \mathscr{CN}(0, \mathbf{I}_{r_k})$.

The BS sends N_s downlink data streams to the K scheduled users. Each user receives one spatial stream, which means $N_s = K$. The BS selects K antennas to transmit the downlink data streams to the scheduled users with $K << N_t$. The BS linearly precodes its downlink data vector.

Thus, the transmitted signal $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ can be written as

$$\mathbf{x} = \mathbf{W}\mathbf{s} = \sum_{k=1}^{K} \sqrt{\rho_k}\mathbf{w}_k s_k \tag{5.6}$$

where $\mathbf{s} \in \mathbb{C}^{K \times 1}$ is the downlink symbol vector transmitted by the BS to the K scheduled users. s_k is the transmitted symbol to the user k with $\mathbb{E}\left\{|s_k|^2\right\} = 1$. ρ_k is the transmit power allocated for user k and $\mathbf{w}_k \in \mathbb{C}^{N_t \times 1}$ denotes the beamformer vector that the BS uses to transmit the symbol s_k to user k, where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k, \ldots, \mathbf{w}_K] \in \mathbb{C}^{N_t \times K}$ is the precoding matrix with unit-norm columns. Hence, the transmit power at the BS can be expressed as $P = \mathbb{E}\left\{\|\mathbf{x}\|^2\right\} = \sum_{k=1}^{K} \rho_k \mathbb{E}\left\{\|\mathbf{w}_k\|^2\right\} = \sum_{k=1}^{K} \rho_k$. The received signal $y_k \in \mathbb{C}$ at user k can be written as

$$y_k = \sqrt{\rho_k}\mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{i=1; i \neq k}^{K} \sqrt{\rho_i}\mathbf{h}_k^H \mathbf{w}_i s_i + n_k \tag{5.7}$$

where $s_k \in \mathbb{C}$ and $n_k \sim \mathscr{CN}(0, \sigma_k^2)$ is the downlink data symbol and the noise associated with user k, respectively. The signal-to-interference-plus-noise ratio (SINR) at user k is given by

$$\text{SINR}_k = \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma^2 + \sum_{i=1; i \neq k}^{K} \rho_i |\mathbf{h}_k^H \mathbf{w}_i|^2} \tag{5.8}$$

We denote the interference term by $\mathbf{I}_k = \sum_{i=1; i \neq k}^{K} \rho_i |\mathbf{h}_k^H \mathbf{w}_i|^2$. The sum rate can be expressed as

$$\begin{aligned} R_{\text{sum}}(\mathscr{A}) &= \sum_{k=1}^{K} \log 2 \left(1 + \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma_k^2 + \sum_{i=1; i \neq k}^{K} \rho_i |\mathbf{h}_k^H \mathbf{w}_i|^2}\right) \\ &= \sum_{k=1}^{K} \log 2 \left(1 + \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma_k^2 + \mathbf{I}_k}\right) \end{aligned} \tag{5.9}$$

5.3 Problem Formulation

In this section, we consider the problem of finding the best subset of BS antennas that maximizes the downlink achievable rate under a total power constraint. The problem can be written as

$$\begin{aligned} &\max_{\mathscr{A}} R_{\text{sum}}(\mathscr{A}) \\ &s.t\,|\mathscr{A}| \geq K \\ &\sum_{k=1}^{K} \rho_k \leq P \end{aligned} \tag{5.10}$$

In this chapter, we assume that the BS has perfect knowledge of the CSI. Moreover, the BS will use zero-forcing (ZF) as a linear precoder. Due to the assumption of perfect CSI available at the BS and since we use the ZF precoder, the downlink achievable sum-rate in (5.9) can be rewritten as

$$R_{\text{sum}}(\mathscr{A}) = \sum_{k=1}^{K} \log 2\left(1 + \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma_k^2}\right) \tag{5.11}$$

Then the problem discussed above can be rewritten as

$$\begin{aligned} &\max_{\mathscr{A}} \sum_{k=1}^{K} \log 2\left(1 + \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma_k^2}\right) \\ &s.t\,|\mathscr{A}| \geq K \\ &\sum_{k=1}^{K} \rho_k \leq P \end{aligned} \tag{5.12}$$

For optimal power allocation, we apply the water-filling algorithm [72] and the optimal power

allocated for user k can be written as

$$\rho_k = \left(\mu - \frac{1}{|\mathbf{h}_k^H \mathbf{w}_k|^2} \right)^+ \tag{5.13}$$

where x^+ is equal to $\max(0,x)$ and μ should satisfy the following equation

$$\sum_{N_t \in \mathbb{A}} \left(\mu - \frac{1}{|\mathbf{h}_k^H \mathbf{w}_k|^2} \right)^+ = P \tag{5.14}$$

The antenna selection problem is difficult to solve because it is a combinational optimization problem. We can use an exhaustive search method such as a brute force approach to find the optimal set of antennas. However, the computational complexity of this method increases exponentially as the number of antennas grows. Hence, we develop a low complexity iterative algorithm to solve the problem and compare the proposed algorithm with JASUS algorithm developed in [111]. The JASUS algorithm and the proposed algorithm are summarized in Algorithm 1 and Algorithm 2, respectively. The JASUS algorithm starts by initializing a set $\mathscr{A}$ with the set of all antennas. Then, the algorithm iterates over all the elements in the set, trying to remove the index of the antenna that degrades the performance gain. The algorithm terminates when it reaches $(M-N)$ iterations. However, the JASUS algorithm uses a sequential backward selection method. In this method, once a feature is discarded in the backward selection, there is no possibility for it to be reconsidered again. We improve the JASUS algorithm by using a forward selection method to obtain a better set of antennas selected. The proposed algorithm starts by initializing a set of antennas required randomly. Then, the algorithm iterates between the current set and the remaining set to obtain the best set that maximize the achievable sum-rate.

Algorithm 5: JASUS [111]

Input : Channel coefficients **H**
Output : The set of antennas $\mathbb{A}$
Initialization :

- $t \longleftarrow 1$
- $\mathbb{A} \longleftarrow \{1,\ldots,M\}$
- while $t < (M-K)$ do
  - maxRate $\longleftarrow 0$
  - for each m in $\mathbb{A}$ do
    - $R_{-m} = Rsum(\mathbb{A}\backslash m)$
    - if $R_{-m} >$ maxRate then
      - maxRate $\longleftarrow R_{-m}$
      - $m_{bad} \longleftarrow m$
    - end if
  - end for
  - $\mathbb{A} \longleftarrow \mathbb{A}\backslash\{m_{bad}\}$
  - $t \longleftarrow t+1$
- end while

5.4 Performance Evaluation

In this section, we investigate the performance of antenna selection in FDD multiuser massive MIMO systems. As a performance metric, we use the downlink achievable sum-rate to compare the results of the proposed algorithm with the JASUS algorithm through Monte Carlo simulations. Simulation results are provided to demonstrate the performance of the proposed algorithm. Simulations were conducted using MATLAB. The antenna spacing for the transmit antennas at the BS is assumed to be 0.5λ. The number of users in the cell is set to be $K = 16$. The number of receive antennas at the user is set to be 1. The number of transmit antennas at the BS is set to be 64. The average transmit SNR is 16 dB. We generate the channel correlation matrix using the one-ring

Algorithm 6: Proposed algorithm

Input : Channel coefficients **H**
Output : The set of antennas $\mathbb{A}$
Initialization :

- select a current set A randomly
- compute the sum-rate R_A
- set the difference between current set and the remaining set as B
- for each m in B do
- for each k in A do
  - update the current set
  - calculate the sum-rate R
  - if $R > R_A$
    - update the current set
    - update R_A
    - update the difference set B
    - else
      - keep the current set
  - end if
- end for
- end for

(OR) model as follows:

$$[\mathbf{R}_k]_{i,j} = \frac{\alpha_k}{2\Delta} \int_{-\Delta+\theta_k}^{\Delta+\theta_k} e^{-j2\pi \frac{d}{\lambda}(i-j)\sin(x)} dx \tag{5.15}$$

where $\alpha_k = 1 \ \forall k$, $\Delta = 12$, and θ_k is uniformly distributed over $[0, 2\pi)$. Figure 5.1 shows the average downlink achievable rate versus the SNR for $K = 16$ users and $N_t = 64$ antennas at the BS. We observe that the proposed algorithm outperform the JASUS algorithm because our proposed algorithm does not ignore any index, which means that once an index is discarded, there is a

possibility to select this index with other possible resulting combinations. In Figure 5.2, we plot the downlink achievable sum-rate versus the number of transmit antennas selected at the BS. In this figure, we set the number of users to be $K = 16$, the number of transmit antennas at the BS is $N_t = 64$, and transmit SNR $= 16$ dB. The figure shows that the proposed algorithm gives better performance compared with the JASUS algorithm, especially when the number of the transmit antennas selected is much less than the number of antennas available at the BS. We expected this result because, as the number of transmit antennas selected increases, the number of combinations searched decreases.

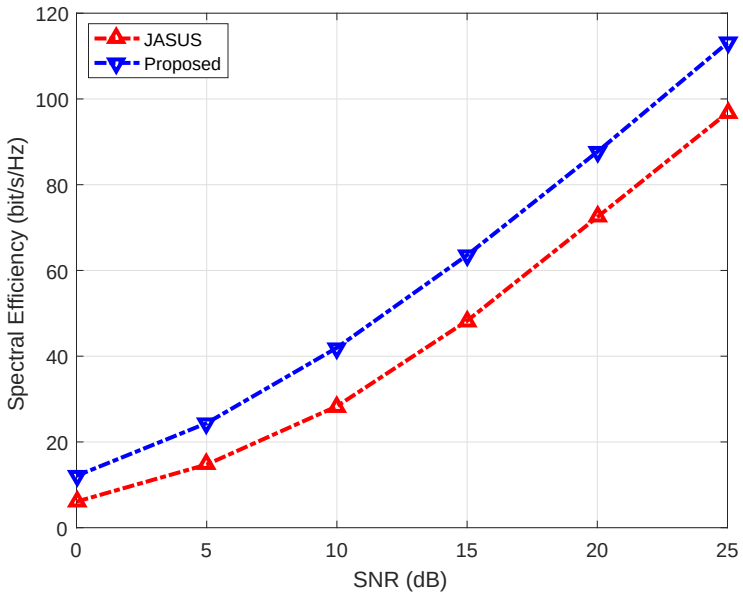

Figure 5.1: The average achievable sum-rate versus SNR for $N_t = 64$ antennas at the BS.

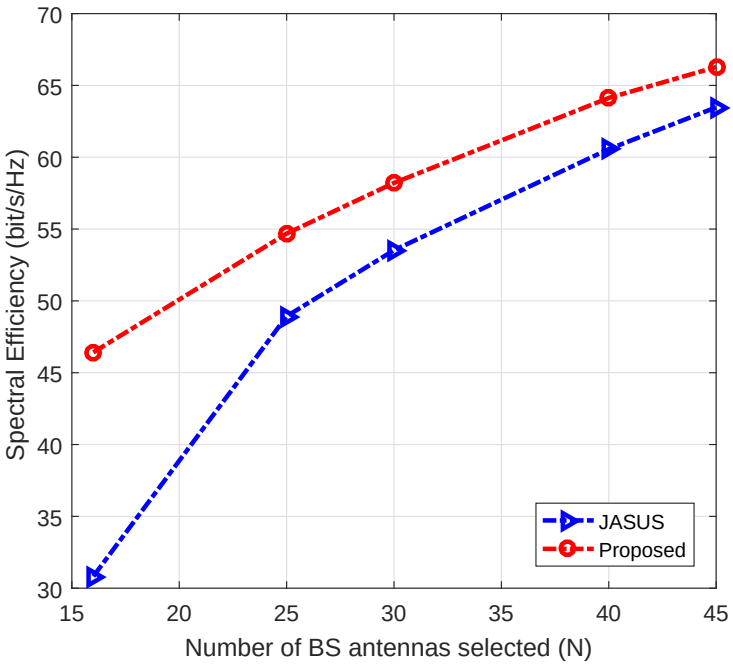

Figure 5.2: The average achievable sum-rate versus the number of transmit antennas selected for $K = 16$, $SNR = 16$ dB and $N_t = 64$ antennas at the BS.

Chapter 6

Conclusions

In wireless communication systems, there is a huge demand for reliability, high data rates, high energy efficiency, high mobility, and high spectral efficiency as it is required by the modern wireless applications. To keep up with the fast wireless data traffic development, a vital objective of next generation of wireless cellular networks is to enhance the data throughput by orders of magnitude; 100X and even 1000X higher throughput are often mentioned as 5G design goals. One of the key technologies for 5G wireless communication is massive MIMO. Massive multiple input multiple output (MIMO) systems are a promising technology for next generation wireless communications due to their ability to increase capacity and enhance both spectrum and energy efficiency. Most existing work on massive MIMO considers time-division duplexing (TDD) systems and relies on channel reciprocity to perform precoding and resource allocation. However, frequency-division duplex (FDD) systems dominate current cellular networks and offer many benefits over TDD systems in delay-sensitive services and traffic applications. To utilize the benefit of massive MIMO systems, accurate downlink channel state information at the transmitter (CSIT) is essential. Conventional approaches to obtain CSIT for FDD multi-user massive MIMO systems require downlink

training and uplink CSI feedback. However, such training results in large overhead for massive MIMO systems because of the large dimensionality of the channel matrix.

In this dissertation, first, we aim to improve the performance of wireless cellular networks by designing an efficient downlink beamforming for FDD massive MIMO systems based on partial channel state information and considering the mismatch error. Second, we estimate the downlink channel using compressive sensing techniques to reduce the downlink training overhead and design an efficient downlink beamforming based on CSI estimated.

In Chapter 2, we designed an efficient downlink beamforming method based on partial CSI. The partial CSI is the uplink dominant DoA which is obtained during uplink training phase. By exploiting the relationship between uplink (UL) direction of arrivals (DoAs) and downlink (DL) direction of departures (DoDs), we derived an expression for estimated downlink DoDs, which will be used for downlink beamforming to compare the performance with traditional method in terms of downlink achievable rate that we derived. Simulation results also verified that, in terms of achievable rate, our proposed method outperform the traditional beamforming method.

In Chapter 3, we introduced a new method to estimate the channel state information for each user in FDD multi-user massive MIMO systems based on a compressive sensing technique. The downlink channel estimation scheme based BFS algorithm for FDD multi-user massive MIMO systems was proposed. We considered a spatially correlated channel and evaluated three models to generate the channel correlation matrix. Moreover, an uplink feedback based limited feedback scheme was proposed. We analyzed the performance of the CSI estimator in terms of normalized mean squared error. To this end, by exploiting the sparsity structure of the downlink channel matrix, we designed an efficient algorithm to acquire CSI. We verified our analysis using simulations. Our simulation results showed that our proposed method outperforms the traditional channel

estimation methods.

In Chapter 4, we proposed a feedback scheme that exploits the sparsity structure of the downlink channel matrix and design a low-complexity beam-forming scheme for FDD multi-user massive MIMO systems with spatially correlated channels. The proposed schemes consider codebook-based feedback, which is adopted in the LTE/LTE-advanced systems. Specifically, after estimating the downlink channel for each user during the downlink training phase, we designed a low-complexity beam-forming scheme for downlink transmission based on the estimated channel matrix. We next updated the channel matrix using limited feedback and calculated the downlink achievable rate as a performance metric for the limited feedback scheme to compare the results of rank 1 and rank 2 feedback through Monte Carlo simulations.

Finally, in Chapter 5, we investigated the problem of transmit antenna selection for FDD multi-user massive MIMO systems and developed a new low complexity algorithm that finds the best set of antennas required at the BS. The proposed algorithm starts by initializing a set of antennas required randomly. Then, the algorithm iterates between the current set and the remaining set to obtain the best set that maximize the downlink achievable rate.

References

[1] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, 2014.

[2] A. Gohil, H. Modi, and S. K. Patel, "5G technology of mobile communication: A survey," in *Intelligent Systems and Signal Processing (ISSP), 2013 International Conference on*. IEEE, 2013, pp. 288–292.

[3] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[4] T. Van Chien and E. Björnson, *Massive MIMO Communications*. Cham: Springer International Publishing, 2017, pp. 77–116. [Online]. Available: https://doi.org/10.1007/978-3-319-34208-5_4

[5] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 574–590, 2018.

[6] Y. Yu, E. Dutkiewicz, X. Huang, and M. Mueck, "Downlink resource allocation for next generation wireless networks with inter-cell interference," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1783–1793, 2013.

[7] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. Ten Brink, I. Gaspar, N. Michailow, A. Festag *et al.*, "5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications." *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, 2014.

[8] "Exploiting caching and multicast for 5G wireless networks, author=Poularakis, Konstantinos and Iosifidis, George and Sourlas, Vasilis and Tassiulas, Leandros," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2995–3007, 2016.

[9] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[10] R. Ferrus, O. Sallent, J. Perez-Romero, and R. Agusti, "On 5g radio access network slicing: Radio interface protocol features and configuration," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 184–192, 2018.

[11] M. Lauridsen, L. C. Gimenez, I. Rodriguez, T. B. Sørensen, and P. E. Mogensen, "From LTE to 5G for Connected Mobility." *IEEE Communications Magazine*, vol. 55, no. 3, pp. 156–162, 2017.

[12] "Cognitive radio for next-generation wireless networks: An approach to opportunistic channel selection in IEEE 802.11-based wireless mesh."

[13] R. Y. Kim and S. Mohanty, "Advanced power management techniques in next-generation wireless networks [Topics in Wireless Communications]," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 94–102, 2010.

[14] Z. Gao, L. Dai, S. Han, I. Chih-Lin, Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 144–153, 2018.

[15] Y. Kim, H. Ji, J. Lee, Y.-H. Nam, B. L. Ng, I. Tzanidis, Y. Li, and J. Zhang, "Full dimension MIMO (FD-MIMO): The next evolution of MIMO in LTE systems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 26–33, 2014.

[16] S. M. Razavizadeh, M. Ahn, and I. Lee, "Three-dimensional beamforming: A new enabling technology for 5G wireless networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 94–101, 2014.

[17] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selén, and J. Sköld, "5G wireless access: requirements and realization," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 42–47, 2014.

[18] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[19] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment and Practice," *IEEE Journal on Selected Areas in Communications*, 2017.

[20] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive mimo communication for future wireless systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 836–869, 2018.

[21] C. Jeong, J. Park, and H. Yu, "Random access in millimeter-wave beamforming cellular networks: issues and approaches," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 180–185, 2015.

[22] C. Park and T. S. Rappaport, "Short-range wireless communications for next-generation networks: UWB, 60 GHz millimeter-wave WPAN, and ZigBee," *IEEE Wireless Communications*, vol. 14, no. 4, pp. 70–78, 2007.

[23] "5G new radio: Waveform, frame structure, multiple access, and initial access, author=Lien, Shao-Yu and Shieh, Shin-Lin and Huang, Yenming and Su, Borching and Hsu, Yung-Lin and Wei, Hung-Yu," *IEEE communications magazine*, vol. 55, no. 6, pp. 64–71, 2017.

[24] K. Xiao, W. Li, M. Kadoch, and C. Li, "On the secrecy capacity of 5G MmWave small cell networks," *IEEE Wireless Communications*, vol. 25, no. 4, pp. 47–51, 2018.

[25] R. C. Daniels and R. W. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Vehicular technology magazine*, vol. 2, no. 3, pp. 41–50, 2007.

[26] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1224–1240, 2015.

[27] B. Li, D. Zhu, and P. Liang, "Small cell in-band wireless backhaul in massive MIMO systems: A cooperation of next-generation techniques," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 7057–7069, 2015.

[28] K. J. Zou, K. W. Yang, M. Wang, B. Ren, J. Hu, J. Zhang, M. Hua, and X. You, "Network synchronization for dense small cell networks," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 108–117, 2015.

[29] 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.

[30] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[31] 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.

[32] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[33] T. L. Marzetta *et al.*, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, p. 3590, 2010.

[34] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[35] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, "Massive MIMO performance—TDD versus FDD: What do measurements say?" *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2247–2261, 2018.

[36] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[37] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 905–923, 2016.

[38] O. Elijah, C. Y. Leow, A. R. Tharek, S. Nunoo, and S. Z. Iliya, "Mitigating pilot contamination in massive MIMO system—5G: An overview," in *2015 10th Asian Control Conference (ASCC)*. IEEE, 2015, pp. 1–6.

[39] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, 2016.

[40] T. L. Marzetta, G. Caire, M. Debbah, I. Chih-Lin, and S. K. Mohammed, "Special issue on massive MIMO," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 333–337, 2013.

[41] X. You and Y. Wang, "Massive mimo channel modeling," Ph.D. dissertation, Master's thesis, Lund University, 2015.

[42] X. Rao and V. K. 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, 2014.

[43] R. Shafin, L. Liu, and J. Zhang, "On the channel estimation for 3D massive MIMO systems," *IEEE COMSOC MMTC E-Letter*, vol. 9, no. 6, pp. 15–19, Nov 2014.

[44] M.-F. Tang, C.-C. Chen, and B. Su, "Downlink path-based precoding in FDD massive MIMO systems without CSI feedback," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016 IEEE*. IEEE, 2016, pp. 1–5.

[45] R. Shafin, L. Liu, J. Zhang, and Y. C. Wu, "Doa estimation and capacity analysis for 3-d millimeter wave massive-mimo/fd-mimo ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6963–6978, Oct 2016.

[46] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.

[47] J.-C. Shen, J. Zhang, E. Alsusa, and K. B. Letaief, "Compressed CSI acquisition in FDD massive MIMO: How much training is needed?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4145–4156, 2016.

[48] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2012, vol. 34.

[49] B. Allen and M. Ghavami, *Adaptive array systems: fundamentals and applications*. John Wiley & Sons, 2006.

[50] H. Xie, F. Gao, S. Zhang, and S. Jin, "UL/DL channel estimation for TDD/FDD massive MIMO systems using DFT and angle reciprocity," in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*. IEEE, 2016, pp. 1–5.

[51] K. Hugl, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channel in FDD systems," *European Cooperation in the field of Scientific and technical research*, May 2002.

[52] X. Zhang, J. Tadrous, E. Everett, F. Xue, and A. Sabharwal, "Angle-of-arrival based beamforming for FDD massive MIMO," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 704–708.

[53] U. Ugurlu, R. Wichman, C. B. Ribeiro, and C. Wijting, "A multipath extraction-based CSI acquisition method for FDD cellular networks with massive antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2940–2953, 2016.

[54] V. K. Lau, S. Cai, and A. Liu, "Closed-loop compressive CSIT estimation in FDD massive MIMO systems with 1 bit feedback," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 2146–2155, 2016.

[55] C. C. Tseng, J. Y. Wu, and T. S. Lee, "Enhanced Compressive Downlink CSI Recovery for FDD

Massive MIMO Systems Using Weighted Block ℓ_1 -Minimization," *IEEE Transactions on Communications*, vol. 64, no. 3, 2016.

[56] M. Li and Y. Lu, "A refined genetic algorithm for accurate and reliable DOA estimation with a sensor array," *Wireless Personal Communications*, vol. 43, no. 2, pp. 533–547, 2007.

[57] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[58] O. Birkenes, "On the limits to determining the direction of arrival for narrowband communication," *Diploma, Australian National University*, 2002.

[59] M. Al-Nuaimi, R. Shubair, and K. Al-Midfa, "Direction of arrival estimation in wireless mobile communications using minimum variance distortionless response," in *The Second International Conference on Innovations in Information Technology (IIT'05)*, 2005, pp. 1–5.

[60] R. Shafin, L. Liu, and J. C. Zhang, "DoA estimation and capacity analysis for 3D massive-MIMO/FD-MIMO OFDM system," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, December 2015, pp. 181–184.

[61] R. Shafin, L. Liu, and J. Zhang, "DoA Estimation and RMSE Characterization for 3D Massive-MIMO/FD-MIMO OFDM System," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.

[62] R. Shafin and L. Liu, "DoA Estimation and Performance Analysis for Multi-Cell Multi-User 3D mmWave Massive-MIMO OFDM System," in *2017 IEEE Wireless Communication and Networking Conference (WCNC)*, March 2017.

[63] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve

the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[64] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[65] 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.

[66] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

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

[68] Z. Han, H. Li, and W. Yin, *Compressive sensing for wireless networks*. Cambridge University Press, 2013.

[69] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Communications Surveys & Tutorials*, 2017.

[70] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," *IEICE transactions on communications*, vol. 96, no. 3, pp. 685–712, 2013.

[71] E. J. Candès *et al.*, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3. Madrid, Spain, 2006, pp. 1433–1452.

[72] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[73] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 884–893, 2006.

[74] E. Bjornson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1807–1820, 2010.

[75] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[76] L. Liu, R. Chen, S. Geirhofer, K. Sayana, Z. Shi, and Y. Zhou, "Downlink MIMO in LTE-Advanced: SU-MIMO vs. MU-MIMO," *IEEE Commun. Mag.*, pp. 140–147, Feb. 2012.

[77] S. Mosleh, J. D. Ashdown, J. D. Matyjas, M. J. Medley, J. Zhang, and L. Liu, "Interference Alignment for Downlink Multi-Cell LTE-Advanced Systems With Limited Feedback," *IEEE Transactions on Wireless communications*, vol. 15, no. 12, pp. 8107–8121, Dec. 2016.

[78] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on communications*, vol. 48, no. 3, pp. 502–513, 2000.

[79] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE transactions on information theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[80] S. Bazzi and W. Xu, "Downlink Training Sequence Design for FDD Multi-User Massive MIMO Systems," *IEEE Transactions on Signal Processing*, 2017.

[81] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems

with spatial channel correlation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2868–2882, 2015.

[82] E. Björnson, M. Kountouris, M. Bengtsson, and B. Ottersten, "Receive combining vs. multi-stream multiplexing in downlink systems with multi-antenna users," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3431–3446, 2013.

[83] S. Chatzinotas, M. A. Imran, and R. Hoshyar, "On the multicell processing capacity of the cellular MIMO uplink channel in correlated Rayleigh fading environment," *IEEE Transactions on Wireless communications*, vol. 8, no. 7, 2009.

[84] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, 2016.

[85] I. Bergel, Y. Perets, and S. Shamai, "Uplink downlink rate balancing and throughput scaling in FDD massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2702–2711, 2016.

[86] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *arXiv preprint arXiv:1304.6690*, 2013.

[87] B. Lee, J. Choi, J.-Y. Seol, D. J. Love, and B. Shim, "Antenna grouping based feedback compression for FDD-based massive MIMO systems," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3261–3274, 2015.

[88] H. Almosa, R. Shafin, S. Mosleh, Z. Zhou, Y. Li, J. Zhang, and L. Liu, "Downlink channel estimation and precoding for FDD 3D Massive MIMO/FD-MIMO systems," in *Wireless and Optical Communication Conference (WOCC), 2017 26th*. IEEE, 2017, pp. 1–4.

[89] J. Fang, X. Li, H. Li, and F. Gao, "Low-rank covariance-assisted downlink training and channel estimation for FDD massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1935–1947, 2017.

[90] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, 2017.

[91] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE communications magazine*, vol. 42, no. 10, pp. 54–59, 2004.

[92] W. Huang, Z. Lu, C. Zhang, Y. Huang, S. Jin, and L. Yang, "Beam-blocked compressive channel estimation for FDD massive MIMO systems," in *Wireless Communications and Networking Conference (WCNC), 2016 IEEE*. IEEE, 2016, pp. 1–6.

[93] C.-C. Tseng, J.-Y. Wu, and T.-S. Lee, "Enhanced Compressive Downlink CSI Recovery for FDD Massive MIMO Systems Using Weighted Block l1-Minimization," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1055–1067, 2016.

[94] N. Jindal, "MIMO broadcast channels with finite rate feedback," *arXiv preprint cs/0603065*, 2006.

[95] X. Su, J. Zeng, J. Li, L. Rong, L. Liu, X. Xu, and J. Wang, "Limited feedback precoding for massive MIMO," *International Journal of Antennas and Propagation*, vol. 2013, 2013.

[96] H. Wang, W. Wang, and Z. Zhang, "On the design of hybrid limited feedback for massive MIMO systems," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 4795–4800.

[97] S. Cho, K. Huang, D. Kim, and H. Seo, "Interference alignment for uplink cellular systems with limited feedback," *IEEE Communications Letters*, vol. 16, no. 7, pp. 960–963, 2012.

[98] J.-S. Kim, S.-H. Moon, S.-R. Lee, and I. Lee, "A new channel quantization strategy for MIMO interference alignment with limited feedback," *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 358–366, 2012.

[99] Y. Zhang, W. Feng, and N. Ge, "Dealing with large overhead in FDD massive MIMO systems: a one-bit feedback scheme," in *2016 23rd International Conference on Telecommunications (ICT)*. IEEE, 2016, pp. 1–5.

[100] W. Shen, L. Dai, Y. Shi, B. Shim, and Z. Wang, "Joint channel training and feedback for FDD massive MIMO systems," *IEEE Transactions on vehicular technology*, vol. 65, no. 10, pp. 8762–8767, 2016.

[101] H. Almosa, S. Mosleh, E. Perrins, and L. Liu, "Downlink Channel Estimation with Limited Feedback for FDD Multi-User Massive MIMO with Spatial Channel Correlation," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, May 2018.

[102] S. Sesia, I. Toufik, and M. Baker, "LTE: The UMTS long term evolution: From theory to practice," *2nd Edition, Wiley*, Aug. 2011.

[103] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, p. 461–471, Feb. 2004.

[104] T. Yoo and A. G. Smith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, p. 528–541, Mar. 2006.

[105] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[106] M. Sadek, A. Tarighat, and A. H. Sayed, "Active antenna selection in multiuser mimo communications," *IEEE Trans. Signal Process.*, vol. 55, no. 4, p. 1498–1510, Apr. 2007.

[107] B. Farhang-Boroujeny, Q. Spencer, and A. L. Swindlehurst, "Layering techniques for space-time communications in multi-user networks," *in Proc. 58th IEEE Veh. Technol. Conf.*, vol. 2, p. 1339–1343, Oct. 2003.

[108] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, 2015.

[109] W. Liu, S. Han, and C. Yang, "Energy efficiency scaling law of massive mimo systems," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 107–121, 2017.

[110] D. Ha, K. Lee, and J. Kang, "Energy efficiency analysis with circuit power consumption in massive MIMO systems," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2013, pp. 938–942.

[111] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, "Joint transmit antenna selection and user scheduling for massive MIMO systems," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2015, pp. 381–386.

[112] Y. Gao and T. Kaiser, "Antenna selection in massive MIMO systems: Full-array selection or subarray selection?" in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2016, pp. 1–5.

[113] T.-H. Tai, W.-H. Chung, and T.-S. Lee, "A low complexity antenna selection algorithm for energy efficiency in massive mimo systems," in *2015 IEEE International Conference on Data Science and Data Intensive Systems*. IEEE, 2015, pp. 284–289.

[114] S. Qin, G. Li, G. Lv, G. Zhang, and H. Hui, "L1/2-regularization based antenna selection for RF-chain

limited Massive MIMO systems," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2016, pp. 1–5.

[115] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, "Feedback Energy Reduction in Massive MIMO Systems," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[116] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *2013 IEEE globecom workshops (GC Wkshps)*. IEEE, 2013, pp. 129–134.

[117] H. Tang and Z. Nie, "RMV antenna selection algorithm for massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 239–242, 2018.

[118] P. V. Amadori and C. Masouros, "Power efficient massive MU-MIMO via antenna selection for constructive interference optimization," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 1607–1612.