Channel Estimation in Massive Multi-User MIMO Systems

Based on Low-Rank Matrix Approximation

Muamer Hawej

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Electrical and Computer Engineering) at

Concordia University

Montreal, Quebec, Canada

June 2019

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Muamer Hawej

Entitled: Channel Estimation in Massive Multi-User MIMO Systems Based on Low-Rank Matrix Approximation

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Electrical and Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____ Chair

Dr. Mehdi Hojjati

____ External Examiner

Dr. Georges Kaddoum

____ External to Program

Dr. Ali Dolatabadi

____ Examiner

Dr. Anjali Agarwal

____ Examiner

Dr. Wei-Ping Zhu

____ Thesis Supervisor

Dr. Yousef R. Shayan

Approved by ____

Dr. Rastko R. Selmic, Graduate Program Director

August 7, 2019 ____

Dr. Amir Asif, Dean

Gina Cody School of Engineering & Computer Science

Abstract

Channel Estimation in Massive Multi-User MIMO Systems Based on Low-Rank Matrix Approximation

Muamer Hawej, **PhD**

Concordia University, **2019**

In recent years, massive Multi-User Multi-Input Multi-Output (MU-MIMO) system has attracted significant research interests in mobile communication systems. It has been considered as one of the promising technologies for 5G mobile wireless networks. In massive MU-MIMO system, the base station (BS) is equipped with a very large number of antenna elements and simultaneously serves a large number of single-antenna users. Compared to traditional MIMO system with fewer antennas, massive MU-MIMO system can offer many advantages such as significant improvements in both spectral and power efficiencies. However, the channel estimation in massive MU-MIMO system is particularly challenging due to large number of channel matrix entries to be estimated within a limited coherence time interval. This problem occurs in a single-cell case where both dimensions of the channel matrix grow large. Also, It happens in the multi-cell setting due to the pilot contamination effect.

In this thesis, the problem of channel estimation in both single-cell and multi-cell time division duplex (TDD) massive MU-MIMO systems is studied. Thus, two-channel estimation namely "nuclear norm (NN)" and "iterative weighted nuclear norm (IWNN)" approximation techniques are proposed to solve the channel estimation problem in both systems.

First, channel estimation in a single-cell TDD massive MU-MIMO system is formulated as a convex nuclear norm optimization problem with regularization parameter γ. In this study, the regularization parameter γ is selected based on the cross-validation (CV) curve method. The simulation results in terms of the normalized mean square error (NMSE) and uplink achievable sum-rate (ASR) are provided to show the effectiveness of the NN proposed scheme compared to the conventional least square (LS) estimator. Then, the IWNN approximation is proposed to improve the performance of the NN method. Thus, the channel estimation in a single-cell TDD massive MU-MIMO system is formulated as a weighted nuclear norm optimization problem. The simulation results show the effectiveness of the IWNN estimation approach compared to the standard NN and conventional LS estimation methods in terms of the NMSE and ASR.

Second, both previous estimation techniques are extended to apply in a multi-cell TDD massive MU-MIMO system to mitigate pilot contamination effect. The simulation results in terms of the NMSE and uplink ASR show that the IWNN scheme outperforms the NN and LS estimations in the presence of high pilot contamination effect.

Finally, a novel channel estimation scheme namely "Approximate minimum mean square error (AMMSE)" is proposed to reduce the computational complexity of the minimum mean square error (MMSE) estimator which was proposed for multi-cell TDD massive MU-MIMO system. Furthermore, a brief analysis of the computational complexity regarding the number of multiplications of the proposed AMMSE estimator is provided. It has been shown that the complexity of the proposed AMMSE estimator is reduced compared to the conventional MMSE estimator. The simulation results in terms of the NMSE and the uplink ASR performances show the proposed AMMSE estimation performance is almost the same as the conventional MMSE estimator under two different scenarios: noise-limited and pilot contamination.

Acknowledgments

I would like to gratefully thank my research supervisor Prof. Yousef R. Shayan, for helping me to complete my Ph.D. thesis. During the five years of my Ph.D., I have seen in him an excellent supervisor who can bring the best out from his students, and a kind human being who is honest, fair and helpful to others. I have learned much from his exceptional knowledge and research enthusiasm.

It is my great pleasure to thank my thesis committee members and external examiners: Prof. Wei-Ping Zhu, Prof. Anjali Agarwal, Prof. Ali Dolatabadi (MIAE) from Concordia University, and Prof. Georges Kaddoum as external examiner from École de Technologie Supérieure (ÉTS), for their time of reviewing and constructive comments of my Ph.D. thesis.

I would like to gratefully acknowledge the Libyan Government, Elmergib University, for the entire financial support during my study period. I would like to gratefully acknowledge Concordia University, Natural Sciences and Engineering Research Council of Canada (NSERC) for the partial financial support during this research work.

Lastly, I would like to thank my wife for many sacrifices she has made to support me in undertaking my doctoral studies. I want to thank my parents, brothers, sisters, sons, and daughters, for their constant love, encouragement, and support in all my pursuits.

Contents

List of Figures x

List of Tables xiii

List of Acronyms and Notations xiv

1 Introduction 1

1.1 Literature Review 2

1.2 Motivation and Objectives 5

1.3 Contributions 6

1.4 Thesis Organization 8

2 Preliminaries and Background 10

2.1 Linear and Inverse Problems 10

2.2 Sparse Signals and lp-norm 12

2.3 Low-Rank Matrix Approximation-Based Sparse Matrix Estimation 12

2.3.1 Nuclear Norm Approximation 14

2.3.2 Iterative Weighted Nuclear Norm Approximation 15

2.4 Reconstruction Algorithms 17

2.4.1 Singular Value Thresholding Algorithm 17

2.5 Massive MU-MIMO System 19

2.5.1 Introduction 19

2.5.2 Uplink Channel Estimation 20

2.5.3 Uplink Data Detection 23

2.6 Computational Complexity of MMSE Channel Estimator 24

2.7 Summary 25

3 Nuclear Norm Approximation-Based Channel Estimation for Single-Cell Massive MU-MIMO System 26

3.1 Introduction 26

3.2 System and Channel Models. 27

3.2.1 System Model 27

3.2.2 Channel Model 28

3.3 LS Channel Estimation 30

3.4 Nuclear Norm (NN) Channel Estimation 31

3.4.1 Complexity analysis of NN Estimation 34

3.4.2 Selection of the Regularization Parameter, $\boldsymbol{\gamma}$, ... 34

3.5 Estimation Performance 35

3.5.1 Normalized Mean Square Error (NMSE) 35

3.5.2 Uplink Achievable Sum-Rate (ASR) 35

3.6 Simulation Results 38

3.7 Conclusion 42

4 Iterative Weighted Nuclear Norm Approximation-Based Channel Estimation for Single-Cell Massive MU-MIMO System 43

4.1 Introduction 43

4.2 Iterative Weighted Nuclear Norm (IWNN) Channel Estimation 44

4.2.1 Selection of Regularization Parameter and Weight Vector 47

4.2.2 Computational Complexity Analysis of IWNN Estimator 48

4.3 Simulation Results 50

4.4 Conclusion 54

5 Pilot Decontamination in Massive Multiuser MIMO Systems Based on Low-Rank Matrix Approximation 55

5.1 Introduction 55

5.2 System and Channel Models. 58

5.2.1 System Model 58

5.2.2 Channel Model 60

5.3 Pilot Contamination Effect on LS Channel Estimation. 61

5.4 NN Channel Estimation Method 62

5.5 IWNN Channel Estimation Method 66

5.6 Complexity Analysis of different Estimators 69

5.7 Estimation Performances 69

5.7.1 Normalized Mean Square Error (NMSE) 70

5.7.2 Uplink Achievable Sum-Rate (ASR) 70

5.8 Simulation Results 72

5.9 Conclusion 82

6 Low-Complexity Channel Estimator for TDD Massive Multiuser MIMO Systems 83

6.1 Introduction 83

6.2 Pilot-based Channel Estimation 84

6.2.1 MMSE Channel Estimator 84

6.2.2 Proposed AMMSE Channel Estimator 85

6.3 Estimator Complexity Analysis 90

6.4 Simulation Results 91

6.4.1 Noise-limited Scenario ... 92

6.4.2 Pilot Contamination Scenario ... 94

6.5 Conclusion ... 98

7 Conclusion and Future Work ... 99

7.1 Conclusion ... 99

7.2 Future Work ... 101

Bibliography ... 102

List of Figures

Figure 2.1: Massive MU-MIMO System [54]. 20

Figure 2.2: Channel Matrix in Single-Cell Massive MU-MIMO System 22

Figure 2.3: A simple illustration where the signal from all users share steering matrix A 22

Figure 3.1: Single-cell massive MU-MIMO systems with base station (BS) antennas M and serving K users.. 27

Figure 3.2: NMSE vs. Regularization parameter, γ, over system parameters, M= 80, ... 38

Figure 3.3: NMSE vs. Regularization parameter, γ over the system parameters M= 80, ... 39

Figure 3.4: Comparison between LS and NN estimation methods in terms of NMSE versus SNR over system parameters, M=80, P=20, K=40, and $\tau \in \{40,64\}$ 40

Figure 3.5: Comparison between LS and NN estimations in terms of NMSE versus SNR over system parameters, M=80, P=20, τ =64, and K$\in$ {40,64} 41

Figure 3.6: Uplink ASR vs. Number of BS antennas, M, over system parameters, K =40 users, P =20, and SNR= 0 dB.. 41

Figure 4.1: Comparison between IWNN, NN, and LS estimations over system parameters, M= 80, $\mathrm{P} = 20, \mathrm{K} = 40, \tau = 40$.. 50

Figure 4.2: Normalized estimation error versus Number of Antennas, M, over system parameters, $\mathrm{P} = 20, \mathrm{K} = 40, \tau = 40$, and SNR= 0 dB 51

Figure 4.3: Uplink ASR vs. Number of BS antennas, M, for a single-cell system with parameters, K =40 users and P =20, and SNR= 0 dB.. 52

Figure 4.4: The speed of convergence of the IWNN algorithm for single-cell system with parameters M=80, $\mathrm{P} = 20, \tau = 40, \mathrm{K} = 40$ at $\mathrm{SNR} = 0$ dB. 53

Figure 5.1: Pilot contamination in multi-cell massive MU-MIMO systems 58

Figure 5.2: NMSE vs. Regularization parameter γ for system parameters, M=80, P=20, τ=40, $Kd = 40$ with βjk=1, SNR=0 dB, and $Kc \in 0, 4{,}10$ with βlk= 0.8 $\forall l \neq j$ ... 73

Figure 5.3: Speed convergence of IWNN algorithm in multi-cell system with parameters M = 80, P = 20, τ = 40, Kd =40 with βjk=1, SNR=0dB, and Kc =4 with different values of βlk $\forall$ $l \neq j$ 74

Figure 5.4: NMSE vs. SNR over system parameters, M = 80, P =20, $\tau = 40$, $Kd = 40$ with βjk=1, and $Kc \in 4{,}10$ with βlk=0.8 $\forall$ $l \neq j$ 75

Figure 5.5: NMSE vs. large-scale fading coefficient, βlk for $Kc \in 4{,}10$ over system parameters, M =80, P=20, τ=40, Kd=40, with βjk=1, and SNR = 0 dB 76

Figure 5.6: NMSE vs. Number of contaminated Kc users with βlk= 0.9 each over system parameters, M = 80, P = 20, τ = 40, Kd = 40, and SNR = 0 dB 77

Figure 5.7: NMSE vs. Number of BS antennas M over systems parameters, P=20, τ=40, Kd=40 with βjk=1 each, Kc=10 with βlk=0.8 each $\forall$ $l \neq j$ and SNR = 0 dB 78

Figure 5.8: Uplink ASR vs. number of BS antennas M for multi-cell system parameters, L=3 cells, P=20, Kd=40 with βjk=1 each, Kc=4 with βlk=0.8 each $\forall$ $l \neq j$ and SNR=0 dB .. 79

Figure 5.9: Uplink ASR vs. large-scale factor βlk for Kc=4 users for multi-cell system with parameters, L=3, $M = 200$, P = 20, Kd = 40 with βjk = 1, SNR = 0 dB 80

Figure 5.10: Uplink ASR vs. number of Kc with βlk= 0.8 $\forall$ $l \neq j$ for system parameters, L=3 cells, M =200, P =20, Kd= 40 with βjk=1, and SNR = 0 dB 81

Figure 6.1: NMSE performance comparison between different estimators under different, M, in the noise-limited scenario (βlk= 0) 92
Figure 6.2: ASR performance comparison between different estimators under different, M, in the noise-limited scenario (βlk= 0). 93
Figure 6.3: NMSE performance comparison between different estimators under different, M, in the weak pilot contamination scenario (βlk= 0.1). 94
Figure 6.4: ASR performance comparison between different estimators under different, M, in the weak pilot contamination scenario (βlk= 0.1). 95
Figure 6.5: NMSE performance comparison between different estimators under different, M, in the strong pilot contamination scenario (βlk=0.9). 96
Figure 6.6: ASR performance comparison between different estimators under different, M, in the strong pilot contamination scenario (βlk=0.9). 97

List of Tables

Table 2.1: Duality concepts of vector cardinality and matrix rank minimization 13

Table 3.1: NN Estimation Algorithm for Single-Cell Massive MU-MIMO System 33

Table 4.1: IWNN Estimation Algorithm for Single-Cell Massive MU-MIMO System 49

Table 5.1: NN Estimation Algorithm for Multi-Cell Massive MU-MIMO System..................... 65

Table 5.2: IWNN Estimation Algorithm for Multi-Cell Massive MU-MIMO System................ 68

Table 6.1: IWNN Approximation Algorithm for MMSE Channel Estimator 89

Table 6.2: Asymptotic Complexities of Different Estimators 90

List of Acronyms and Notations

- **AoA** **A**ngle **o**f **A**rrival
- **AWGN** **A**dditive **W**hite **G**aussian **N**oise
- **ASR** **A**chievable **S**um-**R**ate
- **AMMSE** **A**pproximate **M**inimum **M**ean-**S**quared **E**rror
- **BPSK** **B**inary **P**hase **S**hift **K**eying
- **BER** **B**it **E**rror **R**ate
- **BS** **B**ase **S**tation
- **CV** **C**ross-**V**alidation
- **CS** **C**ompressive **S**ensing
- **CSI** **C**hannel **S**tate **I**nformation
- $\mathcal{CN}$ **C**omplex **N**ormal Distribution
- **EVD** **E**igen **V**alue **D**ecomposition
- **FDD** **F**requency **D**ivision **D**uplex
- **IWNN** **I**terative **W**eighted **N**uclear **N**orm
- **ICI** **I**nter-**C**ell **I**nterference
- **I. I. D.** **I**ndependent and **I**dentically **D**istribution
- **LS** **L**east **S**quare
- **LRMA** **L**ow-**R**ank **M**atrix **A**pproximation
- **MRC** **M**aximum **R**atio **C**ombining
- **MIMO** **M**ultiple-**I**nput-**M**ultiple-**O**utput

- **MU-MIMO** **M**ulti-**U**ser **M**ultiple-**I**nput-**M**ultiple-**O**utput
- **MMSE** **M**inimum **M**ean-**S**quared **E**rror
- **MUI** **M**ulti-**U**ser **I**nterference
- **MN** **M**inimum **N**orm
- **NN** **N**uclear **N**orm
- **NMSE** **N**ormalized **M**ean-**S**quared **E**rror
- **OFDM** **O**rthogonal **F**requency **D**ivision **M**ultiplexing
- **RLS** **R**egularized **L**east **S**quare
- **SDP** **S**emi-**D**efinite **P**rogramming
- **SVD** **S**ingular **V**alue **D**ecomposition
- **SNR** **S**ignal-to-**N**oise-**R**atio
- **SINR** **S**ignal-to-**I**nterference-**N**oise-**R**atio
- **TDD** **T**ime **D**ivision **D**uplex
- **ZF** **Z**ero **F**orcing
- **X:** Upper bold letter for matrix
- **x:** Lower bold letter for column vector
- $\mathbf{X}^T$**:** Transpose of matrix **X**
- $\mathbf{X}^H$**:** Hermitian (Conjugate transpose) of matrix **X**
- diag$[x]$: Diagonal matrix with x elements on its main diagonal
- tr$(\mathbf{X})$: Trace of matrix **X**
- $\|\mathbf{x}\|_2$: Euclidean norm of column vector **x**
- $\|\mathbf{X}\|_F$: Frobenius norm of matrix **X**

Chapter 1

Introduction

The demand for high-speed mobile communication service has been recently increased, and it is expected to continue over the years since the number of users is increasing [1]. With this growth, the radio spectrum has been recognized as one of the very precious resources of nature. However, the power available for wireless communication systems is limited due to battery life and device size. Therefore, many research efforts in academia and industry have been invested in increasing the network capacity and achieving high data rate for all users in the network.

Recently, massive multi-input-multi-output (MIMO) technology (also known as large-scale antenna systems) has attracted significant research interests in wireless communication systems [2]. This technology was first introduced by Thomas Marzetta (2010) and considered as a promising technology for 5G mobile communication system [3]. Compared to the traditional multiuser MIMO system with a few BS antennas, massive MIMO systems have many advantages, such as significant improvements in both spectral and power efficiencies in the network, and low hardware complexity of whole system [4], [5].

One of the recently proposed systems for 5G mobile communication technology is the massive multiuser multi-input-multi-output (MU-MIMO) system [6]. In massive MU-MIMO system, each base station (BS) is equipped with a very large number of antenna elements, M, and simultaneously serves a large number of K single-antenna users. Moreover, the previous results show that as the number of BS antennas increases in this system, both spectral and power efficiencies are quickly improved [6]. These improvements are strongly dependent on the availability of the channel state information (CSI) at the BS which does not hold in a real scenario [7], [8]. In other words, the BS needs to perfectly know the CSI in order to detect the data received in the uplink data phase, and also to perform the beamforming for the downlink. In practice, however, the perfect CSI is not available at the BS. Therefore, it will be estimated either in the uplink pilot transmission phase when the time division duplex (TDD) mode is used for massive MIMO, or it can be obtained by feedback link when the frequency division duplex (FDD) is used

[9], [10]. In the TDD system, the CSI is only needed to be estimated at the BS in the uplink pilot phase, which will be used in the downlink as well [11].

Generally speaking, the channel estimation problem in TDD massive MU-MIMO systems is particularly challenging due to a large number of the channel matrix entries to be estimated within a limited coherence time interval. This problem occurs in a single-cell setting where both dimensions of the channel matrix grow large [12]. It also happens in the multi-cell environment when the same orthogonal pilot sequences are reused by other users in the adjacent cells, which results in the so-called pilot contamination [13]. In other words, the same frequency band is used for all cells due to a limited coherence time interval. However, this effect is having a detrimental impact on the actual achievable spectral and energy efficiencies in real systems [14].

In general, the CSI estimation schemes can be classified as pilot-based and subspace-based methods [15]. In pilot-based methods, such as the least square (LS) and minimum mean-squared error (MMSE) approach, the CSI is estimated during the pilot transmission phase by transmitting pilot sequences from all users to their base stations [16], [17]. The CSI is determined during the uplink data phase when the subspace-based methods are used, such as blind and semi-blind channel estimation approaches.

1.1 Literature Review

Channel estimation in TDD massive MU-MIMO systems is considered as one of the exciting research topics, and various channel estimation methods have been proposed for single-cell and multi-cell environments. In multi-cell TDD massive MU-MIMO systems, several research efforts have been spent in the last ten years towards mitigating pilot contamination effect. In [18], an asynchronous time-shifted pilot protocol is proposed to reduce pilot contamination in TDD massive MU-MIMO system by avoiding the simultaneous transmission of pilot sequences from different users among all cells. The basic idea of the time-shifted pilot is to partition the number of cells into several groups. When the users in one group send the uplink pilot signals, the users in the other groups receive downlink data signals. However, the target base station simultaneously receives the uplink pilot signals and the interfering downlink data signals from other base stations, which results in the channel estimation corruption. Thus, this method may not provide accurate channel estimation due to the higher downlink transmit power compared to the uplink. Pilot decontamination based on the collaboration between all base stations has been proposed for TDD massive MU-MIMO system [19]-[21]. However, these approaches can lead to the complete removal of pilot contamination effect under a specific condition which is hard to implement in massive MIMO systems since the collaboration between all base stations is limited in real systems. Pilot decontamination approach based on a combination of a pilot sequence hopping scheme and a modified Kalman filter has been studied in [22], [23]. However, the channel estimation method is performed at multiple time slots. Therefore, this channel estimation approach has considerable computational complexity since the processing time will be too long.

The subspace-based (Blind and Semi-Blind) channel estimation techniques have been proposed to eliminate the pilot contamination effects in TDD massive MU-MIMO system [24]-[27]. In [24], an eigenvalue decomposition (EVD) method has been proposed for channel estimation where the channel matrix can be correctly estimated from the eigenvectors of the received samples of covariance matrix based on the assumption of system parameters. In [27], the authors propose applying a semi-blind channel estimation method to mitigate the effect of pilot contamination in multi-cell multiuser massive MIMO systems. However, this method is based on estimating the uplink data from different users in the target cell and then obtaining the least square channel estimation by treating the detected uplink data users as pilot symbols. Prominent drawbacks of the subspace-based estimation techniques are their high computational complexity, which will severely limit their applications in massive MIMO systems. Other estimation techniques based on coordinated pilot assignment strategies to mitigate pilot contamination in the multi-cell scenario have been proposed in [28]. In this technique, the authors developed a Bayesian channel estimator approach to minimize the pilot contamination effect in massive MIMO systems. This scheme works well under a specific non-overlap condition on the distributions of the multipath angle of arrivals (AoAs) for the desired and interference channels. In practice, however, this scheme is hard to implement in the real scenario. In [29], [30], the authors have proposed novel estimation algorithms to reduce the pilot contamination in TDD massive MIMO systems by exploiting the path diversity in both angle and power domains. In these algorithms, the channel covariance matrices of desired and interference users have to be perfectly known at each base station in order to remove the pilot contamination problem. In practice, however, it is hard to achieve this condition, especially when the interference links are overlapping with the desired links in both angular and power domains.

Unlike previous channel estimation methods, low-rank matrix approximation (LRMA) methods and compressive sensing (CS) techniques have been applied for various problems of wireless communication systems [31]-[37]. Recently, the compressive sensing technique has been applied as a new framework to address the channel estimation problem in a single-cell TDD massive MU-MIMO system [31]. In this technique, the channel estimation problem was formulated as a convex optimization problem and solved via a quadratic semi-define programming (SDP) solver. Due to the limitation of the SDP solver, this method may not be used in real massive MU-MIMO systems where the number of antenna elements is expected to be a large number. In [36], [37], the authors present a novel channel estimation approach which utilizes the sparsity and common support properties to estimate sparse channels and requires a small number of pilots. However, this approach has shown good estimation performance when it is applied for MIMO-OFDM systems.

In this section, various channel estimation techniques for single-cell and multi-cell TDD massive MU-MIMO systems have reviewed. Also, most of recent channel estimation techniques were proposed to mitigate the pilot contamination problem in multi-cell setting are explained. The literature survey indicates that studying the design of channel estimators for single and multi-cell massive MIMO systems would be the most productive effort. In other words, most of the channel estimation methods proposed for massive MIMO systems are of considerable computational complexity.

1.2 Motivation and Objectives

In this thesis, we are motivated to develop new channel estimation schemes for single-cell and multi-cell TDD massive MU-MIMO systems. Since a poor propagation scattering environment is assumed for both systems, the CSI can be estimated based on the low-rank matrix approximation (LRMA) techniques. In a multi-cell setting, the largest singular values of the estimated channel matrix are usually represented by the desired channel power received at the target base station, while the smallest singular values are represented by the interference and noise terms. This conclusion is motived us to mitigate the pilot contamination interference problem in a multi-cell TDD massive MU-MIMO system. Furthermore, the low-complexity channel estimators are required for real massive MIMO systems. Thus, we are motivated to develop new channel estimation scheme with low-complexity for real massive MU-MIMO system.

The objective of this thesis is to directly address the above practical challenges for single-cell and multi-cell TDD massive MU-MIMO systems. In single-cell case, it is beneficial to design a new channel estimator scheme with the capability to estimate the CSI with a minimum number of training sequences to achieve the desired latency. In a multi-cell setting, on the other hand, we have to develop new channel estimation techniques with the capability to cope with different pilot contamination scenarios. Moreover, the previously proposed MMSE channel estimator technique in [28] suffers from high computational complexity due to the large dimension of the covariance matrix inversion, which is scaled with the number of base station antennas. Hence, we aim to reduce the computational complexity of the MMSE estimator by designing a new channel estimator with low-complexity for massive MIMO systems.

1.3 Contributions

The main contribution of chapter 3 is to develop a new channel estimation scheme capable of estimating the CSI of a single-cell TDD massive MU-MIMO system within the limited coherence time interval. Hence, a novel channel estimation scheme, namely "nuclear norm (NN) approximation based on compressive sensing technique," is proposed. In the NN estimation scheme, the channel estimation is formulated as a convex nuclear norm optimization problem. The regularization parameter, γ, of this optimization problem is selected based on the cross-validation (CV)-curve method. Moreover, the CV-curve method is based on minimizing the normalized mean square error (NMSE) at each tuning parameter value, γ, for specific values of the signal-to-noise (SNR) ratio. The proposed NN estimation method is evaluated by using two different performance criterion, the NMSE, and uplink ASR. The simulation results are provided to show the effectiveness of the NN proposed scheme compared to the conventional least square (LS) estimation approach. The relevant contributions of this study are published in [38], [39].

The main contribution of chapter 4 is to improve the performance of the previously proposed NN channel estimation method for a single-cell TDD massive MU-MIMO system. Hence, a novel channel estimation scheme, namely "iterative weighted nuclear norm (IWNN) approximation," is proposed. In IWNN estimation scheme, the channel estimation is formulated as a weighted nuclear norm optimization problem and solved via the proposed iterative searching algorithm.

Furthermore, the initial values of the regularization parameter at each SNR value are selected based on cross-validation curve method. Also, the computational complexity of the IWNN estimation technique is studied in terms of the number of iterations. The simulation results show the effectiveness of the iterative weighted nuclear norm (IWNN) estimation approach compared to the standard NN and conventional LS estimation methods in terms of the NMSE and ASR performances. The relevant contributions of this study are published in [38], [40].

The main contribution of chapter 5 is to develop a new channel estimation scheme capable of mitigating pilot contamination problem in a multi-cell TDD massive MU-MIMO system. Hence, the applications of the NN and IWNN estimation schemes are extended. Moreover, the appropriate setting of the weight vector of the proposed IWNN has been taken into consideration in order to enhance the sparsity of singular values of the channel matrix. Further, a brief analysis of the computational complexity of the proposed NN and IWNN estimation approaches are analyzed and compared to the LS estimation method. The NMSE and uplink ASR performance criterion is used to evaluate the proposed NN and IWNN estimation methods under different pilot contamination scenarios. The simulation results are provided to show the effectiveness of the proposed IWNN channel estimation in the presence of high pilot contamination interference problem and compared to the NN and LS estimations in terms of the NMSE and uplink ASR. The relevant contributions of this study are published in [41].

The main contribution of chapter 6 is to reduce the computational complexity of the minimum mean square error (MMSE) estimator for multi-cell TDD massive MU-MIMO system. It is noteworthy that, the MMSE has been previously proposed for multi-cell massive multiuser MIMO systems. However, it suffers from high computational complexity due to the large dimension of the covariance matrix inversion, which is scaled with the number of base station antennas. Hence, a novel channel estimation scheme, namely "Approximate minimum mean square error (AMMSE)" is proposed. Moreover, the IWNN approximation based on the low-rank reduction theory is considered to design the proposed AMMSE estimator. A brief analysis of the computational complexity regarding the number of multiplications of the proposed AMMSE estimator is provided. It has been shown that the computational complexity of the proposed AMMSE estimator is reduced from $\mathcal{O}(M^3\tau^3)$ to $\mathcal{O}(M\tau PN)$ Compared to the conventional MMSE estimator. The simulation results show the agreements between the proposed AMMSE estimator and the conventional MMSE estimator in terms of the NMSE and the uplink ASR performances. Moreover, these estimation performances of the proposed AMMSE estimator have been investigated under two different scenarios: noise-limited and pilot contamination. The relevant contributions of this study are published in [42].

1.4 Thesis Organization

The rest of this thesis is organized as follows.

In chapter 2, the necessary background materials for an understanding of the mathematical tools needed for low-rank matrix approximation methods and the relevant background of massive MIMO systems are introduced.

In chapter 3, a novel channel estimation method, namely "nuclear norm (NN)" for single-cell TDD massive MU-MIMO system is analyzed and explained. Moreover, the performance of the proposed NN estimation method in terms of the normalized mean square error (NMSE) and uplink achievable sum-rate (ASR) is simulated over different values of the system parameters.

In chapter 4, the iterative weighted nuclear norm (IWNN)" channel estimation scheme for a single-cell massive MU-MIMO system is proposed to improve the previous NN estimation method. Also, the computational complexity of the IWNN estimation technique in terms of the number of iteration is studied.

In chapter 5, the NN and IWNN estimation techniques are applied to multi-cell TDD massive MU-MIMO systems to mitigate pilot contamination problem. The proposed NN and IWNN estimation performances in terms of the NMSE and uplink ASR are evaluated and compared to the conventional LS method under different pilot contamination scenarios. Moreover, a brief analysis of the computational complexity of the IWNN estimation scheme is analyzed and discussed.

In chapter 6, low-complexity channel estimator, namely “approximate minimum mean square error (AMMSE)” is designed for multi-cell TDD massive multiuser MIMO systems. The computational complexity of the proposed AMMSE estimator regarding the number of multiplications is reduced compared to the conventional MMSE estimator. Also, The performance of the proposed AMMSE channel estimator in terms of the NMSE and the uplink ASR is tested and compared to the conventional MMSE channel estimator under two different scenarios: noise-limited and pilot contamination.

Finally, the conclusions of this thesis and some future work areas are provided in chapter 7.

Chapter 2

Preliminaries and Background.

As mentioned in Chapter 1, the thesis research topics cover some areas in applied mathematics, signal processing, and wireless communications. In the following sections, we provide the necessary background materials for an understanding of the mathematical tools needed for low-rank matrix approximation methods and the relevant background of massive MIMO systems.

2.1 Linear and Inverse Problems

In this section, we begin with the review of a linear system with, m, Equations and, n, unknowns. Suppose that we need to reconstruct an unknown signal vector $\boldsymbol{x} \in \mathbb{R}^{n}$ from the known signal vector $\mathbf{y} \in \mathbb{R}^{m}$ via known matrix $\boldsymbol{A}$, i.e.

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{2.1}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the sensing matrix. In (2.1), $\boldsymbol{x}$ is assumed to be a sparse vector. We first consider the case with $m \geq n$ where the number of equations is greater than or equal to the number of unknowns. In this case, the system of linear equations in (2.1) is called determined system with a unique solution. When a signal vector $\boldsymbol{y}$ is noise-free, and $\boldsymbol{A}$ is an m x n full-rank matrix, the true solution is given by [43], [44] as:

$$\widehat{\boldsymbol{x}}_{LS} = (\boldsymbol{A}^{T}\boldsymbol{A})^{-1}\,\boldsymbol{A}^{T}\boldsymbol{y} \tag{2.2}$$

if the measurements include noise, a familiar way to cope with this problem is to rely on the method of least squares (LS), where the optimization problem is given by [45] as

$$\widehat{\boldsymbol{x}}_{LS} = arg\ \min_{\boldsymbol{x}}\ \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 \tag{2.3}$$

The system of linear equations in (2.1) is called underdetermined system equations with an infinite number of solutions when $m < n$ because of there exist infinitely many vectors in the null space of matrix $\boldsymbol{A}$. In other words, the number of equations in (2.1) is less than the number of unknowns [46], [47].

A common approach to solving the linear underdetermined system equations is the regularization approach, where the solution is chosen by minimizing the norm $\boldsymbol{x}$ [48]-[50]. A typical choice of the norm is the squared l_2-norm $||\boldsymbol{x}||_2^2$, and the regularization problem is formulated into two different cases as follows. In the case of a noise-free system

$$\widehat{\boldsymbol{x}}_{MN} = arg\,min\,||\boldsymbol{x}||_2^2 \quad subject\ to\ \ \boldsymbol{y} = \boldsymbol{Ax} \tag{2.4}$$

where MN stands for "minimum norm.". This problem can be analytically solved by the method of Lagrange multipliers as follows:

$$\mathcal{L}(\boldsymbol{x}) = ||\boldsymbol{x}||_2^2 + \lambda^T(\boldsymbol{Ax} - \boldsymbol{y}) \tag{2.5}$$

where λ is a regularization parameter, and the solution is given by

$$\widehat{\boldsymbol{x}}_{MN} = (\boldsymbol{A^T A})^{-\mathbf{1}}\,\boldsymbol{A^T y} \tag{2.6}$$

In the case of a noisy system, another approach might be the utilization of the regularized LS method which considers the optimization problem of the form

$$\widehat{\boldsymbol{x}}_{RLS} = arg\ \min_{\boldsymbol{x}}\ \ ||\boldsymbol{Ax} - \boldsymbol{y}||_2^2\ + \lambda\,||\boldsymbol{x}||_2^2 \tag{2.7}$$

The solution of (2.7) is given as:

$$\widehat{\boldsymbol{x}}_{RLS} = (\lambda\boldsymbol{I} + \boldsymbol{A^T A})^{-\mathbf{1}}\,\boldsymbol{A^T y} \tag{2.8}$$

where λ is the regularization parameter which is used to control the balance between the squared error and the l_2–norm of the solution in (2.7).

2.2 Sparse Signals and l_p-norm

In the analysis of algorithms for compressed sensing and low-rank matrix approximation methods, we encounter various norms, while we usually use the Euclidean l_p-norm in the conventional problems of communications [15], [50]-[52]. Thus, we first define some norms.

- The l_p- norm of a vector $\mathbf{x} = [x_1 \;\; x_2 \ldots x_n]^T \in \mathbb{C}\,\mathbb{R}^n$ is defined for $p \geq 1$ as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \tag{2.9}$$

where $[.]^T$ is the transpose of the vector $\mathbf{x}$. One can use formula (2.9) to define $\|\boldsymbol{x}\|_0$ which not even a quasi-norm, defined as $\|\boldsymbol{x}\|_0 = |supp\ (\mathbf{x})|$ where $supp\ (\mathbf{x}) = \{i: x_i \neq 0\}$ is the set of nonzero components, and $|supp\ (\mathbf{x})|$ is the cardinality of $supp\ (\mathbf{x})$. A signal $\mathbf{x} \in \mathbb{R}^n$ is said to be sparse (or exactly sparse) vector if most of the elements are precisely equal to zero, i.e., $\|\mathbf{x}\|_0 \ll n$. Some authors refer $\|\boldsymbol{x}\|_0$ to as l_0- norm, and it is a useful symbol to describe how sparse the vector $\mathbf{x}$ is.

- The Frobenius norm (F-norm) of an m x n matrix $\boldsymbol{A}$ is defined as

$$\|\boldsymbol{A}\|_F^2 = Tr(\boldsymbol{A}^H \boldsymbol{A}) = \sum_{i=1}^{r} {\sigma_i}^2(\boldsymbol{A}) \tag{2.10}$$

where σ_i is the i^{th} singular value of matrix $\boldsymbol{A}$ and r is the rank of a matrix $\boldsymbol{A}$.

2.3 Low-Rank Matrix Approximation-Based Sparse Matrix Estimation

In recent years, the recovery of low-rank matrices has seen significant activities in many areas of science and engineering [35], [53]-[58]. This is motivated by recent theoretical results for exact reconstruction guarantees and interesting practical applications where the data resides in a low-dimensional linear subspace [56], [57]. More specifically, the low-rank matrix approximation

(LRMA) method has been a growing interest in reconstructing sparse signals from a small number of incoherent linear measurements under suitable conditions [34], [51], [53], [57], [59], [60]. Moreover, the LRMA is a method to obtain a unique solution from an underdetermined linear system taking advantage of the prior knowledge that the right solution is sparse where the most of elements are precisely or approximately equal to zero.

Recently, the LRMA approach has been developed in [40], [41], [56]-[58], where it is shown that a data matrix can be approximating by one whose rank is less than that of the original matrix via convex optimization. However, it relies on the idea of a conventional compressive sensing technique [33], [60] by applying duality concepts between vector cardinality minimization and matrix rank minimization as shown in Table 2.1 [61].

Table 2.1: Duality concepts of vector cardinality and matrix rank minimization

Parsimony concept	vector cardinality	matrix rank
Hilbert space norm	Euclidean	Frobenius
Sparsity inducing	$l_1 - norm$	nuclear norm
Dual norm	$l_\infty - norm$	operator norm
Convex relaxation	linear programming	semi-definite programming

Mathematically, the LRMA problems involving the estimation of low-rank matrices can be formulated in a common framework as follows. Suppose that we need to estimate a sparse unknown matrix $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ with rank $r \ll \min(m,n)$ from its noisy observation matrix $\boldsymbol{Y} \in \mathbb{C}^{m\times n}$, i.e.,

$$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N} \tag{2.11}$$

where **N** is the additive white Gaussian noise (AWGN) matrix. The estimation of sparse low-rank matrices has been studied and used for various applications such as covariance matrix estimation, subspace clustering, image classification [32], [53], [59], [61], [62]. An approach for estimating the sparse low-rank matrix **X** from its noisy observation matrix **Y** has been proposed in [53] by solving the following optimization problem

$$\begin{aligned} &minimize\ rank\ (\boldsymbol{X}) \\ &subject\ to\ \ \boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N} \end{aligned} \tag{2.12}$$

This optimization problem is NP-hard (i.e., l_0-Norm) and no known polynomial-time algorithms exist to solve it. Therefore, we proceed with convex relaxation based on the nuclear norm to obtain a sparse solution in the next subsections.

2.3.1 Nuclear Norm Approximation

Nuclear norm (NN) approximation technique is a particular case of the LRMA and is considered as the l_1-Norm applied to the non-zero singular values of the low-rank matrix [56], [57]. As mentioned in the subsection above, the optimization problem (2.12) is NP-hard, and no known polynomial-time algorithms exist to solve it. Thus in [31], [61], the convex relaxation based on the nuclear norm has been proposed to solve the optimization problem (2.12) by replacing rank (**X**) by $\|\mathbf{X}\|_*$ as

$$\begin{aligned} &minimize\ \|\boldsymbol{X}\|_* \\ &subject\ to\ \ \|\boldsymbol{Y} - \boldsymbol{X}\|_F^2 < \epsilon \end{aligned} \tag{2.13}$$

where $\|\boldsymbol{X}\|_*$ is the nuclear norm of $\boldsymbol{X}$, $\|.\|_F^2$ denotes the F-norm, and ϵ is the predefined noise threshold. The nuclear norm of $\boldsymbol{X}$ is defined as

$$\|\boldsymbol{X}\|_* = \sum_{i=1}^{r} |\sigma_i(\boldsymbol{X})| \tag{2.14}$$

where σ_i is $the\ i^{th}$ singular value of matrix $\boldsymbol{X}$, and $r = \min(m, n)$ is the rank of the matrix $\boldsymbol{X}$. The conventional approach to obtaining a unique solution to the optimization problem (2.13) is to use the regularization parameter in the framework of the above optimization problem as

$$\widehat{\boldsymbol{X}} = arg\,\min_{X} \left\{ \frac{1}{2} \|\boldsymbol{Y} - \boldsymbol{X}\|_F^2 + \gamma \|\boldsymbol{X}\|_* \right\} \tag{2.15}$$

where $\gamma \succ 0$ is a regularization parameter and is considered as the main part of the minimization problem. Note that the NN optimization problem in (2.15) is convex and its global minimum can be directly obtained using the singular value decomposition (SVD) of the input signal matrix $\mathbf{Y}$, then the solution to the NN problem is given by

$$\widehat{\boldsymbol{X}} = \boldsymbol{U}.\mathcal{S}oft(\boldsymbol{\Sigma};\, \gamma).\boldsymbol{V}^* \tag{2.16}$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix whose entries are the singular values of matrix $\boldsymbol{Y}$, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices. In (2.16), the soft-threshold function is applied to each singular value in $\boldsymbol{\Sigma}$ as

$$\mathcal{S}oft(\sigma_i;\, \gamma) = max(\sigma_i - \gamma, 0) \tag{2.17}$$

2.3.2 Iterative Weighted Nuclear Norm Approximation

In the previous subsection, the NN optimization method has certain limitations because it treats all singular values of the channel matrix equally. In other words, it ignores the prior knowledge of the most significant singular values of the sparse matrix. However, this is significantly restricted in the NN estimation method for many practical problems, such as channel estimation in a wireless communication system where the most entries of the channel matrix are sparse. Therefore, the IWNN approximation methods have been recently applied in diverse contexts in machine learning and signal processing to improve the NN channel estimation [40], [63], [64]. Moreover, it is an interesting topic drawing the attention of many researchers in the closely related field of sparse approximation. In other words, the IWNN aims to estimate the unknown data matrix by assigning different weights to different singular values of the known data matrix using an adaptive regularization parameter threshold γ. Thus, the largest singular values of the data matrix shrink less than the smallest ones since the former represents the most significant data information. In [40], [41], [64], the IWNN approximation recently has been proposed to solve the optimization problem (2.11) by replacing rank $(\boldsymbol{X})$ in (2.12) by $\|\boldsymbol{X}\|_{\boldsymbol{w},*}$ as

$$\begin{gathered} minimize\ \|\boldsymbol{X}\|_{\boldsymbol{w},*} \\ subject\ to\ \ \boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N} \end{gathered} \tag{2.18}$$

where $\|\boldsymbol{X}\|_{\boldsymbol{w},*}$ is the weighted nuclear norm of a matrix $\boldsymbol{X}$, which is defined as

$$\|\boldsymbol{X}\|_{\boldsymbol{w},*} = \sum_{i=1}^{r} |w_i \sigma_i(\boldsymbol{X})| \tag{2.19}$$

where $w_i \geq 0$ is a non-negative weight element assigned to each singular value σ_i of the matrix $\mathbf{X}$. In (2.19), the non-negative weight element brings more parameters to the system model, and therefore it is proposed to enhance the sparsity of the nonnegative singular value solutions of the IWNN estimation by adaptively tuning each weight element, w, by using the following formula:

$$w_i^{t+1} = \frac{1}{\sigma_i^t(\boldsymbol{X}) + \varepsilon} \quad i = 1,2 \ldots \ldots \ldots r \tag{2.20}$$

where σ_i^t is the i^{th} singular value of the approximation channel matrix $\boldsymbol{X}$ in the t^{th} iteration, w_i^{t+1} is its corresponding regularization parameter γ in the $(t+1)$-th iteration, ε is a positive small number to avoid dividing by zero. The noisy version of the optimization problem (2.18) is now the IWNN regularization problem which is given as

$$\widehat{\boldsymbol{X}} = arg \min_{\boldsymbol{X}} \left\{\frac{1}{2} \|\boldsymbol{Y} - \boldsymbol{X}\|_F^2 + \gamma \|\boldsymbol{X}\|_{\boldsymbol{w},*}\right\} \quad (2.21)$$

where $\gamma \succ 0$ is a regularization parameter. It should be noted that the IWNN optimization problem in (2.21) is convex when the weights satisfy $w_i \geq 0$ for $i = 1,2 \ldots . r$, and it has a globally optimal solution

$$\widehat{\boldsymbol{X}} = \boldsymbol{U}.\mathcal{S}oft(\boldsymbol{\Sigma};\ \boldsymbol{w})\,.\boldsymbol{V}^* \quad (2.22)$$

where $\boldsymbol{Y} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^*$ is the SVD of $\boldsymbol{Y}$, and $\mathcal{S}oft(\boldsymbol{\Sigma};\ \boldsymbol{w})$ is the generalized soft-thresholding operator with weight vector $\boldsymbol{w}$. Each weight element is assigned to each singular value in $\boldsymbol{\Sigma}$ as

$$\mathcal{S}oft(\sigma_i;\ w_i) = max(\sigma_i - w_i, 0) \quad (2.23)$$

2.4 Reconstruction Algorithms

Practical algorithms to solve the LRMA problems are singular value decomposition, pivoted QR decomposition, interpolative decomposition, and randomized algorithms such as sampling-based methods and random projection-based methods [39]-[41], [43], [56], [57], [65], [66]. Note that, to avoid high computational complexity, we should investigate the structure of each problem and choose a suitable algorithm to exploit it. In the next subsection, we illustrate the best approximation algorithm of a low-rank matrix called thin singular value thresholding (SVT) algorithm [67].

2.4.1 Singular Value Thresholding Algorithm

This subsection introduces the singular value thresholding (SVT) and discusses some of its basic properties. We begin with the definition of a critical building block, namely " the singular value decomposition (SVD) operator." The SVD is a powerful technique in linear algebra since it gives the best approximation of a low-rank matrix $\boldsymbol{X}\,\mathbb{C}^{m\times n}$ (where m > n) [55]. Moreover, it does not require square matrices, and therefore the number of rows, m, does not have to be equal to the number of columns, n.

To explain the idea behind this algorithm, we consider the singular value decomposition (SVD) of a complex matrix $\boldsymbol{X}\ \mathbb{C}^{m\times n}$ of rank n

$$\boldsymbol{X}_{m\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\,\boldsymbol{V}^{*}_{n\times n} \tag{2.24}$$

where **U** and **V** are unitary matrices, i.e., all of the columns of **U** and **V** are orthogonal to each other. The matrix **Σ** is an $m \times n$ rectangular diagonal matrix whose entries are in descending order, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$, along with the main diagonal.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{2.25}$$

where the number of nonzero singular values of the matrix $\boldsymbol{\Sigma}$ is the rank of the matrix $\boldsymbol{X}\ \mathbb{C}^{m\times n}$. Since $m > n$, one can represent the SVD of matrix $\boldsymbol{X}\ \mathbb{C}^{m\times n}$ as

$$\boldsymbol{X}_{m\times n} = \boldsymbol{U}_{m\times n}\,\boldsymbol{\Sigma}_{n\times n}\,\boldsymbol{V}^{*}_{n\times n} \tag{2.26}$$

Here, $\boldsymbol{\Sigma}_{n\times n} = \mathrm{diag}\,(\sigma_1,\ \sigma_2, \dots \dots, \sigma_n)$ is called thin SVD of $\mathbf{X}_{m\times n}$.

Now, suppose we need to approximate a matrix $\mathbf{X}\ \mathbb{C}^{m\times n}$ with $r = \min(m, n)$ by using thin SVD as

$$\boldsymbol{X}_{m\times n} = \boldsymbol{U}_{m\times r}\,\boldsymbol{\Sigma}_{r\times r}\,\boldsymbol{V}^{*}_{r\times n} \tag{2.27}$$

which can be written as:

$$\boldsymbol{X}_{m\times n} = [u_1 u_2 \dots \dots \dots . u_r] \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_r^* \end{bmatrix} \tag{2.28}$$

Then, the rank approximation of $\mathbf{X}_{\mathrm{m}\times\mathrm{n}}$ is given as

$$\boldsymbol{X}_{m\times n} = \sum_{i=1}^{r} \sigma_i \, u_i v_i^* \tag{2.29}$$

where $u_1, u_2, \dots, u_r$ are columns of $\boldsymbol{U}_{m\times r}$ and $v_1, v_2, \dots, v_r$ are columns of $\boldsymbol{V}_{m\times r}$. It can be seen from (2.29) that the matrix $\boldsymbol{X}$ can be approximated by the low-rank approximation using SVD. The projection of the matrix $\boldsymbol{X}$ onto the space spanned by the top singular r vectors of $\boldsymbol{X}$ is called the rank-r approximation (also known as truncated or partial SVD) of $\boldsymbol{X_r}$. The Eckart-Young theorem in [55] states that the above approximation technique is the best rank-r approximation in the F-norm, and it requires only $\mathcal{O}(rmn)$ floating-point operations (flops). Therefore, we will adopt the SVD approximation technique in our work.

2.5 Massive MU-MIMO System

2.5.1 Introduction

Massive multi-input-multi-output (MIMO) technology (also known as large-scale antenna systems) was first introduced by Thomas L. Marzetta (2010) [3]. This technology has been considered as a promising technology for 5G wireless communication systems [1], [9], [68]. Compared to the existing MIMO systems, massive MIMO systems have many advantages, such as significant improvements in spectral efficiency, energy efficiency, and low-complexity of hardware implementation in the receiver side [2].

One of the recently proposed systems for massive MIMO technology is a massive multiuser MIMO (MU-MIMO) system as shown in Figure 2.1 [6]. In massive MU-MIMO system, each base station (BS) is equipped with a very large number of antennas, and simultaneously serves a large number of single-antenna users. Information theory has demonstrated that by increasing the number of BS antennas and a multiplicity of distant single-antenna users can offer vast improvements in spectral efficiency.

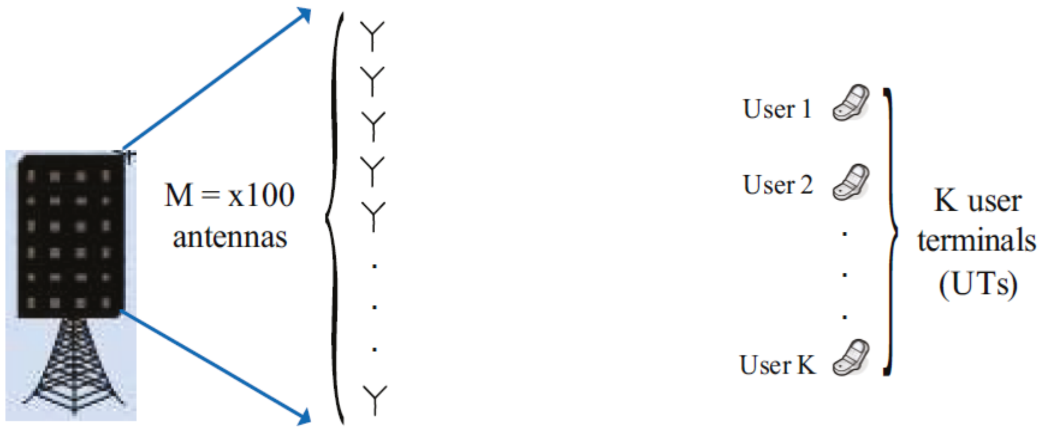

Figure 2.1: Massive MU-MIMO System [54].

However, these advantages can be only achieved if the channel state information (CSI) is perfectly known at the BS. In other words, the interference signals from adjacent cells will be rejected by applying a simple precoding approach under the above assumption. In practice, however, the base station does not know the CSI, and thus it will be estimated either in the uplink training phase when time division duplex (TDD) mode is used for massive MIMO system or can be obtained by feedback link when the frequency division duplex (FDD) mode is used. However, the FDD operation requires two links (training downlink and CSI feedback link) to obtain the CSI at the BS, which results in increased processing time and pilot overhead.

2.5.2 Uplink Channel Estimation

In the TDD system, The CSI is estimated during the uplink pilot phase, where all users from all cells transmit their pilot sequences to their base stations. Consequently, each BS will use the estimated CSI to detect the data received in the uplink phase and perform the precoding for downlink. In existing MIMO systems, the conventional pilot-based channel estimation approaches, such as least square (LS) and minimum mean-squared error (MMSE) estimators are used to estimate the uplink CSI at the base station [16], [28]. The LS estimation technique requires that the length of the training sequences (τ) used by each user to be at least equal to the number of transmit antennas. On the other hand, the MMSE estimation technique requires the covariance matrices of both desired and interference channels to be available at each base station in each coherence time interval. Therefore, the channel estimation problem in TDD massive MU-MIMO systems is particularly challenging due to a large number of channel matrix entries to be estimated during the limited coherence time interval.

Figure 2.2 shows that the large dimension of the channel matrix occurs where both the number of BS antennas and autonomous users in a single-cell massive MU-MIMO system grow large. This effect may be too restrictive in a massive MU-MIMO system, and thus both LS and MMSE estimation methods lead to excessive use of critical communication resources such as energy and spectrum. Moreover, the channel estimation problem also occurs in a multi-cell massive MU-MIMO system when the non-orthogonal pilot sequences are reused in the adjacent cells by other users resulting in so-called pilot contamination effect [14]. This phenomenon has a significant impact on the uplink channel estimation performance, which results in a substantial reduction in the available rates of user terminals. The problem becomes more critical when the gains of cell-edge users are relatively strong as compared to the direct link gains.

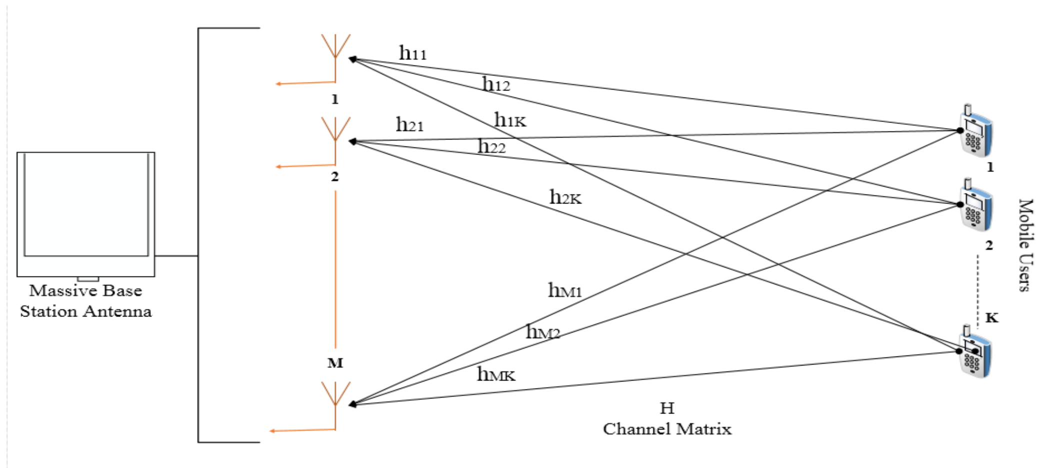

Figure 2.2: Channel Matrix in Single-Cell Massive MU-MIMO System

In this study, we focus only on the poor scattering propagation environment where the number of physical objects is limited [7]. Thus, the actual degree of freedom of the channel matrix is $P(M + K - P)$, not its number of free parameters MK. Figure 2.3 shows the finite scattering propagation multipath channel model where the number of multipath scattering, P, appears in a group with similar delays and angle of arrival (AoA).

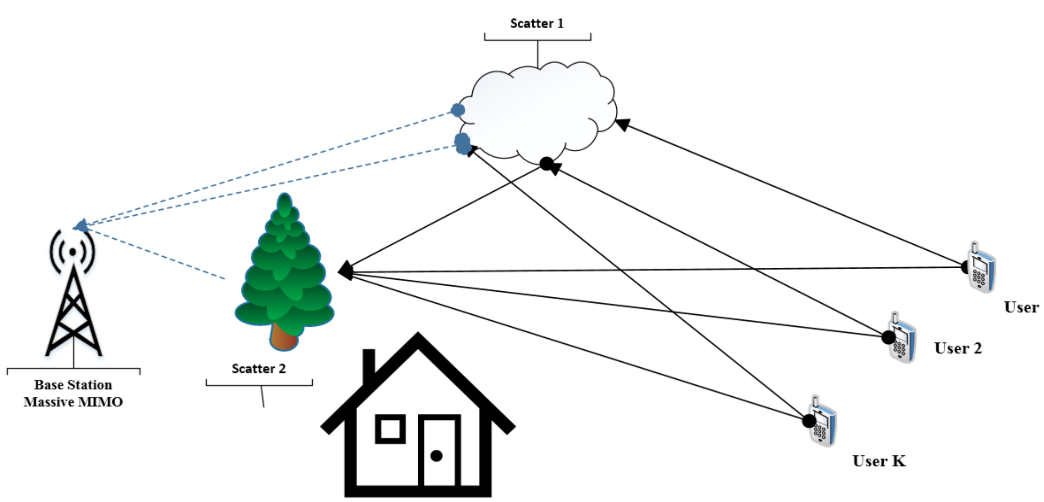

Figure 2.3: A simple illustration where the signal from all users share steering matrix A

2.5.3 Uplink Data Detection

A Linear detection processing such as maximum ratio (MR), zero-forcing (ZF), and minimum mean square error (MMSE) combining schemes are used to detect the uplink data received at the BS [69]-[71]. The BS applies the linear processing combining detector to detect the desired data symbols from the interfering ones. The MR combing is the best solution to detect the uplink data users if the data received at the BS is disturbed by only additive white Gaussian noise (AWGN). In the case of high interference, the optimum combining detectors, such as ZF and MMSE are used at the receiver side. Furthermore, the MMSE detector is more robust than the ZF in the case of high interference.

In this study, the linear MMSE detection scheme with equal power allocation is assumed, which has a detection matrix $\boldsymbol{V}_{\mathrm{MMSE}}$, given as

$$\boldsymbol{V}_{MMSE} = \left(\widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}} + \sigma_n^2 \boldsymbol{I}_K \right)^{-1} \widehat{\boldsymbol{H}}^H \quad (2.30)$$

where $\widehat{\boldsymbol{H}} \triangleq [\, \widehat{\boldsymbol{h}}_1, \widehat{\boldsymbol{h}}_2, \dots\dots \; \widehat{\boldsymbol{h}}_k] \in \mathbb{C}^{M \times K}$ is the estimated channel matrix which is assumed to be estimated during the uplink pilot phase, and σ_n^2 denotes the noise variance. In a single-cell TDD massive MU-MIMO system, the CSI is estimated in the uplink pilot phase and used to perform the data detection in the uplink. To illustrate the basic idea, we consider the uplink data transmission phase where all users K inside the cell simultaneously transmit data symbols, $s_1, s_2 \dots\dots\dots, s_K$, to the BS. Then at the BS, the data symbol of k^{th} user $\hat{s}_k$ is detected by multiplying the total received data symbols with detecting vector $\boldsymbol{v}_k$ of a linear MMSE detector matrix $\boldsymbol{V}_{\mathrm{MMSE}}$ as

$$\hat{s}_k = \boldsymbol{v}_k \boldsymbol{\mathcal{Y}} \quad (2.31)$$

where $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{M \times 1}$ is the total base-band received data symbols at the BS. Now, we consider a multi-cell TDD massive MU-MIMO system with L cells where each BS-l $(1 \leq l \geq L)$ at first detects the signal vector $\hat{s}_k$ by multiplying the total received data symbols with detecting vector $\boldsymbol{v}_{jk}$ of a linear MMSE detector matrix $\boldsymbol{V}_{\text{MMSE}}$ at the target BS-j as

$$\hat{s}_k = \boldsymbol{v}_{jk}\, \boldsymbol{\mathcal{Y}} \tag{2.32}$$

where row vector $\boldsymbol{v}_{jk}$ of matrix $\boldsymbol{V}_{\text{MMSE}}$ can be expressed as

$$\boldsymbol{v}_{jk} = \left(\widehat{\boldsymbol{h}}_{jk}^{H} \ \ \widehat{\boldsymbol{h}}_{jk} + \sigma_n^2\right)^{-1} \widehat{\boldsymbol{h}}_{jk}^{H} \tag{2.33}$$

The uplink achievable sum-rate (ASR) of K users is a performance metric used to evaluate the effectiveness of the proposed estimation method. The achievable uplink rate of K users is computed by using the Shannon capacity as

$$ASR^{uplink} \leq \sum_{k=1}^{K} R_k^{uplink} \leq \sum_{k=1}^{K} log_2(1 + SINR_k) \tag{2.34}$$

where $SINR_k$ is the signal-to-interference-noise-ratio of the k^{th} user at the MMSE detector output. In general, the $SINR_k$ can be computed by using the following formula

$$SINR_k = \frac{Channel\ power\ gain}{MUI + ICI + Noise} \tag{2.35}$$

2.6 Computational Complexity of MMSE Channel Estimator

In literature, the MMSE channel estimator has been previously proposed for multi-cell massive multiuser MIMO systems [28]. The MMSE estimator suffers from high computational complexity due to the large dimension of the covariance matrix inversion which is scaled with the number of base station antennas [72]. In other words, the total number of multiplication operations required to estimate the desired channel matrix by using the MMSE estimator is M^3K^3. Another inherent drawback of the MMSE channel estimator is needed additional information about the statistical distribution of the propagation channels at each base station. However, this information is not available in the real massive MIMO systems.

2.7 Summary

In this chapter, the necessary background materials for an understanding of the mathematical tools needed for low-rank matrix approximation methods have been presented. To start with, the low-rank matrix approximation methods with mathematical representation were introduced. Next, the relevant background of massive MU-MIMO system was presented, and the channel estimation problems in a single and multi-cell TDD massive MU-MIMO systems were explained. Further, the uplink achievable sum-rate capacity of desired users in a single and multi-cell TDD massive MU-MIMO systems was presented. Also, the computational complexity of MMSE channel estimation for TDD massive MU-MIMO system was discussed.

Chapter 3

Nuclear Norm Approximation-Based Channel Estimation for Single-Cell Massive MU-MIMO System

3.1 Introduction

The problem of the uplink channel estimation in a single-cell TDD massive MU-MIMO system is considered as one of the significant challenges due to a large number of the channel matrix entries to be estimated within the limited coherence time interval [1], [3], [6], [73]-[75]. This problem occurs when the base station (BS) antennas and serving users grow large in a given cell. In literature, the conventional least square (LS) and minimum mean square error (MMSE) estimation methods are used to estimate the channel matrix for traditional MIMO systems with a few BS antennas. On the other hand, the LS and MMSE channel estimators may not be directly used in the massive MIMO systems for the following reasons. The LS estimation requires the length of the pilot sequence, τ, used by each user to be at least equal to the total number of users inside the cell, i.e., $\tau \geq K$ [16], [17], while the MMSE estimator has considerable computational complexity since a matrix inversion is scaled with the number of BS antennas [76], [77].

Compressed sensing and low-rank matrix approximation (LRMA) are advanced techniques that have critical applications in many areas of science and engineering [32], [33], [52], [54], [56], [57]. In literature, compressive sensing technique was applied as a new framework for various problems of wireless communication systems, such as the sparse channel estimation problem [31], [33], [54], [60]. Recently, Compressive sensing has been proposed for channel estimation in a single-cell massive MU-MIMO system [54]. However, this estimation technique is only feasible for systems with a small number of users and base station antennas. Therefore, our aim in this research is to develop a new channel estimation scheme capable of estimating the CSI of a single-cell TDD massive MU-MIMO system during the limited coherence time interval. Hence, a novel channel estimation scheme namely " Nuclear Norm Approximation," is proposed. Moreover, the NN estimation proposed method is evaluated by using two different performance criterion, normalized mean square error (NMSE) and uplink achievable sum-rate (ASR).

3.2 System and Channel Models.

3.2.1 System Model

In this study, a single cell TDD massive MU-MIMO system with one BS as shown in Figure 3.1 is considered. The BS is equipped with a very large number of antenna elements M and simultaneously serves a large number of K single-antenna users, i.e., $M \gg K$ [6].

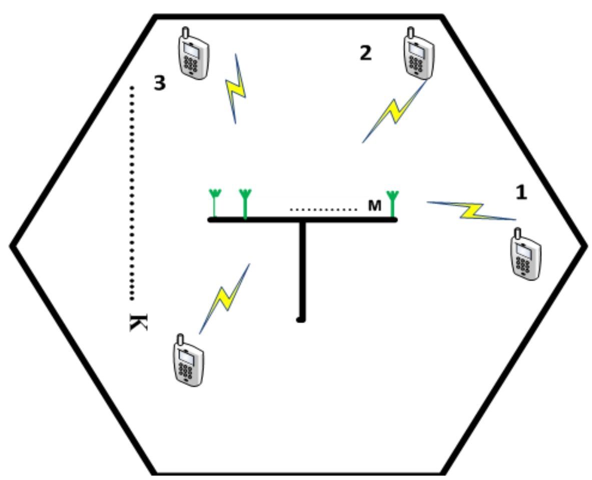

Figure 3.1: Single-cell massive MU-MIMO systems with base station (BS) antennas M and serving K users

We consider the uplink pilot transmission phase where all K users are simultaneously transmitted their signals $\boldsymbol{x}(t)$ to the desired BS. Moreover, the orthogonal Bernoulli pilot sequences with a BPSK modulation scheme are assumed in order to avoid the inter-user interference at the BS, i.e., $x_k(t) \in \{+1, -1\}$. At each time t, the total transmitted signal vector $\boldsymbol{x}(t) \in \mathbb{C}^K$ from K users to the BS is denoted by

$$\boldsymbol{x}(t) = [\ x_1(t),\ x_2(t) \ldots \ldots\ x_K(t)]^T \tag{3.1}$$

and, the received baseband signal vector at the BS is then given as

$$\boldsymbol{y}(t) = \sqrt{\rho_{tr}}\ \boldsymbol{H}\, \boldsymbol{x}(t) + \boldsymbol{n}(t) \tag{3.2}$$

where ρ_{tr} is the transmitted symbol power from each user, and $\boldsymbol{H} \triangleq [\ \boldsymbol{h}_1\ \ \boldsymbol{h}_2 \ldots.\ \boldsymbol{h}_K\ ] \in \mathbb{C}^{M \times K}$ is the complex-valued channel matrix between K users and the BS. In (3.2), $\boldsymbol{n}(t) \in \mathbb{C}^M$ is the complex-valued additive white Gaussian noise (AWGN) vector with zero mean and unit variance σ_n^2, i.e. $\mathcal{CN}(0,\ 1)$.

The total received pilot sequences at the BS can be combined in matrix form as

$$\boldsymbol{Y} = \sqrt{\rho_{tr}}\ \boldsymbol{H}\, \boldsymbol{X} + \boldsymbol{N} \tag{3.3}$$

where $\boldsymbol{X} \triangleq [\,\boldsymbol{x}(1)\ \ \boldsymbol{x}(2) \ldots.\boldsymbol{x}(\tau)\,] \in \mathbb{C}^{K \times \tau}$ is the total transmit pilot sequences from K users to the BS and $\boldsymbol{N} \in \mathbb{C}^{M \times \tau}$ is the spatially and temporally white additive Gaussian noise (AWGN) matrix with zero-mean and element-wise variance σ_n^2, i.e. $\sim \mathcal{CN}(0,\ \boldsymbol{I}_M)$.

3.2.2 Channel Model

In this study, the realistic finite scattering multipath channel model is assumed which has been considered for massive MIMO systems [28], [35], [54]. The channel vector between the k^{th} user and the BS is given by

$$\boldsymbol{h}_k = \frac{\beta_k}{\sqrt{P}} \sum_{p=1}^{P} \boldsymbol{a}(\ \theta_p)\ g_{kp} \tag{3.4}$$

where g_{kp} is the fading gain coefficient between the k^{th} user and the BS associated with each path $p \in 1,2 \dots \dots P$, and β_k is the path loss coefficient between the k^{th} user and the BS which is denoted by

$$\beta_k = \sqrt{\frac{\alpha}{(\ d_k)^{\gamma}}} \tag{3.5}$$

where d_k is the geographical distance between the k^{th} user and the BS, γ is the path-loss exponent, and α is a constant. All path loss coefficients between K users and the BS are assumed to be the same and normalized to unity. In (3.4), $\boldsymbol{a}(\theta_p)$ is the steering vector originating from each k^{th} user to the BS associating with each path $p \in 1,2 \dots \dots P$ which is given by

$$\boldsymbol{a}(\theta_p) = \left[1,\ e^{-j2\pi\frac{D}{\lambda}cos(\theta_p)} . \dots .\ \ e^{-j2\pi(M-1)\frac{D}{\lambda}cos(\theta_p)}\right]^T \tag{3.6}$$

where λ is the signal wavelength, D is the antenna spacing which is assumed to be fixed, and $\theta(p\ ) \in [-\pi/2, \pi/2]$ is a random angle of arrival (AoA) corresponding to each path p.

With the notations above, the channel model for single-cell TDD massive MU-MIMO system can be collectively written in a matrix form as

$$\boldsymbol{H}\ = \boldsymbol{A}\,\boldsymbol{G} \tag{3.7}$$

where $\boldsymbol{A} \triangleq [\ \boldsymbol{a}(\theta_1\ )\ \boldsymbol{a}(\theta_2\ ) \dots \dots \boldsymbol{a}(\theta_P)\ ]$ is a $M\ \times\ P$ matrix containing steering vectors, $\boldsymbol{G} \triangleq [\ \boldsymbol{g}_1\ \ \boldsymbol{g}_2 \dots \dots .\ \boldsymbol{g}_K]$ is the $P\ \times\ K$ matrix of the fading coefficients between the K users and the BS. All elements in each vector $\boldsymbol{g}_k \triangleq [\ g_{k1}\ \ g_{k2} \dots .\ g_{kP}]^T$ are assumed to be independent Rayleigh fading coefficients with zero mean and unit variance, i.e. , $g_{kp} \sim \mathcal{CN}(0,\ 1)$.

3.3 LS Channel Estimation

The conventional LS channel estimator is a pilot-based channel estimator which is used to estimate the CSI during the uplink pilot transmission phase by correlating the received signals at the base station with known orthogonal pilot sequences [16], [17]. Based on the system and channel models presented above, the LS estimate of $\boldsymbol{H}$ is then given by

$$\widehat{\boldsymbol{H}}_{LS} = \frac{1}{\sqrt{\rho_{tr}}} \quad \boldsymbol{Y}\,\boldsymbol{X}^{H}(\boldsymbol{X}\boldsymbol{X}^{H})^{-1} \tag{3.8}$$

where $\widehat{\boldsymbol{H}}_{LS}$ is the least square estimated channel matrix, and $\boldsymbol{X}$ is the orthogonal known pilot matrix, i.e. $\boldsymbol{X}^{H}\boldsymbol{X} = \tau\,\boldsymbol{I}_{K}$. By substituting $\boldsymbol{X}^{H}\boldsymbol{X} = \tau\,\boldsymbol{I}_{K}$ and (2.11) into (3.8), we rewrite (3.8) as

$$\widehat{\boldsymbol{H}}_{LS} = \boldsymbol{H} + \frac{1}{\tau\sqrt{\rho_{tr}}}\,\boldsymbol{N}\,\boldsymbol{X}^{H} \tag{3.9}$$

and then, the optimal estimation error in terms of the F-norm is given as

$$\min_{\boldsymbol{H}} \mathbb{E}\left\{\left\| \widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right\|_{F}^{2}\right\} = \frac{KM}{\tau(\rho_{tr}/\sigma_{n}^{2})} \tag{3.10}$$

As it appears in (3.10), the LS channel estimation performance is limited by noise contamination which results in the poor channel estimation performance. The noise contamination effect occurs when the received signal power at the BS is small. Another inherent drawback of the LS-based channel estimation is the spectral efficiency loss due to the bandwidth consumed by training sequences [4, 8]. By observing the above drawbacks of the LS-based channel estimation, we proceed with the application of LRMA methods to develop a new estimation technique for single-cell TDD massive MU-MIMO systems in the next subsection.

3.4 Nuclear Norm (NN) Channel Estimation

The NN estimation method is a convex optimization problem, and its global solution can be directly obtained using the singular value decomposition (SVD) of the input matrix [56], [57], [61]. Furthermore, the NN is a particular case of the LRMA method which solves the relaxation version of the rank minimization problem. In a real massive MU-MIMO system, the channel matrix model in (3.7) can be considered as a low-rank matrix since it has many sparse singular values. Consequently, the NN channel estimation method based on compressive sensing technique is proposed for TDD massive MU-MIMO system.

Based on the LS channel estimation in (3.9), the channel estimation problem in a single-cell TDD massive MU-MIMO system is formulated as a nuclear norm minimization problem as follows:

$$\widetilde{\boldsymbol{H}} = arg\,\min_{\boldsymbol{H}}\left\{\frac{1}{2}\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right\|_F^2 + \gamma\|\boldsymbol{H}\|_*\right\} \tag{3.11}$$

where $\widetilde{\boldsymbol{H}}$ is a complex approximate channel matrix, and γ is the regularization parameter. In (3.11), $\|\boldsymbol{H}\|_*$ is the nuclear norm of $\boldsymbol{H}$, which is the sum of its singular values. The nuclear norm of $\boldsymbol{H}$ can be defined as:

$$\|\boldsymbol{H}\|_* = \sum_{i=1}^{r}|\sigma_i(\boldsymbol{H})| \tag{3.12}$$

where $\sigma_i(\boldsymbol{H})$ is the i^{th} singular value of the channel matrix $\boldsymbol{H}$ and $r \leq \min(M, K, P)$ is the rank of channel matrix $\boldsymbol{H}$. By substituting (3.12) into (3.11), we rewrite (3.11) as

$$\widetilde{\boldsymbol{H}} = arg\,\min_{\boldsymbol{H}}\left\{\frac{1}{2}\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right\|_F^2 + \gamma\sum_{i=1}^{r}|\sigma_i(\boldsymbol{H})|\right\} \tag{3.13}$$

The property of the squared Frobenius norm of a matrix (.) is defined as

$$\|.\|_F^2 = Tr((.)^H(.)) = \sum_{i=1}^{r} {\sigma_i}^2(.) \tag{3.14}$$

The property in (3.14) is then applied to (3.13) as

$$\widetilde{\boldsymbol{H}} = arg\min_{\boldsymbol{H}} \left\{ \frac{1}{2}\, Tr\left\{ \left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H} \right)^H \left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H} \right) \right\} + \gamma \sum_{i=1}^{r} |\sigma_i(\boldsymbol{H})| \right\} \tag{3.15}$$

which is equivalent to:

$$\widetilde{\boldsymbol{H}} = arg\min_{\boldsymbol{H}} \left\{ \frac{1}{2} \left(Tr\left(\left(\widehat{\boldsymbol{H}}_{LS} \right)^H \left(\widehat{\boldsymbol{H}}_{LS} \right) \right) - 2Tr\left(\left(\widehat{\boldsymbol{H}}_{LS} \right)^H (\boldsymbol{H}) \right) + Tr\left((\boldsymbol{H})^H (\boldsymbol{H}) \right) \right) + \gamma \sum_{i=1}^{r} |\sigma_i(\boldsymbol{H})| \right\} \tag{3.16}$$

In (3.16), the optimization problem achieves the upper bound $Tr\left\{ \left(\widehat{\boldsymbol{H}}_{LS} \right)^H \left(\widehat{\boldsymbol{H}}_{LS} \right) \right\} = \sum_{i=1}^{r} \sigma_i(\widehat{\boldsymbol{H}}_{LS}) \sigma_i(\boldsymbol{H})$ if $\widehat{\boldsymbol{U}}_{LS}^{\boldsymbol{H}} = \boldsymbol{U}$ and $\widehat{\boldsymbol{V}}_{LS} = \boldsymbol{V}^{\boldsymbol{H}}$. As mentioned earlier in the channel model section, this condition can be satisfied in a massive MU-MIMO system when all users inside the cell are sharing the same steering matrix $\boldsymbol{A}$. Based on this conclusion; we may further simplify (3.16) as:

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}) = arg\min_{\boldsymbol{\sigma}(\boldsymbol{H})} \left\{ \frac{1}{2} \left(\sum_{i=1}^{r} {\sigma_i}^2(\widehat{\boldsymbol{H}}_{LS}) - 2 \sum_{i=1}^{r} \sigma_i(\widehat{\boldsymbol{H}}_{LS}) \sigma_i(\boldsymbol{H}) + \sum_{i=1}^{r} {\sigma_i}^2(\boldsymbol{H}) \right) + \gamma \sum_{i=1}^{r} |\sigma_i(\boldsymbol{H})| \right\} \tag{3.17}$$

In vector form, the optimization problem in (3.17) is simplified as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}) = \min_{\boldsymbol{\sigma}(\boldsymbol{H})} \left\{ \frac{1}{2} \left\| \boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}(\boldsymbol{H}) \right\|_2^2 + \gamma \sum_{i=1}^{r} |\sigma_i(\boldsymbol{H})| \right\} \tag{3.18}$$

where $\boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS})$, $\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}})$ are the singular value vectors of the LS and the channel approximation, respectively. The proposed NN estimation method for massive MU-MIMO channel estimation is summarized in Table 3.1.

Table 3.1: NN Estimation Algorithm for Single-Cell Massive MU-MIMO System

1: Apply the SVD method of $\widehat{\boldsymbol{H}}_{LS}$ as

$$[\widehat{\boldsymbol{U}}_{LS}\,\boldsymbol{\Sigma}_{LS}\,\widehat{\boldsymbol{V}}_{LS}] = SVD(\widehat{\boldsymbol{H}}_{LS}) \tag{3.19}$$

where a $M \times K$ diagonal matrix $\boldsymbol{\Sigma}_{LS}$ whose diagonal elements are the singular values of $\widehat{\boldsymbol{H}}_{LS}$ corresponding to the left and right eigenvectors of unitary matrices $\widehat{\boldsymbol{U}}_{LS} \in \mathbb{C}^{M \times M}$ and $\widehat{\boldsymbol{V}}_{LS} \in \mathbb{C}^{K \times K}$.

2: Choose the regularization parameter γ by using Cross-Validation (CV) curve method, as explained above.

3: Solve an optimization problem in (3.18) to obtain the singular values estimation matrix $\widetilde{\boldsymbol{\Sigma}}$, which is defined as.

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} diag(\tilde{\sigma}_1(\widetilde{\boldsymbol{H}}), \ldots\ldots \tilde{\sigma}_r(\widetilde{\boldsymbol{H}}), 0, \ldots\ldots \tilde{\sigma}_K(\widetilde{\boldsymbol{H}}) \\ \boldsymbol{0}_{M-K \times K} \end{pmatrix} \tag{3.20}$$

4: Finally, the estimated channel matrix is determined as

$$\widetilde{\boldsymbol{H}} = \widehat{\boldsymbol{U}}_{LS}\ \widetilde{\boldsymbol{\Sigma}}\ \ \widehat{\boldsymbol{V}}_{LS}^{H} \tag{3.21}$$

3.4.1 Complexity analysis of NN Estimation

The main complexity of the proposed NN estimation method comes from the optimization problem in step 3 which has K real variables and one linear constraint. In comparison with the channel estimation scheme in [31], we have managed to reduce the number of variables from $M\tau$ complex variables to K real variables and one $(M + K) \times (M + K)$ semidefinite constraint to one linear constraint. In contrast, the proposed NN channel estimation scheme has higher complexity compared to the LS estimation, but better estimation performance.

3.4.2 Selection of the Regularization Parameter, γ,

The regularization parameter, γ, in (3.18) is an essential issue for the success of the NN optimization problem which controls the trade-off error between the data fidelity, $\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right\|_F^2$, and the prior information $\|\boldsymbol{H}\|_*$. The discrepancy principle (DP), cross-validation (CV), and L-curve are some of the empirical existing selection methods[47], [78], [79]. The regularization parameter in DP is selected based on the sum of squares of the weighted residuals which is equal to the mean of a chi-square distribution [47]. The L-curve selection method is based on a log-log plot of the solution norm versus the residual norm error for different values of the regularization parameter, γ, [78], [79]. The CV criterion is based on the selection of the regularization parameter that minimizes the normalized mean square error (NMSE) of the optimization problem.

In this study, it is proposed to use a CV-curve method to select the optimal values of the regularization parameter γ. Moreover, the CV-curve method is based on minimizing the normalized mean square error (NMSE) at each tuning parameter value γ for specific values of the signal-to-noise (SNR) ratio. The following formula can be used to select the optimal value of the regularization parameter γ using the CV criteria

$$\gamma = \underset{\gamma \in \{\gamma_1, \gamma_2, \dots \gamma_{max}\}}{arg\,min} NMSE(\gamma) \tag{3.22}$$

where $\gamma_1 \geq 0$ is the initial value of the regularization parameter and γ_{max} is the maximum value of the regularization parameter, which is assumed to be any integer value such that $\gamma_{max} \gg \gamma_1$.

3.5 Estimation Performance

In this section, we need to study the choice of the system parameters for the proposed channel estimator so that it is robust to variation in the SNR and number of base station antennas. In practice, the estimation performance criteria used to evaluate any estimation methods are NMSE, ASR, and BER. In this study, the NMSE and ASR criteria are only used to evaluate the proposed estimation techniques which will be explained in the following two subsections. In [27], it has been shown that the BER performance for massive MIMO systems are improved as the number of base station antenna increases which is similar to the NMSE performance.

3.5.1 Normalized Mean Square Error (NMSE)

The following formula for the NMSE is used to evaluate the proposed estimation which is given by [15], [31]as

$$NMSE(dB) = 10log_{10}\left(\frac{E\left\|\widetilde{\boldsymbol{H}} - \boldsymbol{H}\right\|_F^2}{E\|\boldsymbol{H}\|_F^2}\right) \tag{3.23}$$

where $\boldsymbol{H}$ and $\widetilde{\boldsymbol{H}}$ are the desired channel matrix at the BS-j and its estimate, respectively.

3.5.2 Uplink Achievable Sum-Rate (ASR)

The uplink ASR capacity performance of K users is given by the Shannon capacity formula as

$$ASR \leq \sum_{k=1}^{K} log_2(1 + SINR_k) \tag{3.24}$$

where $\mathrm{SINR_k}$ is the received signal-to-interference-noise ratio of each k^{th} user at the linear detector processing output. In literature, the linear detector processing such as maximum ratio combing (MRC), zero-forcing (ZF), and minimum mean square error (MMSE) detector schemes are used at the BS to detect the uplink data symbol for each k^{th} user [12], [69], [70], [80]. In massive multiuser MIMO systems, however, the MRC detector scheme may not be the best solution due to the strong multiuser interference. On the other hand, at a low signal-to-noise ratio (SNR), ZF detector does not work well due to the noise enhancement [69]. Moreover, the MMSE detector scheme has the capability of excluding the disadvantages of both MCR and ZF detectors which will be adapted in our work.

In this study, it is proposed to use a linear minimum mean square error (MMSE) detector at the BS-j to detect the uplink data symbol for each k^{th} user. Then, the MMSE detector matrix $\boldsymbol{V}_{MMSE} \in \mathbb{C}^{K\times M}$ is given as

$$\boldsymbol{V}_{MMSE} = \left(\widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}} + \sigma_n^2 \boldsymbol{I}_K \right)^{-1} \widehat{\boldsymbol{H}}^H \tag{3.25}$$

where $\widehat{\boldsymbol{H}} \triangleq [\, \widehat{\boldsymbol{h}}_1, \widehat{\boldsymbol{h}}_2, \dots \ \widehat{\boldsymbol{h}}_k] \in \mathbb{C}^{M\times K}$ is the estimated channel matrix which is obtained during the channel estimation phase, and σ_n^2 is the noise variance. To compute the $\mathrm{SINR_k}$ of each k^{th} user in (3.24), we consider the uplink data transmission phase where all users in the cell-j are simultaneously transmitted their data symbols, $s_1, s_2, \dots, s_K$, to the BS-j. During the uplink data transmission phase, the received data vector $\boldsymbol{y}(n) \in \mathbb{C}^{M\times 1}$ at each time symbol, n, is given by

$$\boldsymbol{y}(n) = \sqrt{\rho_d} \sum_{k=1}^{K} \boldsymbol{h}_k \, s_k(n) + \boldsymbol{z}(n) \tag{3.26}$$

where ρ_{d} is the average symbol power used by each user, $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}$ is the channel vector between the k^{th} user and the BS antennas, and $\boldsymbol{z}$ is the AWGN vector with zero mean and unit variance, σ_n^2. Then, the data symbol detected of each k^{th} user at the MMSE detector output is computed by multiplying the k^{th} row vector $\boldsymbol{v}_k$ of $\boldsymbol{V}_{MMSE}$ in (3.25) with the received signal vector $\boldsymbol{\mathcal{Y}}$ in (3.26) as

$$\hat{s}_k(n) = \boldsymbol{v}_k\, \boldsymbol{\mathcal{Y}}(n) \tag{3.27}$$

Then further, we extend (3.27) as

$$\hat{s}_k(n) = \sqrt{\rho_k}\, \boldsymbol{v}_k\, \boldsymbol{h}_k\, s_k(n) + \sum_{i=1, i\neq k}^{K} \sqrt{\rho_i}\, \boldsymbol{v}_k\, \boldsymbol{h}_i s_i(n) + \boldsymbol{v}_k\, \boldsymbol{z}(n) \tag{3.28}$$

In (3.28), the first term represents the received data symbol of the k^{th} user, while the second and third terms are representing the interference from other users and noise, respectively. The $\mathrm{SINR_k}$ of each k^{th} user at the MMSE detector output can be computed as

$$SINR_k = \frac{\rho_k\, |\boldsymbol{v}_k\, \boldsymbol{h}_k|^2}{\sum_{i=1, i\neq k}^{K} \rho_i\, |\boldsymbol{v}_k\, \boldsymbol{h}_i|^2 + |\boldsymbol{v}_k\, \boldsymbol{z}|^2} \tag{3.29}$$

By substituting (3.29) into (3.24), we compute the uplink ASR of K users as

$$ASR \leq \sum_{k=1}^{K} log_2\left(1 + \frac{\rho_k\, |\boldsymbol{v}_k\, \boldsymbol{h}_k|^2}{\sum_{i=1, i\neq k}^{K} \rho_i\, |\boldsymbol{v}_k\, \boldsymbol{h}_i|^2 + |\boldsymbol{v}_k\, \boldsymbol{z}|^2}\right) \tag{3.30}$$

where the k^{th} row vector $\boldsymbol{v}_k$ of $\boldsymbol{V}_{MMSE}$ is defined as

$$\boldsymbol{v}_k = \left(\, \widehat{\boldsymbol{h}}_k{}^H\, \widehat{\boldsymbol{h}}_k + \sigma_n^2\right)^{-\mathbf{1}} \widehat{\boldsymbol{h}}_k{}^H \tag{3.31}$$

3.6 Simulation Results

This section contains the simulation results of the proposed channel estimation schemes for TDD massive MU-MIMO system. We consider a single-cell system with the only base station (BS) which is located at the center of the cell. The BS is equipped with a large number of M=80 antennas and simultaneously serves K=40 single-antenna users. It should be noted that the following assumptions have been taken through all the simulation results presented in this section. The number of multipath originating from each user to the BS is P=20, angle of arrival θ_p associated with each path is $\theta(p) = -\pi/2 + (p-1)\pi/2P$, and the steering vector parameters D and λ are assumed to have D/λ =0.5, where p=1, 2…….P. In the first experiment, the regularization parameter γ of the proposed estimation method is selected based on the CV curve method.

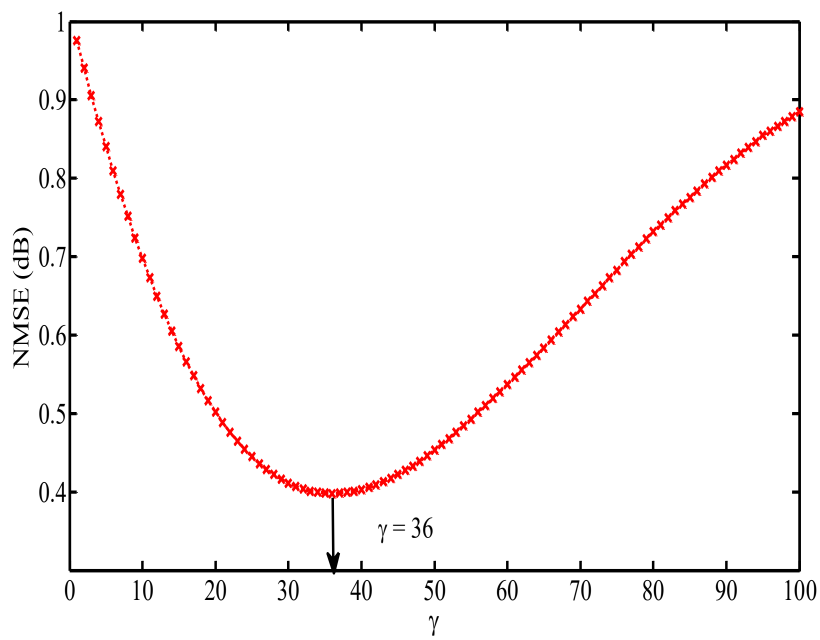

Figure 3.2: NMSE vs. Regularization parameter, γ, over system parameters, M= 80, $\mathrm{P} = 20, \mathrm{K} = 40, \tau = 40$, and SNR= 0dB

Figure 3.2 and Figure 3.3 illustrate two examples of how the optimal regularization parameter γ of the proposed estimation method is chosen at low and high SNR values (0 and 10 dB) which are equal to 36 and 8, respectively

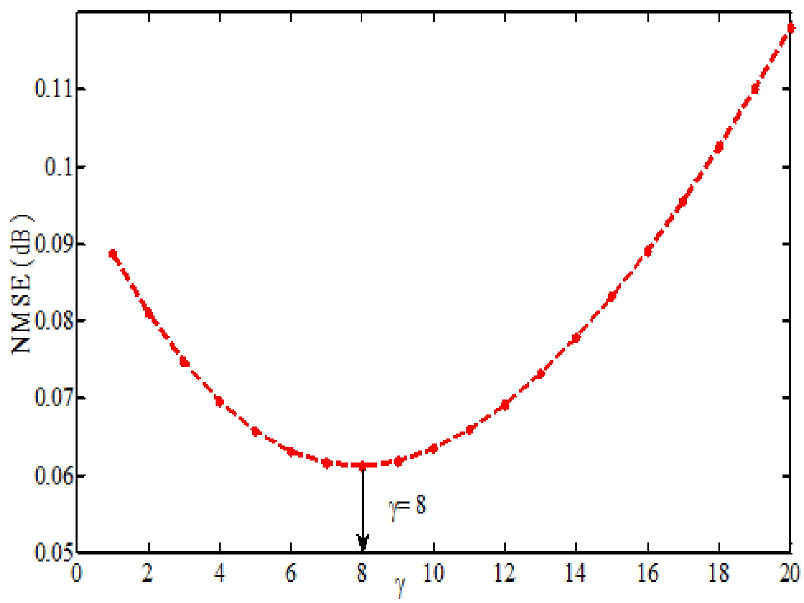

Figure 3.3: NMSE vs. Regularization parameter, γ over the system parameters M= 80, $\mathrm{P} = 20, \mathrm{K} = 40, \tau = 40$, and SNR= 10 dB

In Figure 3.4, we display the NMSE versus the SNR values with a different length of pilot sequence $\tau = 40, 64$, for both LS and the NN estimation methods. It can be seen from Figure 3.4 that when τ increases, both LS and NN channel estimation performance are decreased, but the proposed NN-based method achieves significantly better performance, as expected from the analysis. Then next, the effect of increasing the number of K users is studied under the fixed number of BS antennas, i.e., M=80. Figure 3.5 shows that the NMSE of both estimators increases when we increase the number of K users. Again, the performance of the proposed NN estimation outperforms the LS method for different numbers of K users. Furtheremore, the performance loss K=40 to K=40 of NN is less than that of the LS.

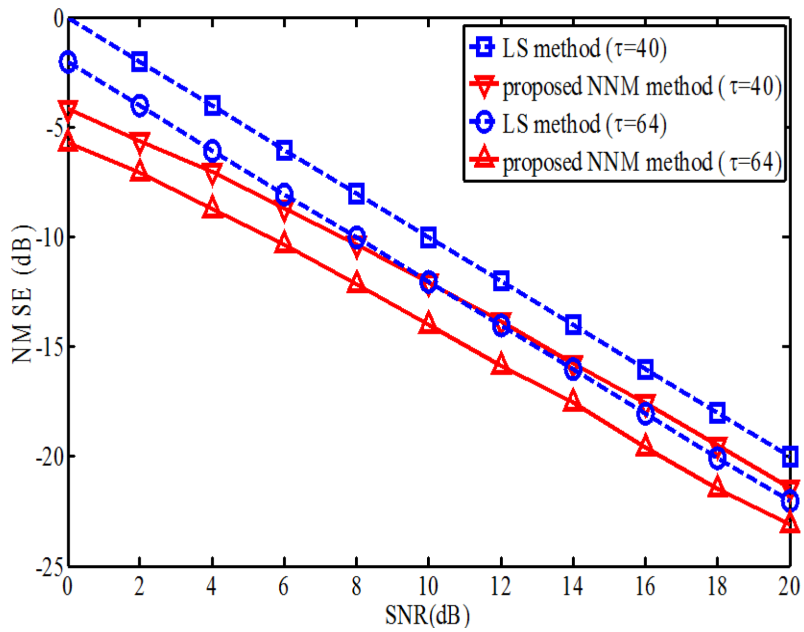

Figure 3.4: Comparison between LS and NN estimation methods in terms of NMSE versus SNR over system parameters, M=80, P=20, K=40, and τ ∈ {40,64}

Finally, we evaluate the effectiveness of the NN proposed estimation method in terms of the uplink ASR performance. Figure 3.6 illustrates the uplink ASR performance of the proposed NN estimator under a different number of BS antennas M and compared to the LS estimation method. For reference, the uplink ASR upper bound obtained with perfect CSI is also simulated. It can be seen from Figure 3.6 that as the number of BS antennas M increases, the uplink ASR of the proposed NN estimation is 2 bps/Hz better than the LS estimation and about 6 bps/Hz less than the perfect CSI.

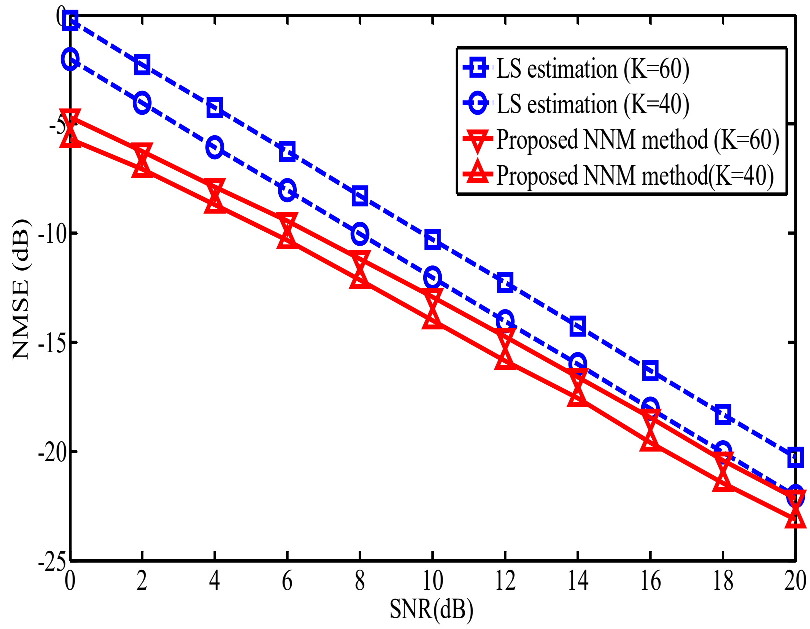

Figure 3.5: Comparison between LS and NN estimations in terms of NMSE versus SNR over system parameters, M=80, P=20, τ =64, and K∊ {40,64}

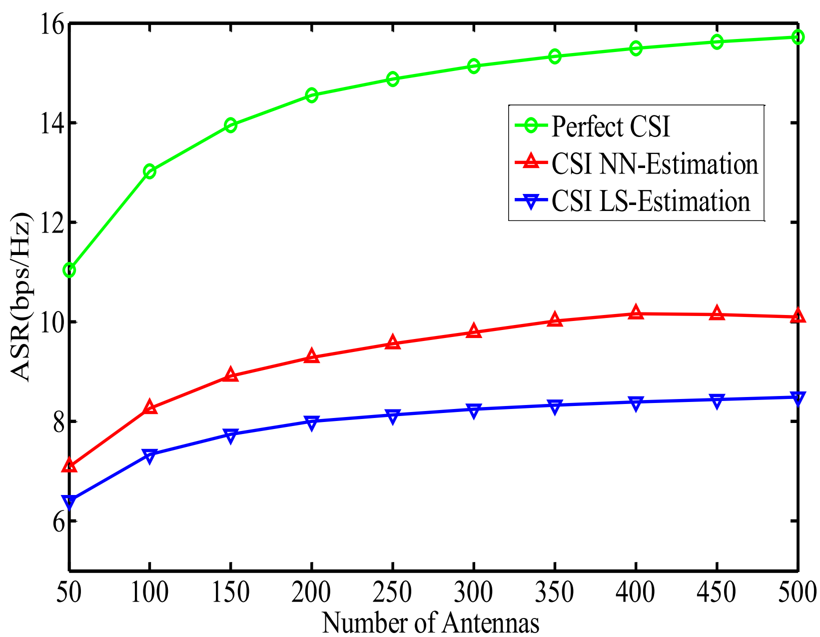

Figure 3.6: Uplink ASR vs. Number of BS antennas, M, over the system parameters, K =40 users, P =20, and SNR= 0 dB

3.7 Conclusion

In this study, a novel channel estimation method namely "nuclear norm (NN) approximation," has been proposed for a single-cell TDD massive MU-MIMO system. The main aim of the proposed scheme is to estimate the large channel matrix entries of single-cell TDD massive MU-MIMO system with a limited number of pilot sequences. Consequently, the channel estimation problem was formulated as a unconstrainted nuclear norm minimization problem and solved via the proposed algorithm. The simulation results show that the performance of the proposed scheme in terms of the NMSE significantly outperforms the traditional LS estimation, which ignores the sparsity feature of the channels.

Chapter 4

Iterative Weighted Nuclear Norm Approximation-Based Channel Estimation for Single-Cell Massive MU-MIMO System

4.1 Introduction

The nuclear norm (NN) channel estimation method was proposed for a single-cell TDD massive MU-MIMO system in the previous chapter [39]. However, the NN estimation has certain limitations because it treats all singular values of the channel matrix equally. In other words, it ignores the prior knowledge on the largest singular values which are representing the most significant singular values of the channel matrix. Moreover, the regularization parameter, γ, used in the NN estimation forces all singular values of the channel matrix equally toward zero which results in degradation of the channel estimation performance. However, this significantly restricts the NN estimation capability in dealing with many practical problems, such as massive MU-MIMO channel estimation problem where the largest singular values have a definite meaning and should be treated differently [54], [56], [57], [59], [64], [81].

In this study, iterative weighted nuclear norm (IWNN) approximation based on compressive sensing technique is proposed for channel estimation in TDD massive MU-MIMO system to improve the previously proposed NN estimation method [40]. To the best of the authors' knowledge, the IWNN channel estimation method has not been used for massive MU-MIMO system to date.

4.2 Iterative Weighted Nuclear Norm (IWNN) Channel Estimation

The main goal of the IWNN estimation approach is to improve the NN channel estimation method by assigning different weights to different singular values of the channel matrix. Moreover, the IWNN aims to estimate the channel matrix by shrinking the largest singular values of the channel matrix less than the smallest ones by using an adaptive regularization parameter, γ, threshold [7], [51], [58], [82]. It should be noted that the smallest singular values of the channel matrix are usually represented by the noise and interference channel power when the locations of the interference users are assumed to be far away from the target base station [83]. More specifically, the IWNN estimation method ignores the smallest singular values and estimates the channel matrix based on the largest singular values of the channel matrix, and thus it is more accurate than the NN estimation method. It is noteworthy that, the IWNN estimation method has been recently utilized in a different context in statistics and signal processing [58], [63], [64]. Therefore, the channel estimation for a single-cell massive MU-MIMO system in Equation (2.20) in the chapter (3) is formulated as an unconstrained weighted nuclear norm regularization problem as follows

$$\widetilde{\boldsymbol{H}} = arg \min_{\boldsymbol{H}} \left\{ \frac{1}{2} \left\| \widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H} \right\|_F^2 + \gamma \|\boldsymbol{H}\|_{\boldsymbol{w},*} \right\} \tag{4.1}$$

where γ is the nonnegative regularization parameter, and $\|\boldsymbol{H}\|_{\boldsymbol{w},*}$ is the weighted nuclear norm of $\boldsymbol{H}$ which is denoted by [13] as

$$\|\boldsymbol{H}\|_{\boldsymbol{w},*} = \sum_{i=1}^{r} |w_i \sigma_i(\boldsymbol{H})| \tag{4.2}$$

where w_i is the non-negative weight element which is assigned to each singular value, $\sigma_i(\boldsymbol{H})$, and $r \leq \min(M, P, K)$ is the rank of the channel matrix, $\boldsymbol{H}$. By substituting the equation (4.2) into (4.1), we can rewrite (4.1) as

$$\widetilde{\boldsymbol{H}} = arg\min_{\boldsymbol{H}}\left\{\frac{1}{2}\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right\|_F^2 + \gamma\sum_{i=1}^{r}|w_i\sigma_i(\boldsymbol{H})|\right\} \tag{4.3}$$

where the first term in (4.3) is the F-norm, while the second term is a ℓ_1-norm penalty function which is very important for the success of the NN estimation.

In (4.3), the regularization parameter γ controls the relative importance between the sparsity of the solution (ℓ_1-norm) term and the fitness to the measurements (F-norm) term. The F-norm property of any matrix (.) can be defined as

$$\|.\|_F^2 = Tr((.)^H(.)) = \sum_{i=1}^{r}\sigma_i{}^2(.) \tag{4.4}$$

By applying this property in (4.3), the Equation (4.3) can be rewritten as:

$$\widetilde{\boldsymbol{H}} = arg\min_{\boldsymbol{H}}\left\{\frac{1}{2}\,Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right)^H\left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}\right)\right\} + \gamma\sum_{i=1}^{r}|w_i\sigma_i(\boldsymbol{H})|\right\} \tag{4.5}$$

which is equivalent to

$$\widetilde{\boldsymbol{H}} = arg\min_{\boldsymbol{H}}\left\{\frac{1}{2}\,Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS}\right)^H\left(\widehat{\boldsymbol{H}}_{LS}\right)\right\} - Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS}\right)^H(\boldsymbol{H})\right\} + \frac{1}{2}\,Tr\{(\boldsymbol{H})^H(\boldsymbol{H})\} + \gamma\sum_{i=1}^{r}|w_i\sigma_i(\boldsymbol{H})|\right\} \tag{4.6}$$

By applying the SVD to $\widehat{\boldsymbol{H}}_{LS}$ and $\boldsymbol{H}$ in (4.6) as $\widehat{\boldsymbol{H}}_{LS} = \widehat{\boldsymbol{U}}_{LS}\,\boldsymbol{\Sigma}_{LS}\,\widehat{\boldsymbol{V}}_{\boldsymbol{LS}}^H$, and $\boldsymbol{H} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^H$, where $\widehat{\boldsymbol{U}}_{LS}$, $\widehat{\boldsymbol{V}}_{LS}$, $\boldsymbol{U}$, and $\boldsymbol{V}$ are unitary matrices, the second term in (4.6) becomes $Tr\left\{\widehat{\boldsymbol{H}}_{\boldsymbol{LS}}^H\,\boldsymbol{H}\right\} = Tr\left\{\widehat{\boldsymbol{V}}_{LS}\,\boldsymbol{\Sigma}^H{}_{LS}\,\widehat{\boldsymbol{U}}_{LS}^H\;\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{H}}\right\}$ if $\widehat{\boldsymbol{U}}_{LS}^{\boldsymbol{H}}\,\boldsymbol{U} = \boldsymbol{I}_M$ and $\widehat{\boldsymbol{V}}_{LS}\,\boldsymbol{V}^{\boldsymbol{H}} = \boldsymbol{I}_M$ which can be rewritten as $\sum_{i=1}^{\mathrm{r}}\sigma_i\left(\widehat{\boldsymbol{H}}_{LS}^{(j)}\right)\sigma_i\left(\boldsymbol{H}_j^{(j)}\right)$. In massive MU-MIMO system, this condition can be satisfied as long as a large number of base station antennas M is assumed [24]. Based on this assumption, the Equation (4.6) can be rewritten as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}) = \underset{\sigma_1, \sigma_2 \ldots\ldots \sigma_r}{arg\,min} \left\{ \frac{1}{2} \left(\sum_{i=1}^{r} {\sigma_i}^2(\widehat{\boldsymbol{H}}_{LS}) - 2 \sum_{i=1}^{r} \sigma_i(\widehat{\boldsymbol{H}}_{LS}) \sigma_i(\boldsymbol{H}) + \sum_{i=1}^{r} {\sigma_i}^2(\boldsymbol{H}) \right) + \gamma \sum_{i=1}^{r} |w_i \sigma_i(\boldsymbol{H})| \right\} \tag{4.7}$$

The optimization problem in (4.7) can be simplified in vector form as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}) = \min_{\boldsymbol{\sigma}(\boldsymbol{H})} \left\{ \frac{1}{2} \left\| \boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}(\boldsymbol{H}) \right\|_2^2 + \gamma\, \boldsymbol{w}^T\, \boldsymbol{\sigma}(\boldsymbol{H}) \right\} \tag{4.8}$$

where the first term in (4.8) is ℓ_2-norm which represents the data-fidelity term (residual error (R)), while the second term is an ℓ_1-norm penalty function. In (4.8), $\boldsymbol{\sigma}(\boldsymbol{H})$ and $\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}})$ are the singular value vectors of the actual and its approximation, respectively. In (4.8), $\boldsymbol{w} = [w_1 \;\; w_2 \ldots \ldots w_r]$ is a non-negative weight vector assigned to each singular value σ_i of the matrix $\boldsymbol{H} \in \mathbb{C}^{M \times K}$.

The weight vector $\boldsymbol{w}$ brings more parameters to the system model, and therefore the appropriate setting of the weights plays a crucial role in the success of the proposed IWNN channel estimation method [64]. In this study, it is proposed to enhance the sparsity of the nonnegative singular value solutions of the IWNN estimation by adaptively tuning each weight element, w, by using the following formula:

$$w_i^{t+1} = \frac{\mu}{\tilde{\sigma}_i^t(\widetilde{\boldsymbol{H}}) + \varepsilon} \qquad i = 1,2 \ldots\ldots\ldots r \tag{4.9}$$

where $\tilde{\sigma}_i^t$ is the i^{th} singular value of the approximate channel matrix $\widetilde{\boldsymbol{H}}$ in the t^{th} iteration, w_i^{t+1} is its corresponding regularization parameter, γ, in the $(t+1)$-th iteration, ε is a small positive number to avoid dividing by zero, and μ is the step size parameter which is used to accelerate time convergence. The proposed iterative algorithm used to solve the WNN optimization problem in (4.8) is summarized in Table 4.1.

4.2.1 Selection of Regularization Parameter and Weight Vector

As mentioned in the previous sections, the regularization parameter γ and weight vector $\boldsymbol{w}$ are important factors for implementing the proposed IWNN estimation method. In other words, the regularization parameter γ is used to control the trade-off error between the data fidelity, $\left\|\boldsymbol{\sigma}\left(\widehat{\boldsymbol{H}}_{LS}\right)-\boldsymbol{\sigma}(\boldsymbol{H})\right\|_2^2$, and the prior information $\boldsymbol{\sigma}(\boldsymbol{H})$ in (4.8). Also, the weight vector $\boldsymbol{w}$ is used to enhance the sparsity of sparse singular value solutions by adaptively tuning weights through the formula given in (2.20). However, the appropriate selection method of the regularization parameter γ and setting the weight vector $\boldsymbol{w}$ in (4.8) are playing a crucial role in the success of the proposed IWNN channel estimation method. Therefore, we use the cross-validation (CV) curve method which is explained in the previous chapter to select the initial value of the regularization parameter, γ, in the proposed algorithm of Table 4.1.

Moreover, the singular values of a matrix are always sorted in non-ascending order, and therefore the large singular values usually correspond to the subspaces of more critical components of the data matrix. Thus, we would better shrink the larger singular values in the WNN estimation method less by assigning smaller weights. Based on this conclusion, the weight vector $\boldsymbol{w}$ in (4.8) is considered to be in a non-descending order since the formula given in (4.9) is used to update the weight vector $\boldsymbol{w}$. It should be noted that a vector of ones is used as the initial weight vector with the expectation that this selection leads to better results, and then it is updated at each iteration, t.

4.2.2 Computational Complexity Analysis of IWNN Estimator

The computational complexity of the IWNN proposed method is correlated with the number of iterations, T, in step 2 in Table 4.1. Thus, it is beneficial to keep this complexity low. In other words, the trade-off between the estimation performance in terms of the residual error (R) and the number of iterations, T, is fundamental to the success of the IWNN proposed estimation method. Moreover, the following formula is used to compute the residual error (R) of the optimization problem in (4.8) as:

$$R = \left\|\boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}(\boldsymbol{H})_{t+1}\right\|_2^2 - \left\|\boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}(\boldsymbol{H})_{t}\right\|_2^2 \tag{4.10}$$

where t is the iteration number.

The LS-based channel estimation has a computational complexity of order $\mathcal{O}(MK)$, while the previously proposed NN estimation method with the total computational complexity of order $\mathcal{O}(M^2KP)$. Compared to the estimation methods above, the IWNN-based channel estimation has the highest complexity of order $\mathcal{O}(M^2KPT)$, where T is the total number of iterations, but better estimation performance.

Table 4.1: IWNN Estimation Algorithm for Single-Cell Massive MU-MIMO System

1: Apply the SVD method of $\widehat{\boldsymbol{H}}_{LS}$ as

$$[\widehat{\boldsymbol{U}}_{LS}\ \boldsymbol{\Sigma}_{LS}\ \widehat{\boldsymbol{V}}_{LS}] = SVD(\widehat{\boldsymbol{H}}_{LS}) \tag{4.11}$$

where a $M \times K$ diagonal matrix $\boldsymbol{\Sigma}_{LS}$ whose diagonal elements are the singular values of $\widehat{\boldsymbol{H}}_{LS}$ corresponding to the eigenvectors of unitary matrices $\widehat{\boldsymbol{U}}_{LS} \in \mathbb{C}^{M \times M}$ and $\widehat{\boldsymbol{V}}_{LS} \in \mathbb{C}^{K \times K}$.

2: Choose the number of iterations T_{step}, and then set t=0, μ=0.5, $\varepsilon = 10^{-5}$ and initial weights $\boldsymbol{w} = [1, 1, \dots . r]$.

3: Choose the initial value of the regularization parameter γ by using Cross Validation (CV) curve method as it is explained in the previous chapter.

4: Solve the optimization problem in (4.8) to obtain the singular values estimation matrix $\widetilde{\boldsymbol{\Sigma}_t}$ as.

$$\widetilde{\boldsymbol{\Sigma}_t} = \begin{pmatrix} diag(\tilde{\sigma}_1(\widetilde{\boldsymbol{H}}), \dots .. \tilde{\sigma}_r(\widetilde{\boldsymbol{H}}), 0, \dots .. \tilde{\sigma}_K(\widetilde{\boldsymbol{H}}) \\ \boldsymbol{0}_{M-K \times K} \end{pmatrix} \tag{4.12}$$

5: t=t+1.

6: Update the weight vector $\widehat{\mathbf{w}}$ by using the formula in (4.9).

7: Repeat steps from 4-6 until convergence to a predefined residual(R) is obtained or when the chosen number of iterations is reached.

8: Finally, the estimated channel matrix is determined as

$$\widetilde{\boldsymbol{H}} = \widehat{\boldsymbol{U}}_{LS}\ \widetilde{\boldsymbol{\Sigma}_t}\ \ \widehat{\boldsymbol{V}}_{LS}^{H} \tag{4.13}$$

4.3 Simulation Results

This section contains the simulation results of the IWNN proposed channel estimation method for TDD massive MU-MIMO system. The same system and channel models presented in the previous chapter for a single-cell TDD massive MU-MIMO system have been considered in this study as well. The initial values of the regularization parameter, γ, of the proposed IWNN estimation method are selected based on the CV curve method which is explained in the previous chapter. Moreover, a vector of ones is used as the initial weight vector of the proposed IWNN estimation method. In this study, two estimation performance criteria namely normalized mean square error (NMSE) and uplink achievable sum-rate (ASR) are used to evaluate the proposed estimation method.

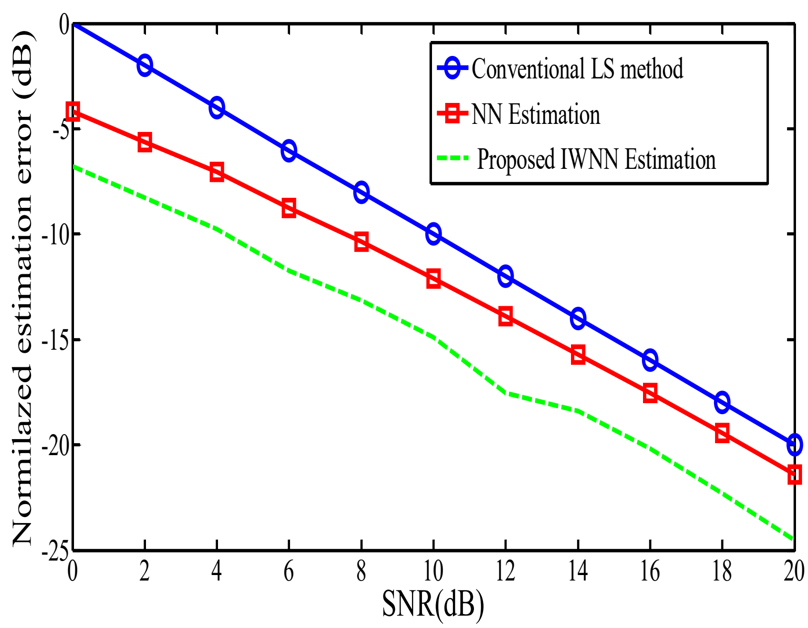

Figure 4.1: Comparison between IWNN, NN, and LS estimations over the system parameters, M= 80, $P = 20, K = 40, \tau = 40$.

In the first experiment, the comparison between the IWNN, NN, and LS estimation methods in terms of the NMSE estimation performance is investigated under different values of the SNR. Figure 4.1 shows that the IWNN estimation has substantial improvement compared to the NN and LS estimations with the cost of the number of iterations, i.e., $l < 10$.

In the second experiment, the NMSE performance of the IWNN proposed method versus the number of BS antennas, M. is studied and compared to the NN and LS estimations. Figure 4.2 shows that, as the number of antennas increases, the IWNN proposed estimation method is improved and outperforms both NN and LS estimations in terms of the NMSE. Moreover, the LS estimation method does not show any improvement as expected from the theoretical analysis.

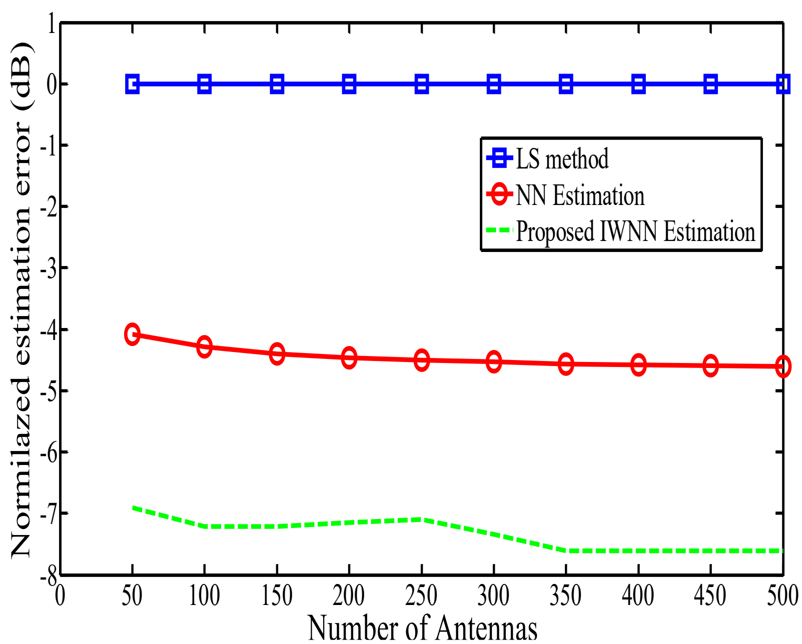

Figure 4.2: Normalized estimation error versus Number of Antennas, M, over system parameters, $P = 20, K = 40, \tau = 40$, and SNR= 0 dB

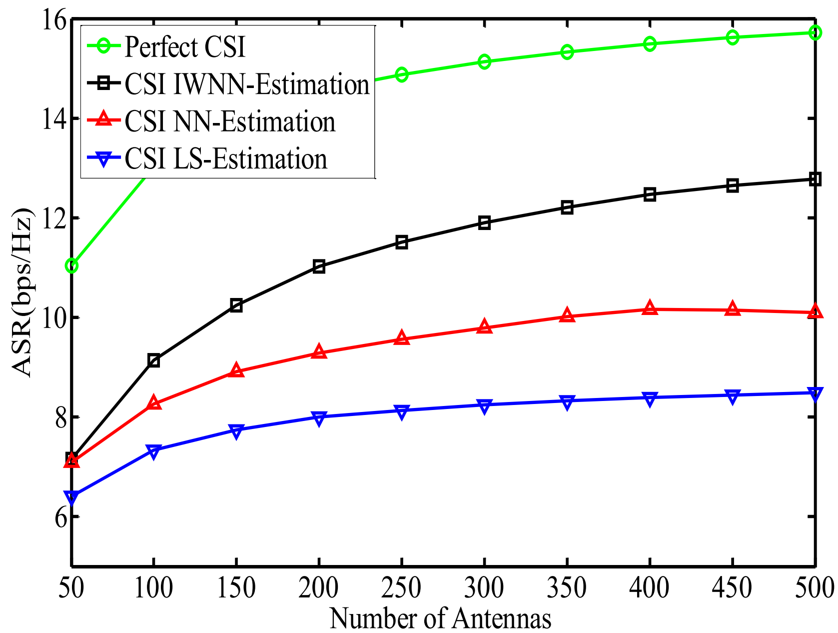

Figure 4.3: Uplink ASR vs. Number of BS antennas, M, for a single-cell system with parameters, K =40 users and P =20, and SNR= 0 dB

In the third experiment, we evaluate the effectiveness of the NN previously proposed estimation method in terms of the ASR performance criterion for TDD single-cell massive MU-MIMO system. To consider the worst-case scenario, the K users are assumed to be distributed on the cell-edge and have the same distance from the BS, i.e., β_k=1. Figure 4.3 shows the uplink ASR performance obtained by different estimation methods under a different number of antennas, M. For reference, the upper bound on uplink ASR is also simulated by designing the MMSE detector matrix with perfect CSI. It can be seen from Figure 4.3 that as the number of antennas, M, increases, the uplink ASR obtained by IWNN estimation method is improved by 4.5 bps/Hz and by 3 bps/Hz compared to the LS and NN estimations, respectively. However, it is about 6 bps/Hz less than the perfect CSI, which is reasonable for uplink massive MU-MIMO system.

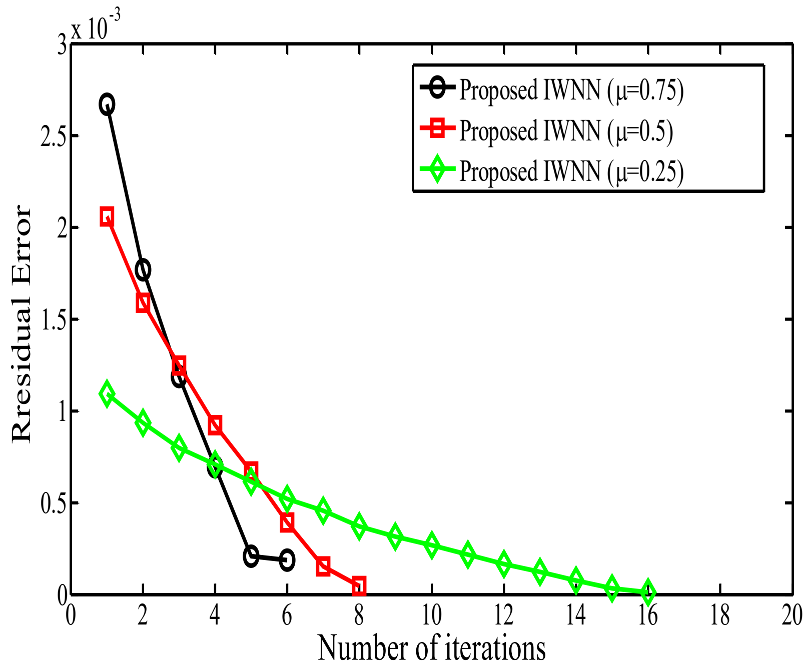

Figure 4.4: The speed of convergence of the IWNN algorithm for single-cell system with parameters M=80, $P = 20, \tau = 40, K = 40$ at $SNR = 0$ dB.

Finally, the computational complexity of the IWNN estimation technique is studied. However, this complexity is related to the number of iterations and therefore, the residual error (R) versus the number of iterations, t, is plotted based on the formula in (4.10). Figure 4.4 shows the residual error versus the number of iterations at SNR= 0dB. It is clear that the residual error depends on the selection of step size μ. Choosing μ too small results in a large number of iterations, but better estimation performance, whereas choosing μ too large results in poor estimation performance. We conclude that the suitable value of the step size μ is 0.5 for the number of iteration, $l < 10$.

4.4 Conclusion

In this study, a novel channel estimation algorithm exploiting the sparsity of the channel matrix is proposed for a single-cell massive MU-MIMO system. The main goal of this estimation approach is to improve the standard nuclear norm (NN) estimation method. Hence, the channel estimation problem was formulated as a weighted nuclear norm (WNN) minimization problem and solved via the proposed iterative algorithm. Furthermore, the initial values of the regularization parameter at each SNR value are selected based on cross-validation curve method. Moreover, the computational complexity of the IWNN estimation technique is also studied in terms of the number of iterations. The simulation results show that the proposed IWNN method with the cost of having a few iterations (i.e., $l < 10$) achieves better estimation performance compared to the standard NN and conventional LS estimation methods in terms of the NMSE and ASR performances.

Chapter 5

Pilot Decontamination in Massive Multiuser MIMO Systems Based on Low-Rank Matrix Approximation

5.1 Introduction

The channel estimation in multi-cell massive MU-MIMO systems is considered one of the significant challenges due to the pilot contamination problem [2], [6], [11], [14]. This problem occurs during the uplink pilot transmission phase when the non-orthogonal pilot sequences are used by other users in the adjacent cells. Moreover, due to the limited coherence time interval, the same frequency band is used for all cells which results in the inter-cell interference at each base station. This interference problem is called pilot contamination, which degrades the channel estimation performance [14].

In literature, several research efforts have been done in the last ten years towards mitigating the pilot contamination in TDD multi-cell massive MIMO systems. In [14], [18], the asynchronous time-shifted pilot protocol is proposed to reduce pilot contamination in TDD massive MU-MIMO system by avoiding the simultaneous transmission of pilot sequences from different users among all cells. However, this method may not provide accurate channel estimation due to the higher downlink transmit power compared to the uplink. Blind and Semi-Blind-based channel estimation approaches are proposed to eliminate the pilot contamination effects in TDD massive MU-MIMO system [24]-[27]. These techniques are based on estimating the channel matrix in the uplink data phase and use it for beamforming at the downlink. In these techniques, the pilot contamination effect is reduced as the length of an uplink received data increases. However, the length of the uplink received data in massive MIMO systems is limited, and therefore it may become difficult to implement in real scenarios. Pilot decontamination based on the collaboration between all base stations has been proposed for TDD massive MU-MIMO system [19]-[21], [28], [73], [75], [84]. However, these approaches can lead to the complete removal of pilot contamination effect under a certain condition which is hard to implement in massive MIMO systems since the collaboration between all base stations is limited in real systems.

In [29], [30], the authors have proposed novel estimation algorithms to reduce the pilot contamination in TDD massive MIMO systems by exploiting the path diversity in both angle and power domains. Under the condition that the channel covariance matrices of desired and interference users are perfectly known at each base stations, the pilot contamination problem is reduced. However, this specific condition may not be achieved in a real scenario since both channel covariance matrices are not available at the base stations. Pilot decontamination approach based on a combination of a pilot sequence hopping scheme and a modified Kalman filter has been studied in [22], [23]. However, the channel estimation method is performed at multiple time slots. Therefore, this channel estimation approach is of a considerable computational complexity since the processing time will be too long.

Unlike previous channel estimation methods, low-rank matrix approximation (LRMA) methods and compressive sensing (CS) techniques have been applied as a new framework for various problems of wireless communication systems such as the sparse channel estimation problem [32]-[35], [39], [40], [52], [54], [60]. For example, compressive sensing has been proposed for the channel estimation problem in TDD multi-cell massive MU-MIMO systems [54]. However, such a method is only feasible for systems with a small number of users and base station antennas.

In our previous work, the NN channel estimation scheme was proposed for a single-cell massive MU-MIMO system in chapter 3. It has been shown that the NN with less number of pilot sequences significantly outperforms the conventional LS estimation. In chapter 4, the channel estimation performance of the NN method has been improved by proposing the IWNN estimation. However, both proposed estimation schemes have shown significant improvements in the channel estimation performance in terms of the NMSE and uplink ASR capacity compared to the conventional LS estimation.

In this work, we propose to extend the applications of the LRMA in multi-cell massive MU-MIMO channel estimation to mitigate pilot contamination problem. Consequently, the nuclear norm (NN) and iterative weighted nuclear norm (IWNN) estimation methods are proposed to deal with the pilot contamination problem in such systems. The main contributions in this work are summarized as follows:

- The NN estimation method is proposed for channel estimation in multi-cell massive MU-MIMO systems to mitigate pilot contamination problem
- Regularization parameter of the NN estimation method is selected based on the cross-validation (CV) curve method [47].
- In the presence of high pilot contamination problem in multi-cell massive MU-MIMO systems, the IWNN estimation method is proposed to improve the NN estimation performance.
- To enhance the sparsity of singular values of channel matrix solutions, the appropriate setting of the weight vector of the iterative algorithm has been taken into consideration.
- The simulation results show the efficiencies of both proposed algorithms compared with the conventional LS method in terms of the NMSE performance.

5.2 System and Channel Models.

5.2.1 System Model

We consider a TDD multi-cell massive MU-MIMO system with L cells where each cell contains one BS which is equipped with a very large number of antennas, M, and simultaneously serves a large number of, K, single-antenna users [6]. In this study, the channel estimation problem is studied when the non-orthogonal pilot sequences are reused by other users in the adjacent cells, giving rise to pilot contamination problem as shown in Figure 5.1.

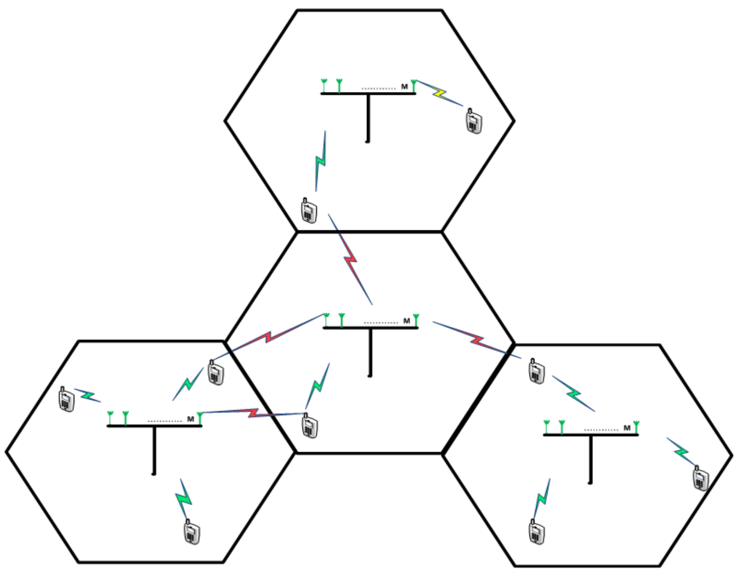

Figure 5.1: Pilot contamination in multi-cell massive MU-MIMO systems

We consider the uplink pilot transmission phase, where all users from all cells simultaneously transmit their pilot signal $\boldsymbol{x}(t)$ to their desired BSs. The orthogonal pilot sequences with a BPSK modulation scheme are assumed to avoid intra-cell interference among K users in a given cell, i.e., $x_k(t) \in \{+1, -1\}$. However, the same orthogonal pilot signals are reused by other users in the adjacent cells which results in pilot contamination problem. To illustrate this idea, we suppose that the BS-j in the cell-j is the target base station unless otherwise specified. The total transmitted signal vector $\boldsymbol{x}(t) \in \mathbb{C}^K$ from all K users to the BS-j is given as

$$\boldsymbol{x}(t) = [\ x_1(t),\ x_2(t) \ldots \ldots .\ x_K(t)]^T \tag{5.1}$$

At each time, t, the received pilot signal $\boldsymbol{y}_j(\mathrm{t}) \in \mathbb{C}^M$ at the BS-j in vector form is denoted by

$$\boldsymbol{y}_j(t) = \sqrt{\rho_{tr}}\boldsymbol{H}_j\, \boldsymbol{x}(t) + \sqrt{\boldsymbol{\rho_{tr}}} \sum_{l \neq j}^{L} \boldsymbol{H}_l\, \boldsymbol{x}(t) + \boldsymbol{n}_j(t) \tag{5.2}$$

where ρ_{tr} is the transmitted symbol power from each user, and $\boldsymbol{n}_j(\mathrm{t}) \in \mathbb{C}^M$ is the complex-valued additive white Gaussian noise (AWGN) vector with zero mean and unit variance σ_n^2, i.e., $\mathcal{CN}(0,\ 1)$. In (5.2), $\boldsymbol{H}_j \in \mathbb{C}^{M \times K_d}$ is the channel matrix between the desired users K_d and the BS-j in the cell-j which is defined as

$$\boldsymbol{H}_j \triangleq \left[\boldsymbol{h}_{j1}\ \boldsymbol{h}_{j2}\ \ldots \ldots \ldots \boldsymbol{h}_{jK_d}\right] \tag{5.3}$$

and $\boldsymbol{H}_l \in \mathbb{C}^{M \times K_c}$ is the channel matrix between the contaminated K_c users in the adjacent cells $l \neq j$ and the BS-j which is defined as

$$\boldsymbol{H}_l \triangleq \left[\boldsymbol{h}_{l1}\ \boldsymbol{h}_{l2}\ \ldots \ldots \ldots \boldsymbol{h}_{lK_c}\right] \tag{5.4}$$

In each coherence time interval, the total pilot sequences received at the BS-j in matrix form can be expressed as

$$\boldsymbol{Y}_j = \sqrt{\rho_{tr}} \sum_{l=j}^{L} \boldsymbol{H}_l\, \boldsymbol{X} + \boldsymbol{N}_j \tag{5.5}$$

where $\boldsymbol{X} \triangleq [\,\boldsymbol{x}(1)\ \ \boldsymbol{x}(2) \ldots .\boldsymbol{x}(\tau)\,] \in \mathbb{C}^{K \times \tau}$ is the total transmit pilot sequences from K users to the BS, and $\boldsymbol{N}_j \in \mathbb{C}^{M \times \tau}$ is the spatially and temporally white additive Gaussian noise (AWGN) matrix with zero-mean and element-wise variance σ_n^2, i.e., $\sim \mathcal{CN}(0,\ \boldsymbol{I}_M)$.

5.2.2 Channel Model

In this study, the realistic finite-dimensional multipath channel model is considered for all channel users (desired and interfering users) which is studied for massive MIMO systems [6], [28], [29], [31], [34], [35], [41], [73]. The channel vector between the k^{th} user and the BS-j in the cell-l is given as

$$\boldsymbol{h}_{lk} = \frac{\beta_{lk}}{\sqrt{P}} \sum_{p=1}^{P} \boldsymbol{a}(\theta_p)\, g_{lkp} \tag{5.6}$$

where g_{lkp} is the fading coefficient between the k^{th} user and the BS-j in the cell-l associated with each path $p \in 1,2 \ldots \ldots P$, and β_{lk} is the path loss coefficient between the k^{th} user and the BS-j in cell-l which is denoted by

$$\beta_{lk} = \sqrt{\frac{\alpha}{(d_{lk})^{\delta}}} \tag{5.7}$$

where d_{lk} is the geographical distance between the k^{th} $u$ser in cell-l and the BS-j, δ is the path-loss exponent, and α is a constant dependent on the prescribed average signal to noise ratio (SNR) at the cell edge. In (5.6), $\boldsymbol{a}(\theta_p) \in \mathbb{C}^M$ is the steering vector originating from each k^{th} user to the BS j associated with each path, p, which is given as

$$\boldsymbol{a}(\theta_p) = \left[1, e^{-j2\pi\frac{D}{\lambda}cos(\theta_p)} \ldots\ldots\ldots\ldots\ldots, e^{-j2\pi(M-1)\frac{D}{\lambda}cos(\theta_p)}\right]^T \tag{5.8}$$

where λ is the signal wavelength, D is the antenna spacing which is assumed to be fixed, and $\theta_p^{(j)} \in [-\pi/2, \pi/2]$ is the random angle of arrival (AoA) corresponding to each path p.

The total channel model including the steering matrix, flat fading matrix and geometric attenuation matrix for all users can be collectively written in a matrix form as

$$\boldsymbol{H}_l = \boldsymbol{A}\,\boldsymbol{G}_l\,\boldsymbol{D}_l^{1/2} \tag{5.9}$$

where $\boldsymbol{A} \triangleq [\,\boldsymbol{a}(\theta_1)\ \boldsymbol{a}(\theta_2)\ \ldots\ldots\boldsymbol{a}(\theta_P)\,] \in \mathbb{C}^{M\times P}$ is a full-rank steering matrix, and $\boldsymbol{G}_l \triangleq [\,\boldsymbol{g}_{l1}\ \boldsymbol{g}_{l2}\ \ldots\ldots\ldots\boldsymbol{g}_{lk}\,] \in \mathbb{C}^{P\times K}$ is a Rayleigh flat fading channel matrix between the K users and the BS-j. The entries in each vector $\boldsymbol{g}_{lk} \triangleq [g_{lk1}\ \ g_{lk2}\ \ldots\ldots g_{lkp}]^T$ are assumed to be independently identically distributed $(i.i.d)$ symmetrical complex Gaussian random variable with zero mean and unit variance, i.e., $\mathrm{g}_{lkp} \sim \mathcal{CN}(0,\ 1)$. In (5.9), $\boldsymbol{D}_l \in \mathbb{C}^{K\times K}$ is a diagonal matrix whose diagonal elements are $[\boldsymbol{D}_l]_{kk} = \beta_{lk}$.

5.3 Pilot Contamination Effect on LS Channel Estimation.

In this section, we study the impact of the pilot contamination problem in multi-cell massive MU-MIMO systems when the LS estimation method is used. In general, the LS method relies on correlating the total received signal $\boldsymbol{Y}^{(j)}$ in (5.5) with known orthogonal pilot sequences $\boldsymbol{X}$ [16], [17]. Hence, the LS estimation matrix at the BS-j is given by

$$\widehat{\boldsymbol{H}}_{LS} = \frac{1}{\sqrt{\rho_{tr}}}\ \boldsymbol{Y}_j\,\boldsymbol{X}^H(\boldsymbol{X}\boldsymbol{X}^H)^{-1} \tag{5.10}$$

where $\widehat{\boldsymbol{H}}_{LS}$ is the LS estimation matrix, and $\boldsymbol{X}$ is the pilot matrix which is assumed to be orthogonal and known at each base station, i.e., $\boldsymbol{X}\boldsymbol{X}^H = \tau\,\boldsymbol{I}_K$. By substituting (5.5) and $\boldsymbol{X}\boldsymbol{X}^H = \tau\,\boldsymbol{I}_K$ into (5.10), we can rewrite (5.10) as

$$\widehat{\boldsymbol{H}}_{LS} = \boldsymbol{H}_j + \sum_{l\neq \mathrm{j}}^{L}\boldsymbol{H}_l + \frac{1}{\tau\sqrt{\rho_{tr}}}\boldsymbol{N}_j\,\mathbf{X}^{\mathrm{H}} \tag{5.11}$$

According to the channel model in (5.9), the normalized mean square error (NMSE) of (5.11) is given as

$$NMSE = \frac{1}{\sum_{k=1}^{K_d} \beta_{jk}} \left(\sum_{l \neq j}^{L} \sum_{k=1}^{K_c} \beta_{lk} + \frac{K_d \, \sigma_n^2}{\tau \ \rho_{tr}} \right) \tag{5.12}$$

As it appears in (5.12), the NMSE performance of the LS estimation is limited by pilot and noise contamination problems. Pilot contamination in the first term of (5.12) is caused by the interference from other users in the adjacent cells, while the noise contamination in the second term in (5.12) occurs when the transmit power from each user inside each cell is small [24]. However, noise contamination can be reduced by increasing the length of the pilot sequence, while pilot contamination cannot [16], [17], [69]. Therefore, in the next two sections, we propose to apply the applications of the LRMA methods to deal with the pilot contamination effect in multi-cell massive MU-MIMO system.

5.4 NN Channel Estimation Method

The NN estimation method is an optimization problem which aims to estimate the non-zero singular values of the low-rank matrix. In this section, the NN channel estimation method is proposed to mitigate pilot contamination problem in multi-cell massive MU-MIMO systems. It should be noted that the NN estimation method was utilized in a different context in statistics and signal processing [32], [35], [51], [53], [59], [61], [67], [81], [83]. However, we have adopted that in our work. As such, the channel estimation problem in (5.11) is formulated as the nuclear norm regularization problem as,

$$\widetilde{\boldsymbol{H}}_j = arg \min_{\boldsymbol{H}_j} \left\{ \frac{1}{2} \left\| \widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j \right\|_F^2 + \gamma \left\| \boldsymbol{H}_j \right\|_* \right\} \tag{5.13}$$

where $\widetilde{\boldsymbol{H}}_j$ is the approximation of the desired channel matrix $\boldsymbol{H}_j$, and γ is a regularization parameter of this optimization problem. In (5.13), $\left\|\boldsymbol{H}_j\right\|_*$ denotes the nuclear norm of the channel matrix $\boldsymbol{H}_j$ which is defined as

$$\left\|\boldsymbol{H}_j\right\|_* = \sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right| \tag{5.14}$$

where $\sigma_i(\boldsymbol{H}_j)$ denotes the i^{th} singular value of the desired channel matrix $\boldsymbol{H}_j$, and r is the rank of $\boldsymbol{H}_j$. The number of multipath, P, originated from each user to the base station is assumed to be small compared to the number of base station antennas, M, and desired users, K_d, and therefore the rank of the desired channel matrix $\boldsymbol{H}_j$ is $r \leq \min(M,\ K_d, P)$ which will be $r = P$. Hence, this channel can be approximated by using LRMA methods. By substituting (5.14) into (5.13), we can rewrite (5.13) as

$$\widetilde{\boldsymbol{H}}_j = arg\min_{\boldsymbol{H}_j}\left\{\frac{1}{2}\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j\right\|_F^2 + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.15}$$

The property of the Frobenius norm of any matrix (.) can be defined as

$$\|.\|_F^2 = Tr((.)^H(.)) = \sum_{i=1}^{r}\sigma_i{}^2(.) \tag{5.16}$$

By applying this property into (5.15), we can rewrite (5.15) as

$$\widetilde{\boldsymbol{H}}_j = arg\min_{\boldsymbol{H}_j}\left\{\frac{1}{2}\,Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j\right)^H\left(\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j\ \right)\right\} + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.17}$$

Moreover, we extend (5.17) to:

$$\widetilde{\boldsymbol{H}}_j = arg\min_{\boldsymbol{H}_j}\left\{\frac{1}{2}\left(Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS}\right)^H\left(\widehat{\boldsymbol{H}}_{LS}\right)\right\} - 2Tr\left\{\left(\widehat{\boldsymbol{H}}_{LS}\right)^H\left(\boldsymbol{H}_j\right)\right\} + Tr\left\{\left(\boldsymbol{H}_j\right)^H\left(\boldsymbol{H}_j\right)\right\}\right)\right.$$
$$\left. + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.18}$$

In (5.18), we apply the SVD into $\widehat{\boldsymbol{H}}_{LS}$ and $\boldsymbol{H}_j$ as $\widehat{\boldsymbol{H}}_{LS} = \widehat{\boldsymbol{U}}_{LS}\,\boldsymbol{\Sigma}_{LS}\,\widehat{\boldsymbol{V}}_{LS}^H$, and $\boldsymbol{H}_j = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^H$, where $\widehat{\boldsymbol{U}}_{LS}$, $\widehat{\boldsymbol{V}}_{LS}$, $\boldsymbol{U}$, and $\boldsymbol{V}$ are unitary matrices, i.e., $\widehat{\boldsymbol{U}}_{LS}^H\,\boldsymbol{U} = \boldsymbol{I}$ and $\widehat{\boldsymbol{V}}_{LS}\,\boldsymbol{V}^H = \boldsymbol{I}$. Then, the second term in (5.18) can be written as $Tr\left\{\widehat{\boldsymbol{H}}_{LS}{}^H\boldsymbol{H}_j\right\} = Tr\left\{\widehat{\boldsymbol{V}}_{LS}\,\boldsymbol{\Sigma}^H{}_{LS}\,\widehat{\boldsymbol{U}}_{LS}^H\;\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^H\right\}$ which is equivalent to: $Tr\left\{\widehat{\boldsymbol{H}}_{LS}\,\boldsymbol{H}_j\right\} = \sum_{i=1}^{r}\sigma_i\left(\widehat{\boldsymbol{H}}_{LS}\right)\sigma_i\left(\boldsymbol{H}_j\right)$. Thus, we rewrite (5.18) as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = arg\min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)}\left\{\frac{1}{2}\left(\sum_{i=1}^{r}\sigma_i{}^2(\widehat{\boldsymbol{H}}_{LS}) - 2\sum_{i=1}^{r}\sigma_i(\widehat{\boldsymbol{H}}_{LS})\sigma_i(\boldsymbol{H}_j) + \sum_{i=1}^{r}\sigma_i{}^2(\boldsymbol{H}_j)\right)\right.$$
$$\left. + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.19}$$

Equation (5.19) can be written as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = \min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)}\left\{\frac{1}{2}\sum_{i=1}^{r}\left(\sigma_i{}^2(\widehat{\boldsymbol{H}}_{LS}) - 2\,\sigma_i\left(\widehat{\boldsymbol{H}}_{LS}\right)\sigma_i\left(\boldsymbol{H}_j\right) + \sigma_i{}^2\left(\boldsymbol{H}_j\right)\right) + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.20}$$

Moreover, we simplify (5.20) as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = \min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)}\left\{\frac{1}{2}\sum_{i=1}^{r}\left(\sigma_i(\widehat{\boldsymbol{H}}_{LS}) - \sigma_i(\boldsymbol{H}_j)\right)^2 + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.21}$$

which is equivalent to

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = \min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)}\left\{\frac{1}{2}\left\|\boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}\left(\boldsymbol{H}_j\right)\right\|_2^2 + \gamma\sum_{i=1}^{r}\left|\sigma_i(\boldsymbol{H}_j)\right|\right\} \tag{5.22}$$

where $\boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS})$ and $\boldsymbol{\sigma}(\boldsymbol{H}_j)$ are the singular value vectors of matrices $\widehat{\boldsymbol{H}}_{LS}$ and $\boldsymbol{H}_j$, respectively. In (5.22), $\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j)$ is the approximate singular value vector of the matrix $\widetilde{\boldsymbol{H}}_j$, and γ is the regularization parameter. The proposed NN estimation method for multi-cell massive MU-MIMO channel estimation is summarized in Table 5.1.

Table 5.1: NN Estimation Algorithm for Multi-Cell Massive MU-MIMO System

1: Apply the SVD method of $\widehat{\boldsymbol{H}}_{LS}$ as

$$SVD(\widehat{\boldsymbol{H}}_{LS}) = [\widehat{\boldsymbol{U}}_{LS}\, \boldsymbol{\Sigma}_{LS}\, \widehat{\boldsymbol{V}}_{LS}] \tag{5.23}$$

where a $M \times K$ diagonal matrix $\boldsymbol{\Sigma}_{LS}$ whose diagonal elements are the singular values of $\widehat{\boldsymbol{H}}_{LS}$ corresponding to the eigenvectors of unitary matrices $\widehat{\boldsymbol{U}}_{LS} \in \mathbb{C}^{M \times M}$ and $\widehat{\boldsymbol{V}}_{LS} \in \mathbb{C}^{K \times K}$.

2: Choose the regularization parameter γ by using Cross-Validation (CV) curve method which is explained in the previous chapter.

3: Solve the optimization problem in (5.22) to obtain the singular values estimation matrix $\widetilde{\boldsymbol{\Sigma}}$, which is defined as

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} diag(\tilde{\sigma}_1(\widetilde{\boldsymbol{H}}_j), \ldots.. \tilde{\sigma}_r(\widetilde{\boldsymbol{H}}_j), 0, \ldots.. \tilde{\sigma}_K(\widetilde{\boldsymbol{H}}_j)) \\ \boldsymbol{0}_{M-K \times K} \end{pmatrix} \tag{5.24}$$

4: Finally, the estimated channel matrix is determined as

$$\widetilde{\boldsymbol{H}}_j = \widehat{\boldsymbol{U}}_{LS}\ \widetilde{\boldsymbol{\Sigma}}\ \widehat{\boldsymbol{V}}_{LS}^H \tag{5.25}$$

5.5 IWNN Channel Estimation Method

The NN estimation method has a fundamental limitation because it treats all singular values of the channel matrix equally with the same threshold. Therefore, it ignores the prior knowledge of the larger singular values of $\boldsymbol{H}_j$ which are more important than the smaller ones since they represent the major energy components of the desired channel matrix**.** On the other hand, the IWNN estimation method treats all singular values of $\boldsymbol{H}_j$ with a different threshold. Moreover, the IWNN shrinks less the largest singular values (desired channels), while shrinking more the smallest ones (interference channels) [40], [56]-[58], [62]. Based on this idea, the IWNN-based channel estimation for multi-cell massive MU-MIMO systems is proposed in this chapter. The IWNN method aims to improve the NN estimation performance by reducing the pilot contamination effect. It should be noted that the IWNN estimation method was utilized in our previous work [28]. However, we have adopted that in this work. Hence, the channel estimation problem in (5.11) can be reformulated as the weighted nuclear norm regularization problem as follows:

$$\widetilde{\boldsymbol{H}}_j = arg\min_{\boldsymbol{H}_j}\left\{\frac{1}{2}\left\|\widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j\right\|_F^2 + \gamma\left\|\boldsymbol{H}_j\right\|_{\boldsymbol{w},*}\right\} \tag{5.26}$$

where $\left\|\boldsymbol{H}_j\right\|_{\mathbf{w},*}$ denotes the weighted nuclear norm of $\boldsymbol{H}_j$ which is defined as

$$\left\|\boldsymbol{H}_j\right\|_{\boldsymbol{w},*} = \sum_{i=1}^{r}\left|w_i\sigma_i(\boldsymbol{H}_j)\right| \tag{5.27}$$

where w_i is the non-negative weight element which is assigned to each singular value σ_i of $\boldsymbol{H}_j$. The weight vector itself brings more parameters in the system model, and therefore the appropriate setting of the weights plays a crucial role in the success of the proposed IWNN method for channel estimation [64]. We propose to enhance the sparsity of singular value solutions by adaptively tuning weights through the following formula which is given as:

$$w_i^{t+1} = \frac{\mu}{\tilde{\sigma}_i^t(\widetilde{\boldsymbol{H}}_j) + \varepsilon} \quad i = 1,2 \dots \dots \dots . r \tag{5.28}$$

where $\tilde{\sigma}_i^t$ is the i^{th} singular value of the approximate channel matrix $\widetilde{\boldsymbol{H}}_j^j$ in t^{th} iteration and ε is a small positive number to avoid dividing by zero. In (5.28), μ is the step size which is utilized to speed up the time convergence of the proposed searching algorithm at each iteration. By substituting (5.27) into (5.26), we can rewrite (5.26) as:

$$\widetilde{\boldsymbol{H}}_j = arg \min_{\boldsymbol{H}_j} \left\{ \frac{1}{2} \left\| \widehat{\boldsymbol{H}}_{LS} - \boldsymbol{H}_j \right\|_F^2 + \gamma \sum_{i=1}^{r} \left| w_i \sigma_i(\boldsymbol{H}_j) \right| \right\} \tag{5.29}$$

To simplify (5.29), the same steps in section 5.4 are followed by applying the Frobenius norm property given in (5.16) into (5.29). Then, we obtain the following WNN optimization problem in a vector form as

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = \min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)} \left\{ \frac{1}{2} \sum_{i=1}^{r} \left(\sigma_i(\widehat{\boldsymbol{H}}_{LS}) - \sigma_i(\boldsymbol{H}_j) \right)^2 + \gamma \sum_{i=1}^{r} \left| w_i \sigma_i(\boldsymbol{H}_j) \right| \right\} \tag{5.30}$$

which is equivalent to:

$$\widetilde{\boldsymbol{\sigma}}(\widetilde{\boldsymbol{H}}_j) = \min_{\boldsymbol{\sigma}(\boldsymbol{H}_j)} \left\{ \frac{1}{2} \left\| \boldsymbol{\sigma}(\widehat{\boldsymbol{H}}_{LS}) - \boldsymbol{\sigma}(\boldsymbol{H}_j) \right\|_2^2 + \gamma \, \boldsymbol{w} \, \boldsymbol{\sigma}(\boldsymbol{H}_j) \right\} \tag{5.31}$$

where $\boldsymbol{w} = [w_1 \;\; w_2 \dots . \dots w_r]^T$ is the non-negative weight vector with elements of $w_i \geq 0, \; i = 1 \dots . . r,$ where each weight element is assigned to each $\sigma_i(\mathbf{H}_\mathrm{j})$. The proposed IWNN estimation method for multi-cell massive MU-MIMO channel estimation is summarized in Table 5.2.

Table 5.2: IWNN Estimation Algorithm for Multi-Cell Massive MU-MIMO System

1: Apply the SVD method of $\widehat{\boldsymbol{H}}_{LS}$ as

$$SVD\left(\widehat{\boldsymbol{H}}_{LS}\right) = [\widehat{\boldsymbol{U}}_{LS}\, \boldsymbol{\Sigma}_{LS}\, \widehat{\boldsymbol{V}}_{LS}] \tag{5.32}$$

where a $M \times K$ diagonal matrix $\boldsymbol{\Sigma}_{LS}$ whose diagonal elements are the singular values of $\widehat{\boldsymbol{H}}_{LS}$ corresponding to the eigenvectors of unitary matrices $\widehat{\boldsymbol{U}}_{LS} \in \mathbb{C}^{M \times M}$ and $\widehat{\boldsymbol{V}}_{LS} \in \mathbb{C}^{K \times K}$.

2: Choose the number of iterations T_{step}, and then set t=0, μ=0.5, $\varepsilon = 10^{-5}$ and initial weights $\boldsymbol{w} = [1, 1, \dots . r]$.

3: Choose the initial value of the regularization parameter γ by using Cross-Validation (CV) curve method as it is explained in the previous chapter.

4: Solve the optimization problem in (5.31) to obtain the singular values estimation matrix $\widetilde{\boldsymbol{\Sigma}_t}$ as

$$\widetilde{\boldsymbol{\Sigma}_t} = \begin{pmatrix} diag(\tilde{\sigma}_1(\widetilde{\boldsymbol{H}}_j), \dots . . \tilde{\sigma}_r(\widetilde{\boldsymbol{H}}_j), 0, \dots . . \tilde{\sigma}_K(\widetilde{\boldsymbol{H}}_j) \\ \boldsymbol{0}_{M-K \times K} \end{pmatrix} \tag{5.33}$$

5: t=t+1.

6: Update each weight element w_i for the weight vector, $\boldsymbol{w}$, in the optimization problem (5.31) by using (5.28).

7: Repeat steps from 4-6 until convergence to a predefined residual(R) is obtained or when the chosen number of iterations is reached.

8: Finally, the estimated channel matrix is determined as

$$\widetilde{\boldsymbol{H}}_j = \widehat{\boldsymbol{U}}_{LS}\ \widetilde{\boldsymbol{\Sigma}_t}\ \widehat{\boldsymbol{V}}_{LS}^H \tag{5.34}$$

5.6 Complexity Analysis of different Estimators

The computational complexity of the LS estimation method comes from multiplying $\boldsymbol{X^H}$ by $\mathbf{Y}_{\mathrm{j}}$ in (8), which has a complexity order $\mathcal{O}(MK)$. The main complexity of the NN estimation method comes from step 1 and step 3 in Table 5.1. In step 1, the SVD of matrix $\widehat{\boldsymbol{H}}_{LS}$ has a complexity order $\mathcal{O}(MK^2)$, and the optimization problem in step 3 has r real variables and one linear constraint which has a complexity order of $\mathcal{O}(r^2)$. Therefore, the NN method has higher computational complexity than the LS method but better estimation performance. On the other hand, the main complexity of the IWNN estimation method in Table 5.2 comes from increasing the number of iterations in step 3, and therefore it has a complexity order of $\mathcal{O}(Tr^2)$, where T is the total number of iterations. Thus, the IWNN has the highest computational complexity compared to the NN and LS estimation methods, but better estimation performance.

One way to reduce the complexity order of the IWNN method is to decrease the number of iterations in by choosing the appropriate value of the step size parameter μ in (21). Note that, selecting the step size parameter μ too large can result in poor estimation performance, where the range of the step size parameter μ is $0 < \mu \leq 1$.

5.7 Estimation Performances

In this section, we study the choice of the system parameters for the proposed channel estimator which are robust to variations in the SNR and number of base station antennas. Thus, the following two estimation performance criteria are used to evaluate the proposed estimation method for multi-cell massive MU-MIMO systems in the presence of the pilot contamination problem.

5.7.1 Normalized Mean Square Error (NMSE)

The following formula of the NMSE performance metric is used to evaluate the proposed channel estimation which is given as

$$NMSE\ (dB) = 10log_{10}\left(\frac{E\left\|\widetilde{\boldsymbol{H}}_j - \boldsymbol{H}_j\right\|_F^2}{E\left\|\boldsymbol{H}_j\right\|_F^2}\right) \tag{5.35}$$

where $\boldsymbol{H}_j$and $\widetilde{\boldsymbol{H}}_j$ are the desired channel matrix at the BS-j and its approximation, respectively.

5.7.2 Uplink Achievable Sum-Rate (ASR)

The second performance metric used to evaluate the proposed channel estimation per cell in multi-cell TDD massive MU-MIMO system is the uplink ASR capacity. To show how this formula is derived, we consider the uplink data transmission phase where all users from all cells simultaneously transmit their data symbols, $s_1, s_2 \ldots, s_K$ to their base stations. However, the same system parameters used in the channel estimation phase are assumed to be used for the uplink data transmission phase as well. To illustrate this idea, we suppose that BS-j is the target BS, K_d is the desired users in the cell-j, and K_c is the interfering users from adjacent cells, i.e., $l \neq$ j. The received data vector $\varphi_{\mathbf{j}}(\mathrm{n}) \in \mathbb{C}^{M \times 1}$ at the target, BS-j is given by

$$\boldsymbol{\varphi}_j(n) = \sqrt{\rho_d} \sum_{k=1}^{K_d} \boldsymbol{h}_{jk} s_{jk}(n) + \sqrt{\rho_d} \sum_{l=1, l\neq j}^{L} \sum_{k=1}^{K_c} \boldsymbol{h}_{lk} s_{lk}(n) \ + \mathbf{z}_j(n) \tag{5.36}$$

where $\sqrt{\rho_\mathrm{d}}\, s_{jk}$ is the transmitted data symbol by the k^{th} user in the cell-j to the BS-j, and ρ_d is the average power used by the k^{th} user. In (5.36), $\mathbf{z}_\mathrm{j}(\mathrm{n}) \in \mathbb{C}^{M \times 1}$ is the AWGN vector with zero mean and element-wise variance, σ_n^2, i.e., $\mathcal{CN}(0,\ \sigma_n^2)$, and $\boldsymbol{h}_{jk} \in \mathbb{C}^{M \times 1}$ is the desired channel vector between the k^{th} user in cell-j and the target BS-j, while $\boldsymbol{h}_{lk} \in \mathbb{C}^{M \times 1}$, $\forall\ l \neq$ j is the interference channel vectors between the k^{th} user in cell-l and the BS-j. It should be noted that the channel model in (5.9) is now reused for the uplink data phase as well since we assume all user's locations in all cells are fixed in each coherence time interval.

In general, the uplink ASR performance of K users per cell is given by the Shannon capacity formula [12] as

$$ASR \leq \sum_{k=1}^{K} log_2(1 + SINR_k) \tag{5.37}$$

where $\mathrm{SINR_k}$ is the received signal-to-interference-noise ratio of each k^{th} user at the linear detector processing output. To compute the $\mathrm{SINR_k}$ for each k^{th} user in (5.37), a linear minimum mean square error (MMSE) detector scheme is assumed to be used at each base station where the k^{th} row vector $\boldsymbol{v}_{jk}$ of the MMSE detector is given by [70] as

$$\boldsymbol{v}_{jk} = \left((\widehat{\boldsymbol{h}}_{jk})^H \, \widehat{\boldsymbol{h}}_{jk} + \sigma_n^2 \right)^{-\mathbf{1}} (\widehat{\boldsymbol{h}}_{jk})^H \tag{5.38}$$

where $\widehat{\boldsymbol{h}}_{jk} \in \mathbb{C}^{M\times 1}$ is the estimated channel vector between the k^{th} user in cell-j and the BS-j, and σ_n^2 is the noise variance. By multiplying each k^{th} row vector $\boldsymbol{v}_{jk}$ in (5.38) by the received signal vector $\varphi_{\mathbf{j}}(\mathrm{n})$ in (5.36), we can detect each data user as

$$\hat{s}_{jk}(n) = \boldsymbol{v}_{jk} \, \boldsymbol{\varphi}_j(n) \tag{5.39}$$

By substituting (5.36) and (5.38) into (5.39), we rewrite (5.39) as

$$\begin{aligned} \hat{s}_{jk}(n) = {} & \sqrt{\rho_d}\, \boldsymbol{v}_{jk} \, \boldsymbol{h}_{jk} s_{jk}(n) + \sum\nolimits_{i=1, i\neq k}^{K} \sqrt{\rho_d}\, \boldsymbol{v}_{jk} \, \boldsymbol{h}_{ji} s_{ji}(n) \\ & + \sqrt{\rho_d} \sum_{l=1, l\neq j}^{L} \sum_{k=1}^{K} \boldsymbol{v}_{jk} \, \boldsymbol{h}_{lk} s_{lk}(n) + \boldsymbol{v}_{jk} \mathbf{z}_j(n) \end{aligned} \tag{5.40}$$

In (3.28), the first term represents the received data symbol of the k^{th} user in the cell-j at the BS-j, and the second term represents the multiuser interference (MUI) in the cell-j. The third and fourth terms represent inter-cell interference (ICI) from the adjacent cells, and the noise term, respectively. Now, the $\mathrm{SINR_k}$ for each k^{th} user in (5.37) can be calculated as

$$SINR_k = \frac{\rho_d \left|\boldsymbol{v}_{jk}\, \boldsymbol{h}_{jk}\right|^{\mathbf{2}}}{\sum_{i=1, i\neq k}^{K} \rho_d \left|\boldsymbol{v}_{jk}\, \boldsymbol{h}_{jk}\right|^2 + \rho_d \sum_{l=1, l\neq j}^{L} \sum_{k=1}^{K} \left|\, \boldsymbol{v}_{jk}\, \boldsymbol{h}_{lk}\right|^2 + \left|\boldsymbol{v}_{jk}\, \boldsymbol{z}_j\right|^{\mathbf{2}}} \tag{5.41}$$

By substituting (3.29) into (3.24), we compute the total uplink ASR capacity of K users as

$$ASR \leq \sum_{k=1}^{K} log_2\left(1 + \frac{Channel\ power\ gain}{MUI + ICI + Noise}\right) \tag{5.42}$$

where the channel power gain=$\rho_d \left|\boldsymbol{v}_{jk}\boldsymbol{h}_{jk}\right|^2$, MUI=$\sum_{i=1, i\neq k}^{K} \rho_d \left|\boldsymbol{v}_{jk}\, \boldsymbol{h}_{ji}\right|^2$, ICI= $\rho_\mathrm{d} \sum_{l=1, l\neq \mathrm{j}}^{L} \sum_{k=1}^{K} \left|\, \boldsymbol{v}_{jk}\, \boldsymbol{h}_{lk}\right|^2$, and the noise term = $\left|\boldsymbol{v}_{jk}\, \boldsymbol{z}_j\right|^2$.

5.8 Simulation Results

This section contains the simulation results of the IWNN proposed channel estimation method for TDD massive MU-MIMO system with $L = 3$ cells. Each cell has one BS with $M = 80$ antennas and simultaneously serves $K_d = 40$ desired users. We address the channel estimation problem where the strongest received signals at the target BS-j are significant, and the rest are small and can be neglected. In all experiments presented in this section, we assume that the desired channel is estimated for more than 1000 exact channel realizations, and the steering vector has the following parameters, $\theta(p\,) = -\pi/2 + (p-1)\pi/2P$, and D/λ =0.5 where p=1, 2……20. Also, the length of each pilot sequence is assumed to be τ =40 symbols, and the total number of arbitrary multipath is considered to be $P = 20$. With the assumptions above, the desired channel matrix $\boldsymbol{H}_j^{(j)}$

has asymptotically low-rank as long as P is small relative to M and P, i.e., $r \leq \min(M, K_d,\ P)$= 20. Two different number of contaminated users are assumed, i.e. $K_c \in \{4,10\}$ in order to study the effect of increasing the number of interfering users at each adjacent cell-edge. We set the path loss coefficients β_{lk}= 0.8 for all K_c users in all adjacent cells $l \neq$ j, and β_{jk}=1 for all K_d users in the cell-j.

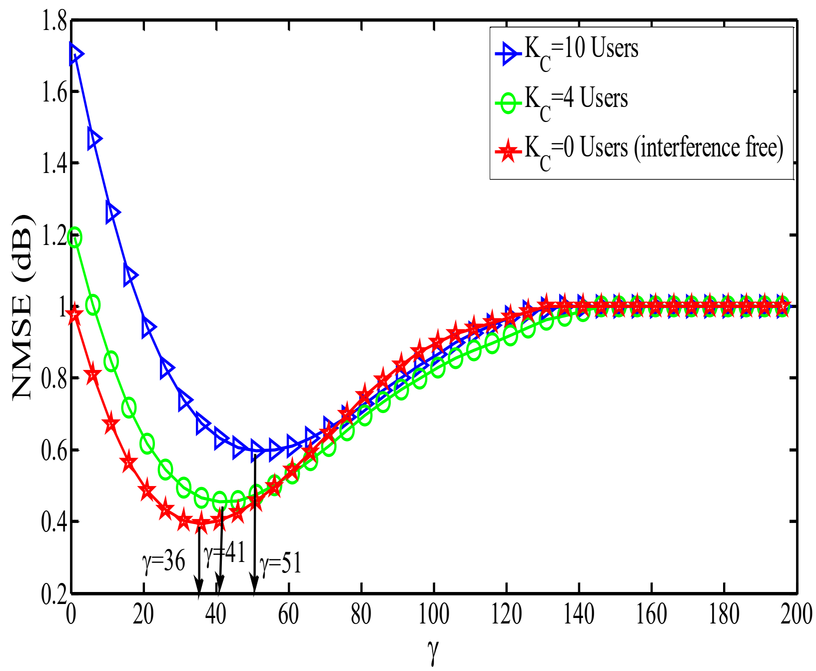

Figure 5.2: NMSE vs. Regularization parameter γ for system parameters, M =80, P=20, τ=40, $Kd = 40$ with β_{jk}=1, SNR=0 dB, and $Kc \in \{0, 4, 10\}$ with β_{lk}= 0.8 $\forall l \neq j$

The regularization parameter, γ, in (5.22) and (5.31) is selected based on the cross-validation (CV) curve method [47]. The CV criterion is based on the selection of the optimal value of the regularization parameter that minimizes the NMSE for each particular value of the signal-to-noise ratio (SNR). By using the following formula, we plot the NMSE versus the regularization parameter γ at each specific value of SNR.

$$\gamma = \underset{\gamma \in \{\gamma_1, \gamma_2, \ldots \gamma_{max}\}}{arg\,min} NMSE(\gamma) \tag{5.43}$$

Then, the optimal value of the regularization parameter is selected at the minimum value of the NMSE. Figure 5.2 shows an illustrative example, how the optimal regularization parameter γ is selected at SNR= 0 dB for three different contaminated users scenarios, i.e., $K_c \in \{0, 4, 10\}$. Then, the experiment is repeated for each value of the SNR with the same scenario.

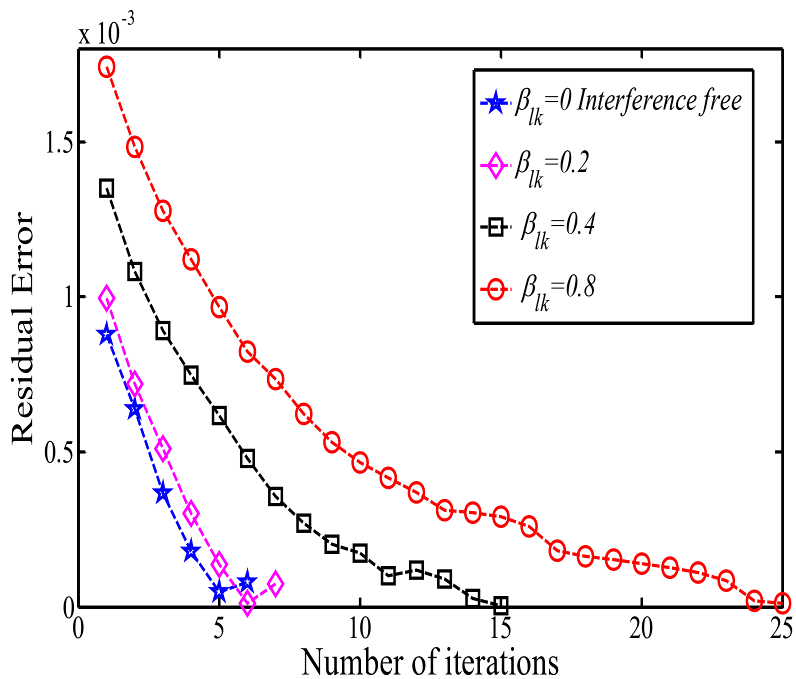

Figure 5.3: Speed convergence of IWNN algorithm in multi-cell system with parameters M = 80, P = 20, τ = 40, Kd =40 with β_{jk}=1, SNR=0dB, and Kc =4 with different values of β_{lk} $\forall\, l \neq j$

Next, we study the computational complexity of the proposed IWNN algorithm in terms of the number of iterations, T, and the residual error, R. It is noteworthy that, noise and pilot contamination problems produce the residual error. Thus, the residual error versus the number of iterations is simulated over the system parameters M=80, P = 20, K_d = 40, τ = 40 and K_c= 4 at different values of large-scale fading coefficients β_{lk}= 0, 0.2, 0.4, and 0.8. Figure 5.3 shows the residual error versus the number of iterations for different cases of the pilot contamination problem. It is clear that the residual estimation error is quickly decreased to zero with the cost of the number of iteration t <10 and t <25 for the low and high pilot contamination problem, respectively.

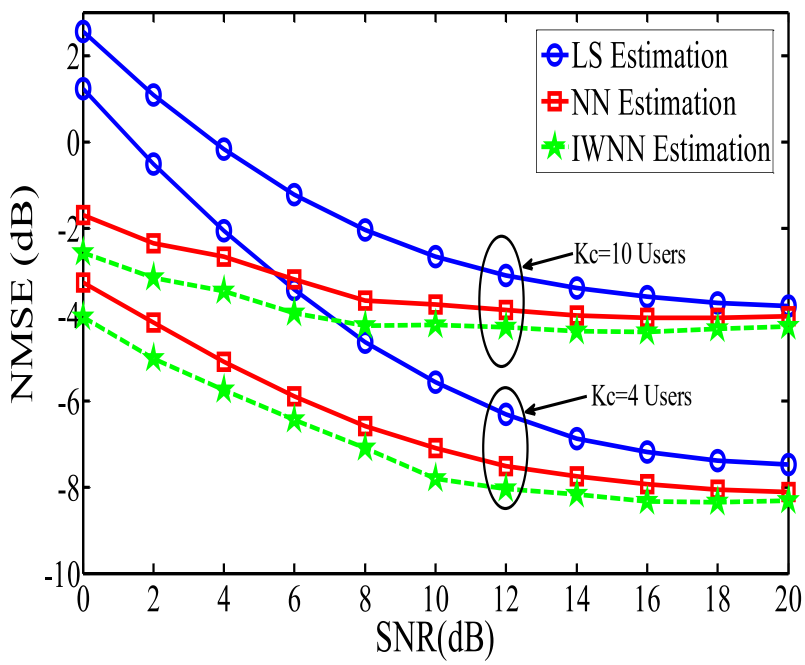

Figure 5.4: NMSE vs. SNR over system parameters, M = 80, P =20, $\tau = 40$, $K_d = 40$ with β_{jk}=1, and $K_c \in \{4{,}10\}$ with β_{lk}=0.8 $\forall\ l \neq j$.

In Figure 5.4, the NMSE versus SNR for the proposed estimation approaches under the different number of contaminated users, i.e., $K_c \in \{4{,}10\}$ is investigated and compared to the conventional LS estimation. As we can see from Figure 5.4, the LS estimator has poor estimation performance due to the pilot contamination problem, while the NN and IWNN proposed methods have better estimation performance. Moreover, the IWNN estimation method has shown substantial improvement over the NN estimation method.

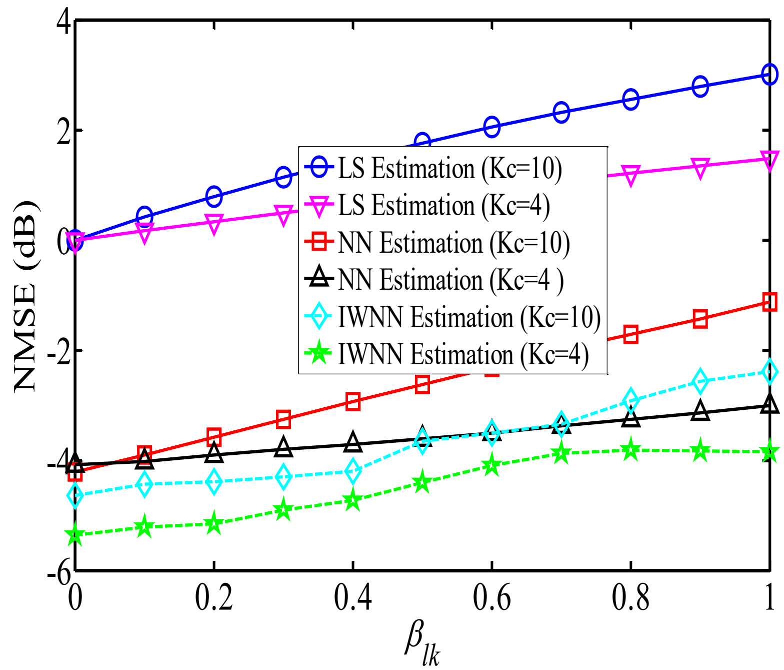

Figure 5.5: NMSE vs. large-scale fading coefficient, β_{lk} for $K_c \in \{4,10\}$ over system parameters, M =80, P=20, τ=40, K_d=40, with β_{jk}=1, and SNR = 0 dB

In Figure 5.5, the effect of pilot contamination in terms of the large-scale fading coefficient β_{lk} of the interfering users in the adjacent cells is studied. Figure 5.5 demonstrates the NMSE performance versus large-scale fading coefficients β_{lk} of two different number of contaminated users, i.e., $K_c \in \{4,10\}$. It can be seen from Figure 5.5 that, the NMSE performances of the LS, NN, and IWNN estimation methods are degraded as β_{lk} increases. However, both proposed estimation approaches exhibit a better estimation performance compared to the LS method in the presence of low and high pilot contamination problem. In other words, the IWNN estimation demonstrated its ability to estimate the desired channel matrix under high pilot contamination problem originated from K_c=10 cell-edge users.

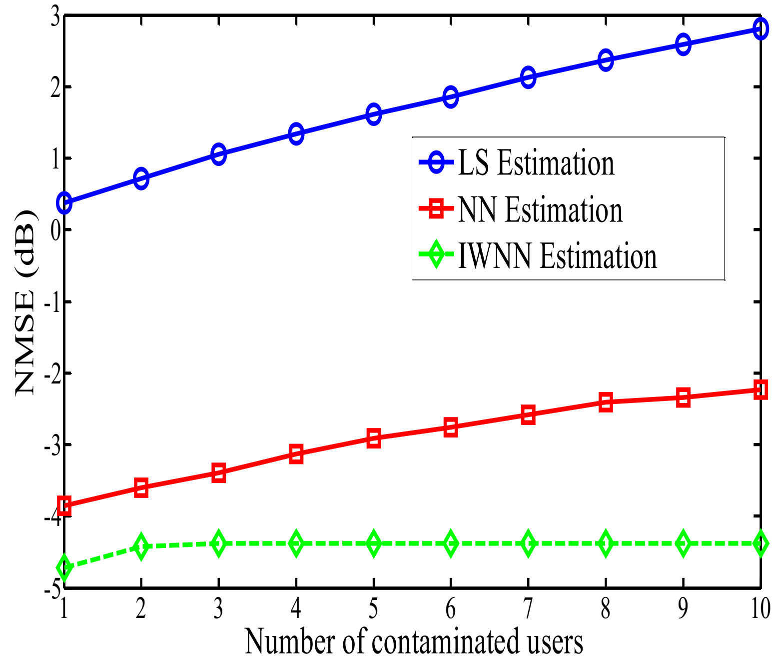

Figure 5.6: NMSE vs. Number of contaminated K_c users with β_{lk}= 0.9 each over system parameters, M = 80, P = 20, τ = 40, K_d = 40, and SNR = 0 dB

In Figure 5.6, the effect of increasing the number of contaminated K_c users with β_{lk}= 0.9 for $l \neq$ j on the channel estimation performance is investigated. It can be seen from Figure 5.6 that as the number of contaminated users increases, the estimation error is also increased in both LS and NN estimations. On the other hand, the IWNN estimation is not changed with K_c, which means that it can mitigate the high pilot contamination problem.

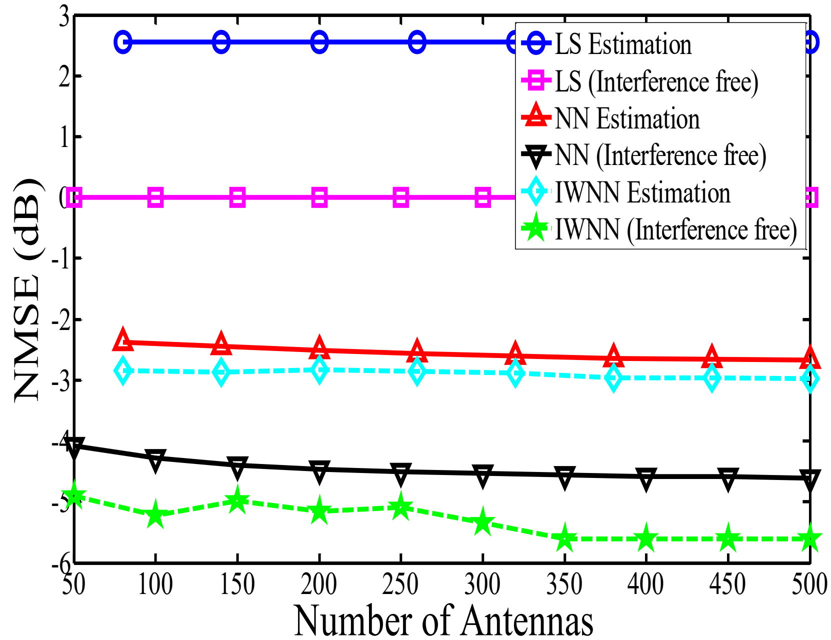

Figure 5.7: NMSE vs. Number of BS antennas M over systems parameters, P=20, τ=40, K_d=40 with β_{jk}=1 each, K_c=10 with β_{lk}=0.8 each $\forall\ l \neq j$ and SNR = 0 dB

In Figure 5.7, the effect of increasing the number of antennas on the NMSE and ASR estimation performances of the proposed estimations is studied. In Figure 5.7, we display the NMSE versus the number of BS antennas with and without interference for the LS, proposed NN, and IWNN estimation methods. It can be seen from Figure 5.7 that as the number of antennas M increases, the performance of the LS estimator is quickly saturated (due to pilot contamination effect), while the performances of our proposed NN and IWNN estimation methods are improved. Compared to the LS and NN estimation methods, the IWNN estimation method provides the highest estimation performance with the cost of a small number of iterations.

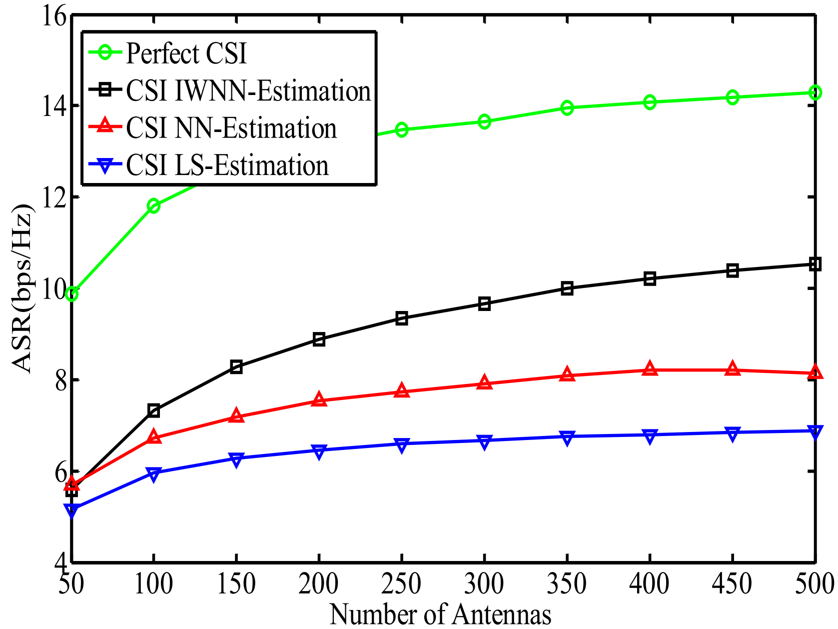

Figure 5.8: Uplink ASR vs. Number of BS antennas, M, for a multi-cell system parameters, L=3 cells, P =20, K_d= 40 with β_{jk}=1 each, K_c=4 with β_{lk}=0.8 each $\forall\ l \neq j$ and SNR=0 dB

Figure 5.8 illustrates the uplink ASR performance obtained by LS, NN, and IWNN estimation methods under a different number of antennas, M. It can be seen in Figure 5.8 that as the number of antennas, M increases, the uplink ASR performances obtained by LS and NN estimation methods are quickly saturated. On the other hand, as the number of antennas, M, increases, the uplink ASR obtained by IWNN estimation method is improved.

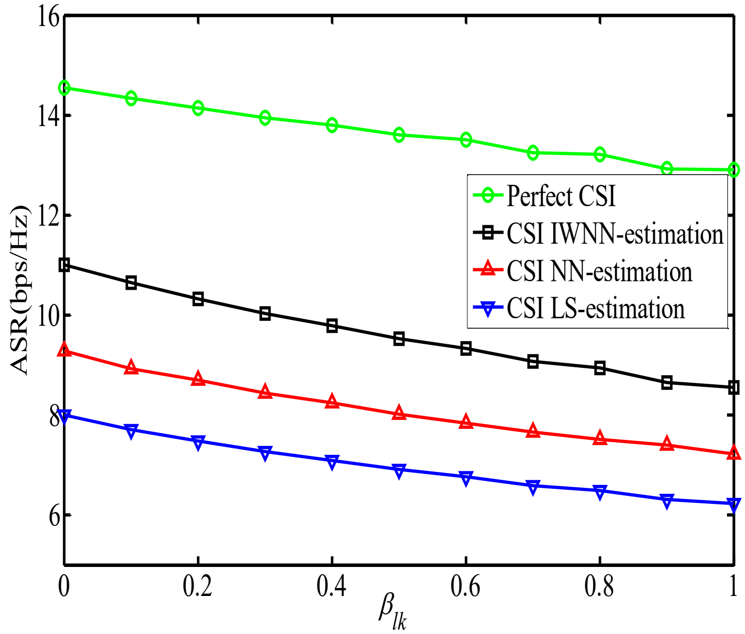

Figure 5.9: Uplink ASR vs. large-scale factor β_{lk} for K_c = 4 users for multi-cell system with parameters, L=3, $M = 200$, P = 20, K_d = 40 with β_{jk} = 1, SNR = 0 dB

Based on the fact that the effect of pilot contamination problem increases as the large-scale factor$\beta_{lk} \forall\, l \neq$ j increases, we are interested in studying this effect on the uplink ASR performance. Figure 5.9 demonstrates the uplink ASR performance obtained with different channel estimation methods under different values of the factorβ_{lk}. It can be seen from Figure 5.9 that the uplink ASR obtained by IWNN estimation method outperforms the ASR obtained by NN and LS estimation methods.

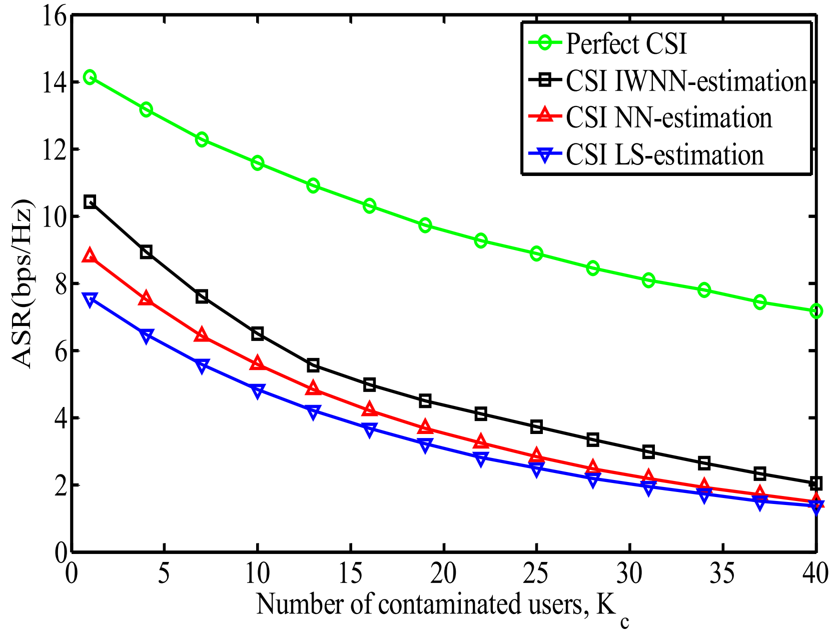

Figure 5.10: Uplink ASR vs. number of K_c with β_{lk}= 0.8 $\forall \ l \neq j$ for system parameters, L=3 cells, M =200, P =20, K_d= 40 with β_{jk}=1, and SNR = 0 dB

Finally, we examine the impact of increasing the number of contaminated users K_c in the adjacent cells on the uplink ASR performance. It can be observed from Figure 5.10 that as K_c increases the uplink ASR performances obtained by LS and NN estimation methods degrade. For example, when $K_c \geq 25$, the uplink ASR performances obtained by NN and LS are almost the same. In contrast, the uplink ASR obtained by IWNN estimation method has less degradation compared to the one obtained by NN and LS methods (due to the iterative searching algorithm proposed for WNN estimation method).

5.9 Conclusion

In this chapter, two-channel estimation algorithms exploiting the sparsity of the channel matrix have been proposed for multi-cell massive MU-MIMO systems. The main goal of these estimation approaches is to mitigate pilot contamination problem. In the first method, the channel estimation problem is formulated as the nuclear norm (NN) minimization problem. Furthermore, the iterative weighted nuclear norm (IWNN) estimation method is proposed to improve the NN estimation performance. The regularization parameter of both NN and IWNN optimization methods is selected based on the cross-validation (CV) curve method. Then further, a brief analysis of the computational complexity of the proposed NN and IWNN estimation approaches are analyzed and compared to the LS estimation method. It is shown that the conventional LS estimation has the lowest complexity with poor estimation performance, while the IWNN has the highest complexity with more estimation performance accuracy.

Moreover, two estimation performance metrics namely normalized mean square error (NMSE) and uplink achievable sum-rate (ASR)" are used to evaluate the proposed NN and IWNN estimation methods under different pilot contamination scenarios. The simulation results show that as the number of base station antennas increases, the NMSE and uplink ASR performances of our proposed NN and IWNN channel estimation approaches are improved compared to the conventional LS method. Furthermore, the IWNN estimation method demonstrates substantial improvement over the NN method in the presence of high pilot contamination problem with the cost of having a small number of iterations.

Chapter 6

Low-Complexity Channel Estimator for TDD Massive Multiuser MIMO Systems

6.1 Introduction

This chapter addresses the problem of minimum mean square error (MMSE) channel estimator in time division duplex (TDD) massive multiuser multi-input multi-output (MIMO) systems. It is noteworthy that, the MMSE channel estimator has been previously proposed for multi-cell massive multiuser MIMO systems [28]. The MMSE estimator suffers from high computational complexity due to the large dimension of the covariance matrix inversion which is scaled with the number of base station antennas [72]. Another inherent drawback of the MMSE channel estimator is the need for additional information about the statistical distribution of the propagation channels (i.e., covariance matrices of the desired and interfering channel users) at each base station. However, this information is not available in the real massive MIMO systems.

Consequently, we propose an alternative scheme namely "Approximate minimum mean square error (AMMSE)" channel estimator by using the low-rank matrix approximation technique to reduce the computational complexity of the MMSE estimator. Our contributions in this work are summarized as:

- Iterative weighted nuclear norm (IWNN) approximation method is proposed to design a novel AMMSE channel estimator for multi-cell TDD massive multiuser MIMO system.
- The computational complexity of the proposed AMMSE estimator is analyzed and compared to the conventional LS and MMSE estimators.

- Normalized mean square error (NMSE) and uplink achievable-sum rate (ASR) performance metrics are used to evaluate the proposed AMMSE estimator under the noise and pilot contamination scenarios.
- Finally, the simulation results show the agreements between the proposed AMMSE and conventional MMSE channel estimators in terms of the NMSE and ASR performances.

6.2 Pilot-based Channel Estimation

First, the MMSE-based channel estimation for a multi-cell TDD massive multiuser MIMO system is explained. Second, the low-complexity AMMSE proposed channel estimator is presented and analyzed. Finally, we investigate the performance of the AMMSE proposed channel estimator in a multi-cell TDD massive multiuser MIMO system with and without pilot contamination problem.

6.2.1 MMSE Channel Estimator

A linear MMSE technique is considered as one of the optimal channel estimators for massive MIMO systems to reduce the pilot contamination effect. Therefore, it has been perversely proposed for a multi-cell TDD massive MU-MIMO system [28], [29]. Moreover, a linear MMSE channel estimator relies on two key ideas to complete removal of pilot contamination effects in the multi-cell massive MIMO systems. The first one is the exploitation of the channel covariance information of both desired and interfering users under a specific condition on the covariance matrices. The second key idea is the use of a covariance-aware pilot assignment strategy within the channel estimation phase to satisfy the requirement in the first one.

To explain above, we consider a multi-cell TDD massive multiuser MIMO system and channel models in the previous chapter. In [28], the expression of a linear MMSE channel estimator for the desired channel matrix at the target BS-j is given as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{R}_j \left(\boldsymbol{R}_j + \sum_{l \neq j}^{L} \boldsymbol{R}_l + \frac{\sigma_n^2}{\tau\, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \widehat{\boldsymbol{H}}_{LS} \tag{6.1}$$

where $\widehat{\boldsymbol{H}}_{LS} \in \mathbb{C}^{M \times K_d}$ is the least square (LS) estimation of the channel matrix, ρ_{tr} is the average signal power used by each K users in all cells, τ is the length of the pilot sequence, and σ_n^2 is the noise variance. In (6.1), $\boldsymbol{R}_j \in \mathbb{C}^{M \times M}$ is the covariance matrix of the desired channel users, while $\boldsymbol{R}_l \in \mathbb{C}^{M \times M}$ is the covariance matrices of the interfering channels from all adjacent cells $l \neq$ j.

As it appears in (6.1), a linear MMSE estimator suffers from high computational complexity due to the large dimension of the covariance matrix inversion which is scaled with the number of base station antennas. Another inherent drawback of the MMSE channel estimator is the need for additional information about the covariance matrices $\boldsymbol{R}_j$ and $\boldsymbol{R}_l\ \forall\ l \neq j$ at each BS which are not available in the real massive MIMO systems. Therefore, we further proceed with the iterative nuclear norm approximation method to design an alternative channel estimator scheme namely "Approximate minimum mean square error (AMMSE)" to reduce the computational complexity of the conventional MMSE channel estimator.

6.2.2 Proposed AMMSE Channel Estimator

The eigenvalue decomposition (EVD) method and low-rank reduction theory are applied to design low-complexity AMMSE channel estimator for multi-cell massive MU-MIMO system. The EVD in [24] is applied to the covariance matrices $\boldsymbol{R}_j$ and $\boldsymbol{R}_l$ in the MMSE channel estimator in (6.1) as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j\ \boldsymbol{\Sigma}_j\ (\boldsymbol{U}_j)^H \left(\boldsymbol{U}_j\ \boldsymbol{\Sigma}_j\ (\boldsymbol{U}_j)^H + \sum_{l \neq j}^{L} \boldsymbol{U}_l\ \boldsymbol{\Sigma}_l\ (\boldsymbol{U}_l)^H + \frac{\sigma_n^2}{\tau\, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \widehat{\boldsymbol{H}}_{LS} \tag{6.2}$$

where $\boldsymbol{\Sigma}_j$ is a diagonal matrix containing the eigenvalues $\sigma_{j1} \geq \sigma_{j2} \geq \cdots \geq \sigma_{jM}$ of the desired channels and $\boldsymbol{\Sigma}_l \; \forall \; l \neq \mathrm{j}$ is a diagonal matrix containing the eigenvalues $\sigma_{l1} \geq \sigma_{l2} \geq \cdots \geq \sigma_{lM}$ of the interference channels. In (6.2), $\boldsymbol{U}_j \in \mathbb{C}^{M\times M}$ and $\boldsymbol{U}_l \in \mathbb{C}^{M\times M}$ are unitary eigenvector matrices of desired and interference channels, respectively. For the worst-case scenario, we assume that all channel users from all cells share the same angle of arrivals. Under this assumption, the correlation among these channels is increased. Therefore, all channel users will have the same steering matrix with$\boldsymbol{U}_j = \boldsymbol{U}_l$ and different eigenvalues. Thus, the Equation in (6.2) is simplified as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \; \boldsymbol{\Sigma}_j \, \boldsymbol{U}_j{}^H \left(\boldsymbol{U}_j \left(\boldsymbol{\Sigma}_j \; + \sum_{l\neq j}^{L} \; \boldsymbol{\Sigma}_l \; \right) \boldsymbol{U}_j{}^H + \; \frac{\sigma_n^2}{\tau \, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \widehat{\boldsymbol{H}}_{LS} \tag{6.3}$$

By using the matrix inversion identity $\boldsymbol{A}(\boldsymbol{BA}+\boldsymbol{I})^{-\mathbf{1}}$=$(\boldsymbol{AB}+\boldsymbol{I})^{-\mathbf{1}}\boldsymbol{A}$, the Equation (6.3) can be written as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \; \boldsymbol{\Sigma}_j \left(\boldsymbol{U}_j{}^H \boldsymbol{U}_j \left(\boldsymbol{\Sigma}_j \; + \sum_{l\neq j}^{L} \; \boldsymbol{\Sigma}_l \; \right) + \; \frac{\sigma_n^2}{\tau \, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \boldsymbol{U}_j{}^H \, \widehat{\boldsymbol{H}}_{LS} \tag{6.4}$$

In a massive MIMO system with a very large number of base station antennas, M, we have

$$\boldsymbol{U}_j{}^H \boldsymbol{U}_j = \boldsymbol{I}_M \tag{6.5}$$

By substituting (6.5) into (6.4), we rewrite (6.4) as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \; \boldsymbol{\Sigma}_j \left(\boldsymbol{\Sigma}_j + \sum_{l\neq j}^{L} \boldsymbol{\Sigma}_l \; + \frac{\sigma_n^2}{\tau \, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \boldsymbol{U}_j{}^H \, \widehat{\boldsymbol{H}}_{LS} \tag{6.6}$$

which is equivalent to

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \; \boldsymbol{\Sigma}_j \, (\boldsymbol{\Sigma}_{LS})^{-1} \, \boldsymbol{U}_j{}^H \, \widehat{\boldsymbol{H}}_{LS} \tag{6.7}$$

where $\boldsymbol{\Sigma}_{LS}$ is an LS diagonal matrix containing the eigenvalues $\widehat{\sigma}_{j1} \geq \widehat{\sigma}_{j2} \geq \cdots \geq \widehat{\sigma}_{jM}$ on its diagonal which is defined as

$$\boldsymbol{\Sigma}_{LS} = \boldsymbol{\Sigma}_j + \sum_{l \neq j}^{L} \boldsymbol{\Sigma}_l + \frac{\sigma_n^2}{\tau\, \rho_{tr}} \boldsymbol{I}_M \tag{6.8}$$

Next, equation (6.7) can be simplified as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \;\Delta\; \boldsymbol{U}_j{}^H \;\widehat{\boldsymbol{H}}_{LS} \tag{6.9}$$

where Δ is a diagonal matrix which can be expressed as

$$\begin{aligned}\Delta &= \boldsymbol{\Sigma}_j \left(\boldsymbol{\Sigma}_j + \sum_{l \neq j}^{L} \boldsymbol{\Sigma}_l + \frac{\sigma_n^2}{\tau\, \rho_{tr}} \boldsymbol{I}_M \right)^{-1} \\ &= diag\left(\frac{\sigma_{j1}}{\sigma_{j1} + \sum_{l \neq j}^{L} \sigma_{l1} + \frac{\sigma_n^2}{\tau\, \rho_{tr}}}, \ldots\ldots, \frac{\sigma_{jM}}{\sigma_{jM} + \sum_{l \neq j}^{L} \sigma_{lM} + \frac{\sigma_n^2}{\tau\, \rho_{tr}}} \right)\end{aligned} \tag{6.10}$$

To reduce the computational complexity of the MMSE in (6.9) and introduce the AMMSE proposed channel estimator, the theory of low-rank reduction in [15] is applied to (6.9) as

$$\widehat{\boldsymbol{H}}_{MMSE} = \boldsymbol{U}_j \begin{bmatrix} \Delta_r & 0 \\ 0 & 0 \end{bmatrix} (\boldsymbol{U}_j)^H \;\widehat{\boldsymbol{H}}_{LS} \tag{6.11}$$

where Δ_r is the $r \times r$ upper left corner of the matrix Δ with entries

$$\delta_i = \begin{cases} \dfrac{\sigma_{ji}}{\sigma_{ji} + \sum_{l \neq j}^{L} \sigma_{li} + \frac{\sigma_n^2}{\tau\, \rho_{tr}}}, & i = 1,2, \ldots\ldots, r \\ 0, & i = r+1, \ldots, M \end{cases} \tag{6.12}$$

where r is the rank of the desired channel covariance matrix.

It should be noted that the subspace eigenvector matrix $\boldsymbol{U}_j$ in (6.11) can be approximated by using a subset of discrete Fourier transform (DFT) basis [30], [34]. This DFT basis can be selected based on a small number of channel observations. However, the diagonal matrix Δ_r is only required to be known at the target BS-j to complete the AMMSE proposed estimator design. Thus, the IWNN approximation based on the low-rank reduction theory is applied in this work. It is noteworthy that, the IWNN approximation has been recently proposed for massive MU-MIMO channel estimation in our previous work [40], [41], [58]. Hence, the LS eigenvalue estimation problem in (6.8) is formulated as an unconstrained weighted nuclear norm (WNN) optimization problem as follows

$$\tilde{\Delta}_r = arg\, \min_{\boldsymbol{\Sigma}_j^j} \left\{ \frac{1}{2} \left\| \boldsymbol{\Sigma}_{LS} - \boldsymbol{\Sigma}_j \right\|_F^2 + \gamma \sum_{i=1}^{r} \left| w_i\, \sigma_{ji} \right| \right\} \tag{6.13}$$

where $\tilde{\Delta}_r$ is the diagonal matrix with approximate eigenvalues $\tilde{\sigma}_{j1} \geq \tilde{\sigma}_{j2} \geq \ldots \ldots \geq \tilde{\sigma}_{jr}$ at the main diagonal, and γ is the regularization parameter which is used to control the trade-off between the data fidelity, $\frac{1}{2} \left\| \boldsymbol{\Sigma}_{LS} - \boldsymbol{\Sigma}_j \right\|_F^2$, and the penalty function, $\sum_{i=1}^{r} \left| w_i\, \sigma_{ji} \right|$. Moreover, the regularization parameter γ can be computed by using the formula [54] which is given as

$$\gamma = \sqrt{2(M + LK)\tau\, \rho_{tr} \sigma_n^2} \tag{6.14}$$

In (6.13), w_i is the weight element which can be obtained by using the formula in [64] which is given as:

$$w_i^{t+1} = w_i^t + \frac{\mu}{\tilde{\sigma}_i^t + \varepsilon} \quad i = 1,2 \ldots . r \tag{6.15}$$

where $\tilde{\sigma}_i^t$ is the estimated eigenvalue of $\tilde{\Delta}_r$ in the t^{th} iteration, μ is the step-size parameter which is used to accelerate time convergence. To avoid dividing (6.15) by zero, a small positive number ε is added. The proposed IWNN approximation method used to design the AMMSE channel estimator is summarized in Table 6.1.

Table 6.1: IWNN Approximation Algorithm for MMSE Channel Estimator

1: Choose the number of iterations, T, μ = 0.5, ε=10^{-5}, and then set t=0.

2: Using the LS channel estimation approach which is explained in the previous chapter, compute the LS covariance matrix as

$$\widehat{\boldsymbol{R}}_{LS} = E\{(\widehat{\boldsymbol{H}}_{LS})(\widehat{\boldsymbol{H}}_{LS})^{H}\} \tag{6.16}$$

3: Apply the EVD of $\widehat{\mathbf{R}}_{LS}$, to obtain the LS diagonal matrix $\boldsymbol{\Sigma}_{LS}$, as

$$\widehat{\boldsymbol{R}}_{LS} = \boldsymbol{U}_{LS}\ \boldsymbol{\Sigma}_{LS}\ \boldsymbol{U}_{LS}{}^{H} \tag{6.17}$$

4: Select the initial weight elements as

$$w_i^0 = \frac{1}{\hat{\sigma}_{ji}^{LS}}\ \ ,\ \ where\ \ i = 1, 2, \dots \dots \dots \dots, M \tag{6.18}$$

5: Using the formula in (6.14), compute the regularization parameter, γ.

6: To obtain, $\tilde{\Delta}_r$, solve the optimization problem in (6.13).

7: $t = t + 1$

8: Update the weight elements by using the formula in (6.15).

9: Repeat steps from 6-8 until the selected T is reached.

10: Finally, compute the desired channel matrix by substituting (6.13) into (6.11).

6.3 Estimator Complexity Analysis

The main complexity of the MMSE channel estimator comes from a significant dimension of matrix inversion in (6.1) [72], [77]. It is scaled with the length of the pilot sequence, τ, and the number of BS antennas, M. In other words, the total number of multiplications required to estimate the desired channel matrix by using the MMSE estimator is $M^3\tau^3$. Compared to the MMSE channel estimator, the total number of multiplications needed to estimate the desired channel matrix by using the proposed AMMSE channel estimator is only $M\tau PT$, where T is the total number of iterations. As a result, we have managed to reduce the number of multiplications using the proposed AMMSE channel estimator from $M^3\tau^3$ to $M\tau PT$. However, the proposed AMMSE estimator can only be considered as low-complexity channel estimator if $T \leq 100$. As such, the computational complexity in terms of the total number of iterations is studied. The asymptotic complexities and the estimation performances of the conventional LS, MMSE, and the proposed AMMSE estimators are summarized in Table 6.2.

Table 6.2: Asymptotic Complexities of Different Estimators

Channel Estimators	Computational Complexity	Estimation Performance
LS	$\mathcal{O}(M\tau)$	Low
AMMSE	$\mathcal{O}(M\tau PT)$	Moderate
MMSE	$\mathcal{O}(M^3\tau^3)$	High

6.4 Simulation Results

This section contains the simulation results of the proposed AMMSE channel estimator for multi-cell TDD massive MU-MIMO system. However, these simulation results are compared to the conventional LS and MMSE estimators under two different scenarios: noise-limited and pilot contamination. To investigate above, we consider the same system and channel models for multi-cell massive MU-MIMO system with L= 3 cells as explained in the previous chapter. In this study, we assume each cell containing the total number of $K = 20$ users, and one BS with $M = 100$ antennas. We assume the length of the pilot sequence is $\tau = 20$, and the number of multipath is P=10. The SNR is set to 0 dB, and the steering vector parameters are selected to be D/λ = 0.5, and $\theta_p = -\pi/2 + (p-1)\pi/2P$, where p=1, 2…...P. The two following performance criteria are used to evaluate the proposed AMMSE channel estimator in different interference scenarios. The first criteria is a normalized mean square error (NMSE) which is given by [28] as

$$NMSE(dB) = 10log_{10}\left(\frac{E\left\|\widehat{\boldsymbol{H}}_{MMSE} - \boldsymbol{H}_j\right\|_F^2}{E\left\|\boldsymbol{H}_j\right\|_F^2}\right) \tag{6.19}$$

where $\mathbf{H}_j$ and $\widehat{\boldsymbol{H}}_{MMSE}$ are the desired channel and its estimate by using the conventional MMSE or AMMSE estimator, respectively. The second criteria is the uplink achievable sum-rate (ASR) of K users which is given as

$$ASR \le \sum_{k=1}^{K} log_2(1 + SINR_k) \tag{6.20}$$

where $\mathrm{SINR_k}$ is the received signal-to-interference-plus-noise ratio of the k^{th} user at the linear detector output. When the linear maximum ratio combing (MRC) detector is assumed at the BS-j, the weight vector of MRC is defined as $\boldsymbol{v}_{jk} = \widehat{\boldsymbol{h}}_{jk}$. It should be noted that the $\mathrm{SINR_k}$ of each k^{th} user in (6.20) can be calculated by using Equation (3.41) in the previous chapter.

6.4.1 Noise-limited Scenario

In this scenario, we study the behaviors of the proposed AMMSE estimator for multi-cell massive MU-MIMO system under a noisy setting. The noise-limited scenario is given by setting β_{lk}= 0 for all interfering channel users in the adjacent cells, while β_{jk}=1 for all desired channel users in the target cell-j.

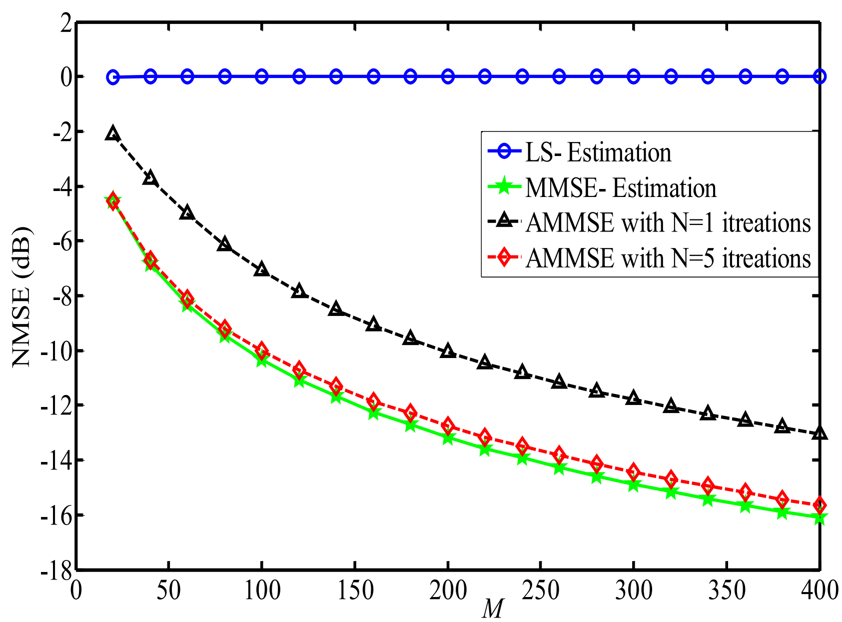

Figure 6.1: NMSE performance comparison between different estimators under different, M, in the noise-limited scenario (β_{lk}= 0)

In Figure 6.1, the NMSE performance of the proposed AMMSE estimator is evaluated under a different number of BS antennas M, and compared to the conventional LS and MMSE estimators. It can be seen from Figure 6.1 that the NMSEs of the AMMSE and MMSE estimators are converged to zero as M increases, while the LS estimator does not show any change. Moreover, the NMSE of the proposed AMMSE estimator is only about 0.5 dB less than the MMSE estimator when $M \geq 100$.

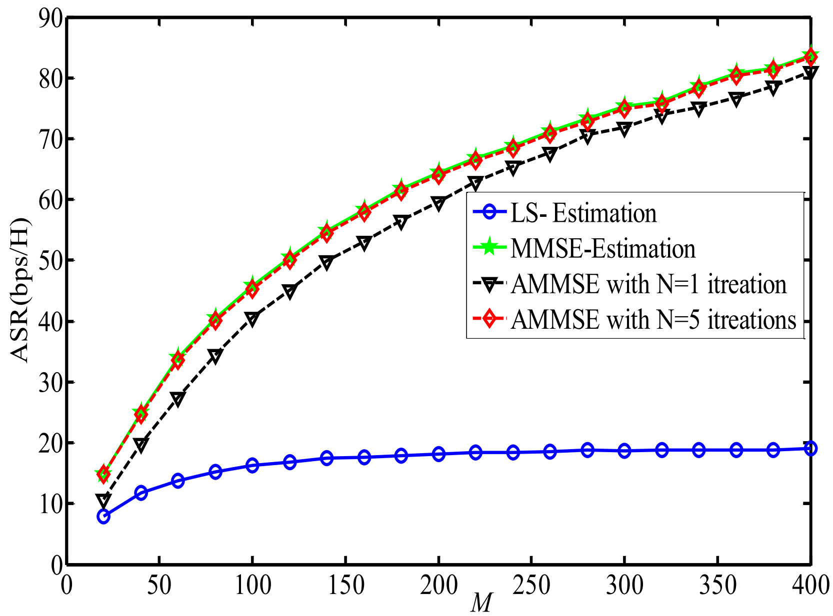

Figure 6.2: ASR performance comparison between different estimators under different, M, in the noise-limited scenario (β_{lk}= 0).

In Figure 6.2, the ASR performance of the proposed AMMSE estimator is investigated under a different number of BS antennas M, and compared to the conventional LS and MMSE estimators. It can be seen from this figure that as M increases, the uplink ASRs of the proposed AMMSE and MMSE estimators are almost the same and improved compared to the LS estimation performance. On the other hand, the uplink ASR of the LS estimator is quickly saturated as $M \geq 100$.

As mentioned earlier in section 6.3, the main computational complexity of the proposed estimator comes from the number of iterations, T. Thus, we now turn our attention to study the computational complexity of the proposed AMMSE estimator. It can be seen from Figure 6.1 and

Figure 6.2 that after only N=5 iterations, the AMMSE estimator is quickly converged to the MMSE estimator which means that the AMMSE estimator has low complexity.

6.4.2 Pilot Contamination Scenario

In this scenario, we study the behavior of the proposed AMMSE estimator under the weak and strong pilot contamination effects. The weak pilot contamination is given by setting β_{lk}= 0.1 for all interfering channel users in the adjacent cells, while β_{lk}= 0.9 for the strong pilot contamination case. For all desired channel users in the target cell-j, β_{jk}=1 and SNR = 0 dB have remained the same as in the noise-limited scenario. First, the impact of the weak pilot contamination on the proposed AMMSE estimator is studied and compared to various estimation performances under a different number of BS antennas, M.

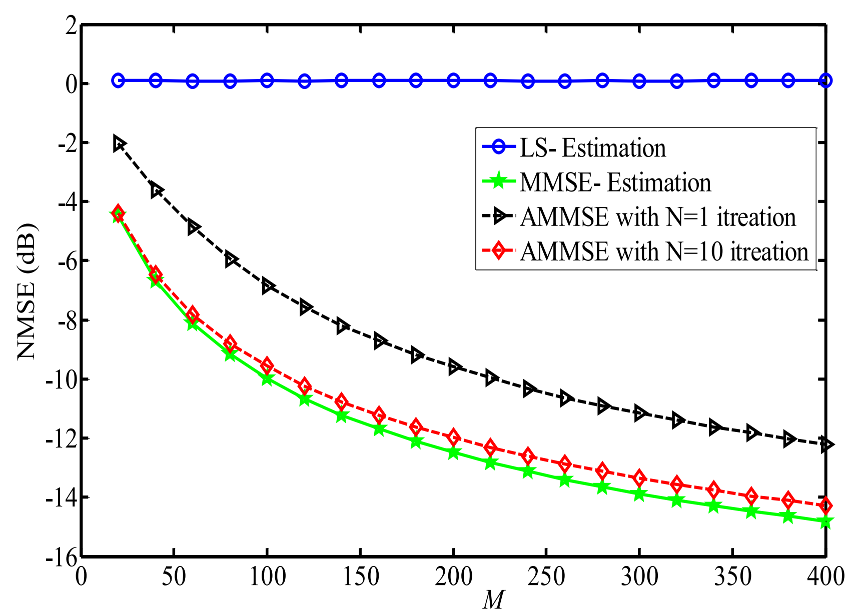

Figure 6.3: NMSE performance comparison between different estimators under different, M, in the weak pilot contamination scenario (β_{lk}= 0.1).

Figure 6.3 displays the comparison between the conventional LS, MMSE, and the proposed AMMSE estimators in terms of the NMSE. It can be seen from this figure that as M increases, the performance of the AMMSE estimator is about 3 bps/Hz less than the MMSE and about 12 bps/Hz better than the LS estimator when $T = 1$ iteration. By increasing the number of iterations from $T = 1$ to $T = 10$ in the proposed algorithm, the performance of the proposed AMMSE estimator is improved by 2 bps/Hz compared to the algorithm with $T = 1$.

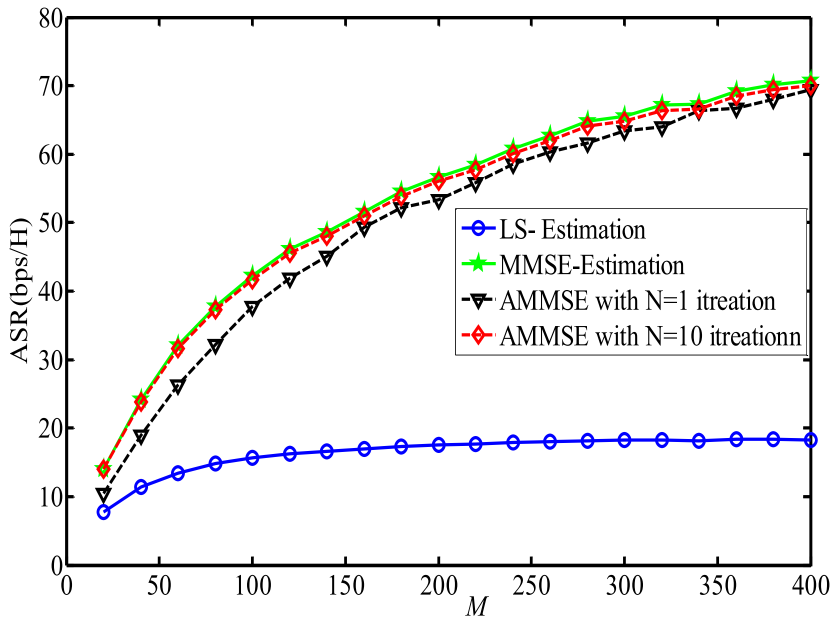

Figure 6.4: ASR performance comparison between different estimators under different, M, in the weak pilot contamination scenario (β_{lk}= 0.1).

Moreover, the uplink ASR performance of the proposed AMMSE is investigated in the presence of weak pilot contamination and compared to the LS and MMSE estimation performances. It can be seen from Figure 6.4 that as M increases, the uplink ASR of the proposed AMMSE is almost the same as MMSE estimation performance even with a small number of iterations, T. On the other hand, the uplink ASR of the proposed AMMSE achieves significantly better performance compared to the LS estimator.

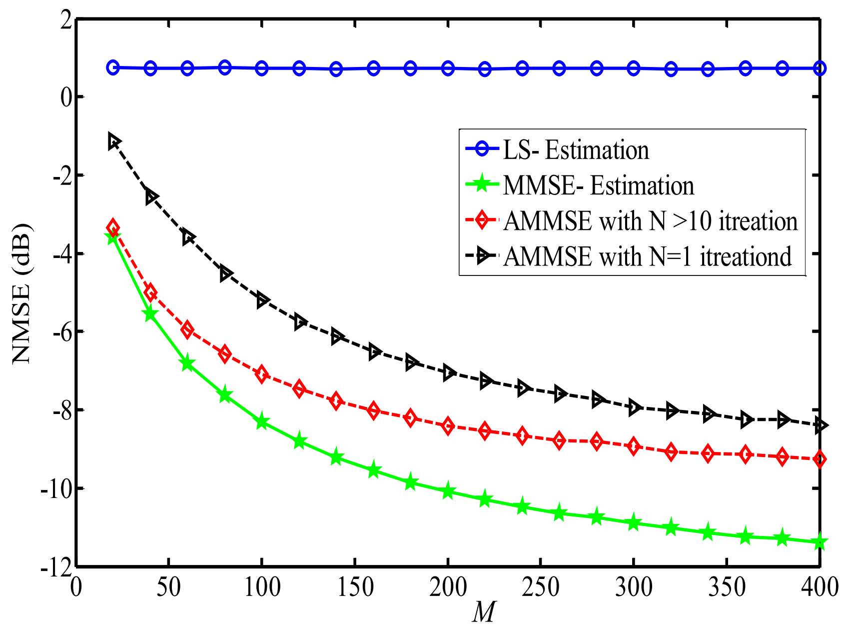

Figure 6.5: NMSE performance comparison between different estimators under different, M, in the strong pilot contamination scenario (β_{lk}=0.9).

Finally, the impact of strong pilot contamination on the proposed AMMSE estimator under a different number of BS antennas, M is studied. Figure 6.5 shows the NMSE performance of different estimators versus the number of base station antennas, M. It can be seen from this figure that when T=1 iteration and M increases, the NMSE of the proposed AMMSE estimator is about 3 dB less than the MMSE estimator and 11 dB better than the LS estimator. Moreover, when $M \leq 100$ and $T \geq 10$, the NMSE performance of the proposed estimator is converged to the MMSE estimator, while it diverges as M increases.

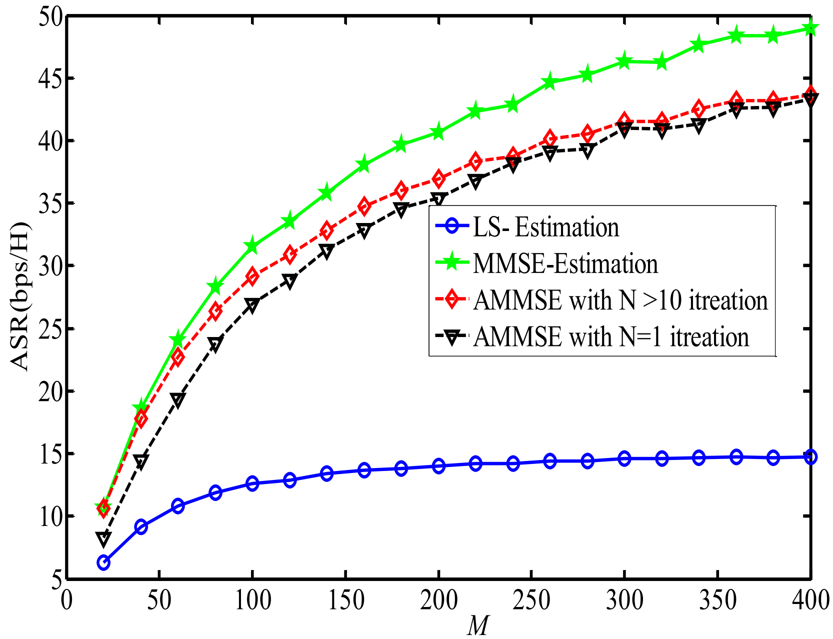

Figure 6.6: ASR performance comparison between different estimators under different, M, in the strong pilot contamination scenario (β_{lk}=0.9).

Figure 6.6 illustrates the ASR performance of different estimators versus the number of base station antennas, M. It can be seen from this figure that when $M \leq 100$ and $N \geq 10$ iterations, the ASR performance of the proposed AMMSE estimator is only about 2 bps/Hz less than the MMSE estimator. However, as M increases, the ASR performance of the proposed AMMSE estimator is improved by 5 bps/Hz compared to the MMSE estimator with a small number of iterations. In contrast, the uplink ASR obtained by the LS channel estimation is quickly saturated for $M \geq 100$.

6.5 Conclusion

In this work, a novel low-complexity channel estimator namely “approximate minimum mean square error (AMMSE)” for multi-cell TDD massive MU-MIMO systems has been proposed. The IWNN approximation based on the low-rank reduction theory is proposed to design the proposed AMMSE estimator. Compared to the conventional MMSE estimator, the computational complexity of the proposed AMMSE estimator regarding the number of multiplications is reduced from $\mathcal{O}(M^3\tau^3)$ to $\mathcal{O}(M\tau PN)$. The simulation results show that the proposed estimator and the conventional MMSE estimator have almost the same performance in terms of the NMSE and uplink ASR performances. These estimation performances of the proposed AMMSE estimator have been investigated under two different scenarios: noise-limited and pilot contamination. Moreover, the AMMSE channel estimation performance outperforms the LS estimation in terms of the NMSE and the uplink ASR.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

This thesis has been concerned with the development of new channel estimation techniques for single-cell and multi-cell TDD massive MU-MIMO networks. More specifically, the proposed estimation schemes are based on the applications of low-rank matrix approximation (LRMA) techniques. We have also provided analysis and simulation results to show the performance improvements of the proposed channel estimation schemes in both single-cell and multi-cell systems. Two performance criteria namely " normalized mean square error (NMSE) and uplink achievable sum-rate (ASR)" are used to evaluate the proposed estimation schemes under different interference scenarios. A brief analysis of the complexities of the proposed schemes regarding the number of iterations is provided to confirm the reduction of the computational complexity of the proposed systems.

In Chapter 3, a novel channel estimation approach namely "nuclear norm (NN) approximation," has been proposed for a single-cell TDD massive MU-MIMO system. The main aim of the proposed scheme is to estimate the channel matrix entries with a limited number of pilot sequences. Hence, the channel estimation problem is formulated as a unconstrainted nuclear norm minimization problem and solved via the proposed algorithm. The simulation results show that the performance of the proposed scheme in terms of the NMSE and uplink ASR significantly outperforms the traditional LS estimation, which ignores the sparsity feature of the channels.

In Chapter 4, the proposed iterative weighted nuclear norm (IWNN) approximation scheme has been proposed to improve the previously proposed nuclear norm (NN) estimation method. The accuracy of the proposed iterative algorithm has been controlled by an appropriate setting of the weight element which is assigned to each singular value of the channel matrix. The simulation results show that the proposed IWNN method with the cost of having a few iterations results in significant improvements regarding the NMSE and uplink ASR over the LS and NN estimation schemes. Furthermore, the computational complexity of the IWNN estimation technique is also studied in terms of the number of iterations.

Chapter 5 has extended the applications of the LRMA to reduce the pilot contamination problem in a multi-cell massive MU-MIMO system. Hence, the NN and IWNN estimation schemes exploiting the sparsity of the channel matrix have been proposed. The simulation results show that both the performances of the proposed NN and IWNN in terms of the NMSE and uplink ASR are improved compared to the conventional LS method. Furthermore, the IWNN estimation scheme demonstrates substantial improvement over the NN estimation under different pilot contamination scenarios.

In Chapter 6, a low-complexity channel estimator namely "approximate minimum mean square error (AMMSE)" for multi-cell TDD massive MU-MIMO systems has been proposed. The computational complexity of the proposed AMMSE estimator regarding the number of multiplications has been reduced by using IWNN reduction scheme, which is based on the low-rank reduction theory. The simulation results show the performance agreements between the proposed estimator and the conventional MMSE estimator in terms of the NMSE and the uplink ASR performances under two different scenarios: noise-limited and pilot contamination. The AMMSE channel estimation performance outperforms the LS estimation.

7.2 Future Work

Based on our study of massive MU-MIMO channel estimation, a few potential research topics can be identified.

- Our work only deals with the flat fading channel estimation for single-carrier massive TDD MU-MIMO systems. However, in practice, it is natural to consider the same scenarios in a multi-carrier system, i.e., the OFDM-based scheme [85].
- In this thesis, the proposed IWNN estimation scheme for both single-cell and multi-cell massive TDD MU-MIMO systems is amplitude-based projection. An interesting topic now is to develop a robust channel estimation scheme that effectively combines projections in both angular and amplitude domains. Specifically, it is reasonable to consider the channel estimation based on the joint angle of arrival (AoA) estimation scheme and singular value decompensation (SVD) method.
- The channel estimation problems for both single-cell and multi-cell massive TDD MU-MIMO systems are formulated as a convex regularization problem with one penalty function each. The recent results show that the better estimation performances of a sparse low-rank matrix can be obtained by using the non-regularization estimation scheme. Some works have recently considered this approach to estimate a sparse low-rank matrix from its noisy observation in matrix completion problems [56], [57], which can be applied in the proposed estimation methods of this thesis.
- A low-complexity alternative channel estimator (AMMSE) is proposed in this thesis for only a single-carrier multi-cell TDD massive MU-MIMO system. However, in practice, it is natural to consider the same channel estimator in a multi-carrier system, i.e., AMMSE-OFDM-based channel estimator scheme.

Bibliography

[1] B. Raaf, W. Zirwas, K.-J. Friederichs, E. Tiirola, M. Laitila, P. Marsch, and R. Wichman, "Vision for beyond 4G broadband radio systems," in *Proc. IEEE PIMRC'11,* Toronto, Canada, 2011, pp. 2369-2373.

[2] L. Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin, R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Topics Signal Process.,* vol. 8, *(5),* pp. 742-758, 2013.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.,* vol. 9, *(11),* pp. 3590-3600, 2010.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," in *Proc. 49th Allerton Conference on Communication, Control, and Computing,* Monticello, IL; United States, 2011, pp. 1272-1279.

[5] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.,* vol. 17, *(1),* pp. 574-590, 2018.

[6] 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 Process. Mag.,* vol. 30, *(1),* pp. 40-60, 2013.

[7] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multi-cell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Communn.,* vol. 61, *(6),* pp. 2350-2361, 2013.

[8] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.,* vol. 63, *(11),* pp. 3917-3923, 2015.

[9] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, and H. Tullberg, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.,* vol. 52, *(5),* pp. 26-35, 2014.

[10] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next-generation wireless systems," *IEEE Commun. Mag.,* vol. 52, *(2),* pp. 186-195, 2014.

[11] K. Zheng, L. Zhao, J. Mei, B. Shao, W. b Xiang, L. Hanzo, "Survey of Large-Scale MIMO Systems," *IEEE Communications Surveys & Tutorials,* vol. 17, *(3),* pp. 1738-1760, 2015.

[12] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.,* vol. 54, *(2),* pp. 114-123, 2016.

[13] O. Elijah, C.Y. Leow, T.A. Rahman, S. Nunoo, S.Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Communications Surveys & Tutorials,* vol. 18, *(2),* pp. 905-923, 2016.

[14] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, "Pilot contamination reduction in multi-user TDD systems," in *2010 IEEE International Conference on Communications,* Cape Town, South Africa, 2010, pp. 1-5.

[15] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* MA: Addison-Wesley: 1991.

[16] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.,* vol. 54, *(3),* pp. 884-893, 2006.

[17] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Fortieth Asilomar Conference on Signals, Systems and Computers,* Pacific Grove, CA, USA, 2006, pp. 359-363.

[18] L. Cottatellucci, R. R. Müller, and M. Vehkapera, "Analysis of pilot decontamination based on power control," in *IEEE Vehicular Technology Conference (VTC Spring),* Dresden, Germany, 2013, pp. 1-5.

[19] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large-scale antenna systems," in *2012 IEEE International Symposium on Information Theory Proceedings,* Cambridge, MA, USA, 2012, pp. 1137-1141.

[20] L. Li, A. Ashikhmin, and T. Marzetta, "Pilot contamination precoding for interference reduction in large scale antenna systems," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton),* Monticello, IL, USA, 2013, pp. 226-232.

[21] D. Neumann, A. Gruendinger, M. Joham, and W. Utschick, "Pilot coordination for large-scale multi-cell TDD systems," in *Proc. 18th Int. ITG Workshop Smart Antennas (WSA),* Erlangen, Germany, 2014, pp. 1-6.

[22] J. H. Sørensen and E. de Carvalho, "Pilot decontamination through pilot sequence hopping in massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM),* 2014, pp. 3285-3290.

[23] L. Sanguinetti, A. A. D'Amico, M. Morelli, and M. Debbah, "Random access in uplink massive MIMO systems: How to exploit asynchronicity and excess antennas," in *Proc. IEEE Global Commun. Conf. (GLOBECOM),* 2016, pp. 1-5.

[24] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multi-cell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE Int. Conf. Acoust. , Speech, Signal Process. (ICASSP),* Kyoto, Japan, 2012, pp. 3249-3252.

[25] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Analysis of blind pilot decontamination," in *Proc. 47th Asilomar Conf. Signals, Syst. Comput (ACSSC).* Pacific Grove, CA, USA, 2013, pp. 1016-1020.

[26] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.,* vol. 8, *(5),* pp. 773-786, 2014.

[27] D. Hu, L. He, and X. Wang, "Semi-blind pilot decontamination for massive MIMO systems," *IEEE Trans. Wireless Commun.,* vol. 15, *(1),* pp. 525-536, 2016.

[28] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.,* vol. 31, *(2),* pp. 264-273, 2013.

[29] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Müller, and G. He, "Robust pilot decontamination based on joint angle and power domain discrimination," *IEEE Trans. Signal Process.,* vol. 64, *(11),* pp. 2990-3003, 2016.

[30] H. Yin, L. Cottatellucci, D. Gesbert, R. R. , Müller, and G. He, "Pilot decontamination using combined angular and amplitude based projections in massive MIMO systems," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC),* Stockholm, Sweden, 2015, pp. 216-220.

[31] S. L. H. Nguyen and A. Ghrayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *2013 IEEE Wireless Communications and Networking Conference (WCNC),* Shanghai, China, 2013, pp-2890-2895.

[32] B. Negahban and M. J. Wainwright, "Estimation of (near) low-rank matrices with noise and high-dimensional scaling," *The Annals of Statistics,* vol. 39, *(2),* pp. 1069-1097, 2011.

[33] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications.* 2012.

[34] Z. Chen and C. Yang, "Pilot decontamination in wideband massive MIMO systems by exploiting channel sparsity," *IEEE Trans. Wireless Commun,* vol. 15, *(7),* pp. 5087-5100, 2016.

[35] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access,* vol. 4, pp. 7313-7321, 2016.

[36] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.,* vol. 63, *(1),* pp. 104-118, 2015.

[37] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, " Efficient collaborative sparse channel estimation in massive MIMO," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* Brisbane, QLD, Australia, 2015, pp. 2924-2928.

[38] M. Hawej and Y. R. Shayan, "Evaluation of massive MU-MIMO channel estimation based on uplink achievable-sum rate criteria," in *2019 IEEE 32nd Annual Canadian Conference on Electrical and Computer Engineering (CCECE),* Edmonton, Canada, 2019, pp. 1-5.

[39] M. Hawej and Y. R. Shayan, "Compressive sensing based nuclear norm minimization method for massive MU-MIMO channel estimation," in *2018 IEEE 31st Annual Canadian Conference on Electrical and Computer Engineering (CCECE),* Quebec City, Canada, 2018, pp. 1-5.

[40] M. Hawej and Y. R. Shayan, "Iterative weighted nuclear norm minimization-based channel estimation for massive multi-user MIMO systems," in *IEEE 88th Vehicular Technology Conference (VTC),* Chicago, USA, 2018, pp. 1-5.

[41] M. Hawej and Y. R. Shayan, "Pilot decontamination in massive multi-user MIMO systems based on low-rank matrix approximation," *IET Commun.,* vol. 13, *(5),* pp. 594-600, 2018.

[42] M. Hawej and Y. R. Shayan, "Low-Complexity channel estimation for TDD massive MU-MIMO systems," *Submitted to IEEE Transaction on Vehicular Technology,* 2019.

[43] V. Sundarapandian, *Numerical Linear Algebra.* PHI Learning Pvt. Ltd, 2008.

[44] R. A. Horn, C. R. Johnson, *Matrix Analysis.* New York NY USA: Cambridge Univ. Press, 2012.

[45] B. N. Datta, *Numerical Linear Algebra and Applications.* SIAM, 2010.

[46] H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering.* 2011.

[47] C. R. Vogel, *Computational Methods for Inverse Problems.* 2002.

[48] W. C. Karl, *Regularization in Image Restoration and Reconstruction.* Handbook of Image and Video Processing, 2000.

[49] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York, NY, USA: Cambridge University Press., 2004.

[50] B. Jiang, Y. F. Liu, and Z. Wen, "Lp-norm regularization algorithms for optimization over permutation matrices," *SIAM Journal on Optimization,* vol. 26, *(4),* pp. 2284-2313, 2016.

[51] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-Norm Minimization problems when the solution may be sparse " *IEEE Trans. Inf. Theory,* vol. 54, *(11),* pp. 4789-4812, 2008.

[52] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Proces. Mag,* vol. 24, *(4),* pp. 118-121, 2007.

[53] S. D. Babacan, M. Luessi, R. Molina, and A. K. Katsaggelos, "Sparse Bayesian methods for low-rank matrix estimation," *IEEE Trans. Signal Process.,* vol. 60, *(8),* pp. 3964-3977, 2012.

[54] S. L. H. Nguyen, "Compressive Sensing for Multi-channel and Large-scale MIMO Networks," *Ph. D. Dissertation, Dept. Elect. and Comp. Eng. , Concordia Univ.,* 2013.

[55] N. K. Kumar and J. Schneider, "Literature survey on low rank approximation of matrices," *Linear and Multilinear Algebra,* vol. 65, *(11),* pp. 2212-2244, 2016.

[56] A. Parekh and I. Selesnick, "Enhanced low-rank matrix approximation," *IEEE Signal Process. Lett.,* vol. 23, *(4),* pp. 493-497, 2016.

[57] A. Parekh and I. Selesnick, "Improved sparse low-rank matrix estimation," *Digital Signal Processing,* vol. 139, pp. 62-69, 2017.

[58] N. S. M. Vanidevi, "Fast iterative WSVT algorithm in WNN minimization problem for multiuser massive MIMO channel estimation," *International Journal of Communication Systems,* vol. 31, *(1),* pp. 1-9, 2018.

[59] S. Wright, R. Nowak, and M. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.,* vol. 57, *(7),* pp. 2479-2493, 2009.

[60] C. R. Berger, Z. H. Wang, J. Z. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Comm. Magazine,* vol. 48, *(11),* pp. 164-174, 2010.

[61] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Review,* vol. 52, *(3),* pp. 471-501, 2010.

[62] K. Mohan and M. Fazel, "Reweighted nuclear norm minimization with application to system identification," in *American Control Conference (ACC),* Baltimore, MD, USA, 2010, pp. 2953-2959.

[63] Z. Lu, "Iterative reweighted minimization methods for L-p regularized unconstrained nonlinear programming," *Mathematical Programming,* vol. 147, *(2),* pp. 277-307, 2014.

[64] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),* Columbus, Ohio, 2014, pp. 2862-2869.

[65] T. F. Chan, "Rank revealing QR-Factorizations," *Linear Algebra and its Applications,* vol. 88, pp. 67-82, 1987.

[66] G. W. Stewart, *Matrix Algorithms, Vol 1: Basic Decompositions,* Siam, 1998.

[67] J-F. Cai, E. J. Cand`es, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization,* vol. 20, *(4),* pp. 1956-1982, 2010.

[68] F. Boccardi, R.W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.,* vol. 52, *(2),* pp. 74-80, 2014.

[69] T. Bai and R. W. Heath, "Analyzing uplink SINR and rate in massive MIMO systems using stochastic geometry," *IEEE Trans. Commun.,* vol. 64, *(11),* pp. 4592-4606, 2016.

[70] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun,* vol. 15, *(2),* pp. 1293-1308, 2016.

[71] H. Q. Ngo, M. Matthaiou, T. Q. Duong, and E. G. Larsson, "Uplink performance analysis of multiuser MU-SIMO systems with ZF receivers," *IEEE Trans. Veh. Technol.,* vol. 62, *(9),* pp. 4471-4483, 2013.

[72] V. Savaux, Y. Louët, and F. Bader, "Low-complexity approximations for LMMSE channel estimation in OFDM/OQAM," in *23rd International Conference on Telecommunications (ICT),* Thessaloniki, Greece, 2016, pp. 1-5.

[73] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Channel estimation and linear precoding in multiuser multiple-antenna TDD systems," *IEEE Trans. Veh. Technol.,* vol. 60, *(5),* pp. 2102-2116, 2011.

[74] A. Puglielli, N. Narevsky, P. Lu, et al., "A scalable massive MIMO array architecture based on common modules," in *IEEE International Conference on Communication Workshop (ICCW),* London, UK, 2015, pp. 1310-1315.

[75] J. Hoydis, S. Ten Brink, M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.,* vol. 31, *(2),* pp. 160-171, 2013.

[76] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.,* vol. 8, *(5),* pp. 815-830, 2014.

[77] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM Channel Estimation by Singular Value Decomposition," *IEEE Trans. Commun.,* vol. 46, *(7),* pp. 931-939, 1998.

[78] P. C. Hansen, and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Scientific Computing,* vol. 14, *(6),* pp. 1497-1503, 1993.

[79] Q. Leming and R. Partha, "Determination of regularization parameter using L-curve by the LARS-LASSO algorithm," in *Proceedings of the 2006 Joint Statistical Meetings,* Seattle, WA, USA, 2006, pp. 1-5.

[80] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive MIMO linear precoding: A survey," *IEEE Systems Journal,* vol. 12, *(4),* pp. 3920-3931, 2018.

[81] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Science,* vol. 2, *(1),* pp. 183-202, 2009.

[82] S. Ramani, Z. Liu, J. Rosen, J. F. Nielsen, and J. A. Fessler, "Regularization Parameter Selection for Nonlinear Iterative Image Restoration and MRI Reconstruction Using GCV and SURE-Based Methods," *IEEE Transactions on Image Processing,* vol. 21, *(8),* pp. 3659-3672, 2012.

[83] M. Rudelson, and R. Vershynin, "Smallest singular value of a random rectangular matrix," *Communications on Pure and Applied Mathematics,* vol. 62, *(15),* pp. 141-155, 2009.

[84] H. Huh, S. H. Moon, Y. T. Kim, I. Lee, and G. Caire, "Multi-cell MIMO downlink with cell cooperation and fair scheduling: A large system limit analysis," *IEEE Trans. Inf. Theory,* vol. 57, *(12),* pp. 7771-7786, 2011.

[85] Y. Zhang, D. Wang, J. Wang, and X. You, "Channel estimation for massive MIMO-OFDM systems by tracking the joint angle-delay subspace," *IEEE Access,* vol. 4, pp. 10166-10179, 2016.