Random Numerical Linear Precoding and Channel Estimation in Massive MIMO Systems

By

Emmanuel Wanyama Mukubwa

Thesis presented in fulfilment of the requirements for the degree of Doctor of Engineering in Electronic Engineering in the Faculty of Engineering and the Built Environment at Durban University of Technology

April 2021

Abstract

The information growth we have experienced in the immediate past and which continues to increase has consequently brought about the big data era and when pooled with the vast increase in subscriber numbers has led to an ever-escalating demand for more efficient and high-capacity communication systems. The affinity for higher capacity and efficient networks has necessitated the initiation of wireless fifth generation (5G) networks. Among the key technologies underlying the wireless 5G network are massive Multiple-Input Multiple-Output (MIMO) and Cloud Radio Access Network (C-RAN) which enhances spectral efficiency, energy efficiency, security and robustness but suffers from pilot contamination and fronthaul finite capacity.

There have been several attempts to minimize pilot contamination in massive MIMO system through linear precoding. But for those precoding schemes with good performance, they suffer from intricate problem of matrix inversion owing to large antenna numbers inherent in massive MIMO system, yet they do not render themselves readily to hardware parallelization. Also, channel state information estimation remains a challenge within massive MIMO networks. While the finite fronthaul capacity remains a bottleneck in C-RAN network systems. This study presents the formulation of iterative linear precoder that is efficiently parallelizable with efficient channel estimators for massive MIMO and massive MIMO partially centralised C-RAN networks.

The channel precoder was formulated and adapted using the iterative linear Rapid Numerical Algorithm (RNA). This model was then extended to include coordination among multicell massive MIMO system with receive combining computational complexity and efficiency evaluation. RNA model is again used to formulate improved linear and semi-blind channel estimators for massive MIMO systems in combination with the Fast Data Projection Method (FDPM). The semi-blind channel estimator is combined with compressed data channel estimator then extended based on Givens transformations and Data Projection Method (DPM) for massive MIMO partially centralised C-RAN networks. And finally, the estimation of the signal-to-interference-to-noise ratio, bit error rates, spectral efficiency, energy efficiency and normalised mean square error for the respective modelled components was realized. The models above were simulated using MATLAB for the analysis and validation.

The TDD downlink massive MIMO system was considered with varying immediate channel state information qualities for the single cell and multicell systems. For single cell system, there was optimal performance with regard to the signal-to-interference-to-noise ratio and the bit error rate when rapid numerical algorithm was used to implement the matrix inversion process in comparison to existing methods. It also rendered the precoding process highly parallelizable further reducing the complexity. For instance, for base transceiver station with 128 antennas serving 32 user terminals at signal-to-interference-to-noise ratio = 20 the average per user terminal rate was: RNA = 5 bit/sec/Hz, Regularized Zero Forcing (RZF) = 5 bit/sec/Hz and Truncated polynomial Expansion (TPE at $J = 2$) = 2.9 bit/sec/Hz. For the case of the Bit Error Rate (BER), for base transceiver station with 128 antennas serving 32 user terminals at signal-to-interference-to-noise ratio = 10 the BER was: RNA = 1, Regularized Zero Forcing (RZF) = 1 and TPE ($J = 2$) = 5. For the multicell massive MIMO, it was found that the performance of rapid numerical algorithm implementation gave a good spectral efficiency and energy efficiency performance in comparison to existing methods while lowering the complexity further through parallelization. The compressed data channel estimator gave comparable performance for the spectral efficiency and normalized mean square error when compared to the improved linear channel estimators. The semi-blind channel estimators for both massive MIMO and massive MIMO partially centralised C-RAN outperformed the linear channel estimators as well as the compressed data channel estimator.

These results demonstrate that rapid numerical algorithm can effectively eliminate the intricate matrix inversion associated with linear precoding while rendering itself to efficient parallelization. It also shows that the compressed data channel estimator optimally estimates the channel covariance matrix while reducing the amount of channel state information transmitted in estimation process. The semi-blind channel estimators have the optimal performance with regard to the normalised mean square error. It was also illustrated that the Givens transformation based semi-blind estimator outperforms the FDPM based semi-blind channel estimator.

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Durban University of Technology will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Submitted by:

17/08/2020

....

E. W. Mukubwa Date

Student Number: 21650054

Approved for Final Submission by:

....

Supervisor: Dr. A. O. Sokoya Date

Acknowledgements

I would like to express my gratitude to the following people and organisations for their contribution towards this research:

- To God be the glory for His unending provision
- My first words of gratitude to my academic supervisor Dr. Oludare A. Sokoya for his comprehensive and invaluable guidance in completing this work.
- Durban University of Technology and the Department of Electronics and Computer Engineering for the financial support during my Doctoral studies without which I could not have managed to carry out this research.
- To all those of you who could have contributed in one way or another in making this research a success I register my sincere appreciations.

"*An academic dialect is perfected when its terms are hard to understand and refer only to one another*" Mason Cooley.

Dedication

To Gustave Maghanu, Xanana Mwiruni, Taffi Naliaka and Zaneta Khalayi your love and trust in me is the springboard of my unwavering strength.

In memory of my late daughter Makinia

Table of Contents

Abstract i

Declaration iii

Acknowledgements iv

Dedication v

Table of Contents vi

List of General Notations xi

List of Acronyms xiii

List of Figures xvi

List of Tables xix

CHAPTER 1 Introduction 1

1.1 Motivation and Background to the Study 1

1.2 Problem Formulation 2

1.3 Research Objectives of the Study 3

1.3.1 Scope of the Study 4

1.4 Organization of the Thesis 4

1.5 Contribution of this Study 5

1.6 Publications 5

1.6.1 Journal Articles 5

1.6.2 Conference Papers 6

CHAPTER 2 Theory of Massive Multiple-Input Multiple-Output Systems 7

2.1 Introduction 7

2.2 Channel Hardening 9

2.3 Pilot Contamination 9

2.3.1 Non-orthogonal Pilot Schemes 10

2.3.2 Hardware Impairment 11
2.3.3 Non-Reciprocal Transceivers 12
2.4 Antenna Arrays 12
2.5 Massive MIMO Channel Measurements 15
2.6 Massive MIMO Channel Characterization 16
2.6.1 Correlation-Based Stochastic Models (CBSMs) 17
2.6.1.1 i.i.d. Rayleigh Channel Model 17
2.6.1.2 Kronecker Based Stochastic Model (KBSM) 17
2.6.1.3 Weichselberger Model 18
2.6.1.4 Virtual Channel Representation (VCR) Model 18
2.6.2 Geometrical-Based Stochastic Models (GBSMs) 19
2.6.2.1 2D GBSM 19
2.6.2.2 3D GBSM 19
2.6.3 General Massive MIMO Channel Model 21
2.7 Massive MIMO Physical Layer Security 21
2.8 Massive MIMO Precoding 23
2.8.1 Linear Precoding 24
2.8.2 Non-Linear Precoding 26
2.8.3 Hybrid Precoding 28
2.8.4 Simulation and Analysis of Selected Precoding Models 29
2.9 Massive MIMO Channel Estimation 36
2.9.1 Training Based Estimation 36
2.9.2 Blind Estimations 38
2.9.3 Semi-Blind Estimations 39
2.9.4 Simulation and Analysis of Selected Channel Estimation Models 40
2.10 Massive MIMO Partially Centralized Cloud Radio Access Network (MPC-RAN) 42
2.11 Summary 44

CHAPTER 3 Rapid Numerical Algorithms Precoding: Single Cell 46
3.1 Introduction 46
3.2 System Model Single Cell 47
3.2.1 Conventional RZF Precoding 48
3.2.2 TPE Precoding 49
3.2.3 Rapid Numerical Algorithms (RNA) Precoding 51
3.2.3.1 RNA Inverse Formulation 51
3.2.3.2 Initial RNA Inverse Formulation 55
3.2.3.3 Criterion to Stop RNA Iterative Procedure 57
3.3 Complex Operations Analysis 58
3.4 Performance Analysis 59
3.4.1 Achievable Rate 60
3.4.2 Bit Error Rate 60
3.5 Simulation Results and Discussion 61
3.6 Conclusion Single Cell Precoding 65
CHAPTER 4 Rapid Numerical Algorithms Precoding: Multicell 66
4.1 Introduction 66
4.2 System Model for Multicell 67
4.2.1 Model of Imperfect Covariance Channel Information 69
4.3 Multicell Precoding 70
4.3.1 RZF Precoding 70
4.3.2 TPE Precoding 70
4.3.3 RNA Precoding 71
4.4 Multicell Linear Precoding with Coordination 74
4.5 Energy Efficiency 78
4.5.1 Receive Combining Computational Complexity 81
4.5.2 Receive Combining Power Computation 83

4.5.3 Throughput and Energy Efficiency 83
4.6 Simulation Results and Discussion 84
4.7 Conclusion Multicell Precoding 91
CHAPTER 5 Channel Estimation and Analysis 93
5.1 Introduction 93
5.2 System Model for Channel Estimation 95
5.3 Improved MMSE and Semi-Blind Channel Estimation 95
5.3.1 Conventional MMSE Channel Estimation 95
5.3.2 Improved MMSE Channel Estimation 96
5.3.2.1 SMW-Based MMSE Estimation 96
5.3.2.2 RNA-Based MMSE Estimation 98
5.3.2.3 Spectral Efficiency 101
5.3.3 Semi-Blind FDPM-based channel estimation 102
5.3.4 Simulation Results and Discussion 104
5.4 Channel Estimation in MPC-RAN 109
5.4.1 System Model 110
5.4.2 Compressed Data CSI Model 110
5.4.2.1 Approximation of $\phi j, k$ 111
5.4.2.2 Approximation of Rj, k 113
5.4.2.3 Channel Estimate Approximation 114
5.4.2.4 Spectral Efficiency Estimate Approximation 114
5.4.2.5 Simulation Results and Discussion 115
5.4.3 GDPM Based Semi-Blind Model for Channel Estimation 118
5.4.3.1 Serial Givens Data Projection Method (SGDPM) 120
5.4.3.2 Parallel Givens Data Projection Method (PGDPM) 121
5.4.3.3 Simulation Results and Discussion 124
5.5 Conclusion on Channel Estimation 128

CHAPTER 6 Conclusion and Recommendations 130
6.1 Introduction 130
6.2 Conclusions 130
6.3 Recommendations 131
Bibliography 133
Appendix A 157
Appendix A.1: Proof of Theorem 3.1 157
Appendix A.2: Proof of Theorem 3.2 158
Appendix A.3: Initial Inverse Formulation Proof 158
Appendix A.4: Strassen's Algorithm 159
Appendix B 161
Appendix B.1: Matrix Analysis 161
Appendix C 163
Appendix C.1: Parallel Computation of Covariance Matrix 163
Appendix C.2: Computational Complexity 164

List of General Notations

Linear algebra

x	scalar
$\boldsymbol{x}$	vector
$\boldsymbol{X}$	matrix
$diag\{x_1, x_2, \dots, x_K\}$	identity matrix of size $K \times K$
$diag\{\boldsymbol{X}\}$	diagonal matrix with entries $x_1, x_2, \dots, x_K$
$[\boldsymbol{X}]_{i,j},\ \boldsymbol{X}_{i,j}$	$(i,j)th$ entry of matrix $\boldsymbol{X}$
$\boldsymbol{x}_i$	ith entry of vector $\boldsymbol{x}$
$\boldsymbol{X}^T$	transpose of $\boldsymbol{X}$
$\boldsymbol{X}^H$	Hermitian of $\boldsymbol{X}$
$\boldsymbol{X}^*$	complex conjugate of $\boldsymbol{X}$
$tr(\boldsymbol{X})$	trace of $\boldsymbol{X}$
$det(\boldsymbol{X})$	determinant of $\boldsymbol{X}$
$\|\boldsymbol{X}\|_2$	Spectral norm of matrix $\boldsymbol{X}$
$\|\boldsymbol{X}\|_F$	Frobenius norm of matrix $\boldsymbol{X}$
$\|\boldsymbol{x}\|_q$	L_q norm of vector $\boldsymbol{x}$
$\mathbb{D}(x)$	square diagonal matrix with the main diagonal having the elements of $\mathbf{X}$
$\mathbb{D}(\boldsymbol{X})$	square diagonal matrix with its main diagonal having only the diagonal elements of $\boldsymbol{X}$

Analysis

$\mathbb{R}, \mathbb{C},$	the real and complex numbers
$\mathbb{C}^{M\times K}$	set of matrices with size $M \times K$
$\mathbb{C}^{M\times 1}$	set of vectors with size M
$\|x\|$	absolute
$\Re\{x\}$	real part of x
$\Im\{x\}$	imaginary part of x
$\mathcal{O}(K)$	big-O notation

Probability related

X	scalar random variable
μ	general measure
$\mathbb{E}[X]$	expectation of X
$var[X]$	variance of X
$\xrightarrow{a.s.}$	almost sure convergence
$\sim$	distributed as for example $X \sim \mathcal{CN}(0,1)$
$\mathcal{CN}(k, \Phi)$	complex Gaussian distribution with mean k and covariance Φ

List of Acronyms

5G	fifth generation
2D	two dimensional
3D	three dimensional
3GPP	3rd generation partnership project
AOA	angle of arrival
AOD	angle of departure
ARCoSaMP	adaptive regularized compressive sampling matching pursuit
AWGN	additive white Gaussian noise
BBU	baseband processing unit
BCS	Bayesian compressed sensing
BER	bit error rate
BF	beamforming
BOOMP	block optimized orthogonal matching pursuit
BS	base station
BSs	base stations
CBSMS	correlation-based stochastic models
COST	cooperation in science and technology
CP	circuit power
C-RAN	cloud radio access network
CS	compressed sensing
CSI	channel state information
CSM	Cholesky and Sherman-Morrison
DFT	discrete Fourier transform
DL	downlink
DPC	dirty paper coding
DPM	data projection method
EE	energy efficiency
EM	expectation-maximization
ETP	effective transmit power
EVD	Eigenvalue decomposition

EW-MMSE	Element wise minimum mean square error
FDD	frequency division duplex
FDPM	fast data projection method
FSCAPI	fast single compensation approximated power iteration
GBSMS	geometric-based stochastic models
GDPM	Givens data projection method
HL-THP	hybrid-linear Tomlinson-Harashima precoding
ICA	independent component analysis
i.i.d	independent and identically distributed
KBSM	kronecker based stochastic model
LMMSE	linear minimum mean square error
LOS	line of sight
LS	least square
L-BFGS	limited-memory Broyden-Fletcher-Goldfarb-Shanno
MAP	maximum a-posteriori
METIS	mobile and wireless communications Enablers for the Twenty-twenty Information Society
MF	matched filter
MIMO	multiple-input multiple-output
MMSE	minimum mean square error
mmWave	millimetre wave
MPCs	multipath components
MPC-RAN	massive MIMO partially centralized cloud radio access network
MSE	mean square error
MRT	maximum ration transmission
NLOS	non-line of sight
NMSE	normalized mean square error
OPLS	orthogonal projection-based least square
PC	power consumption
PEACH	polynomial ExpAnsion channel
PGDM	parallel Givens data projection method
PHY	physical
PZF	phased zero forcing

QAM	quadrature amplitude modulation
RF	radio frequency
RNA	rapid numerical algorithms
RRH	remote radio heads
RZF	regularized zero forcing
SABMP	Support agnostic Bayesian matching pursuit
SDMA	space division multiple access
SE	spectral efficiency
SGDPM	serial Givens data projection method
SINR	signal-to-interference-and-noise ratio
SMW	Sherman-Morrison-Woodbury
SNR	signal-to-noise ratio
SSCE	structured sparse channel estimation
SVD	single value decomposition
TDD	time division duplex
THP	Tomlinson-Harashima precoding
TPE	truncated polynomial estimation
UC	unique combination
UL	uplink
UT	user terminal
UTs	user terminals
VCR	virtual channel representation
VP	vector perturbation
WINNER	wireless world initiative new radio
WR	whitening rotation
ZF	zero forcing

List of Figures

Figure 1.1: Pilot contamination [11], [12]. 2
Figure 1.2: Transmission of beam-formed DL signal [8]. 3
Figure 2.1: TDD mode transmission cycle [19]. 8
Figure 2.2: Pilot contamination in massive MIMO system with multicell [8], [9]. 10
Figure 2.3: Antenna array configuration [27], [39]. 14
Figure 2.4: Massive MIMO with distributed jamming [83]. 22
Figure 2.5: Massive MIMO downlink system model [101]. 24
Figure 2.6: The system model. 25
Figure 2.7: System model for conventional THP [113]. 27
Figure 2.8: Block diagram representation of a generically Vector Perturbation system [114]. 28
Figure 2.9: Average per UT rate vs. SINR for different orders of J in TPE precoding ($M = 512, K = 128,\ \tau = 0.15$). 32
Figure 2.10: Average per UT rate vs. SINR for varying CSI errors at the BS ($M = 512, K = 128,\ J = 4$). 32
Figure 2.11: Average per UT rate vs. SINR for different orders of J in TPE precoding ($M = 512, K = 64,\ \tau = 0.15$). 33
Figure 2.12: Average per UT rate vs. SINR for varying CSI errors at the BS ($M = 512, K = 64,\ J = 4$). 34
Figure 2.13: BER vs. SINR for different orders of J in TPE precoding ($M = 512, K = 128, \tau = 0.15$). 35
Figure 2.14: BER vs. SINR for different orders of J in TPE precoding ($M = 512, K = 64,\ \tau = 0.15$). 35
Figure 2.15: The NMSE corresponding to approximation of spatially correlated channel with 41
Figure 2.16: Average NMSE vs. total training power at 24 transmitting antennas and 4 receiving antennas. 41
Figure 2.17: Comparison of NMSE vs. number of transmitting antennas. 42
Figure 2.18: Cloud radio access network [161]. 43

Figure 2.19: MPC-RAN architecture [166]. 43
Figure 3.1: Average per UT rate vs. SNR for different orders of J in TPE precoding ($M = 128, K = 32,\ \tau = 0.15$). 62
Figure 3.2: Average per UT rate vs. SINR for varying CSI errors at the BS ($M = 128, K = 32, J = 4$). 62
Figure 3.3: Average per UT rate vs. SINR for different orders of J in TPE precoding ($M = 128, K = 16,\ \tau = 0.15$). 63
Figure 3.4: Average per UT rate vs. SINR for varying CSI errors at the BS ($M = 128, K = 16, J = 4$). 63
Figure 3.5: BER vs. SINR for different orders of J in TPE precoding ($M = 128, K = 32,\ \tau = 0.15$). 64
Figure 3.6: BER vs. SINR for different orders of J in TPE precoding ($M = 128, K = 16,\ \tau = 0.15$). 64
Figure 4.1: Achievable SE per cell vs. number of BS antennas with a reuse factor of f = 1.. 85
Figure 4.2: Achievable SE per cell vs. number of BS antennas with a reuse factor of f = 2... 86
Figure 4.3: Achievable SE per cell vs. number of BS antennas with a reuse factor of f = 4... 86
Figure 4.4: Achievable SE per cell vs. number of BS antennas with coordination and a reuse factor of 1. 87
Figure 4.5: Achievable SE per cell vs. number of BS antennas with coordination and a reuse factor of 4. 88
Figure 4.6: Average per UT rate vs. SINR for CSI error of $\tau = 0.45$ at the BS ($M = 128, K = 6,\ J = 3$). 88
Figure 4.7: Total CP vs. number of BS antennas. 90
Figure 4.8: EE per cell vs. Capacity per cell for a reuse factor of 1. 90
Figure 4.9: EE per cell vs. Capacity per cell for a reuse factor of 4. 91
Figure 5.1: Achievable NMSE vs. number of BS antennas for conventional MMSE estimator ($f = 1, 2, 4$). 105
Figure 5.2: Achievable NMSE vs. number of BS antennas for SMW-based MMSE estimator ($f = 1, 2, 4$). 106
Figure 5.3: Achievable NMSE vs. number of BS antennas for RNA-based MMSE estimator ($f = 1, 2, 4$). 106

Figure 5.4: Achievable NMSE vs. number of BS antennas for FDPM-based semi-blind estimator ($f = 1, 2, 4$). 107
Figure 5.5: Comparison of achievable NMSE vs. number of BS antennas for all the estimators ($f = 1$). 108
Figure 5.6: Comparison of achievable NMSE vs. number of BS antennas for all the estimators ($f = 4$). 109
Figure 5.7: Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 1. 116
Figure 5.8: Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 4. 116
Figure 5.9: The normalized MSE vs. number of RRH antennas with a reuse factor of 1. 117
Figure 5.10: The normalized MSE vs. number of RRH antennas with a reuse factor of 4. .. 117
Figure 5.11: The column-wise Givens rotation and general Givens rotation. 123
Figure 5.12: Achievable NMSE vs. number of RRH antennas for PGDPM-based semi-blind estimator ($f = 1, 2, 4$). 124
Figure 5.13: Comparison of achievable NMSE vs. number of RRH antennas for RNA, FDPM and PGDPM estimators ($f = 1$). 125
Figure 5.14: Comparison of achievable NMSE vs. number of RRH antennas for RNA, FDPM and PGDPM estimators ($f = 4$). 125
Figure 5.15: The normalized MSE vs. number of RRH antennas with a reuse factor of 1. .. 127
Figure 5.16: The normalized MSE vs. number of RRH antennas with a reuse factor of 2. .. 127
Figure 5.17: The normalized MSE vs. number of RRH antennas with a reuse factor of 4. .. 128

List of Tables

Table 2.1: Massive MIMO channel models. 17

Table 4.1: Average down link sum SE for M = 128 and K = 6 for diverse pilot reuse factors f. The highest value for each scheme is in bold face. The values are based on Figures 1, 2 and 3. 87

Table 4.2: CP model parameters. 89

Table 4.3: Cost parameters. 89

Table 5.1: FDPM sub-space tracking algorithm. 103

Table 5.2: Performance of conventional MMSE estimator. 107

Table 5.3: Performance of RNA estimator. 107

Table 5.4: Performance of SMW estimator. 107

Table 5.5: Performance of FDPM estimator. 108

Table 5.6: SGDPM sub-space tracking algorithm. 120

Table 5.7: SGDPM sub-space tracking algorithm parameters. 120

Table 5.8: PGDPM sub-space tracking algorithm. 123

Table 5.9: Performance of PGDPM estimator. 124

CHAPTER 1

Introduction

1.1 Motivation and Background to the Study

The data blast witnessed in the recent past brought about the Big Data era which, coupled with the immense rise in subscriber numbers, resulted in a growing demand for communication networks [1]. This has necessitated the advent of wireless fifth generation (5G) networks to enhance capacity in communication networks. Among the key technologies underlying wireless 5G network is massive Multiple-Input Multiple-Output (MIMO) which enhances spectral efficiency, energy efficiency, security and robustness but suffers from pilot contamination [2], [3].

Massive MIMO is the technological framework that scales up MIMO in the order of hundreds of antenna arrays to support tens of User Terminals (UTs) simultaneously in the same time-frequency resources [4]. For massive MIMO networks an orthogonal pilot signal (training sequence) is preferably allocated to each User Terminal (UT). The maximum possible number of potential orthogonal pilot signal that can exist is limited by the coherent time ratio. Therefore, the usable supply of such orthogonal pilot signals is quickly depleted in the multicell systems.

To simultaneously cater for many UTs, the limited orthogonal pilot signals must be reused within neighboring cells at a specified reuse factor. Hence, we have pilot contamination as a result of interference that emanates from this reuse of pilots. The interference brings in a non-orthogonal pilot information to the present user. Thus, the propagated pilot signal is bound to be impaired by the known and unknown interferences. The intra- and inter-cell interference are the common channel impairments [5]–[7].

For a given instance in time the Base Station (BS) learns the channel response of its corresponding by estimation based on pilot signals received during Uplink (UL) transmission. But this pilot signals suffers from noise and inter-cell interference in the process of being received by the BS. As illustrated in Figure 1.1, three UTs are simultaneously transmitting culminating to the signal received by the BS being a superposition of their corresponding

transmitted signals. Consequently, resulting in contamination of the desired pilot signal [8]–[10].

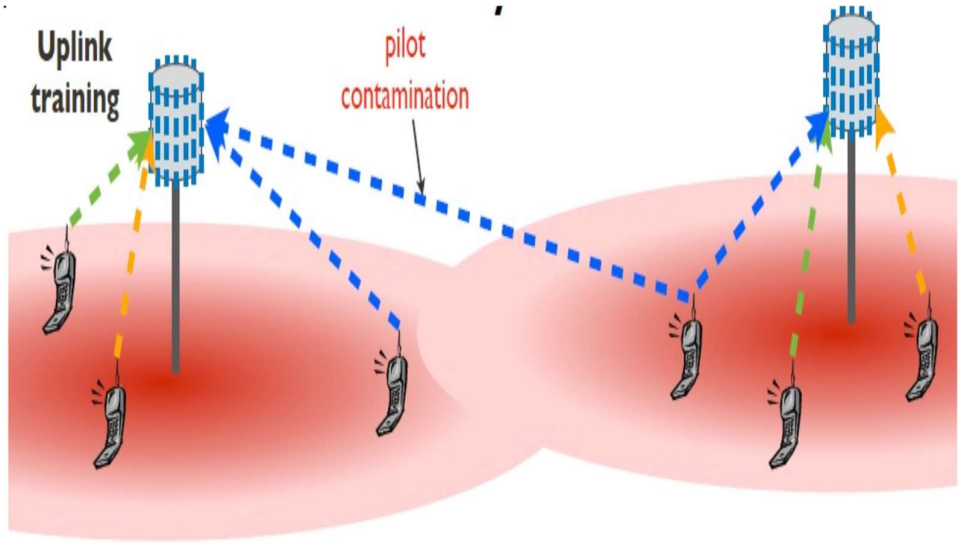

Figure 1.1: Pilot contamination [11], [12].

1.2 Problem Formulation

During channel approximation for the corresponding UT, it is a difficult task for the BS to distinctively isolate the received signals from the two UTs under consideration. Consequently, the signal causing interference mimics a coloured noise degrading the estimation accuracy of the channel. Also, the BS is forced to approximate the channel as a superposition of the desired UT and undesired UT pilot signals [8], [9]. Afterwards payload data is transmitted by the desired UT where the BS coherently combines the received signal with the approximated channel, but a portion of the undesired signal is involuntarily and coherently involved in the combining process [8]–[10]. Putting in focus the large number (M) of transmitting antennas associated with massive MIMO [13], this condition is exacerbated as the receive combining array gain enhances both the desired and undesired signal power in proportion to M. During Downlink (DL) transmission the BS will involuntarily transmit the beam-formed DL signal to the desired UT and undesired UT as exemplified in Figure 1.2 [8].

It is worth noting that pilot contamination does not only occur when the interfering UT transmits the same pilot signal as the desired UT, but wherever a non-orthogonal interfering signal is send by any UT it will result in pilot contamination [8]. It is evident that the UL and DL performance of massive MIMO is degraded by pilot contamination. This has occasioned a number of techniques being proposed to minimize the effect of pilot contamination [14].

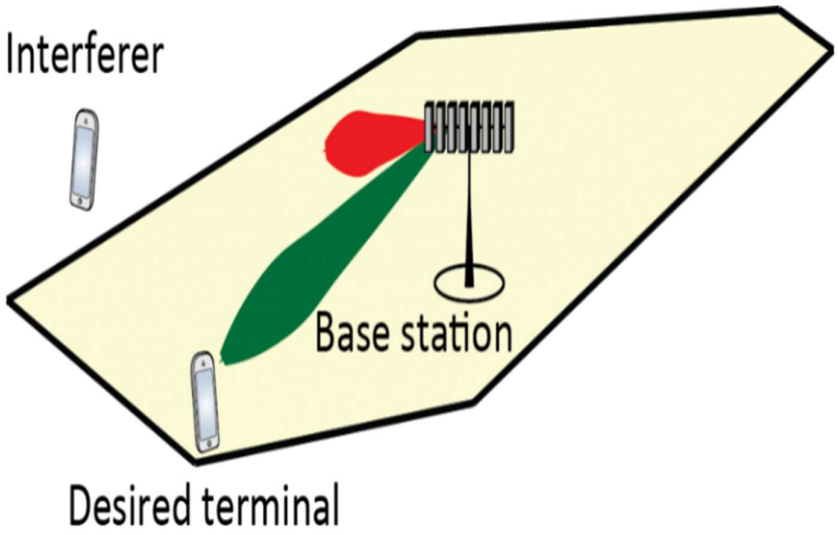

Figure 1.2: Transmission of beam-formed DL signal [8].

The massive MIMO encoding techniques play a significant role in mitigating pilot contamination. Among the encoding techniques, precoding which requires the Channel State Information (CSI) at the transmitter has come out as a probable encoding technique for massive MIMO. The existing linear precoding methods suffers from complex matrix inversion and do not render themselves to efficient parallelization to enhance and fasten the precoding process. Another component of interest is the channel estimation techniques. We need channel estimates that are of good quality and this relies on the techniques we employ. The available channel estimation techniques either suffers from matrix inversion complexity or they do not render themselves to efficient parallelization to hasten the channel estimation process or compression of channel estimation data to mitigate fronthaul finite capacity problem [9].

Therefore, this research seeks to develop a linear precoding method that will resolve the matrix inversion problem and with efficient parallelization capability for massive MIMO system. We also seek to formulate a semi-blind channel estimation model to enhance and hasten the channel estimation process. These models for precoding and channel estimation will seek to achieve pilot contamination mitigation, enhance energy and Spectral Efficiency (SE) of massive MIMO system.

1.3 Research Objectives of the Study

The objectives of this study are:

a. Formulate the iterative linear precoding method for a single cell in massive MIMO system.

b. Formulate the stopping criterion of the iterative linear precoder.
c. Formulate the initial value of the iterative linear precoder.
d. Perform complex analysis of the iterative linear precoder.
e. Introduce coordination between the BSs to iterative linear precoder for multicell massive MIMO system and model the receive combining computational complexity.
f. Formulate an improved Minimum Mean Square Error (MMSE) channel estimator and combine it with a less complex sub-space estimator to realize a semi-blind channel estimator.
g. Consider massive MIMO-Cloud Radio Access Network (C-RAN) and adapt a compressive data channel estimator. Then model a sub-space channel estimator with efficient parallelization and combine it with the adapted linear compressive data channel estimator to realize a semi-blind channel estimator.
h. Validate the models formulated in **a. – g**.

1.3.1 Scope of the Study

The scope of this work focusses on developing iterative linear precoder for massive MIMO, then adapt it to develop a robust semi-blind channel estimator using data compression technique. This work will cover the random numerical algorithms, semi-blind techniques, compression techniques and parallelization techniques.

1.4 Organization of the Thesis

This section gives a brief overview of how the thesis is organized. Each of the chapters contained in this thesis provides an insight into contents, the work realized and how the objectives were accomplished in this research. To start with, **Chapter 1** serves as the introduction of the thesis in general and presents: motivation and background of the study, problem formulation, objectives of the study, the scope of the study, contribution of this study to the body of existing knowledge and the publications. The rest of the chapters are outlined as follows:

Chapter 2 gives a brief background of massive MIMO and the corresponding important aspects as well as the enabling technologies. It addresses the channel hardening, pilot contamination, massive MIMO channel aspects and ends with extension of massive MIMO to C-RAN. **Chapter 3** discusses rapid numerical algorithm for the single cell massive MIMO system and parallelization and compares it to existing precoding technologies. **Chapter 4**

discusses the rapid numerical algorithm for multicell massive MIMO system, mathematical formulation, and the methods to use in analysing the energy efficiency. **Chapter 5** describes an implementation of channel in massive MIMO-C-RAN estimation based on compressed data, formulate semi-blind channel estimation model, and compares it existing methods. **Chapter 6** presents conclusions and the recommendations concerning this study.

1.5 Contribution of this Study

The contribution of this study to the knowledge in massive MIMO field is as enumerated below:

a. We model the rapid numerical algorithm to circumvent the matrix inversion and make the precoding algorithm render itself to high parallelization at hardware implementation level. This achieved by vectorising the matrix inversion process and reducing it to a simple multiplication problem.
b. Adaptation of rapid numerical algorithm to multicell massive MIMO with coordination and development of receive combining complexity to evaluate energy efficient of the model against the conventional models.
c. Formulation of a semi-blind channel estimation model from hybridization of improved MMSE and fast data projection method with improved performance.
d. Adaptation of compressed data channel estimator to the massive MIMO-C-RAN and formulating of a sub-space estimator that offers efficient parallelization then combine them together to form a semi-blind channel estimator for massive MIMO-cloud-RAN.
e. Contribute to the improvement of the quality of massive MIMO system and respective application network because of **a** - **d**.

1.6 Publications

The articles published in the course of this research are listed as follows.

1.6.1 Journal Articles

a. Emmanuel Mukubwa, Oludares Sokoya, “Efficient Channel Estimation in Massive MIMO Partially Centralized Cloud-Radio Access Network Systems,” International Journal of Embedded and Real-Time Communication Systems (IJERTCS). (Accepted on 23/06/2020).

b. Emmanuel Mukubwa, Oludares Sokoya. 2020, "Performance Analysis of Linear Precoders with Imperfect Channel Covariance Information for Multicell System," IET Journal of Engineering, vol. 2020, issue 8, Pg. 697 – 705.
c. Emmanuel Mukubwa, Oludares Sokoya. 2019, "Efficient and low-complexity matrix inversion scheme for massive MIMO systems using rapid numerical algorithms," IET Journal of Engineering, vol. 2019, issue 10, Pg. 7244 – 7249.

1.6.2 Conference Papers

a. Emmanuel Mukubwa, Oludares Sokoya "Analysis of Channel Estimation Performance in MPC-RAN: Improved MMSE and Compressed Data Techniques." 26th IEEE Conference of Open Innovations Association FRUCT, 24-26 April 2020, Yaroslavl, Russia.
b. Emmanuel Mukubwa, Oludares Sokoya "Comparison of Improved MMSE and the Semi-Blind Channel Estimation Methods." International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD2020), 6-7 August 2020, Durban, South Africa.
c. Emmanuel Mukubwa, Oludares Sokoya, Dimov Ilcev, "2018. Comparison and Analysis of Massive MIMO Linear Precoding Schemes in the Downlink," IEEE AFRICON 2017, 18-20 September 2017, Cape Town, South Africa.
d. Emmanuel Mukubwa**,** Oludares Sokoya, Dimov Ilcev, "2018. Performance Analysis and Comparison of Massive MIMO Precoding Schemes," SATNAC 2017 conference, 03-10/09/2017 September 2017, Barcelona, Spain.

CHAPTER 2

Theory of Massive Multiple-Input Multiple-Output Systems

2.1 Introduction

The quantity of "conversations" (voice or information) that can hypothetically be conducted in each region in all valuable radio spectrum has increased tremendously over the years. This has resulted in an ever-increasing demand for higher data rates, which is expected to continue in the near future [15]. While the regular users are keeping on grasping current remote applications, for example, social networking, mobile and internet money services, and video streaming, this exponential increment in wireless information traffic compels the telecommunication service providers to constantly develop their current system hardware and put in new system infrastructure [16]. To this end, massive MIMO has emerged as a staple of modern wireless cuisine.

In massive MIMO systems, hundreds or more antennas are bundled on to one base station [17]. The enormous increment in the quantity of antennas presents new difficulties for transceiver design and realization, but additionally introduces attractive compensation in relation to communication and signal processing [18]. To go around the design and realization of massive MIMO transceivers, a number of antenna array realization has been suggested in [18]–[22]. For attractive compensation, if the quantity of antennas at the base station is much bigger than the quantity of users in the network, straightforward conjugate Beamforming (BF) precoding (DL) as well as Matched Filter (MF) recognition (UL) at the base station prompt close to ideal performance and consequently low complex signal processing at the base station and at the user terminals. Besides, as the number of antennas at the base station becomes vast the noise and the small-scale fading factors are averaged out [23], this called channel hardening.

To deal with signalling overhead for massive MIMO systems due to CSI acquirement, Time Division Duplex (TDD) (Figure 2.1) process is preferred, because in Frequency Division Duplex (FDD) operation the measure of CSI feedback increments with the quantity of antennas at the base station. The figure shows that phase 1 and phase 2 represents uplink transmission and involves the user terminals. During phase 1 every user terminal uses T_u time slots to

transmit to its respective base station. Then during phase 2, a precise training sequence with a length of τ is sent to the respective base station by the corresponding user terminal. Based on which channels are estimated by the base station. T_t provides the uplink transmission cycle duration with μ being the duration in T_t set aside for channel estimation. During phase 3, the channel estimate in phase 2 is then utilized by the base stations to send downlink data back to the user terminals in a duration of T_d time slots [24].

A noteworthy shortcoming in massive MIMO systems emanates from pilot sullying. Pilot sullying or contamination is inherent in massive MIMO systems due to reuse of identical pilot groupings in various cells. The reuse is inevitable since for a specified pilot sequence span, the quantity of linearly autonomous pilot sequences is restricted. However, various effective systems have been proposed to mitigate pilot contamination [25], [26].

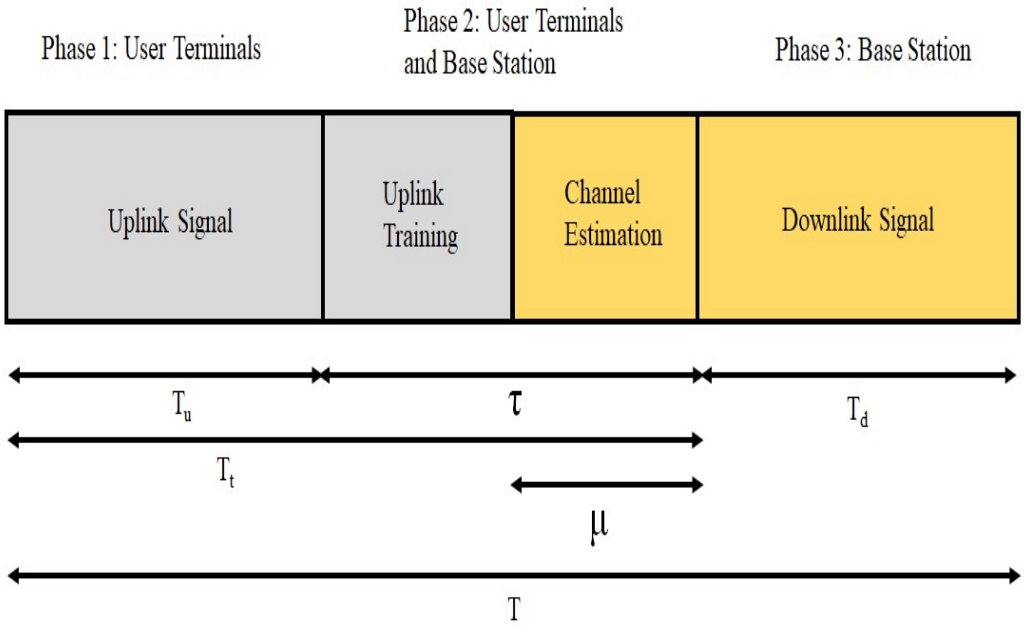

Figure 2.1: TDD mode transmission cycle [24].

Thus, this chapter reviews the Channel hardening in massive MIMO, Pilot contamination, antenna arrays, massive MIMO channel measurements, massive MIMO channel characterization, massive MIMO physical layer security, massive MIMO precoding, massive MIMO channel estimation, and massive MIMO partially centralized cloud radio access networks.

2.2 Channel Hardening

Small scale fading stands out among the impairments to wireless communication systems. Infinitesimal variations encountered by the wave in the propagation environment induces arbitrary variations within the channel gain leading to small scale fading. Consequently, the channel is rendered unreliable since in the instance when a very small channel gain occurs the received signal is in error. To mitigate small scale fading, multiple channels having independent realizations are used to transmit the signal with enhanced diversity. The use of multiple antennas at the BS or UT becomes very attractive since they offer spatial diversity. For the case of massive MIMO this spatial diversity leads to a phenomenon called channel hardening. It simply implies that a fading channel in the presence of channel hardening embodies a non-fading channel. The random fluctuations are still present on the channel however the effect on the transmitted signal is insignificant [27].

According to [27], the quotient of the supported instantaneous data rate and the averaged fading data rate gives a measure of channel hardening achieved. If the channel hardening has been realized, the relative fluctuations will be small. Channel hardening poses several practical implications. The first is the attaining of an almost deterministic channel that enhances reliability leading to a lower latency. The next is the absence of scheduling diversity; you cannot schedule users since the random fluctuations are small within the propagation environment. The estimation of the present realization of $\|\boldsymbol{h}\|^2$, with $\boldsymbol{h}$ being the channel description, does not avail much as it is close to its relative average value. This has great significance in massive MIMO as it alleviates the need for downlink pilots [28], [29]. Consequently, it allows the use of TDD in this work which allows us to acquire CSI in the UL and use it in the DL due to channel reciprocity.

2.3 Pilot Contamination

The main impediment to massive MIMO performance is pilot contamination that results in degraded channel estimation. Pilot contamination only occurs in TDD massive MIMO multi-cell systems during the uplink as a result of interference from the other users' pilot information from neighbouring cells, hardware impairment and non-reciprocal transceivers. The interference brings in a non-orthogonal pilot information to the present user [30], [31]. The Figure 2.2 depicts the pilot contamination in multicell massive MIMO system. The interfering pilots are shown by dotted lines while the intended pilots are shown by the solid lines.

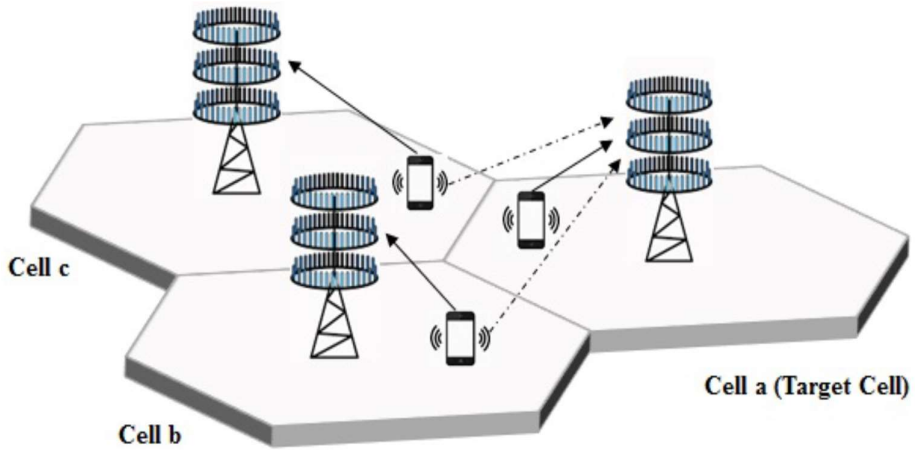

Figure 2.2: Pilot contamination in massive MIMO system with multicell [8], [9].

More specifically, when the service-array correlates its received pilot signal with pilot sequence associated with a particular terminal, it obtains a channel estimation that is contaminated by a linear combination of channels to the other terminals that share the same pilot sequence. Down-link beam forming based on the contaminated channel estimation results in interference that is directed to those terminals that share the same pilot sequence. Similar interference is associated with up-link transmissions of data. This directed interference grows with the number of service-antennas at the same rate as the desired signal [17]. The existing linear precoding methods suffers from complex matrix inversion and do not render themselves to efficient parallelization to enhance and fasten the precoding process. This then informed the first objective of this thesis. Next we review the various sources of pilot contamination.

2.3.1 Non-orthogonal Pilot Schemes

In [30] and [31] mutually orthogonal pilots was assumed for a multicell massive MIMO system with same shared frequency by all cells and neglected the intra-cell interference. For a frequency reuse factor of 1, inter-cell interference comes into play affecting the pilot signals and culminating to pilot contamination in the system from neighbouring cells. The system is an uplink multicell massive MIMO with $L > 1$ cells and K single antenna UTs in each cell that transmits signals to a BS in corresponding cell having M antennas at the same time. The propagation matrix, $\boldsymbol{G}$, has dimensions $M \times 1$ for each cell, [18], [32], [33]. The TDD protocols are synchronized across cells to simultaneously transmit pilot signals and data to all cells. The received complex baseband signal $\boldsymbol{y}_l \in \mathbb{C}^M$ at the lth BS in the jth cell was evaluated as

$$\boldsymbol{y}_l = \sqrt{P_u}\sum_{j=1}^{L}\boldsymbol{G}_{l,j}\boldsymbol{X}_j^H + \boldsymbol{N}_l \tag{2.1}$$

$\boldsymbol{N}_l \in \mathbb{C}^M$ represents the noise experienced at *l*th BS when pilots are transmitted and $\boldsymbol{X}_j \in \mathbb{C}^K$ is training vector transmitted by UTs in all cells. At the BS, the intended pilot signals are correlated with all the received unintended pilot signals as the UTs from neighbouring cells add to pilot contamination [17], [30]. The channel matrix as estimated at the *l*th BS is evaluated as

$$\widehat{\boldsymbol{G}}_{l,l} = \sqrt{P_u}\boldsymbol{G}_{l,l} + \sqrt{P_u}\sum_{j=1}^{L}\boldsymbol{G}_{l,j}\boldsymbol{X}_j^H\boldsymbol{X}_j^* + \boldsymbol{N}_l \tag{2.2}$$

with $\mathbf{X}_j^*$ representing a complex conjugate of $\mathbf{X}_j$ and P_u is the uplink average transmit power at each UT.

2.3.2 Hardware Impairment

The impact of massive MIMO impairment due to hardware is extensively covered in [34]–[36]. According to [34], impairments due to hardware results into an error between the theoretical transmit signal and the real generated signal and also a distorted signal at reception. These impairments can be categorized as phase noise, amplifier nonlinearity, quadrature imbalance and quantization errors. The accuracy of channel estimation is degraded by these impairments which might perpetually culminate to pilot contamination and hence poor performance of massive MIMO system. For every component, the non-ideal characteristics are modelled to come up with an overall compensation algorithm to mitigate hardware impairment. The data transmission and channel estimation based on pilots are performed on the downlink channel and the resultant ideal model for the received signal at the UT is

$$\boldsymbol{z} = \boldsymbol{h}^H\boldsymbol{x} + n \tag{2.3}$$

where $\boldsymbol{h}$ is channel matrix, $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$ is deterministic pilot signal or stochastic data signal with zero-mean. n is ergodic stochastic process $n = n_{noise} + n_{interf}$ with independent noise from receiver, n_{noise} and interference, n_{interf} from other UTs transmitting concurrently. Considering noise distortion, the model of non-ideal DL system for the hardware is expressed as [34]:

$$\boldsymbol{z} = \boldsymbol{h}^H(\boldsymbol{x} + \boldsymbol{\eta}_t^{BS}) + \boldsymbol{\eta}_r^{UT} + n \tag{2.4}$$

where $\boldsymbol{\eta}_t^{BS} \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{\eta}_t^{UT} \in \mathbb{C}^{M\times 1}$ are stochastic processes describing residual transceiver impairments of the hardware in the transmitter at BS and the hardware in the receiver at the UT, respectively. Similarly, the model of non-ideal UL system for the hardware is expressed as

$$\boldsymbol{z} = \boldsymbol{h}^H(\boldsymbol{s} + \boldsymbol{\eta}_t^{UT}) + \boldsymbol{\eta}_r^{BS} + n \tag{2.5}$$

where $\boldsymbol{s} \in \mathbb{C}^{M\times 1}$ is deterministic pilot signal or stochastic data signal with zero-mean.

2.3.3 Non-Reciprocal Transceivers

Forward and backward physical channels in TDD are set to function on identical carrier frequency rendering them reciprocal. Normally the power amplifiers, the low noise amplifiers, and the effective electromagnetic channel for a point-to-point TDD are taken to be identical. When exploiting channel reciprocity, residual offset frequency impact is great. Even with a small offset in terms of few hertz, it quickly accumulates and render both the downlink and uplink channels non-reciprocal within a limited time [37], [38].

2.4 Antenna Arrays

The performance of all types of MIMO systems strongly depends on the properties of the antenna arrays and of the propagation environment in which the system operates. The complexity of the propagation environment, in combination with the ability of antenna arrays to exploit this complexity, limits achievable system performance. As the number of antenna elements in networks increases, we encounter both opportunities and challenges [18]. Precoding of an antenna array is often said to direct the signal from the antenna array towards one or more receivers [18]. This then highlights the need to review the antenna array formation to be adopted in this research.

The configurations of antenna array may directly affect the channel properties, furthering the performance of massive MIMO system [32]. Realistic implementation of enormous number of antennas expected in massive MIMO, which are confined in specified dimensions of antenna array, brings about lessened inter-element spacing which negatively impacts system

performance due to spatial correlation [39]. The models of spatial correlation in wireless systems become crucial for precise hypothetical performance investigation.

The spatial correlation approximation for the uniform rectangular and cylindrical antenna arrays, using a 3D channel model are developed in [19]. It is shown that, for a huge spatial correlation as a result of small angle spreading negatively affects the convergence of properties of massive MIMO as the antennas increase in number. Moreover, the topology of antennas in massive MIMO has negligible impact. It is found that correlation has a detrimental effect when the channel is Independent and Identically Distributed (i.i.d.). To decrease correlation and consequently near the i.i.d. case by metric convergence, inter-element spacing in the antenna arrays is increased in relation with the angle spread. In Figure 2.3 below the various antenna arrays topologies are shown.

In [20] it was found that if you compact elements that are single-polarized with distances between them less than $\lambda/3$ brings no benefits, and in essence this realization is evidently inferior to elements with dual-polarization spaced $\lambda/2$ apart, because of their diminished mutual couplings and due to the diversity induced by polarization. This is the consequence of the fact that as the inter-element spacing reduces, the array can be effectively seen as a single antenna [18]. Although horizontal or level linear arrays give unsurpassed performances, 2D arrays can achieve a huge percentage of these performances with decreased horizontal dimensions although with an eminent increment on the aggregate number of radiating elements.

Massive MIMO arrays are likely to be realized in a 2D or 3D array structure, rather than a linear array given in [40]. When the antenna elements in the linear array have indistinguishable gain patterns, the array experiences the issue of front-back uncertainty and though has a superior angular resolution it is only in the azimuth [18], [32]. One disadvantage of having intense array realization using 2D or 3D lies in the inherent amplification of coupling effects because of the expansion in the quantity of adjacent antennas. In 3D arrays, just the antennas situated on the 3D array surface add to the enhancement of information capacity [22], which in actuality restricts the convenience of concentrated 3D array realization. This emanates from the fundamental illustration of Maxwell's equations, which demands that electromagnetic field in the volume of a 3D array is totally depicted by the field on its surface, and in this manner no additional information can be removed from components inside the 3D array.

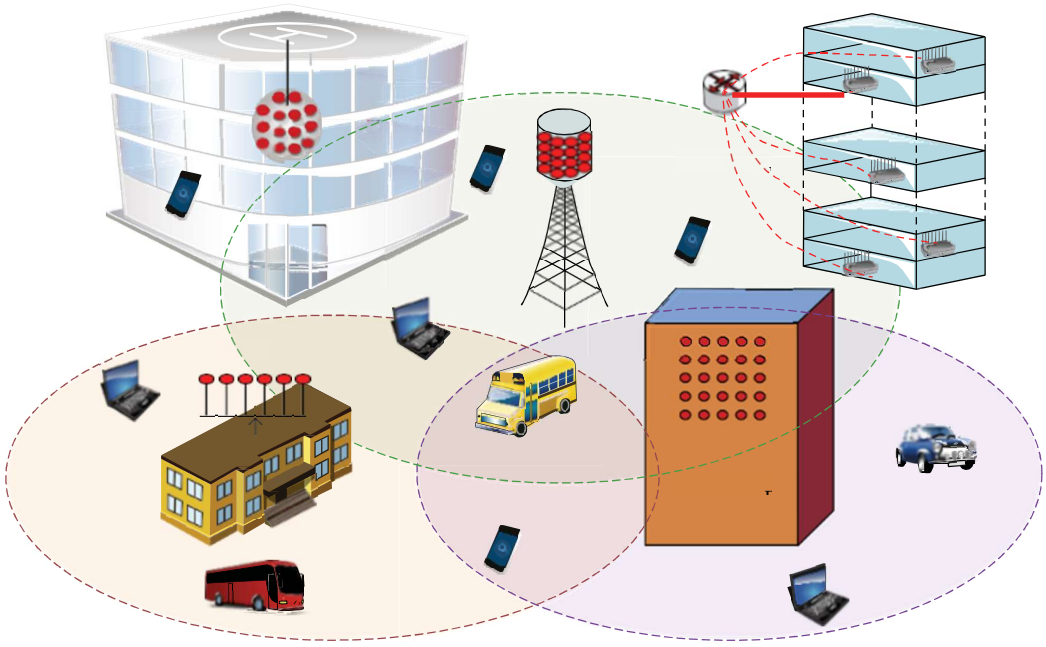

Figure 2.3: Antenna array configuration [32], [41].

Trade-offs in implementation of massive MIMO was analysed in [42]. It was proposed that the array be designed as a grid of the same common modules that will host a small number of antennas. This is envisioned to deal with the challenges related with the huge quantity of elements. They are then interconnected to their nearest neighbours digitally through a mesh network that provides connection to a central processor. It also did suggest the optimum trade-off between radiated power and overhead incurred for additional elements to the array.

Irregular antenna arrays were investigated in [43]. In comparison with regular antenna array, the irregular antenna array contributed towards improvement of the achievable rate when the antennas were greater than or equal to a specified threshold. These results and way of thinking could be extended to massive MIMO to investigate how irregularity in antenna array affects the system performance.

The efficient use of Millimetre-wave (mmWave) frequencies in massive MIMO would possibly need beamforming and beam steering. The downside of mmWave bands: greater path loss at certain frequencies, like 60 GHz, due to smaller antenna aperture and atmospheric absorption. It greatly limits the available propagation distances. Using mmWave demands use of directional antennas with high gain. The base stations here are likely to use hundreds of antennas, while user terminals use fewer than 16 antennas. It is envisaged that antenna arrays will be very small at these frequencies, permitting the use of a large number of elements [44].

Fading remains a problem due to extreme sensitivity of the channel to little changes within the environment for mmWave systems.

The use of more antennas at the UTs was explored by [45]. It was shown that linear detectors can sufficiently handle multi-antenna UTs in massive MIMO. The spectral efficiency increases with the number of user antennas per terminal N. If the number of users per cell is *K*, when the value *NK* is fixed the highest spectral efficiency is achieved by having NK single antenna users. From this, more user antennas work to enhance spatial multiplexing in networks with fewer users. For the purpose of this research, we adopt linear antenna arrays for simplicity in simulation process.

2.5 Massive MIMO Channel Measurements

It has been shown theoretically that the spectral, power, and energy efficiency of communication systems would be greatly enhanced by massive MIMO [13]. So far these significantly hypothetical results are anchored around several critical postulates concerning wireless channel that need to be adequately verified through measurement campaigns. The measurements carried out by several researchers appears to validate the view that channel characteristics such as received power, or the direction of arrival evolve along the massive MIMO array. This further vindicates the long standing assumption that massive MIMO channel is non-stationary [33], [46]–[50]. There have been attempts to show that the sum rate of the massive MIMO systems increases with the increase in the number of transmit antennas [51].

The clustering results in delay domain were found similar in different frequency bands [49]. It is reported in [46]–[50] that the power variations were observed in the cluster and this power variations are different for different frequency bands. The total number of clusters is higher in Non-Line Of Sight (NLOS) condition than in Line Of Sight (LOS) condition. Again, the total number of clusters is higher when using omni-directional antenna than in the case of directional antenna. The dynamic range of directional antenna is lower than that of the omni-directional antenna in NLOS condition and it's a little higher in LOS condition [49].

When the length of observed clusters increases, the number of clusters diminishes. However, no significant difference observed when different frequencies or polarizations are used. Clusters can contribute over a short part of the array or over the entire array and consequently their power contribution varies significantly [49]. The delay spread has a noticeable difference in LOS and NLOS conditions, with lower values observed in LOS condition. The delay spread at higher frequencies is noticeably lower than at lower frequencies, while there is no clear

impact of polarization. Delay spreads for directional and omni-directional antennas are comparable in LOS and NLOS conditions. The received energy in directional antennas is a little more spread out in time which means it has slightly higher delay spread than the omni-directional antenna. But this can also result from misalignment of the transmitter and receiver antennas for directional antennas [49].

The angle spreads are lower in LOS case than NLOS when you consider diverse frequency bands and polarization. The higher values of angle spreads are received in lower frequency bands with significant impact observed in LOS and minimal in NLOS. Also the angle spread for directional antenna are visibly lower in comparison to omni-directional antenna and there is minimal difference in LOS and NLOS conditions for both the antennas [49]. Clusters undergo large scale fading which is observed as a deviation of cluster power from linear regression line. The shadowing factor is higher in lower frequency bands as compared to higher frequency bands with no impact from polarization [46], [47], [49]. It is said that there is significant correlation between channel vectors at diverse positions, which solely determine the interference level in the system. There is a correlation between the angle spread and the delay spreads that is dependent on the propagation conditions, frequency band, and the directivity of the antennas in use [49], [51]. The correlation is also found to be dependent on the aperture size [50], [52]. In conclusion it is observed that an increase in the number of antennas improves channel orthogonality, but this can only hold to a certain degree [51], [52]. This part serves to support our assumption of large antenna arrays and the fact that for the given coherent time the massive MIMO channel though non-stationary can be treated as stationary due to slow change in channel characteristics experienced in each coherent time. This forms the basis of our assumptions in chapters to come.

2.6 Massive MIMO Channel Characterization

Several massive MIMO channel models have been put forward and the existing ones improved and optimized through measurement to reflect practical massive MIMO propagation environment. According to [28] and [52] the massive MIMO channel models can be categorized into Correlation-Based Stochastic Models (CBSMs) and Geometrical-Based Stochastic Models (GBSMs) as in Table 2.1. The CBSMs are said to have a lower complexity with limited accuracy and are used in theoretical analysis of massive MIMO models. The GBSMs models may accurately reflect the realistic channel properties although with high computational complexity and are highly suitable for massive MIMO channel models.

Table 2.1: Massive MIMO channel models.

Modelling method	Category
CBSM	❖ i.i.d. Rayleigh channel ❖ General correlation channel model ➢ Kronecker model ➢ Mutual coupling channel model • Weichselberger model • Virtual channel representation
GBSM	❖ 2D channel Model ❖ 3D channel model

2.6.1 Correlation-Based Stochastic Models (CBSMs)

These are the theoretical models employed in evaluating the massive MIMO system performance. CBSMs capture the correlation properties of channels among antenna elements on the transmitter and/or receiver side except for the classic i.i.d. Rayleigh fading channel model.

2.6.1.1 i.i.d. Rayleigh Channel Model

The independent and identically distributed Rayleigh fading channel model is broadly received for the hypothetical investigation of the massive MIMO systems. But the i.i.d. Rayleigh fading channel model assumes there is no correlation and common coupling between transmitter antennas or receiver antennas, making them appropriate for massive MIMO distributed antennas as opposed to co-located antennas. The massive MIMO system performance was analysed using the i.i.d. Rayleigh fading channels in [54], [55]. Where it was pointed out that matrices of massive MIMO channel can be analysed easily by use of central limit theorem and random matrix theory since channel coefficients are i.i.d. in nature. Further, in [55] and [56] it is shown that the eigenvalues distribution of the product of the channel matrix with its Hermitian transpose converges to Marčenko-Pastur law as the size of massive MIMO channel matrix becomes large.

2.6.1.2 Kronecker Based Stochastic Model (KBSM)

According to [58], [59] KBSM is employed to model the massive MIMO system channel with correlation between antennas. The total spatial correlation matrix can be expressed as the kronecker product of one-sided transmit and receive correlation matrices as in [60], which are not coupled in any case. This model is preferred in modelling the channel covariance. Due to its simple implementation as well as incorporating antenna correlation, KBSM is commonly

applied in capacity and performance analysis of massive MIMO systems. However, KBSM will lead to underestimation of mutual information in massive MIMO since it cannot generate an actual channel [61]. It has been already reported in [46], [62] that individual antennas on a massive antenna array may observe different set scatterers. These scatterers may appear or disappear as you move along the array axis leading to different antennas observing different scatterers for massive MIMO case. The KBSM with Birth-Death process on Array Axis (KBSM-BD-AA) is developed in [63] to capture this effect for massive MIMO channels. It was found that the birth-death process on array axis minimizes the correlation between antennas as antennas in a large array may observe different scatterers.

2.6.1.3 Weichselberger Model

The Weichselberger model assumes separate spatial correlation at transmit and receive link ends but unlike KBSM introduces a joint correlation between transmit and receive eigenbases which is obtained from measurement [64]–[66]. The correlation between transmit and receive eigenbases is dependent on the environment and thus can reflect the scattering structure of the environment as well as the spatial structure of the channel. Since Weichselberger model is applied in wireless communication systems design; it is of great importance to build useful model parameters that can lead to adaptive system structure. The Weichselberger model optimizes the trade-off between accuracy and complexity in massive MIMO channel models. As opposed to KBSM, the Weichselberger model is well placed to be applied in co-located antenna scenarios when the coupling has to be considered between transmit and receive eigenbases [67].

2.6.1.4 Virtual Channel Representation (VCR) Model

According to [66] in VCR the MIMO channel is modelled in beam space as opposed to eigenspace and employs Discrete Furrier Transform matrices (DFT). The channel capacities of VCR model have been investigated although it has a pitfall that it only supports single polarized Uniform Linear Arrays [68], [69]. The real antenna configuration influences the angular resolution and consequently the ‘correctness’ of the VCR model. The increase in number of antennas decreases the angular bins and this has the effect of increasing the accuracy of the VCR model. This point out that VCR model might play a significant role in massive MIMO [66], [70].

2.6.2 Geometrical-Based Stochastic Models (GBSMs)

In [71], [72], the interaction process and the Interacting Objects (IOs) geometric location are characterized and modelled. The Multipath Components (MPCs) parameters are obtained from the interfering objects through a simple ray tracing method. The concept of clusters is employed where the MPCs tend to depart and arrive in clusters. This simplifies the GBSM as cluster parameters changes are minimal or none when transmitter and/or receiver move over a large area [71], [72]. The GBSM brings around a number of resourceful capabilities like ease of modelling and implementing directional properties, capturing of time variation in a channel through e.g., cluster visibility regions and decoupling the antenna influence from the propagation channel hence the possibility of simulating different antenna types [72].

2.6.2.1 2D GBSM

The usual shape Two-Dimensional (2D) GBSMs employed in MIMO channels presume that effective scatterers situate on normal shapes such as rings, ellipses, and rectangles as stated in [67]. A novel 2D non-stationary wideband multi-confocal ellipse model is proposed for massive MIMO channel modelling in [73]. The spherical Wavefront effect as well as non-stationery properties on the array and time axes are captured in this model, while Angle Of Arrivals (AOAs) and Angle Of Departures (AODs) are assumed dependent in the ellipse model. WINNER models inspired 2D non-stationery wideband GBSM with the capability of capturing the key characteristics of massive MIMO is proposed in [74]. The model introduces parabolic wavefront through a second-order approximation to spherical wavefront. This approximation leads to linear angular drifts across the array which results in reduced theoretical and computational complexity in comparison to spherical wavefronts [74], [75].

2.6.2.2 3D GBSM

According to [76] scatterers can disperse in vertical plane. This means that the effects of elevation angles need to be considered to capture the full channel characteristics. This renders the 2D GBSMs insufficient to model massive MIMO practical channel and hence the need for the development of 3D GBSMs.

Several conventional 3D massive MIMO models are readily available in literature. The COST 2100 channel model that can replicate the stochastic properties of multi-link MIMO channels over time, frequency and space is developed in [77]. This model was found to be generic and flexible making it suitable to model multi-user or distributed MIMO scenarios. The extension to COST 2100 for massive MIMO application was proposed in [78]. The extended COST 2100

sought to realize 3D propagation, polarization, cluster behaviour along the large array axis, and multipath component gain functions for closely spaced users. These proposed model extensions were implemented and found to be capable of replicating channel statistics in terms of user separability, sum-rates, and temporal behaviour.

Theoretical non-stationery 3D wideband twin-cluster channel model for massive MIMO systems with carrier frequencies in the order of gigahertz is proposed in [79]. Here the near-field effects rather than the far-field effects are taken into consideration. These effects are the spherical wavefront assumption and the birth-death process to model non-stationary properties of clusters like appearance and disappearance on both array and time axes. The impact of elevation angles to correlation functions was also discussed.

The METIS channel models in [80] consists of a map-based model, a stochastic model and a hybrid model which is a combination of the two. The METIS stochastic channel model is a further extension of the WINNER/3GPP model to enhance its deployment in massive MIMO channel modelling. The hybrid model avails a flexible and scalable channel modelling framework. The spherical wavefront and other extensions are as in the previous 3D models.

The 3D unified GBSM framework model is discussed and implemented in [81]. This a hybrid of WINNER II channel model [82] and Saleh-Velenzuela channel model [83]. It was observed that elevation angles can cause a huge impact on channel capacities and polarization of antennas has the effect of halving the dimension of an antenna array while reducing the channel capacity.

A 3D wideband multi-confocal ellipsoid model for MIMO system is proposed in [84]. The 3D wideband multi-confocal ellipsoid model describes the channel in both azimuth and elevation directions, including delay, Doppler frequency, Angle Of Departure (AOD), and the Angle Of Arrival (AOA). It is observed that the 3D wideband multi-confocal ellipsoid model captures the effect of clusters more accurately. The extension of the 3D wideband multi-confocal ellipsoid model to support massive MIMO would be worth looking into.

In this research we adopt the i.i.d. channel model for the simulation and analysis of the precoding and channel estimation techniques. This is picked since most work related to this research have assumed the i.i.d. channel for the hypothetical investigation of the massive MIMO systems.

2.6.3 General Massive MIMO Channel Model

If we consider an uplink massive MIMO system with K single antenna user terminals transmit signals to a BS equipped with M antennas simultaneously. The propagation matrix, $\boldsymbol{G}$, with dimensions $M \times K$, is given as the product of a $M \times K$ matrix, $\boldsymbol{H}$, which represents small scale fading, and $K \times K$ diagonal matrix, $\boldsymbol{D}^{1/2}$, which has a $K \times 1$ vector, β, in its diagonal elements for the large scale fading coefficients [18], [32], [33]. At the point when the NLOS channel is considered, the channel model can be summed up by

$$\boldsymbol{G} = \boldsymbol{H}\boldsymbol{D}^{1/2} \tag{2.6}$$

where $\boldsymbol{D} = diag\{\beta_1, \beta_2, \dots, \beta_K\}$ and $\beta_K = \emptyset d_K^{-\alpha} \xi_K$. $\emptyset$ is the constant related to the antenna and carrier frequency, d_K is the distance between the base station and the kth user terminal, α is the path loss exponent, and ξ_K is the log-normal shadow fading with $10log_{10}\xi_K \sim \mathcal{N}(0, \sigma_{sh}^2)$. We use this channel model in this work as we assume NLOS conditions.

2.7 Massive MIMO Physical Layer Security

Wireless channel is inherently open subjecting the link between two parties who are legitimately communicating to attacks from an adversary. The third-party attacks can be classified into two groups: jamming and eavesdropping [85]. An example of distributed jamming is shown in Figure 2.4. For the case of eavesdropping the adversary seeks to establish the message that was communicated between the two legitimate parties. While in the case of jamming, the adversary assumes the position of a malicious transmitter thereby advancing noise on the wireless channel rendering the legitimately received message hard and sometimes totally impossible to be recovered [85].

Traditionally, cryptographic strategies have been employed at application layer to cater for information security of the wireless communication systems. Of late information security of wireless communication systems has been implemented at the Physical (PHY) layer [86]. Practically PHY layer implementation can be approached from two perspectives; it can be implemented to reinforce the already existing information security strategies or supplant them entirely and hence eliminate the tedious systems and protocols used in dispersion and supervision of the key [87].

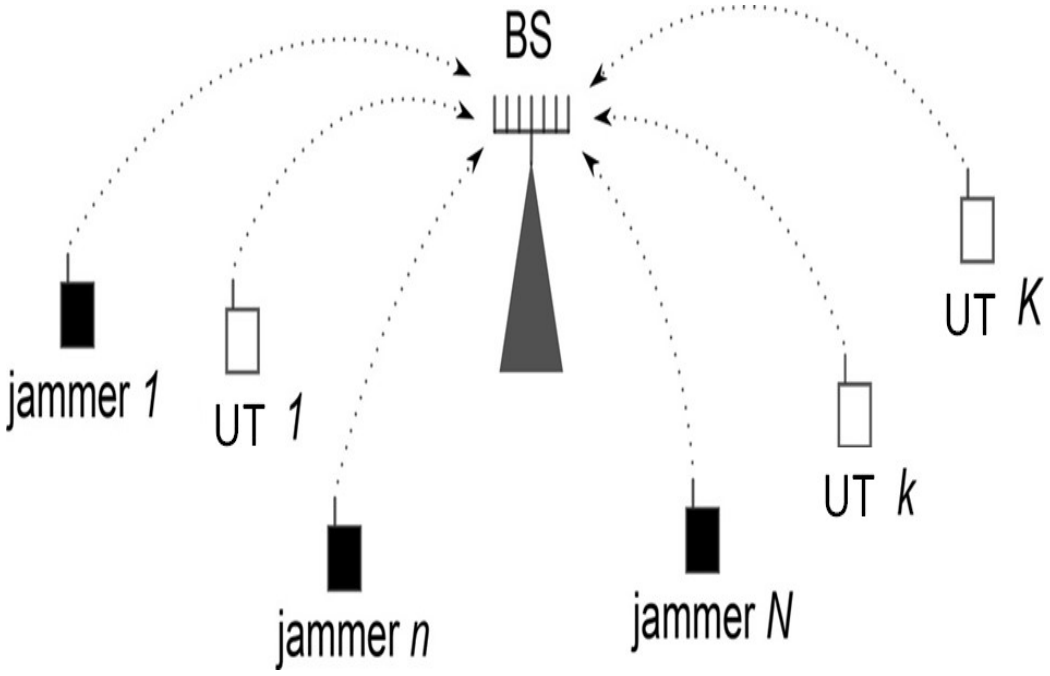

Figure 2.4: Massive MIMO with distributed jamming [85].

The PHY layer security systems have no restriction imposed on eavesdropper computationally. The authentication at PHY layer has good prospect of being used in identification of some malicious attacks for energy limited cases as its faster and lightweight in execution [88]. Information theoretic secrecy forms one of the primary subjects concerning PHY layer security. Perfect secrecy can be realised based on two ways. First, the number of messages need to be finite with number of possible keys being the same. Then there should be no extra information made available concerning the transmitted message Q from the signal E received by the eavesdropper [88]. This leads to the view of the PHY layer security as not requiring secrete keys prior to transmission.

The exploitation of multiple antennas and intentional inclusion of artificial noise in transmission has long been exploited to facilitate PHY layer security implementation in conventional MIMO [89]. Massive MIMO and PHY layer security combination gives high prospects since we have high degrees of freedom in massive MIMO with simple designs of transceiver in comparison to conventional MIMO. Analysis of the secrecy of PHY layer was presented in [90] for regularized channel inversion precoding, an exhaustive examination and design of PHY layer security involving massive MIMO networks has not been adequately addressed. The questions surrounding the generation of artificial noise and effective precoding to secure massive MIMO not forgetting the pilot contamination effects on PHY security need to be explored.

It is shown in [91], [92] that massive MIMO is inherently robust for the case of passive eavesdropping attacks. But it particularly singles out pilot contamination that attacks the channel estimation process and show that it reduces the achievable secrecy capacity as well as being difficult to detect. According to [93], Massive MIMO is inherently robust to attacks for no training-phase jamming as well as eavesdropping on both data and training. It is further shown that with a minimum number of BS antennas it is possible to realise data security without a Wyner encoding for a specified rate of leaking data to the adversary. The attainable rate of secrecy for active or passive eavesdropper is discussed in [94] and a scheme to detect active attacks leveraging on pilot contamination is presented.

Existing studies on physical layer security often assume the availability of perfect channel state information (CSI) and overlook the importance of channel training needed for obtaining the CSI [95]–[97]. An active eavesdropper UT can attack the training phase in wireless communication to improve its eavesdropping performance. This points at a new security attack from the pilot contamination phenomenon, which targets at systems using reverse training to obtain the CSI at the transmitter for precoder design. This attack changes the precoder used by the legitimate UT in a controlled manner to strengthen the signal reception at the eavesdropper UT during data transmission [98]. This informs the reason why this section is included here as another source of pilot contamination through attack on the physical layer.

2.8 Massive MIMO Precoding

MIMO encoding envisage the mapping of symbols, say, from modulation alphabet to symbols to be transmitted over multiple transmit antennas. The MIMO encoding techniques include the spatial multiplexing, space-time coding, [99], [100] and spatial modulation [101]. These three coding techniques do not require any knowledge of the CSI at the transmitter. The MIMO encoding that requires the CSI knowledge at the transmitter is called MIMO precoding. The precoding operation at the transmitter improves the performance of MIMO wireless systems. The application of these MIMO techniques to massive MIMO systems creates a mine of practical interests. It has been demonstrated that for a multiuser MIMO system with M transmit antennas and K user antennas, the capacity grows linearly with min (M, K) [102]. This capacity scaling can be applied to massive MIMO systems where M transmitting antennas communicates with K UTs.

To achieve an optimal transmission scheme dirty paper coding (DPC) is employed, where the transmitter jointly encodes the data symbols for all users using perfect knowledge of CSI [103],

[104]. The transmission scheme based on DPC comes with complexity that prohibits its practical realization. This then necessitates the reliance on transmit-side pre-processing techniques which are either linear precoding or non-linear precoding. Theoretically, the precoding is also referred to as Space Division Multiple Access (SDMA).

2.8.1 Linear Precoding

Linear precoding includes a linear change of the information symbols implied for the clients on the downlink utilizing a precoding matrix. A block diagram in the Figure 2.5 below shows a massive MIMO system on the downlink.

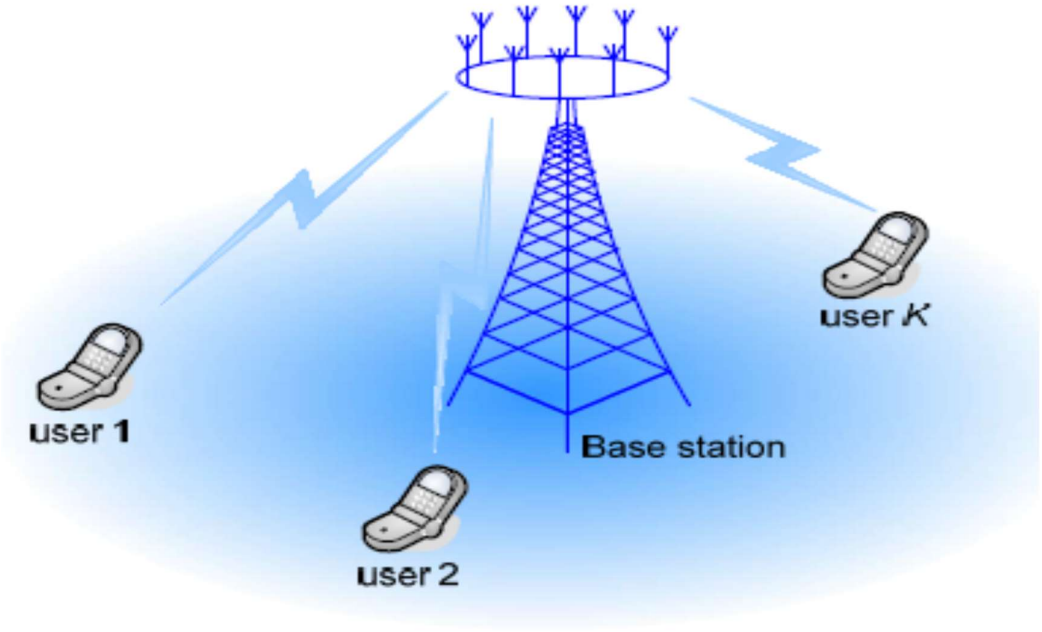

Figure 2.5: Massive MIMO downlink system model [105].

The BS with *M* transmit antennas simultaneously transmits to *K* UTs. If we consider un-coded massive MIMO downlink system, the block diagram representation of such a system is illustrated as in Figure 2.6.

The data symbol for the *k*th UT, $\boldsymbol{x}_k \in \mathbb{C}^{K\times 1}$, is transformed by the precoding matrix $\boldsymbol{A}$ and $\mathbb{E}[|\boldsymbol{x}_k|^2] = 1$. The transmitted signal vector from the BS can be written as

$$\boldsymbol{u} = \sum_{k=1}^{K} \boldsymbol{A}_{m,k}\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x} \tag{2.7}$$

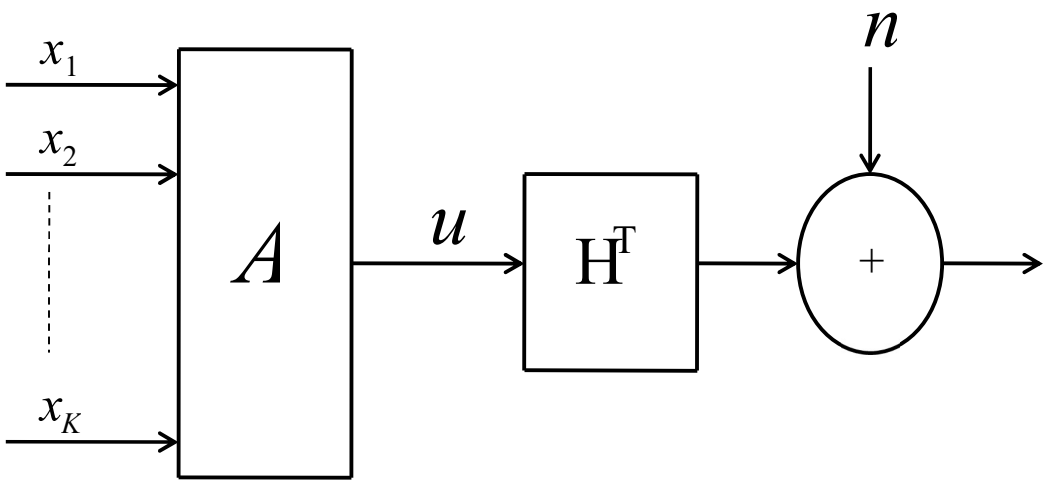

Figure 2.6: The system model.

where $\boldsymbol{A} \in \mathbb{C}^{M\times K}$ is precoding matrix and $\boldsymbol{x} = [x_1\ x_2 \dots x_K]$ is the information vector. The transmit vector u average transmit power is constrained by $\mathbb{E}[|\boldsymbol{u}|^2] = tr(\boldsymbol{A}^*\ \boldsymbol{A}) = P_{tr}$. The channel matrix $\boldsymbol{H} \in \mathbb{C}^{M\times K}$, and the received signal vector at kth UT is given as

$$\boldsymbol{y}_k = \boldsymbol{h}_k^T\boldsymbol{u} + n_k \tag{2.8}$$

where n_k is the noise vector at the kth UT and $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}$ represents the random channel vector between the BS and the kth UT. Then the total signal vector received is given as

$$\boldsymbol{y} = \boldsymbol{H}^T\boldsymbol{u} + n \tag{2.9}$$

where $\boldsymbol{u} = [u_1 u_2 \dots u_K]$ and $\boldsymbol{H} = [h_1\ h_2 \dots h_K] \in \mathbb{C}^{M\times K}$. The linear precoding techniques are pegged on different performance criteria as discussed in the literature [21], [106]–[109]. These include the ZF, Maximum Ratio Transmission (MRT) also known as MF, Regularized Zero Forcing (RZF), and MMSE.

In massive MIMO downlink framework, MMSE precoding serves as the optimal linear precoding scheme. This system is created by the MSE method. Due to average power at each transmitting antenna being constrained; Lagrangian optimizing strategy is utilized for acquiring this precoder. The optimal MMSE precoder is obtained as [110], [111]

$$\boldsymbol{A}_{MMSE} = \frac{1}{\beta}\boldsymbol{H}^*\left(\boldsymbol{H}^T\boldsymbol{H}^* + \frac{K}{P_{tr}}\boldsymbol{I}_K\right)^{-1} \tag{2.10}$$

with β being the scalar of Wiener Filter and given as

$$\beta = \sqrt{\frac{tr(\boldsymbol{B}\boldsymbol{B}^*)}{P_{tr}}} \tag{2.11}$$

and

$$\boldsymbol{B} = \boldsymbol{H}^* \left(\boldsymbol{H}^T \boldsymbol{H}^* + \frac{K}{P_{tr}} \boldsymbol{I}_K \right)^{-1} \tag{2.12}$$

ZF precoding is a linear precoding in which the individual user terminals have one antenna and they experience no multi-user interference. The assumption is that the ZF precoder implements a pseudo-inverse of the channel matrix. This precoding is obtained as [111]–[113]

$$\boldsymbol{A}_{ZF} = \frac{1}{\beta} \boldsymbol{H}^* (\boldsymbol{H}^T \boldsymbol{H}^*)^{-1} \tag{2.13}$$

with β being the scalar of Wiener Filter given as in equation (2.11) and

$$\boldsymbol{B} = \boldsymbol{H}^* (\boldsymbol{H}^T \boldsymbol{H}^*)^{-1} \tag{2.14}$$

The MRT precoding maximizes the Signal-to-Noise Ratio (SNR) and is one of the common schemes in linear precoding. For it to function well, it requires the BS in the massive MIMO system to radiate low signal power. This precoder is given as [111], [113]

$$\boldsymbol{A}_{MRT} = \frac{1}{\beta} \boldsymbol{H}^* \tag{2.15}$$

with β being the scalar of Wiener Filter given as in equation (2.11) and $\boldsymbol{B} = \boldsymbol{H}^*$

2.8.2 Non-Linear Precoding

Linear precoders delivers low complexity though with a lag in performance in comparison to the non-linear precoders, more so when we have a small set of users [114]. The two common non-linear precoding schemes are the Tomlinson-Harashima Precoding (THP) and the Vector Perturbation (VP) precoding.

The non-linear precoding scheme THP employs modulo operation on the transmitter side to limit the power transmitted. The transmitted signal at the receiver has a non-linear distortion

and a similar modulo operation is used to annul this distortion. The system model for THP is depicted in Figure 2.7 [115], [116]

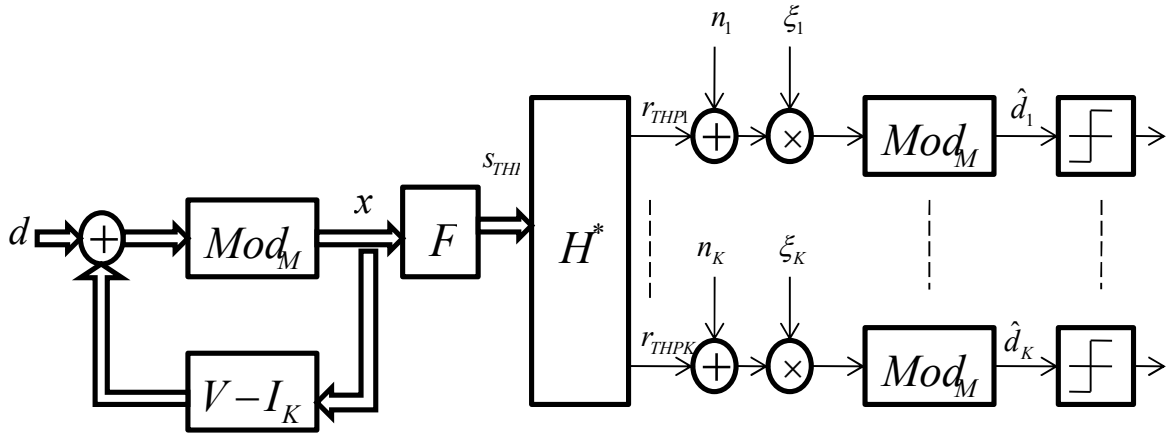

Figure 2.7: System model for conventional THP [116].

The transmit signal of the THP is given as

$$\boldsymbol{S}_{THP} = \boldsymbol{F}\boldsymbol{x} \tag{2.16}$$

where the unitary feed-forward filter $\boldsymbol{F}$ is gotten through QR-decomposition of the channel matrix $\boldsymbol{H} = \boldsymbol{F}\breve{\boldsymbol{V}}$, and the elements of vector $\boldsymbol{x}$ are obtained as

$$[\boldsymbol{x}]_k = Mod_M\left([\boldsymbol{d}]_k - \sum_{l=1}^{k-1}[\boldsymbol{V}-\boldsymbol{I}_K]_{k,l}[\boldsymbol{x}]_l\right), \forall k \in \{1,\dots,K\} \tag{2.17}$$

and the modulo function $Mod_M(\boldsymbol{x})$ is expressed as

$$Mod_M(\boldsymbol{x}) = \boldsymbol{x} - 2\sqrt{M}\left[\frac{1}{2} + \Re\left\{\frac{\boldsymbol{x}}{2\sqrt{M}}\right\}\right] - 2j\sqrt{M}\left[\frac{1}{2} + \Im\left\{\frac{\boldsymbol{x}}{2\sqrt{M}}\right\}\right] \tag{2.18}$$

where $\Re\{.\}$ and $\Im\{.\}$ denote the real and imaginary parts of a complex-valued variable. The feedback matrix $\boldsymbol{V}$ in (2.17) is obtained by $\boldsymbol{V} = \Xi\,\breve{\mathbf{V}}^*$, where $\Xi = \mathrm{diag}\left(\xi_1,\dots,\xi_K\right)$ and $\xi_K = 1/\left[\breve{\mathbf{V}}\right]_{k,k}, \forall k \in \{1,\dots,K\}$. When the signal arrives at the kth UT, it is multiplied by ξ_k before passing through the Mod_M-module and then applied to the Quadrature Amplitude Modulation (QAM) demodulator.

VP precoding scheme is a sub-optimal non-linear precoding technique that surpasses the performance of linear precoding techniques [117], [118]. The Gaussian integer vector is employed in perturbing the data vector in VP [119]. This minimizes the total power, and then

a linear front-end precoder is utilized in precoding of the perturbed data vector as shown in Figure 2.8.

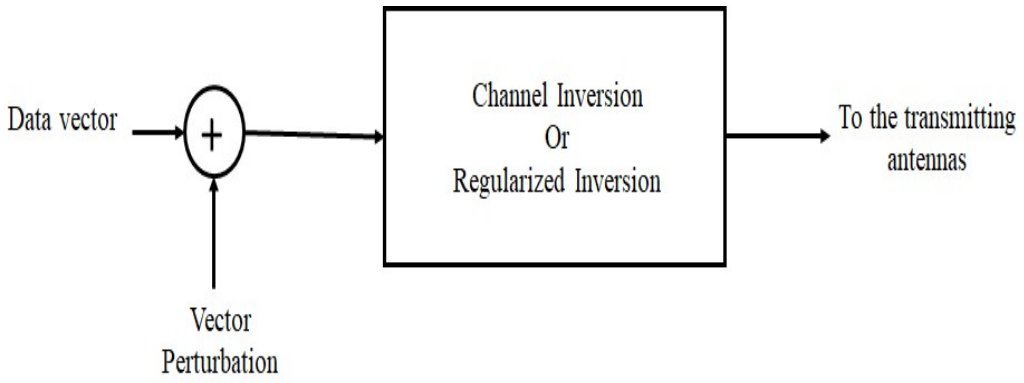

Figure 2.8: Block diagram representation of a generically Vector Perturbation system [117].

The BS transmits K complex symbols $\boldsymbol{x}_k \in \boldsymbol{A}$ at the same time to the K UTs, where $\boldsymbol{A}$ is the symbol alphabet. With assumption that $0 \notin \boldsymbol{A}$, we let $\boldsymbol{x} \triangleq (x_1 \dots x_k)^T$ to be information vector gotten by collection of all symbols. Then the transmit vector is obtained as [120], [121]

$$\boldsymbol{s} = \sqrt{\frac{P_{tr}}{\Gamma(\boldsymbol{z})}}\boldsymbol{H}^{\dagger}(\boldsymbol{x} + \tau\boldsymbol{z}) \tag{2.19}$$

where, $\boldsymbol{H}^{\dagger} = \boldsymbol{H}^{*}(\boldsymbol{H}\boldsymbol{H}^{*})^{-1}$ represents the right pseudo-inverse, τ denotes a real-valued scaling factor that is fixed, and $\boldsymbol{z} \triangleq (z_1 \dots z x_k)^T \in \boldsymbol{G}^K$ denoting a perturbation vector with elements that are Gaussian integers. The choice of the scaling power factor $\Gamma(\boldsymbol{z})$ so that the instantaneous power constraint, $\|\boldsymbol{s}\|^2 \leq P_{tr}$, is satisfied by the equality

$$\Gamma(\boldsymbol{z}) \triangleq \left\|\boldsymbol{H}^{\dagger}(\boldsymbol{x} + \tau\boldsymbol{z})\right\|^2 \tag{2.20}$$

2.8.3 Hybrid Precoding

The hybrid precoders use a combination of both the linear and non-linear precoders into one precoding scheme. It tends to enhance the good attributes of both the linear and non-linear precoders with the view to enhance performance and reduce complexity. The Phased Zero Forcing (PZF) is proposed in [122]. PZF applies the phase-only control at RF domain followed by low-dimensional ZF precoding as per the effective channel seen from baseband. It is shown that PZF results in a highly attractive performance. The phased hybrid precoder is also employed for use in wideband massive MIMO systems in [123]. It is shown that for wideband

a system, performance of hybrid precoder is pegged on employed frequency-domain scheduling method and number of available RF chains. The performance of wideband hybrid precoder compares to that of narrowband hybrid precoder for the case when greedy scheduling is used with many RF chains. The Hybrid Linear-Tomlinson-Harashima precoder (HL-THP) which constitutes an inner linear precoder with outer THPs is discussed in [124]. HL-THP groups the UTs and the intra-group interference is mitigated by the THP while the inter-group interference is mitigated by the inner linear precoder. HL-THP gives higher average achievable rate per UT and a reduced Bit Error Rate (BER) in comparison to linear precoders and compares favourably to the conventional THP.

2.8.4 Simulation and Analysis of Selected Precoding Models

We review the comparison in performance amongst the linear precoders: MMSE, Cholesky and Sherman-Morrison (CSM), and Truncated Polynomial Expansion (TPE) [21], [106], [110], [111], [125].

The linear precoding techniques are pegged on different performance criteria as discussed in the literature [21], [106]–[109]. In massive MIMO downlink framework, MMSE precoding serves as the optimal linear precoding scheme. This system is created by the mean square error (MSE) scheme. Due to average power for every BS transmitting antenna being controlled; Lagrangian optimizing strategy is utilized for acquiring MMSE precoder. The representation of the optimal MMSE precoder is discussed in [110], [111].

Though, conventional linear precoders employed in massive MIMO networks, for example MMSE precoder, give optimal performance its normally associated with inherently high computational complexity emanating from the inversion of large dimensional matrix. The MMSE precoder is modelled as follows.

$$\hat{A}_{MMSE} = \beta_{MMSE}.\hat{H}^{H}.\left(\hat{H}.\hat{H}^{H} + \sigma^{2}.n_{t}.I_{K}\right)^{-1} \tag{2.21}$$

To take care of this issue, the Cholesky-decomposition and Sherman-Morrison lemma was used to formulate CSM-based precoder and afterwards applied on the inversion of matrices by leveraging on asymptotically orthogonal channel property inherent to massive MIMO networks [125].

The CSM-based scheme is used to evade intricate matrix inversion in MMSE precoding. The first step in applying this method is to test the matrix $W_{MMSE} = \left(\hat{H}.\hat{H}^{H} + \sigma^{2}.n_{t}.I_{K}\right)^{-1}$ if it is a positive definite Hermitian matrix. Applying Cholesky decomposition to W_{MMSE} [126]

$$W_{MMSE} = L.L^H \tag{2.22}$$

where L is a lower triangular matrix. This implies that

$$W_{MMSE}^{-1} = (L^H)^{-1}.L^{-1} \tag{2.23}$$

Here the computation of the inversion of matrix W_{MMSE} can be transformed into computation of the inversion of matrix L. To sustain low computation complexity the matrix L is decomposed further into

$$L = B + L' \tag{2.24}$$

where $B = diag(L)$ and $L' = (l'_1, l'_2, \dots, l'_{K-1}, 0)$ is a matrix based on L with the diagonal elements of L set to zero. The matrix is further written as

$$L = B + l'_1.e_1 + l'_2.e_2 + \cdots + l'_{K-1}.e_{K-1} + 0.e_K \tag{2.25}$$

Where e_i is the ith row of the identity matrix I_K, l'_i is the ith column of L'. Applying the Sherman-Morrison lemma [125], [127]

$$L^{-1} = (F_{K-2} + l'_{K-1}.e_{K-1})^{-1} = (F_{K-2})^{-1} + \frac{(F_{K-2})^{-1} l'_{K-1}.e_{K-1}(F_{K-2})^{-1}}{1+l'_{K-1}.(F_{K-2})^{-1}.e_{K-1}} \tag{2.26}$$

where $F_{K-2} = B + \sum_{i=1}^{K-2} l'_i.e_i$. Thus the inversion of F_{K-2} should continue

$$F_{K-2}^{-1} = (F_{K-3} + l'_{K-2}.e_{K-2})^{-1} = (F_{K-3})^{-1} + \frac{(F_{K-3})^{-1} l'_{K-2}.e_{K-2}(F_{K-3})^{-1}}{1+l'_{K-2}.(F_{K-3})^{-1}.e_{K-2}} \tag{2.27}$$

where $F_{K-3} = B + \sum_{i=1}^{K-3} l'_i.e_i$. This is curried on until we compute

$$F_1^{-1} = (F_0 + l'_1.e_1)^{-1} = (F_0)^{-1} + \frac{(F_0)^{-1} l'_1.e_1(F_0)^{-1}}{1+l'_1.(F_0)^{-1}.e_1} \tag{2.28}$$

where $F_0 = B$. Since the B matrix is diagonal matrix, its inversion is easy and efficient. Thus, the inversion of L can be computed by iterating $K-1$ times.

Drawing inspiration from the excellent performance exhibited by conventional precoders like MMSE, TPE was proposed in [128] to substitute matrix multiplication and inversion which gave rise to TPE precoder. The TPE precoder comes in handy for hardware implementation in

real-time as well as reduction in first symbol transmission delay. The availability of hardware resources predetermines the polynomial degree thereby facilitating probable transition from MRT to MMSE smoothly. The TPE precoder is modelled as follows

$$A_{TPE} = \sum_{i=0}^{J-1} \omega_i (H.H^H)^i H^H \tag{2.29}$$

where $\omega_0, \dots, \omega_{J-1}$ are scalar coefficients. The bracketed term in gives potential expression for ω_i, they are not optimal values for $J < \infty$, neither do these values satisfy the power constraint $\frac{1}{K} tr(AA^H) = P$. Then J is the TPE order.

If assume that the matrix A_K is the precoder matrix of the different precoders, with the column a_k removed, the signal-to-interference-and-noise ratio (SINR) at the kth UT becomes.

$$SINR_k = \frac{h_k a_k a_k^H h_k^H}{h_k^H A_k A_k^H h_k + \sigma^2} \tag{2.30}$$

We use MATLAB as the simulation tool for comparison and analysis. We provide the comparison between TPE precoding in [21], the CSM-based precoding in [125], and the well-known MMSE precoding in one cell. This comparison is carried out for M BS antennas and K UTs ($M = 512, K = 128$ and $M = 512, K = 64$) massive MIMO systems.

Figure 2.9 shows comparison of the average bitrates between CSM-based precoding, the TPE precoding and MMSE precoding. Considering Figure 2.9, a few observations can be made. The CSM-based precoding together with MMSE precoding have a better performance in comparison to the TPE precoding. At the point when J has a low value the TPE precoder gives a poor performance, but with increasing value of J, the bitrates of TPE precoder improves and nearly catches up with that of MMSE precoder and CSM-based precoder [109], [128]. The consequence of this improved performance is in the enhanced demand for hardware resources. However, since the TPE precoder approximates the MMSE precoder and CSM-based precoder its performance never surpasses their corresponding performance [109], [128], [129].

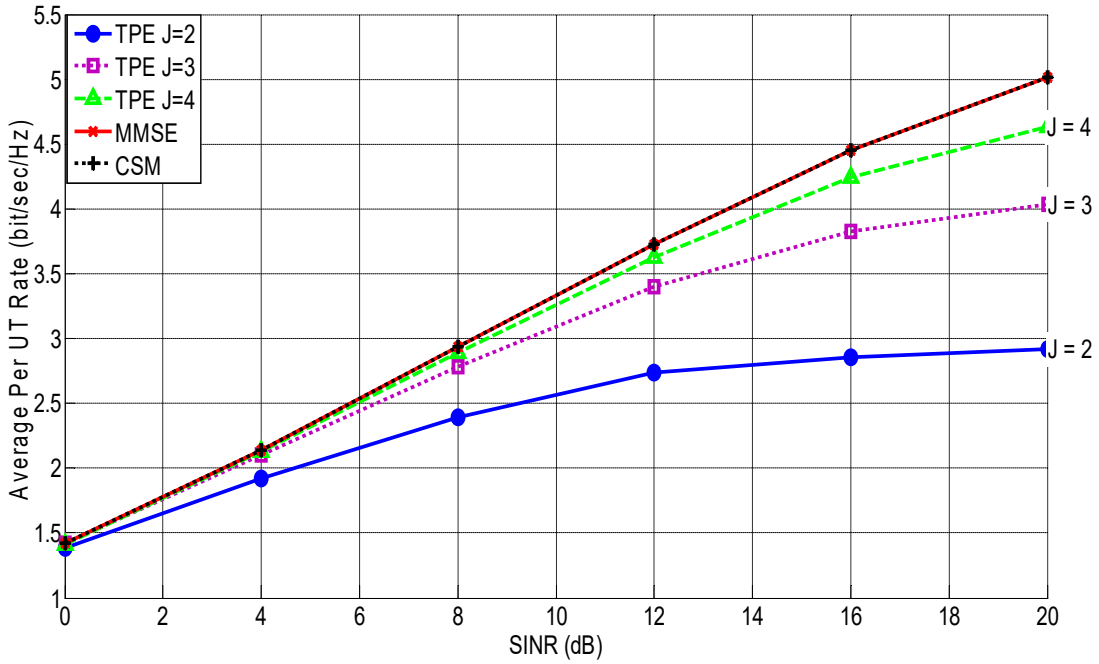

Figure 2.9: Average per UT rate vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding ($\boldsymbol{M = 512, K = 128,\ \tau = 0.15}$).

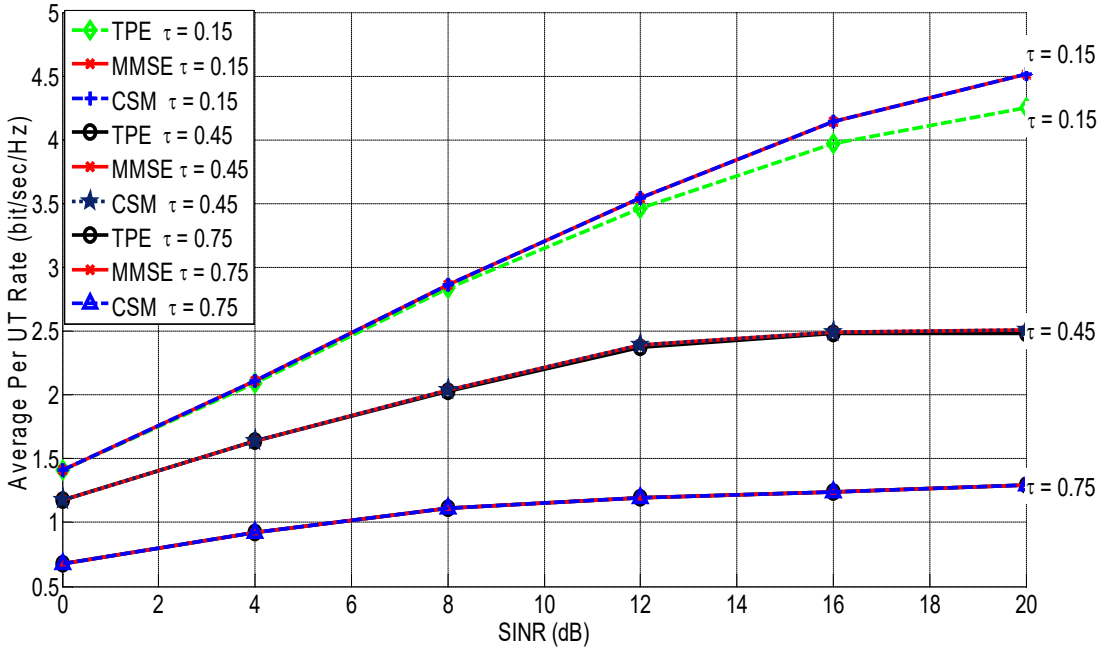

Figure 2.10: Average per UT rate vs. SINR for varying CSI errors at the BS ($M = 512, K = 128,\ J = 4$).

From Figure 2.10 below, the performance of the MMSE precoder, CSM-based precoder and the TPE precoder are analysed under three different levels of channel knowledge: $\tau \in \{0.15, 0.45, 0.75\}$ to obtain the corresponding SINR.

Then the SINR achieved by every precoder is used to compute the achievable rate from Shannon theorem. We consider an additive white Gaussian noise channel, so that the channel capacity is given by

$$R = log_2(1 + SINR) \qquad (bits/sec/Hz) \tag{2.31}$$

From which the achievable rate of the kth UT for massive MIMO downlink system can be written as

$$R_k = \mathbb{E}[log_2(1 + SINR_k)] \qquad (bits/sec/Hz) \tag{2.32}$$

With the worst CSI ($i.e.\ \tau = 0.75$) the MMSE precoder, CSM-based precoder and TPE precoder have nearly similar bitrates. Moreover, this similarity in bitrates is observed under low SINRs regardless of the value of τ.

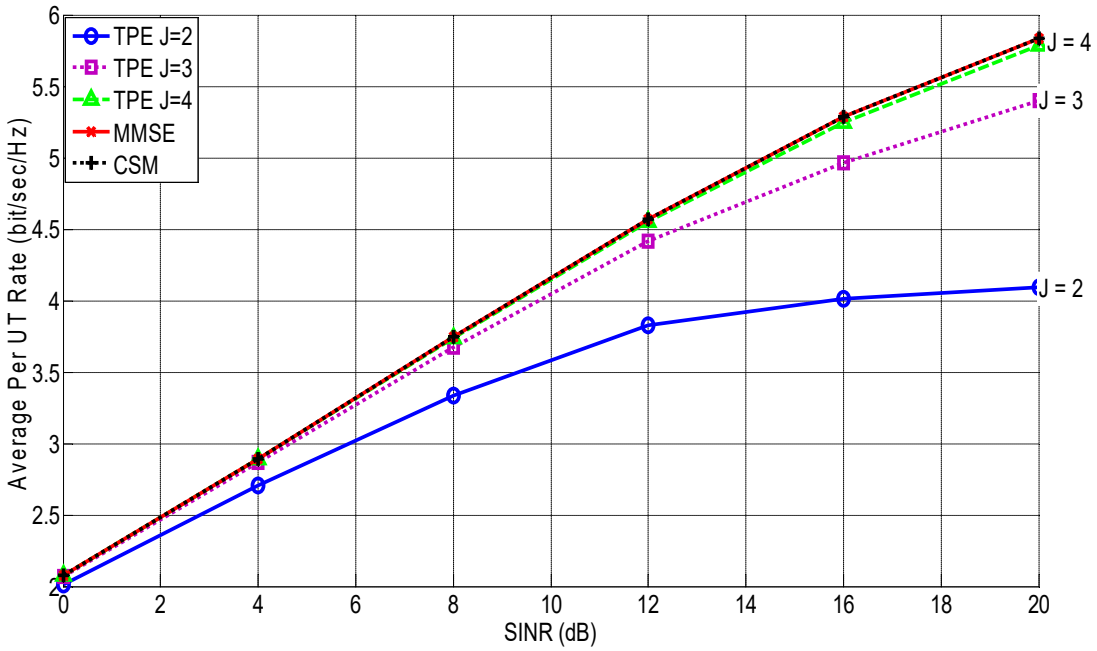

Figure 2.11: Average per UT rate vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding ($\boldsymbol{M = 512, K = 64,\ \tau = 0.15}$).

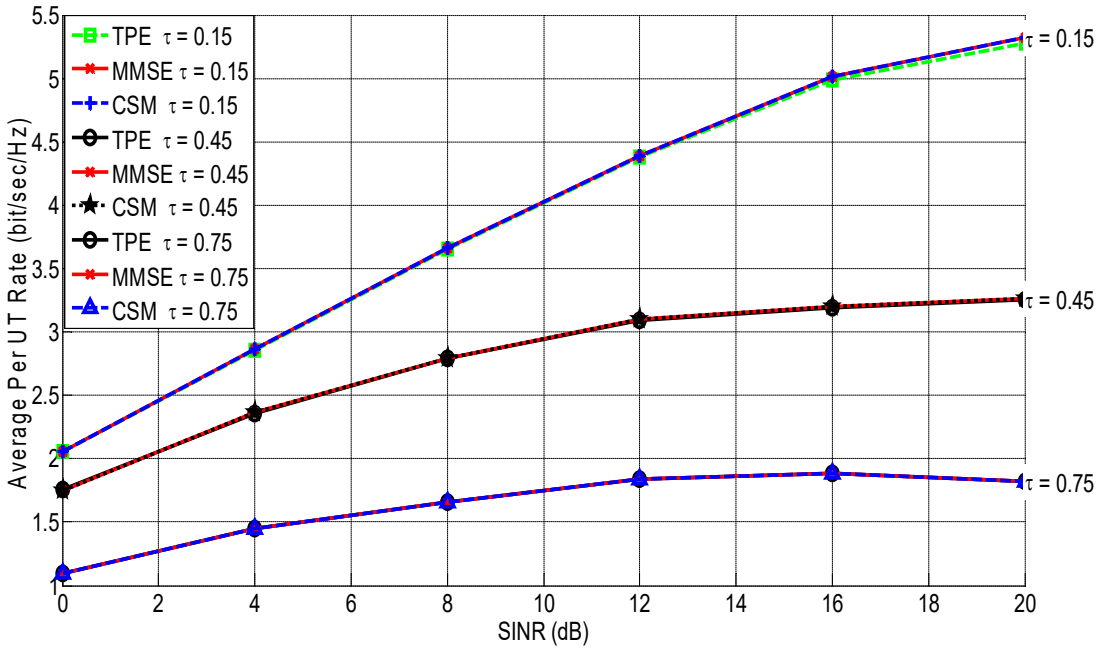

Figure 2.12: Average per UT rate vs. SINR for varying CSI errors at the BS ($\boldsymbol{M = 512, K = 64, J = 4}$).

Figure 2.11 and Figure 2.12 above demonstrates enhanced performance in bitrate when we decrease the UTs antennas (K) to 64 down from 128 with BS antennas (M) held constant at 512. This agrees with the theory of massive MIMO. Thus, increasing M infinitely with K kept smaller the massive MIMO system performance is greatly improved and nears the hypothetical performance of massive MIMO. Additionally, the TPE precoder performance significantly improves nearing that of the MMSE precoder and the CSM-based precoder. Consequently, we can enhance the performance of TPE precoder by significantly increasing the number of BS antennas (M) without necessarily increasing the order J of the polynomial. Hence, massive MIMO systems can utilize the TPE precoding scheme and achieve a good throughput if the number of BS antennas are made significantly greater in comparison to the UTs.

In Figure 2.13 below the BER performance when $M = 512$ and $K = 128$ for the three precoding schemes is shown. At the point when the SINR is low, all the three precoders give a high BER which reduces with increasing SINR and eventually reaches zero. With a lower J, the TPE precoder gives a high BER which decreases as J increases and when $J = 4$, its BER performance nears that of the MMSE precoder and the CSM precoder. This is normal since TPE precoder approximates the MMSE precoder, and with a lower J the approximation is poor, and it improves as J increases and eventually nears the MMSE bit error rate (BER) performance.

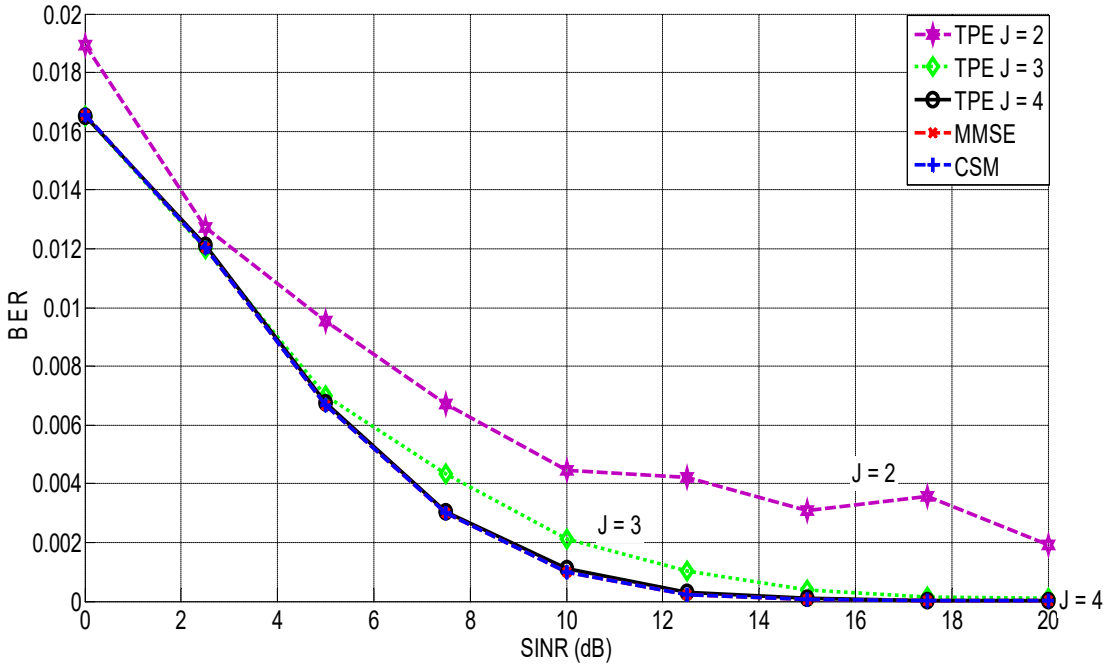

Figure 2.13: BER vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding ($\boldsymbol{M = 512, K = 128}$, $\boldsymbol{\tau = 0.15}$).

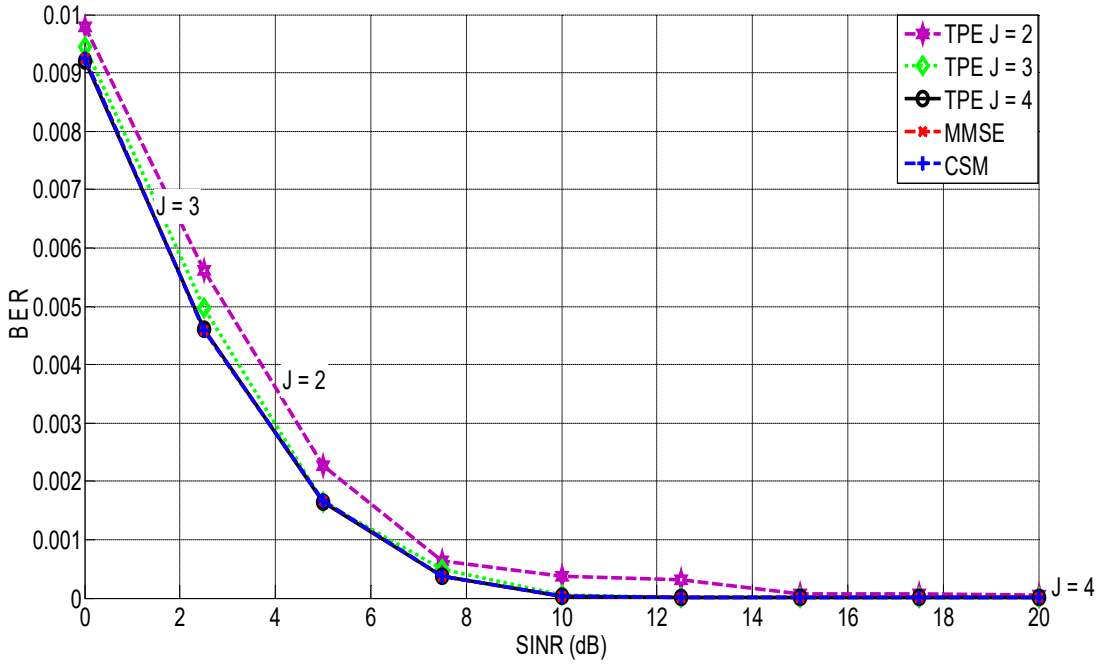

Figure 2.14: BER vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding ($\boldsymbol{M = 512, K = 64}$, $\boldsymbol{\tau = 0.15}$).

In Figure 2.14 below the BER performance when $M = 512$ and $K = 64$ for the three precoding schemes is shown. An improved performance is observed in the BER of all the three

precoding schemes when BS antennas are held constant and the UTs are halved. It can be observed that the BER performance of TPE precoder compares relatively well to that of the MMSE precoder even at a lower J. The reduction in UTs antennas is synonymous to increase in the BS antennas while holding the number of UTs antennas constant. This reduces the BER by enhancing the massive MIMO channel performance through exploitation of the massive MIMO diversity.

2.9 Massive MIMO Channel Estimation

For wireless communication systems, the CSI of each user is estimated by the BS using pilot sequences, and the approximated CSI is used to perceive uplink data and to create precoders for downlink transmission. But the maximum number of orthogonal pilot sequences is restricted [130]. This dictates that in multi-cell massive MIMO the pilot sequences are re-used in adjacent cells, which results in non-orthogonal pilot sequences being used in different cells. Consequently, the estimated CSI at the BS is contaminated by these identical pilot sequences from neighbouring cells. These inaccurate CSI estimates will affect the downlink beamforming and overall system performance negatively [131]. Diverse methods of pilot contamination mitigation have been proposed and implemented. The three key frontiers of pilot contamination mitigation are the pilot assignment, the channel estimation methods and precoding process [132]. Channel estimation methods can either be training based estimation, blind estimation or semi-blind estimation [133].

2.9.1 Training Based Estimation

In training based estimation we have the Least Square (LS) estimator, the Maximum A-Posteriori (MAP) estimator, Element Wise MMSE estimator (EW-MMSE) and the MMSE estimator [132], [133]. The LS is based on training data devoid of any prior information. If we have prior information, then the MAP estimator is used. The MMSE takes the statistical channel information into account when approximating the CSI [132]. In real networks, the number of pilot symbols is considerably smaller than the overall number of users in the network. Hence training-based approximation is subject to pilot contamination due to need for reuse of pilot sequences.

In [134] a training sequence is proposed that promises accurate MIMO CSI estimation. It is meant to deal with three key challenges: arbitrary channel and noise statistics with no specific

models, limitation on properties of the transmit signal, and signal design for massive MIMO. A multi-cell joint MMSE channel estimation is proposed in [135] and written as below.

$$\boldsymbol{Y}_l = \sqrt{P_r}\boldsymbol{\varphi}\boldsymbol{G}_l\boldsymbol{H}_l + \boldsymbol{N}_l \tag{2.33}$$

where $\boldsymbol{\varphi}$ represents the pilot, matrix used in channel estimation process. The scheme employs interference cancellation and joint processing. This in turn brings about the trade-off between the effectiveness and efficiency of the channel estimation at reasonable computational cost leading to improvement in overall system performance [135], [136].

An efficient distributed MMSE algorithm that exploits the strong spatial correlation among the elements of the antenna array to attain close to optimal channel estimates with low complexity is proposed in [136]. The method involved solving of reduced dimensional MMSE for each antenna and then repetitively sharing the information by collaboration among neighbouring array elements. To enhance the channel estimates, data-aided estimation technique was used. Stochastic geometry is invoked in quantifying the pilot sullying.

The Minimum Variance Unbiased (MVU) and the low-complexity channel estimator known as the Polynomial ExpAnsion Channel (PEACH) are suggested in [137]. The PEACH estimator approximates the MMSE estimator by substituting a polynomial expansion for the inverse of the equation. The coefficients of polynomials are designed to obtain minimal MSE for any fixed polynomial order, thus maintaining low complexity.

A frame structure that is staggered for massive MIMO single-cell with knowledge that specific users transmit training pilots at unoverlapped time is suggested in [138]. It was assumed that users need not to be coordinated to send pilots, and providing orthogonal pilots is therefore unnecessary. Two channel estimation methods were proposed to suppress interference i.e., Linear MMSE (LMMSE) and Orthogonal Projection-Based LS (OPLS) for staggered frame structure. It was in view of reducing MSE and estimating the channel response portion of the user that is orthogonal to other users.

Pilot sullying as Compressed Sensing (CS) problem is suggested in [139]. Within the beam domain the channel component is realized as a Gaussian mixture enabling the reconstruction of channel components with the best MSE performance using probabilistic Bayesian inferences [140], [141]. This exploits the inherent channel sparsity and requires quite low pilot overhead. The estimation process starts with a coarse estimation stage followed by a refinement stage to establish transmit or receive spatial frequencies. The use of joint sparsity to substantially reduce the pilot number for channel estimation based on CS is depicted in [142][143].

It was presented in [142] that an increase in BS antenna numbers increases the joint recovery probability of nonzero channel entries. The reliable solution to the model was acquired through the application of Block Optimized Orthogonal Matching Pursuit (BOOMP). While the channel estimation in [143] is based on weighted Homotopy approach which obtains a respectable channel estimation performance by utilizing sparse nature of MIMO channels. It embodies this through an information exchange strategy to greatly improve estimation accuracy and limit the required number of pilot signal through joint channel estimation.

In [144], filtering channel estimation based on a Wiener multidimensional filtering channel estimation algorithm is provided. This suggested a space-time frequency correlation function to decrease the complexity of the algorithm and to reduce the number of pilot signals by half equally. Adaptive Regularized Compressive Sampling Matching Pursuit (ARCoSaMP) algorithm is suggested in [145]. This achieves reconstruction precision by tailoring the support set and leveraging the method of regularization. By this the second range of atoms is realized even when the channel sparsity is uncertain in the support set.

2.9.2 Blind Estimations

Previously, blind estimation relied on asymptotic orthogonality of both channel vectors and data symbols sequences for the vast number of antennas and vast number of received uplink data signals [133]. To correctly separate channels to different users with blind channel estimation we need large antenna numbers and vast information samples for a single coherence interval [133].

In [146]it is argued that the fast-fading coefficient matrix in the channel is the covariance eigenvector matrix for the pilot sample obtained when the number of BS antennas is substantially increased. The study in [147]proposed Diagonal jacket-based estimation approach with the least square iterative projection to speed up channel estimation and reduce pilot sullying. Then the channel matrix is transformed to a lower-dimensional matrix by multiplying the obtained pilot sample in the left with a normalized base vector matrix being transposed in conjugate. CSI is acquired in [25]by applying the technique of subspace estimation using Eigenvalue Decomposition (EVD) on the covariance matrix of the samples obtained, but this only yields a scalar ambiguity. The notational distinction between preferred and interfering channels is not essential in blind channel estimation.

In [26], [148]–[150], the power control and power controlled blind channel estimation technique for cellular systems is discussed. To overcome this, MAP is employed in subspace

channel estimation [133], [151] to enhance performance although with increased complexity. With a limited number of uplink data, blind estimation method can deliver poor results as the subspace estimates might be inaccurate.

2.9.3 Semi-Blind Estimations

Semi-blind channel estimation performance increases with a growing number of antennas and uplink data signals which make it suitable for massive MIMO systems. Semi-blind estimator offers a strict improvement when a smaller number of uplink data signals are available. In combining the training signals and the uplink data signals in estimation process, we simply extend the training phase by the data phase with only statistical information [133]. Different semi-blind estimation approaches have been fronted by several researchers.

According to [152] Subspace-based semi-blind estimator is suggested, which takes advantage of the approximate orthogonality of channel vectors to resolve the ambiguity totally without losing the CSI. In [153] the Projection Approximation Subspace Tracking (PAST) is presented. It offers lower complexity with good tracking capabilities. The semi-blind estimator based on Fast Single Compensation Approximated Power Iteration (FSCAPI) is implemented in [154] with promise of fast convergence and good orthogonality. The Independent Component Analysis (ICA) is proposed in [155]. The approach proposed does not include cell cooperation or priori knowledge about the respective channels associated with all cells. The target cell UL data is first derived from the observed mixture signal, and the channel LS estimate is then obtained based on the demodulated UL data.

An enhanced least square-scaled least square channel estimation method is used in [156] to enhance the traditional semi-blind channel estimation. This is performed in two stages for improved channel estimation: first the channel matrix noise interference is minimized and then the original channel estimate is increased. Semi-blind estimator is implemented in [157] that rely on optimum pilots. First the channel matrix column space is approximated, and then with the aid of optimum pilots estimate the channel matrix. Though this work was only implemented for uncorrelated channels.

In [133] it is suggested that the MAP be used employing the Limited-memory Broyden-Fletcher-Goldfarb-Shanno Algorithm (L-BFGS). This is a limited-memory quasi-Newton method that exhibits low-complexity. Whitening Rotation (WR) semi-blind algorithm for channel estimation is investigated in [158]. To enhance the accuracy of the WR algorithm, it is combined with the conventional Discrete Fourier Transform (DFT) based channel estimator.

2.9.4 Simulation and Analysis of Selected Channel Estimation Models

In this simulation we compare the Normalised MSE (NMSE) of the MMSE, EW-MMSE and the LS channel estimators under same assumptions for the case of a spatially correlated channel as in Figure 2.15 below. The LS offers the highest NMSE and MMSE offers the least NMSE. It can be observed that LS is a suboptimal estimator. The EW-MMSE has an NMSE close to that of MMSE when SNR is low, but as SNR increases then the EW-MMSE NMSE nears that of the LS estimator. The NMSE of these channel estimation techniques are presented as follows.

$$MSE\left(A_{li}^{j}\right)_{MMSE} = tr\left(R_{li}^{j}\right) - 2\sqrt{p_{li}}\tau_p \Re\left(tr\left(R_{li}^{j}A_{li}^{j}\right)\right) + \tau_p tr\left(A_{li}^{j}\left(\Psi_{li}^{j}\right)^{-1}\left(A_{li}^{j}\right)^{H}\right) \tag{2.34}$$

where A_{li}^{j} is a deterministic matrix that specifies the estimation scheme, R_{li}^{j} is the correlation matrix, Ψ_{li}^{j} is the sum of correlation matrix of the UTs that utilize the same pilot sequence, i is the UT, l is the neighbouring cells to cell j, τ_p is the pilot sequence samples, and p_{li} is the UL transmit power. For the EW-MMSE we have

$$MSE_{EW-MMSE} = tr\left(R_{li}^{j}\right) - \sum_{m=1}^{M} \frac{p_{li}\tau_p\left(\left[R_{li}^{j}\right]_{mm}\right)^2}{\sum_{l',\ i'\epsilon\, p_{li}} p_{l'i'}\tau_p\left[R_{l'i'}^{j}\right]_{mm} + \sigma_{UL}^2} \tag{2.35}$$

where $\left[R_{l'i'}^{j}\right]_{mm}$ is the estimation error variance. For the LS

$$MSE_{LS} = tr\left(\frac{p_{li}\tau_p}{p_{l'i'}}R_{l'i'}^{j} + \frac{\sigma_{UL}^2}{p_{li}\tau_p}I_{M_j}\right) \tag{2.36}$$

These are then the formulation used to realize the Figure 2.15.

The simulation algorithms and parameters used to obtain Figure 2.16, are obtained from [137] and used as presented in this article without variations except that we used total training power as the reference parameter.

In Figure 2.16, we assume a system with 24 transmitting antennas and 4 receive antennas. The NMSE performance of the MMSE, the MVU and the PEACH channel estimators is contrasted using minimized training matrices based on MSE. It is observed that the MMSE estimator

gives the best performance followed by the PEACH estimator while the MVU estimator gave the worst performance. The MVU estimator lacks the knowledge of the channel statistics leading to poor estimation performance.

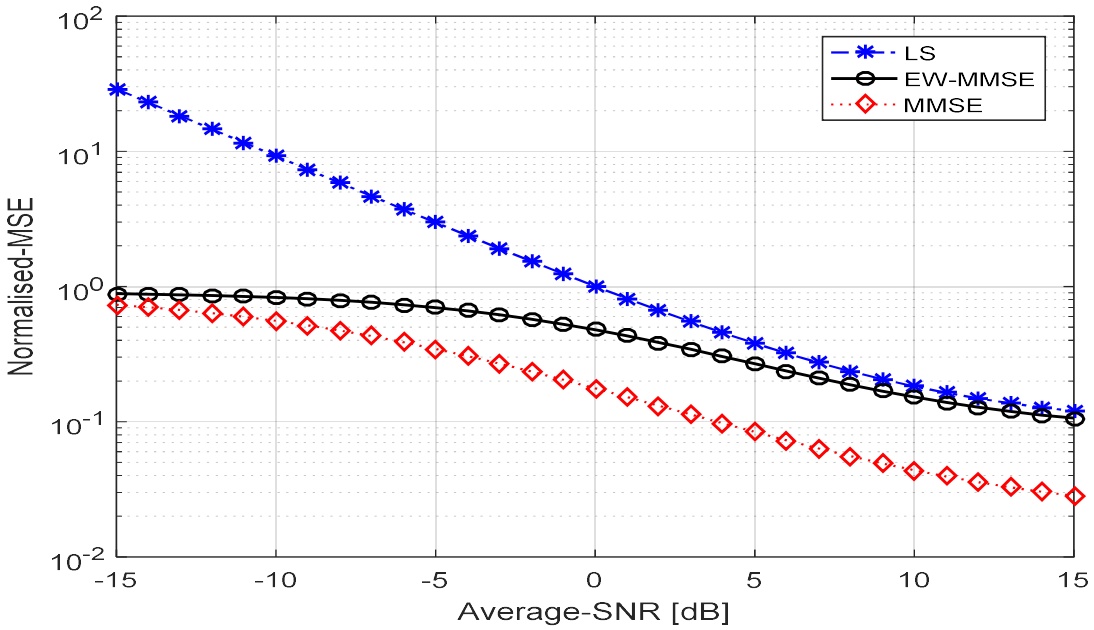

Figure 2.15: The NMSE corresponding to approximation of spatially correlated channel with $\sigma_{\varphi} = 9^0$.

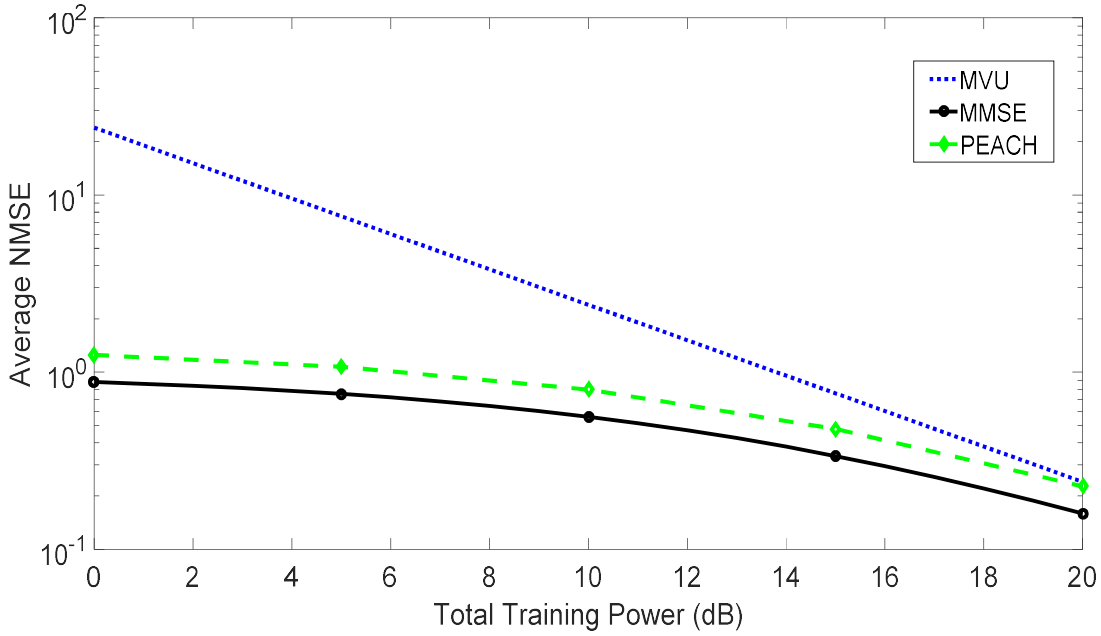

Figure 2.16: Average NMSE vs. total training power at 24 transmitting antennas and 4 receiving antennas.

The simulation algorithms and parameters used to obtain Figure 2.17, are obtained from [153], [154] and used as presented in this articles without variations and we used total training power as the transceiver antennas as the reference parameter.

In Figure 2.17 below, the NMSE performance of the MMSE, FSCAPI and PAST channel estimators is analysed against the varying number of transmitting antennas. It can be observed that the FSCAPI gives the lowest NMSE while the MMSE gives the highest NMSE. Thus, Semi-blind estimators give better approximation of the channel than the linear estimators [137], [153].

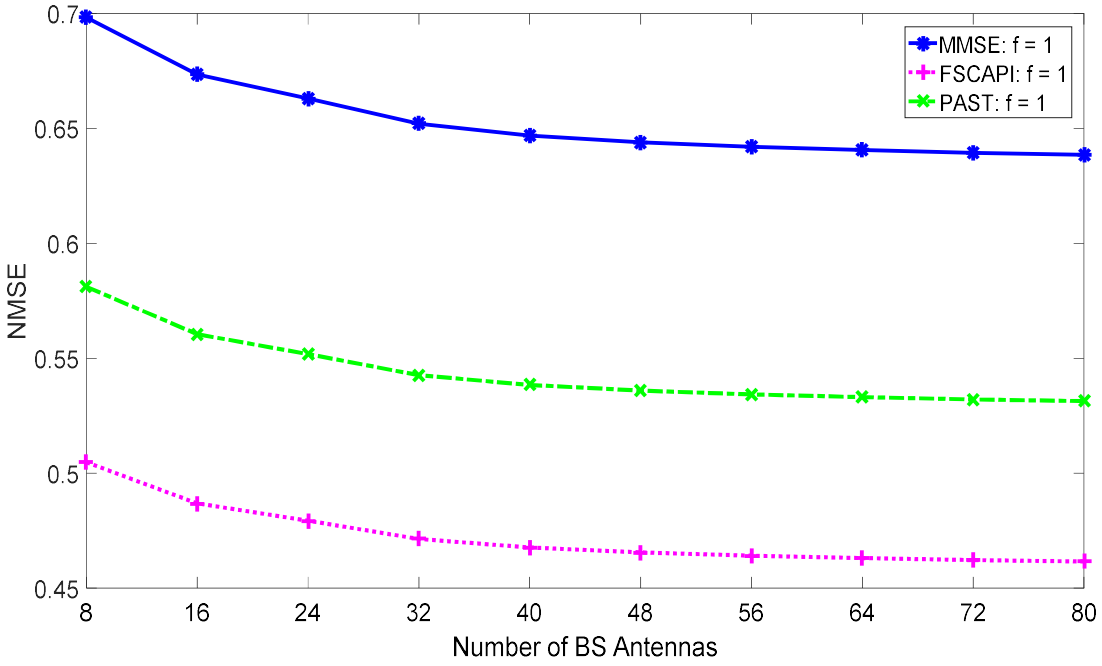

Figure 2.17: Comparison of NMSE vs. number of transmitting antennas.

2.10 Massive MIMO Partially Centralized Cloud Radio Access Network (MPC-RAN)

The mobile internet technology has tremendously increased forcing the telecom operators to review their operating expenses upwards though with dwindling income. This points to an increase in BS to shoulder the increased UT needs which can be too costly in deploying and managing. To this end, C-RAN comes in handy in meeting these UT demands by relying on a new innovative architecture through partial centralization of BS with cooperative solution amongst many operators. This offers reduced cost, enhanced energy efficiency with a centralized network [159].

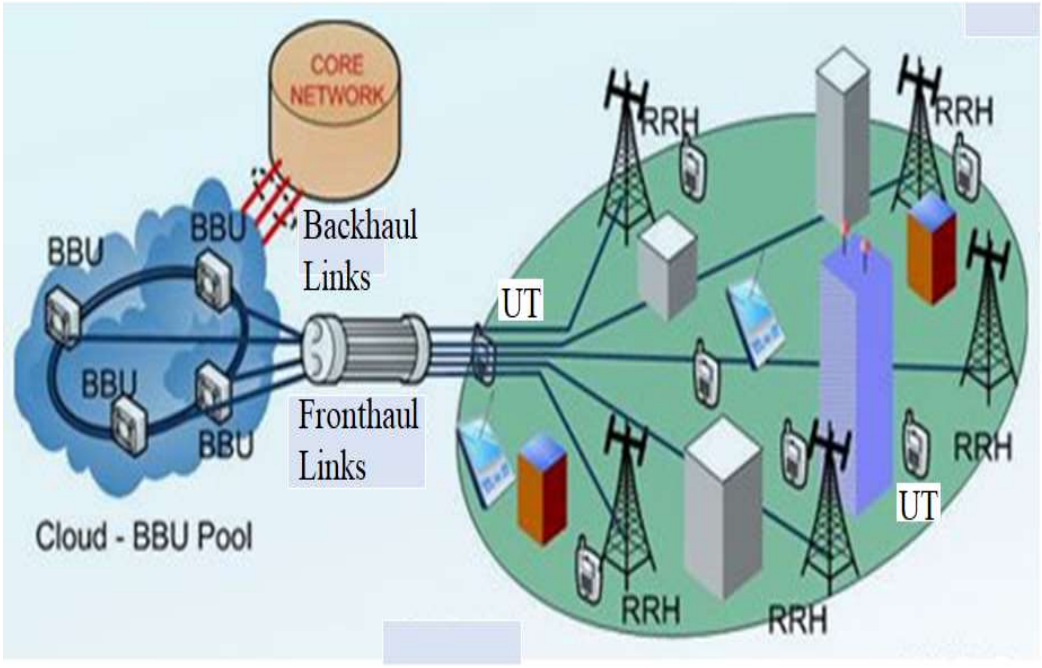

Figure 2.18: Cloud radio access network [160].

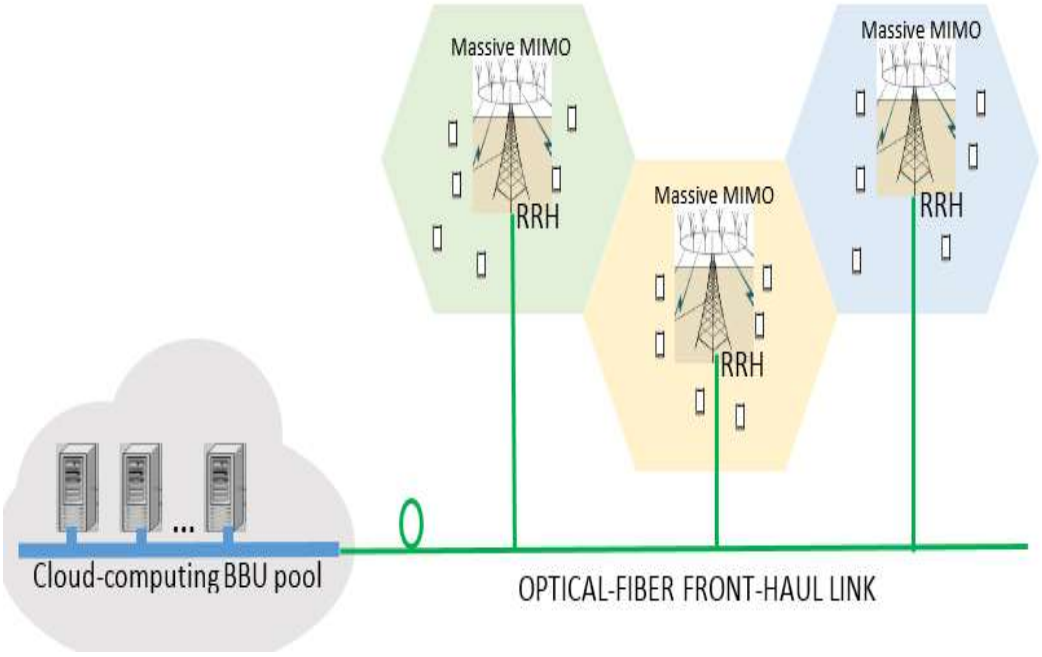

Figure 2.19: MPC-RAN architecture [161].

The portion of the RAN that can be moved to central location by the carriers is the Baseband Processing unit (BBU) which then can made to serve several Remote Radio Heads (RRH). The RRUs are connected to the respective BBUs through high bandwidth links to fulfil the demands of the numerous RRHs. These links are referred to as the fronthaul which can be implemented

by a variety of technologies, including optical fibre, cellular or millimetre wave communication, [162], [163] independently or as a hybrid. The architecture of the C-RAN is depicted as in Figure 2.18.

The key enabling technologies for 5G networks has been singled out to be the C-RAN and the massive MIMO as they promise to lower operational cost and enhance performance. When massive MIMO is utilized in the RRH, fronthaul becomes the limiting factor because of its inherent finite capacity [164]. One of the foreseen solutions to fronthaul finite capacity is to split functions so that some are performed at the RRH and others at the BBU. Considering this suggested architecture, the RRH are charged with performing basic functions like beamforming and the BBU is left to carry out the digital functions including channel estimation. This then renders the fronthaul traffic to be mainly dependent on UT data rates and not on the number of antennas [165], [166]. This leads to the Massive MIMO Partially Centralized C-RAN (MPC-RAN) [167]. The MPC-RAN is illustrated in Figure 2.19 below.

When the partial centralization is combined with distributed cooperation for the case where RRHs are inter-connected it greatly mitigates capacity constraint and time latency on MPC-RANs fronthaul. Thus the common notion is to configure the topology to be adaptive in a way to strike a common balance between the fronthaul constraints and the distributed cooperative processing complexity [163].

It was mentioned earlier that channel estimation remains to be performed at the BBU in MPC-RAN. Thus, it will be thoughtful to see how to reduce the pilot training data for channel estimation. One way is to use compressed primary pilot data before transmission to the BBU. Another fact is to realise that the BBU has enormous computation resources (multicore systems), and thus the channel estimation process can be quickened by use of parallelization processing to take advantage of these resources. Therefore, a channel estimation process that is highly parallelizable will be of great importance to enhance the channel estimation process. These forms the basis of the issues around channel estimation that will be investigated in this work.

2.11 Summary

Massive MIMO provides an appealing opportunity for 5G network systems and beyond by availing increased throughput. The channel hardening phenomenon as a result of spatial multiplexing minimizes small scale fading within system network increasing the throughput. Though massive MIMO suffers from pilot contamination, an array of techniques including

antenna array configurations, channel measurements, channel characterization, channel estimation as well as encoding has been envisioned to mitigate it. The implementation of PHY layer security is envisioned to be exploited as a complement to existing security mechanisms or to replace them completely. The MPC-RAN with partial centralization avails the much-needed capacity efficiency enhancement in 5G networks since it renders the fronthaul traffic to be mainly dependent on UT data rates and not on the number of antennas.

CHAPTER 3

Rapid Numerical Algorithms Precoding: Single Cell

3.1 Introduction

In present wireless communication, MIMO processes have risen considerably because it can increase the capacities and quality of wireless devices in total [168]. Precoding is an inherent process in DL massive MIMO, where each data flag is spatially coordinated in the direction of its anticipated UT. However, the momentary CSI required in massive MIMO precoding plan need to be highly precise [114] which makes it bulky to realize in the long run. Channel variations in massive MIMO are rapid and approximately in the order of a tenth of the frequency and time. On the off chance that coding is implemented across intervals of numerous channels, it was shown for a massive MIMO system with M transmitting antennas with K UTs, the capacity increases as a linear variant of the mean (M, K) $\log_2(1+\text{SNR})$ [18], [102]. The enhancement in capacity is implemented in massive MIMO system with M BS antennas transmitting to K UTs for the TDD case.

The BS is made to communicate with the UT within the same time-frequency resource for further enhancing the capacity. The realization of this capacity enhancement is inherently associated with inter-user interference that degrades the massive MIMO system performance. The DPC is engaged in an attempt to minimize the inter-user interference where the information symbols are encoded jointly by the transmitter for all the UTs [103], [169]. The practical implementation of DPC based transmission system is hampered by its inherent complexity. The pre-processing technologies known as precoding schemes performed on the transmit-side provides the much-needed alternative. For the massive MIMO system, optimal mitigation of pilot contamination can be realized using basic precoding schemes such as Eigen BF and MF [170], although they need the inversion of channel matrix containing all users. Further research points to the need to incorporate linear precoding processes like RZF for practical realization of BS antennas [171]. The RZF precoding suffers from enormous problem of matrix inversion which translates to a high computing complexity [106], [110], [111], [128].

Direct, expansions, and iterative methods are three main categories to calculate large matrix inverse for linear precoding. Direct methods suffer from high complexity as it depends mainly on transferring the matrix to be inverted into a multiplication of simple matrices like QR and Chelosky decomposition [172]. The second category, expansion methods, transfers the inverse of a matrix into a series of matrix vector products like Neumann series (NS) [173], although NS has slow convergence rate. Iterative methods belong to the family of solving linear equations such as Richardson method [174], conjugate gradient (CG) method [175], successive over relaxation (SOR) [176], symmetric successive over relaxation (SSOR) [177], Gauss-Seidel (GS) method [178], and the rapid numerical algorithm (RNA) [179]. They have acceptable performance in massive MIMO systems. The RNA find the inverse of A^{-1} in less number of iterations and with high accuracy [180], the reason we chose to implement it in this work.

In this work we propose a class of rapid numerical algorithm (RNA) precoders to substitute the intricate matrix inversion process in conventional linear precoders. The use of RNA in encoding has been implemented in [11]. Different from [11], the truncated polynomial estimation (TPE) is combined with the iterative method of Schulz to perform matrix inversion. It is demonstrated further that the first three terms provide the fastest iterative process convergence for matrix inversion. Consequently, the complexity is reduced but with a performance that compares favourably with that of RZF as opposed to [21]. This renders the RNA precoders scalable while minimizing computational complexity with comparable performance to that of existing linear precoders. We consider a channel model with transmit correlation embodied at the BS. Performance of the precoder schemes is evaluated in relation to Signal-to-Interference-and-Noise Ratio (SINR), SE and BER.

First, we model the iterative linear precoder and then adopt it individually to the various linear precoding schemes. We simulate the channel model in MATLAB then evaluate the SINR, SE and BER in relation to perfect and imperfect CSI to evaluate the performance over the modelled channel for massive MIMO DL system.

3.2 System Model Single Cell

We examine a DL massive MIMO system having K UTs each with a single antenna and communicate concurrently with M antennas at the BS. We consider un-coded massive MIMO downlink system. The kth UT data symbol, $\boldsymbol{x}_k \in \mathbb{C}^{K\times 1}$, is modified through the precoding matrix $\boldsymbol{A}$ and $\mathbb{E}[|\boldsymbol{x}_k|^2] = 1$. The transmitted signal vector from the BS can be written as

$$\boldsymbol{u} = \sum_{k=1}^{K} \boldsymbol{A}_{m,k}\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x} \tag{3.1}$$

where $\boldsymbol{A} \in \mathbb{C}^{M\times K}$ is the matrix of precoding and $\boldsymbol{x} = [x_1\ x_2 \dots x_K]$ represents a data vector. The transmit vector u average transmit power is constrained by $\mathbb{E}[|\boldsymbol{u}|^2] = tr(\boldsymbol{A}^*\ \boldsymbol{A}) = P_{tr}$. The channel matrix $\boldsymbol{H} \in \mathbb{C}^{M\times K}$, and the received kth UT signal vector is given as

$$\boldsymbol{y}_k = \boldsymbol{h}_k^T\boldsymbol{u} + n_k \tag{3.2}$$

where n_k represents the noise vector of the kth UT and $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}$ is the BS random channel vector with the kth UT. Then the received signal vector in total is expressed as

$$\boldsymbol{y} = \boldsymbol{H}^T\boldsymbol{u} + n \tag{3.3}$$

where $\boldsymbol{u} = [u_1 u_2 \dots u_K]$ and $\boldsymbol{H} = [h_1\ h_2 \dots h_K] \in \mathbb{C}^{M\times K}$. From [21], [106]–[109], diverse performance specification are used to characterize linear precoding schemes.

3.2.1 Conventional RZF Precoding

RZF and MMSE precoding are some of the optimal linear precoding schemes in the massive MIMO downlink network. The best and technically feasible method of precoding used today is obviously RZF precoding [118]. The average power of each antenna transmitting at BS is regulated and consequently the RZF precoder is realized with the help of Lagrangian optimizing strategy. The optimized RZF precoder is realized by [111] as

$$\boldsymbol{A}_{RZF} = \beta_{RZF}\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H + \xi\boldsymbol{I}_K)^{-1} = \beta_{RZF}\boldsymbol{H}^H\boldsymbol{W}_{RZF}^{-1} \tag{3.4}$$

where $\boldsymbol{W}_{RZF} = (\boldsymbol{H}\boldsymbol{H}^H + \xi\boldsymbol{I}_K)$. ξ represents the regularizing factor which is selected adaptively in relation to varying CSI [125]. And then β_{RZF} is calculated as:

$$\beta_{RZF} = \sqrt{\frac{K}{tr(\boldsymbol{W}_{RZF}\boldsymbol{W}_{RZF}^H)}} \tag{3.5}$$

From which, we can express the information signal $\boldsymbol{u}$ by

$$\boldsymbol{u} = \boldsymbol{A}_{RZF}\boldsymbol{x} = \beta_{RZF}\boldsymbol{H}^H\boldsymbol{W}_{RZF}^{-1}\boldsymbol{x} \tag{3.6}$$

From the above equations it can be demonstrated that the corresponding channel matrix can be expressed as

$$\boldsymbol{H}_{equivalent} = \boldsymbol{H}\boldsymbol{A}_{RZF} = \beta_{RZF}\boldsymbol{H}\boldsymbol{H}^{H}\boldsymbol{W}_{RZF}^{-1} \tag{3.7}$$

The information pertaining the instantaneous channel implementation $\widehat{\boldsymbol{h}}_k$, with $k = 1, \dots, K$, for each UT is imperfect at the transmitter. Using the Gauss-markov theorem from [182] we obtain

$$\widehat{\boldsymbol{h}}_k = \sqrt{1-\tau^2}\boldsymbol{h}_k + \tau\boldsymbol{n}_k \tag{3.8}$$

where $\boldsymbol{h}_k$ represents the real channel and $\boldsymbol{n}_k \sim C\mathcal{N}\left(\boldsymbol{O}_{M\times 1}, \frac{1}{K}\Phi\right)$ accounts for independent error. The immediate CSI quality is captured through the scalar variable $\tau \in [0,1]$. And $\boldsymbol{H} = [\boldsymbol{h}_1\ \boldsymbol{h}_2 \dots \boldsymbol{h}_K] \in \mathbb{C}^{M\times K}$ represents the joint imperfect information for all user channels. Therefore, the RZF precoding assuming imperfect channel is given by

$$\widehat{\boldsymbol{A}}_{RZF} = \beta_{RZF}\widehat{\boldsymbol{H}}^{H}\left(\widehat{\boldsymbol{H}}\widehat{\boldsymbol{H}}^{H} + \boldsymbol{\xi}\boldsymbol{I}_K\right)^{-1} \tag{3.9}$$

If we let the matrix $\boldsymbol{A}_K$ to be the matrix $\boldsymbol{A} = \boldsymbol{A}_{RZF}$ where column $\boldsymbol{a}_k$ is eliminated, the resultant SINR for the kth UT can be evaluated as

$$SINR_k = \frac{\boldsymbol{h}_k\boldsymbol{a}_k\boldsymbol{a}_k^{H}\boldsymbol{h}_k^{H}}{\boldsymbol{h}_k^{H}\boldsymbol{A}_k\boldsymbol{A}_k^{H}\boldsymbol{h}_k + \sigma^2} \tag{3.10}$$

with $\boldsymbol{a}_k^{H}, \boldsymbol{h}_k^{H}$ and $\boldsymbol{A}_k^{H}$ representing one-to-one Hermitian transpose of $\boldsymbol{a}_k, \boldsymbol{h}_k$ and $\boldsymbol{A}_K$.

As a requirement the random precoding matrix is re-computed at each coherence time in RZF precoding. This precoding method relies on inversion of matrices that presents computational demands when considering large system [183]; the operations needed in total scales by $\mathcal{O}(K^2M)$ and the inversion algorithms available are complex to realize during hardware implementation. In RZF precoding interference mitigation leverages on the inversion of matrices pointing to the need to bring forward precoding schemes with less complexity and efficient in suppressing the interference [107], [109], [183].

3.2.2 TPE Precoding

It has been demonstrated in [21], [184] that every Hermitian matrix inversion can generally be represented by a matrix polynomial given by

$$\boldsymbol{W}_{RZF}^{-1} = \alpha\big(\boldsymbol{I} - (\boldsymbol{I} - \alpha\ \boldsymbol{W}_{RZF})\big)^{-1} = \alpha\sum_{i=0}^{\infty}(\boldsymbol{I} - \alpha\ \boldsymbol{W}_{RZF})^{i} \tag{3.11}$$

The contribution of the lower terms is significant because for $(\boldsymbol{I} - \alpha\ \boldsymbol{W}_{RZF})^{i}$ the eigenvalues tend to zero as i is incremented. Consequently, this permits the estimation of TPE using the first K terms. We employ this expression in approximating the RZF precoding using matrix polynomial [128]. From $\boldsymbol{A}_{RZF}$ given in equation (3.4), we have,

$$\begin{aligned}\beta_{RZF}\boldsymbol{H}^{H}(\boldsymbol{H}\boldsymbol{H}^{H} + \xi\boldsymbol{I}_{K})^{-1} \approx \textstyle\sum_{i=0}^{J-1}\Big(\beta_{RZF}\alpha\sum_{r=i}^{J-1}\binom{r}{i}(1 - \\ \alpha\xi)^{r-i}(-\alpha)^{i}\Big)(\boldsymbol{H}\boldsymbol{H}^{H})^{i}\boldsymbol{H}^{H}.\end{aligned} \tag{3.12}$$

By inspection, the equation (3.12) yields a precoding matrix structured given by

$$\boldsymbol{A}_{TPE} = \sum_{i=0}^{J-1}\omega_{i}(\boldsymbol{H}\boldsymbol{H}^{H})^{i}\boldsymbol{H}^{H} \tag{3.13}$$

where $\omega_0, \dots, \omega_{J-1}$ are scalar coefficients. The feasible expression of ω_i is the bracketed term in equation (3.12), for these values optimality is not obtained for the case when $J < \infty$, and again the power constraint $\frac{1}{K}tr(\boldsymbol{A}.\boldsymbol{A}^{H}) = P_{tr}$ is not satisfied. Thus J determines the order in the TPE precoder. Denoting a transmit signal corresponding to a user on channel z as $\boldsymbol{u}^{(z)}$, it permits us to use TPE precoding and express the transmit signal by [128]

$$\boldsymbol{u}^{(z)} = \boldsymbol{A}_{TPE}\boldsymbol{x}^{(z)} = \sum_{i=0}^{J-1}\omega_{i}\widehat{\boldsymbol{u}}_{i}^{(z)} \tag{3.14}$$

where $\boldsymbol{x}^{(z)}$ gives the vectorised information, symbols used at a channel z with

$$\widehat{\boldsymbol{u}}_{i}^{(z)} = \begin{cases}\boldsymbol{H}\boldsymbol{x}^{(z)}, & i = 0 \\ \boldsymbol{H}\Big(\boldsymbol{H}^{H}\widehat{\boldsymbol{u}}_{i-1}^{(z)}\Big), & 1 \le i \le J-1\end{cases} \tag{3.15}$$

Thus the J values of the TPE precoder can be iteratively calculated using equation (3.15) [184]. Then computing SINR as presented in equation (3.10), we obtain that $\boldsymbol{a}_k = \boldsymbol{A}\boldsymbol{e}_k$ and $\boldsymbol{h}_k^{H}\boldsymbol{A}_k\boldsymbol{A}_k^{H}\boldsymbol{h}_k = \boldsymbol{h}_k^{H}\boldsymbol{A}\boldsymbol{A}^{H}\boldsymbol{h}_k - \boldsymbol{h}_k^{H}\boldsymbol{a}_k\boldsymbol{a}_k^{H}\boldsymbol{h}_k$, with $\boldsymbol{e}_i$ being the ith row of the identity

matrix $\boldsymbol{I}_K$. Substituting equation (3.13) into equation (3.10), the SINR corresponding to TPE precoder can be expressed as

$$\boldsymbol{SINR}_k = \frac{\boldsymbol{w}^H \boldsymbol{G}_k \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{C}_k \boldsymbol{w} - \boldsymbol{w}^H \boldsymbol{G}_k \boldsymbol{w} + \sigma^2} \tag{3.16}$$

where $\boldsymbol{w} = \left[\omega_0, \dots, \omega_{J-1}\right]^T$ and the entries of the matrices $\boldsymbol{G}_k, \boldsymbol{C}_k \in \mathbb{C}^{J \times J}$ are given as

$$[\boldsymbol{G}_k]_{i,m} = \boldsymbol{h}_k^H (\boldsymbol{H}\boldsymbol{H}^H)^i \boldsymbol{H}\boldsymbol{e}_k \boldsymbol{e}_k^H \boldsymbol{H}^H (\boldsymbol{H}\boldsymbol{H}^H)^m \boldsymbol{h}_k \tag{3.17}$$

$$[\boldsymbol{C}_k]_{i,m} = \boldsymbol{h}_k^H (\boldsymbol{H}\boldsymbol{H}^H)^{i+m+1} \boldsymbol{h}_k \tag{3.18}$$

where $i = 0, \dots, J-1$ and $m = 0, \dots, J-1$.

The value J determines the performance in TPE precoders, with poor performance corresponding to smaller values of J and enhanced performance to larger values of J. But larger values of J corresponds to an increase in hardware needed to realize the TPE precoder and consequently an increase in cost. This served as the basis for formulating the RNA precoder.

3.2.3 Rapid Numerical Algorithms (RNA) Precoding

3.2.3.1 RNA Inverse Formulation

From (3.4) the inversion of $\boldsymbol{W}_{RZF}$ is expected prompting a high computational complexity since it is an enormous matrix of dimension $K \times K$. The RNA-based scheme is utilized to sidestep complex matrix inversion in RZF precoding. The initial phase in applying this technique is to set the matrix $\boldsymbol{W}_{RZF} = \boldsymbol{W}_{RNA}$. We then use the Cholesky decomposition on $\boldsymbol{W}_{RNA}$ [126]

$$\boldsymbol{W}_{RNA} = \boldsymbol{L}\boldsymbol{L}^H \tag{3.19}$$

with $\boldsymbol{L}$ being the lower triangular matrix. Then it follows that

$$\boldsymbol{W}_{RNA}^{-1} = (\boldsymbol{L}^H)^{-1} \boldsymbol{L}^{-1} \tag{3.20}$$

Hence the calculation of the inverse of $\boldsymbol{W}_{RNA}$ matrix is basically reduced to calculation of the inverse of matrix $\boldsymbol{L}$. Therefore, the computation of asymmetric matrix inverse is fundamentally reduced into computing an inverse of a lower triangular matrix. With the knowledge that a lower triangular matrix inverse is also a lower triangular, the intricacy of obtaining the matrix

decreases systematically. To support a low complexity in computing we decompose the matrix $\boldsymbol{L}$ into

$$\boldsymbol{L} = \boldsymbol{B} + \boldsymbol{L}' \tag{3.21}$$

with $\boldsymbol{B} = diag(\boldsymbol{L})$ and $\boldsymbol{L}' = (\boldsymbol{l}'_1, \boldsymbol{l}'_2, \dots, \boldsymbol{l}'_{K-1}, 0)$ is an $\boldsymbol{L}$ based matrix with the diagonal elements in $\boldsymbol{L}$ replaced with zero.

We let $\boldsymbol{F} = \boldsymbol{L}$ be a non-singular matrix with dimensions $K \times K$ whose inverse we are required to compute. Also, we let $\boldsymbol{R}_k$ be the approximated inverse at the kth iteration. The residual matrix gives us an indication on the discrepancy of the calculated inverse from the actual inverse of the matrix $\boldsymbol{F}$. The residual matrix $\boldsymbol{E}_k$ is given by

$$\boldsymbol{E}_k = \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k \tag{3.22}$$

On computing the first inverse $\boldsymbol{R}_k$ the resultant residual matrix $\boldsymbol{E}_k$ is obtained. Rearranging equation (3.22) we obtain,

$$\boldsymbol{F}\boldsymbol{R}_k = \boldsymbol{I} - \boldsymbol{E}_k \tag{3.23}$$

$$(\boldsymbol{F}\boldsymbol{R}_k)^{-1} = (\boldsymbol{I} - \boldsymbol{E}_k)^{-1} \tag{3.24}$$

$$\boldsymbol{F}^{-1} = \boldsymbol{R}_k(\boldsymbol{I} - \boldsymbol{E}_k)^{-1} \tag{3.25}$$

Which is a power series in $\boldsymbol{E}_k$. Hence,

$$\begin{aligned} \boldsymbol{F}^{-1} &= \boldsymbol{R}_k(\boldsymbol{I} - \boldsymbol{E}_k)^{-1} \\ &= \boldsymbol{R}_k \sum_{k=0}^{\infty} (\boldsymbol{E}_k)^k = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{E}_k + \boldsymbol{E}_k^2 + \cdots) \end{aligned} \tag{3.26}$$

Truncating the infinite series to the first two terms gives,

$$\boldsymbol{F}^{-1} = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{E}_k) \tag{3.27}$$

$$\boldsymbol{F}^{-1} = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k) = \boldsymbol{R}_k(2\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k) \tag{3.28}$$

Which can also be expressed as,

$$\boldsymbol{R}_k = \boldsymbol{R}_{k-1}(2\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1}) \tag{3.29}$$

Where the inverse in the next iteration is given by $\boldsymbol{R}_k$. The expression in (3.29) is a matrix inversion scheme known as Schulz iterative method [185], [186]. In [187] it was shown that considering the first three terms of the iterative process for matrix inversion yields the fastest convergence. Hence,

$$\begin{aligned}
\boldsymbol{R}_k &= \boldsymbol{R}_{k-1}(\boldsymbol{I} + \boldsymbol{E}_{k-1} + \boldsymbol{E}_{k-1}^2) = \boldsymbol{R}_{k-1}\big(\boldsymbol{I} + \boldsymbol{E}_{k-1}(\boldsymbol{I} + \boldsymbol{E}_{k-1})\big) \\
&= \boldsymbol{R}_{k-1}\Big(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big)\Big) \\
&= \boldsymbol{R}_{k-1}(\boldsymbol{I} + 2\boldsymbol{I} - \boldsymbol{I}\boldsymbol{F}\boldsymbol{R}_{k-1} - 2\boldsymbol{I}\boldsymbol{F}\boldsymbol{R}_{k-1} + (\boldsymbol{F}\boldsymbol{R}_{k-1})^2) \\
&= \boldsymbol{R}_{k-1}\big(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1}(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big)
\end{aligned} \tag{3.30}$$

The method in equation (3.30) was first put forward by Amat in [188], and it has been demonstrated that the sequence converges to $\boldsymbol{F}^{-1}$ as per the Theorem 3.1.

Theorem 3.2. We define $\boldsymbol{F} = [\boldsymbol{f}_{ij}]$as any given non-singular matrix. If $\boldsymbol{R}_0$ represents the initial inverse estimation satisfying,

$$\|\boldsymbol{E}_0\| \triangleq \|\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_0\| < 1 \tag{3.31}$$

it follows that the iterative expression (3.30) converges and $\boldsymbol{R}_k$ cubically converges to $\boldsymbol{F}^{-1}$.

Proof: See Appendix A.1.

Theorem 3.3. We define $\boldsymbol{F} = [\boldsymbol{f}_{ij}]$ as any given non-singular matrix. Given that $\boldsymbol{F}\boldsymbol{R}_0 = \boldsymbol{R}_0\boldsymbol{F}$ is valid, the sequence $\{\boldsymbol{R}_k\}$ in (3.27) can alternatively be rewritten as $\boldsymbol{F}\boldsymbol{R}_k = \boldsymbol{R}_k\boldsymbol{F}$ which holds for all $\boldsymbol{k} = \boldsymbol{1}, \boldsymbol{2}, \ldots$

Proof: See Appendix A.2.

If the total number of terms is set to P, then applying equations (3.22) and (3.26) the inverse can be obtained by re-writing the iterative expression as

$$\boldsymbol{R}_k = \boldsymbol{R}_{k-1}\big(\boldsymbol{I} + \boldsymbol{E}_{k-1} + \boldsymbol{E}_{k-1}^2 + \cdots + \boldsymbol{E}_{k-1}^{p-1}\big) \tag{3.32}$$

For successive iterations, the inverse can be written

$$\boldsymbol{R}_1 = \boldsymbol{R}_0\left(\boldsymbol{I} + \boldsymbol{E}_0 + \boldsymbol{E}_0^2 + \cdots + \boldsymbol{E}_0^{p-1}\right) \tag{3.33}$$

$$\boldsymbol{R}_2 = \boldsymbol{R}_1\left(\boldsymbol{I} + \boldsymbol{E}_0^p + \left(\boldsymbol{E}_0^p\right)^2 + \left(\boldsymbol{E}_0^p\right)^3 \ldots + \left(\boldsymbol{E}_0^p\right)^{p-1}\right) \tag{3.34}$$

$$\boldsymbol{R}_3 = \boldsymbol{R}_2\left(\boldsymbol{I} + \left(\boldsymbol{E}_0^{p^2}\right) + \left(\boldsymbol{E}_0^{p^2}\right)^2 + \left(\boldsymbol{E}_0^{p^2}\right)^3 \ldots + \left(\boldsymbol{E}_0^{p^2}\right)^{p-1}\right) \tag{3.35}$$

and in general

$$\begin{aligned}\boldsymbol{R}_k = \boldsymbol{R}_{k-1}\Big(\boldsymbol{I} &+ \left(\boldsymbol{E}_0^{p^{k-1}}\right) + \left(\boldsymbol{E}_0^{p^{k-1}}\right)^2 + \left(\boldsymbol{E}_0^{p^{k-1}}\right)^3 \ldots \\ &+ \left(\boldsymbol{E}_0^{p^{k-1}}\right)^{p-1}\Big)\end{aligned} \tag{3.36}$$

If the substitution for $\boldsymbol{R}_{k-1}$ is performed repeatedly it results in

$$\begin{aligned}\boldsymbol{R}_k = \boldsymbol{R}_0\left[\boldsymbol{I} + \boldsymbol{E}_0 + \boldsymbol{E}_0^2 + \cdots + \boldsymbol{E}_0^{p-1}\right]&\left[\boldsymbol{I} + \boldsymbol{E}_0^p + \left(\boldsymbol{E}_0^p\right)^2 + \left(\boldsymbol{E}_0^p\right)^3 \ldots\right. \\ &\left. + \left(\boldsymbol{E}_0^p\right)^{p-1}\right] \ldots \left[\boldsymbol{I} + \left(\boldsymbol{E}_0^{p^{k-1}}\right)\right. \\ &\left. + \left(\boldsymbol{E}_0^{p^{k-1}}\right)^2 + \left(\boldsymbol{E}_0^{p^{k-1}}\right)^3 \ldots\right. \\ &\left. + \left(\boldsymbol{E}_0^{p^{k-1}}\right)^{p-1}\right]\end{aligned} \tag{3.37}$$

$$= \boldsymbol{R}_0 \prod_{h=0}^{k-1}\left[\boldsymbol{I} + \left(\boldsymbol{E}_0^{p^h}\right) + \left(\boldsymbol{E}_0^{p^h}\right)^2 + \left(\boldsymbol{E}_0^{p^h}\right)^3 \ldots + \left(\boldsymbol{E}_0^{p^h}\right)^{k-1}\right]$$

Consequently, if we let $p = 3$ we obtain

$$\boldsymbol{R}_k = \boldsymbol{R}_0 \prod_{h=0}^{k-1}\left[\boldsymbol{I} + \left(\boldsymbol{E}_0^{p^h}\right) + \left(\boldsymbol{E}_0^{p^h}\right)^2\right] \tag{3.38}$$

The expression in (3.38) has great significance in computing $\boldsymbol{R}_k$, in that the computation of the inverse in the current iteration makes use of the inverse in the previous iteration together with two new terms involving the initial residual matrices. To compute these terms, we rely on multiplication of the previously obtained error matrices. Consequently, we reduce the task of inverting a matrix into one of executing a matrix multiplication.

Starting from (3.21), we set $\boldsymbol{R}_0 = \boldsymbol{B}^{-1}$ and hence obtain the residual matrix as

$$\boldsymbol{E}_0 = \boldsymbol{I} - \boldsymbol{L}\boldsymbol{R}_0 \tag{3.39}$$

from which the first iteration can be expressed as

$$\begin{aligned} \boldsymbol{R}_1 &= \boldsymbol{R}_0\left(\boldsymbol{I} + \left(\boldsymbol{E}_0^{3^0}\right) + \left(\boldsymbol{E}_0^{3^0}\right)^2\right) \\ &= \boldsymbol{R}_0(\boldsymbol{I} + (\boldsymbol{E}_0) + (\boldsymbol{E}_0)^2) \\ &= \boldsymbol{R}_0\big(\boldsymbol{I} + (\boldsymbol{E}_0) + (\boldsymbol{E}_0)(\boldsymbol{E}_0)\big) \end{aligned} \tag{3.40}$$

and consequently the kth iteration is computed in a similar manner as follows

$$\boldsymbol{R}_k = \boldsymbol{R}_0 \prod_{h=0}^{k-1}\left[\boldsymbol{I} + \left(\boldsymbol{E}_0^{3^h}\right) + \left(\boldsymbol{E}_0^{3^h}\right)^2\right] \tag{3.41}$$

The inversion of $\boldsymbol{L}$ can be realized by iterating $K-1$ times. From which $\boldsymbol{L}^{-1} = \boldsymbol{R}_k$ and the inverse of $\boldsymbol{W}_{RNA} = \boldsymbol{W}_{RZF}$ is calculated according to (3.20).

3.2.3.2 Initial RNA Inverse Formulation

For a matrix $\boldsymbol{L}$ the initial estimation of its inverse, $\boldsymbol{R}_0$, is said to be good if the resultant residual matrix, $\boldsymbol{E}_0$, is convergent. But we have that

$$\boldsymbol{E}_0 = \boldsymbol{I} - \boldsymbol{L}\boldsymbol{R}_0 \tag{3.42}$$

From which,

$$(\boldsymbol{L}\boldsymbol{R}_0)^{-1} = (\boldsymbol{I} - \boldsymbol{E}_0)^{-1} \tag{3.43}$$

We have that

$$\boldsymbol{I} = (\boldsymbol{I} - \boldsymbol{E}_0)(\boldsymbol{I} - \boldsymbol{E}_0)^{-1} \tag{3.44}$$

Taking this with respect to some given matrix norm,

$$\|(\boldsymbol{I} - \boldsymbol{E}_0)^{-1}\| \leq 1 + \|\boldsymbol{E}_0\|\|(\boldsymbol{I} - \boldsymbol{E}_0)^{-1}\| \tag{3.45}$$

If we divide each side by $\|(\boldsymbol{I} - \boldsymbol{E}_0)^{-1}\|$, we get

$$1 \leq \frac{1}{\|(\boldsymbol{I}-\boldsymbol{E}_0)^{-1}\|} + \|\boldsymbol{E}_0\| \tag{3.46}$$

Starting from (3.42), we can show that

$$\boldsymbol{L}^{-1}\boldsymbol{R}_0{}^{-1} = (\boldsymbol{I}-\boldsymbol{E}_0)^{-1} \tag{3.47}$$

and

$$\boldsymbol{L}^{-1} = \boldsymbol{R}_0(\boldsymbol{I}-\boldsymbol{E}_0)^{-1} \tag{3.48}$$

Taking this with respect to some given matrix norm, we get

$$\begin{aligned} \|\boldsymbol{L}^{-1}\| &= \|\boldsymbol{R}_0(\boldsymbol{I}-\boldsymbol{E}_0)^{-1}\| \\ &\leq \|\boldsymbol{R}_0\|\|(\boldsymbol{I}-\boldsymbol{E}_0)^{-1}\| \end{aligned} \tag{3.49}$$

Applying (3.46) in (3.49) leads to

$$\|\boldsymbol{L}^{-1}\| \leq \|\boldsymbol{R}_0\|\|(\boldsymbol{I}-\boldsymbol{E}_0)^{-1}\| \leq \frac{\|\boldsymbol{R}_0\|}{1-\|\boldsymbol{E}_0\|} \tag{3.50}$$

Since

$$\|(\boldsymbol{I}-\boldsymbol{E}_0)^{-1}\| \leq \frac{1}{1-\|\boldsymbol{E}_0\|} \tag{3.51}$$

It can be deduced from (3.50) that

$$1-\|\boldsymbol{E}_0\| \leq \frac{\|\boldsymbol{R}_0\|}{\|\boldsymbol{L}^{-1}\|} \tag{3.52}$$

Hence

$$\mathscr{b}(1-\|\boldsymbol{E}_0\|) \leq \frac{\|\boldsymbol{R}_0\|}{\|\boldsymbol{L}^{-1}\|} \qquad \mathscr{b} \geq 1 \tag{3.53}$$

From which the residual matrix norm can be computed as

$$\|\boldsymbol{E}_0\| = 1 - \mathscr{b}'\frac{\|\boldsymbol{R}_0\|}{\|\boldsymbol{L}^{-1}\|} \qquad \mathscr{b}' = \frac{1}{\mathscr{b}} \leq 1 \tag{3.54}$$

To have a valid initial inverse estimation $\boldsymbol{R}_0$, it is required that the residual matrix $\boldsymbol{E}_0$ is convergent. To satisfy this condition, we need to have

$$\mathscr{b}' \frac{\|\boldsymbol{R}_0\|}{\|\boldsymbol{L}^{-1}\|} \leq 1 \tag{3.55}$$

And we can calculate the initial inverse as

$$\boldsymbol{R}_0 = \beta \boldsymbol{L} \tag{3.56}$$

And

$$\beta = \frac{1}{\|\boldsymbol{L}^{-1}\|^2} \tag{3.57}$$

The proof is provided for in the Appendix A.3.

3.2.3.3 Criterion to Stop RNA Iterative Procedure

We need to obtain the desired results with a limit on the iterations to be executed through a criterion that is employed to stop the iterative procedure. We derive the criterion to limit the number of iterations. We can write that for the kth iteration, the residual matrix can be written as

$$\boldsymbol{E}_h = \left\| \boldsymbol{E}_0^{3^h} \right\| \tag{3.58}$$

If we evoke the initial inverse formulation above a convergent residual matrix is obtained from the iterative procedure, as follows

$$\|\boldsymbol{E}_0\| = q \quad \text{where} \quad q < 1 \tag{3.59}$$

The applying the properties of norms, (3.58) can be written as

$$\|\boldsymbol{E}_k\| \leq \|\boldsymbol{E}_0\|^{3^k} \leq q^{3^k} < 1 \tag{3.60}$$

The deviation between the actual inverse and the computed inverse can be expressed as

$$\boldsymbol{L}^{-1} - \boldsymbol{R}_k = \boldsymbol{L}^{-1}(\boldsymbol{I} - \boldsymbol{L}\boldsymbol{R}_k) \tag{3.61}$$

Taking this with respect a matrix norm leads to

$$\|\boldsymbol{L}^{-1} - \boldsymbol{R}_k\| \leq \|\boldsymbol{L}^{-1}\| \|\boldsymbol{I} - \boldsymbol{L}\boldsymbol{R}_k\| \leq \|\boldsymbol{L}^{-1}\| \|\boldsymbol{E}_k\| \tag{3.62}$$

and thus

$$\frac{\|\boldsymbol{L}^{-1} - \boldsymbol{R}_k\|}{\|\boldsymbol{L}^{-1}\|} \leq \|\boldsymbol{E}_k\| \tag{3.63}$$

Combining (3.60) and (3.63) we get the following

$$\frac{\|\boldsymbol{L}^{-1} - \boldsymbol{R}_k\|}{\|\boldsymbol{L}^{-1}\|} \leq q^{3^k} \tag{3.64}$$

The residual matrix is readily computed once the initial inverse is obtained. From which the value $\|\boldsymbol{E}_0\| = q$ can be derived. The number of terms in the infinite series is known and can be substituted in (3.64) to determine the number of iterations given the acceptable deviation for the formulated inverse.

3.3 Complex Operations Analysis

The SINRs associated with RZF precoding converge to the deterministic equivalent in (3.5) for the large- (*M, K*) regime. It is important to note that precoding matrices are random quantities whose computation need to be performed every time the channel information is refreshed. With coherence time being typically in the range of a few milliseconds, the inverse of the large dimensional matrix in (3.4) need to be re-computed hundreds of times per second. The required arithmetic operations for inverting a matrix scales cubically with the matrix rank, hence this matrix operation becomes intractable for large scale systems considered in massive MIMO [107], [183], [189]. To minimize the complexity in implementation and retain relative performance to that of RZF precoding, RNA precoding is fronted for a single-cell massive MIMO system. The RNA precoding offers two main benefits over RZF precoding: a) there is a uniform spread of computational operations over time because for every iteration there is only one matrix multiplication; b) the precoding computation is split into several simple matrix-vector multiplications that can be highly parallelized and can be implemented using a massive amount of simple application-specific circuits.

Starting from (3.36), p terms can be computed iteratively for the RNA precoding. Consequently, a multistage hardware implementation is feasible with numerous processing cores being used to parallelize the computation. Hence (3.32) can be expressed as

$$\boldsymbol{R}_k = \boldsymbol{R}_{k-1} + \boldsymbol{R}_{k-1}\boldsymbol{E}_{k-1} + [\boldsymbol{R}_{k-1}\boldsymbol{E}_{k-1}]\boldsymbol{E}_{k-1} + [\boldsymbol{R}_{k-1}\boldsymbol{E}_{k-1}^2]\boldsymbol{E}_{k-1} + \cdots + \left[\boldsymbol{R}_{k-1}\boldsymbol{E}_{k-1}^{p-2}\right]\boldsymbol{E}_{k-1} \tag{3.65}$$

Hence, a single matrix multiplication is needed for each term to be computed. Thus, this expression needs $(2K^3 - K^2)$ operations for the matrix multiplication to be computed. But the total number of terms to be computed is *p-1*, hence we need $(p-1)(2K^3 - K^2)$ operations to compute the terms. Additionally, every iteration for computation of the residual matrix based on (3.27) from preceding iterations comprises $2K^3$ operations. Hence, the total number of operations for computing the inverse in every iteration can be expressed $2pK^3$. If the number of iterations is set to ƌ the corresponding number of operations needed are expressed as

$$op.count = 2ƌpK^3 \tag{3.66}$$

Hence, performing ƌ iterations for a three-term polynomial requires the following operations

$$op.count(N) = 6ƌK^3 - K^2(ƌ - 1) \tag{3.67}$$

Consequently, the required operations scales as $\mathcal{O}(\mathrm{K}^3)$ and it mirrors that of RZF precoding when $\mathrm{M} >> \mathrm{K}$ [129], a typical scenario in massive MIMO. However, we reduce the operation counts by $K^2(ƌ - 1)$ as opposed in RZF precoding. Again, if we employ Strassen's matrix multiplication algorithm (See Appendix A.4), the operations arising from matrix multiplication are reduced further [190]. In addition, the RNA algorithm renders itself to parallel implementation as per equation (3.37) and this becomes a vital aspect in massive MIMO. Hence, if we assume that $z \ll K$ and we use K/z parallel cores to realize RNA precoding then the number of operations will decrease and scale as $\mathcal{O}\left(\frac{z}{K}\mathrm{K}^3\right) = \mathcal{O}(z\mathrm{K}^2)$. Hence, the complexity of RNA precoding becomes lesser in comparison to RZF precoding. Bearing in mind that the key to suppressing interference in RZF precoding relies on matrix inversion, then RNA precoding suppresses interference efficiently, yet it is less complicated.

3.4 Performance Analysis

We examine the key performance indicators/parameters SINR, BER and data rate in relation to perfect and imperfect CSI. We analyse the trade-offs amongst these parameters for the massive MIMO downlink for the three linear precoding schemes discussed earlier.

3.4.1 Achievable Rate

Based on Shannon theorem, we consider an additive white Gaussian noise channel, whose capacity can be expressed as [191]

$$\gamma = log_2(1 + \boldsymbol{SINR}) \qquad (bits/sec/Hz) \tag{3.68}$$

Consequently, the rate that can be achieved by the kth UT is expressed as in (3.43) below for wireless massive MIMO downlink system [108]

$$\boldsymbol{\gamma}_k = \mathbb{E}[log_2(1 + \boldsymbol{SINR}_k)] \qquad (bits/sec/Hz) \tag{3.69}$$

3.4.2 Bit Error Rate

The parameter BER is a key indicator on information interface performance in most communication systems. The occurrence of errors forms one of the guideline parameters concerning data transmission; this makes BER one of the key parameters. BER provides information on various aspects of data transmission like exchange speed and power, among others. For binary modulation the BER is bounded by a single Q-function and expressed as [192]

$$\boldsymbol{BER}_{\infty} = Q\left(\sqrt{E(\boldsymbol{SINR}_k)_{\infty}}\right) = \int^{\infty} \frac{1}{\sqrt{E(\boldsymbol{SINR}_k)_{\infty}}} e^{-\frac{j^2}{2}} dj \tag{3.70}$$

where $E(\boldsymbol{SINR}_k)_{\infty}$ is the asymptotic first moment of $\boldsymbol{SINR}_k$. A noteworthy approximation of the true BER for Gray coded transmission of 16-QAM constellation is [38]

$$\boldsymbol{BER}_{16-QAM} = \frac{3}{8} erfc\left(\sqrt{\frac{1}{10}\boldsymbol{SINR}_k}\right) + \frac{1}{4} erfc\left(\sqrt{\frac{9}{10}\boldsymbol{SINR}_k}\right) - \frac{1}{8} erfc\left(\sqrt{\frac{5}{2}\boldsymbol{SINR}_k}\right) \tag{3.71}$$

with $erfc = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-j^2} dj$ being the complementary error function.

3.5 Simulation Results and Discussion

This section uses simulated results from a 128 × 32 massive MIMO framework and a 128 × 16 massive MIMO framework of the proposed RNA precoding to evaluate the related average achievable rate and the BER. The RZF precoding using the definite matrix inversion is employed as the benchmark. RZF precoding, RNA precoding, and TPE precoding performance is analysed for the single-cell system.

The average bit rates are presented in Figure 3.1 for RZF precoding, RNA precoding and the TPE precoding for comparison. From this figure a number of observations can be deduced. The RZF precoding together with RNA precoding show practically identical performance and which surpasses that of the TPE precoding. When the polynomial power, J, of the TPE precoder is incremented (which corresponds to increment in complexity), it enhances the corresponding bit rates and nears that of the RNA precoder. In any case, this comes at the expense of expanded hardware resources. Hence, the RNA-based precoder outperforms the TPE precoder under identical performance environment.

In Figure 3.2, the RNA precoding, RZF precoding, and the TPE precoding are evaluated under three variations of the channel knowledge levels: $\tau \in \{0.15, 0.45, 0.75\}$ with $J = 4$. When the CSI is poor ($i.e.\ \tau = 0.75$) the RNA precoding, RZF precoding and TPE precoding have relatively similar performance in terms of the bit rates. Moreover, this similarity in bit rates can also be seen at low SNRs regardless of the value of τ.

When the number of UTs is reduced from 32 to 16 with the BS antennas held constant at 128, The bit rates are significantly enhanced as observed in Figure 3.3 and Figure 3.4. Which concurs well with the theory of massive MIMO [18], [102]. But another important observation that can be made is that the performance of the TPE precoder improves significantly and nears that of the RZF precoder, and RNA precoder, though still less than RNA precoder. This shows that if we increase the number of transmitter antennas, we can significantly improve the performance of TPE precoder even without increasing the order J of the polynomial. This shows that we can apply the TPE precoding scheme only in massive MIMO systems with large number of BS antennas in comparison to UTs and achieve quite good throughput. As opposed to RNA-based precoder that can be applied in any of the BS-UTs combination in massive MIMO. It is also worth noting that the RNA precoder performances is close to that of RZF precoder. This can be explained form the fact that RNA precoder is an approximation of the RZF precoder which means it cannot outperform the RZF precoder but if it approximates it well then, the performance of the two precoders will be close and this is the case in this work.

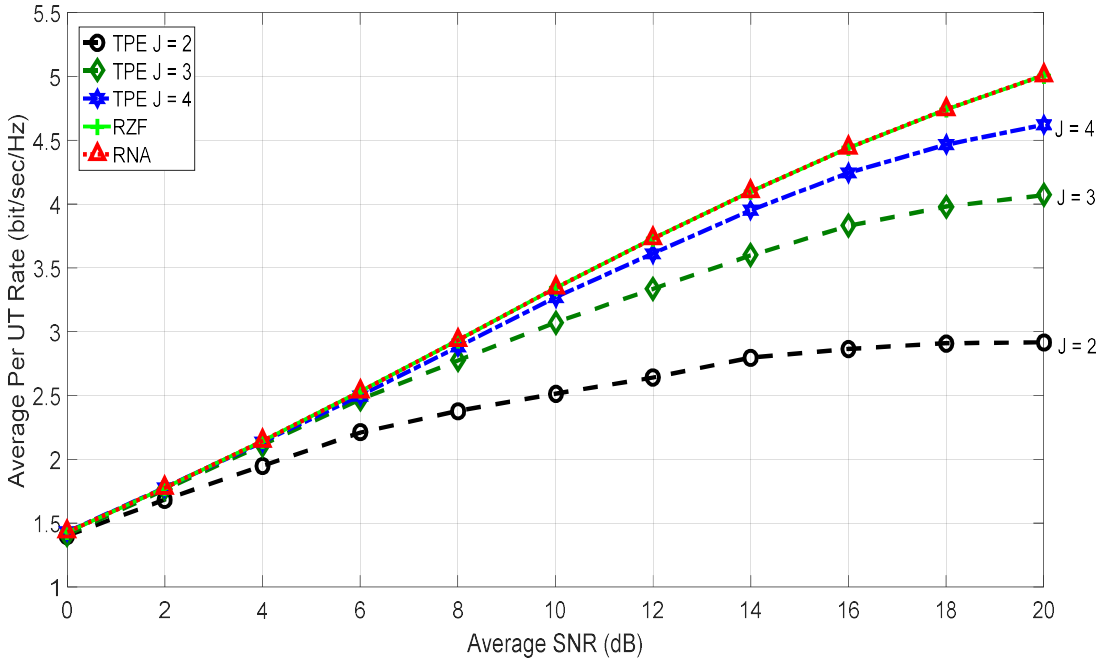

Figure 3.1: Average per UT rate vs. SNR for different orders of J in TPE precoding *($M = 128, K = 32,\ \tau = 0.15$).*

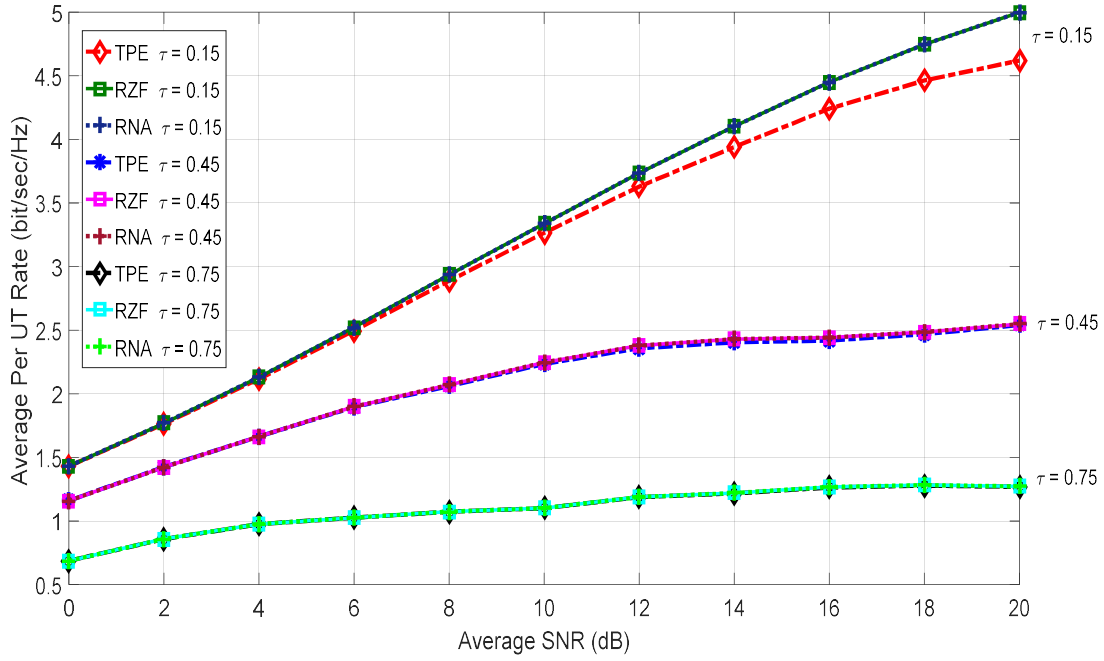

Figure 3.2: Average per UT rate vs. SINR for varying CSI errors at the BS ***($M = 128, K = 32,\ J = 4$)***.

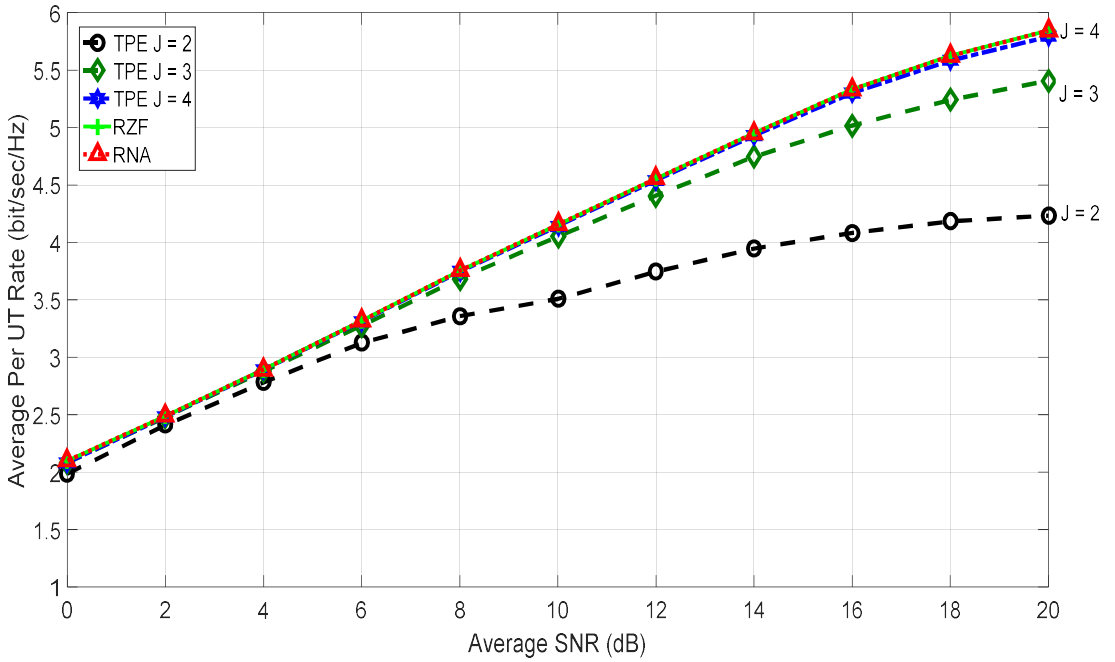

Figure 3.3: Average per UT rate vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding $(\boldsymbol{M = 128, K = 16,\ \tau = 0.15})$.

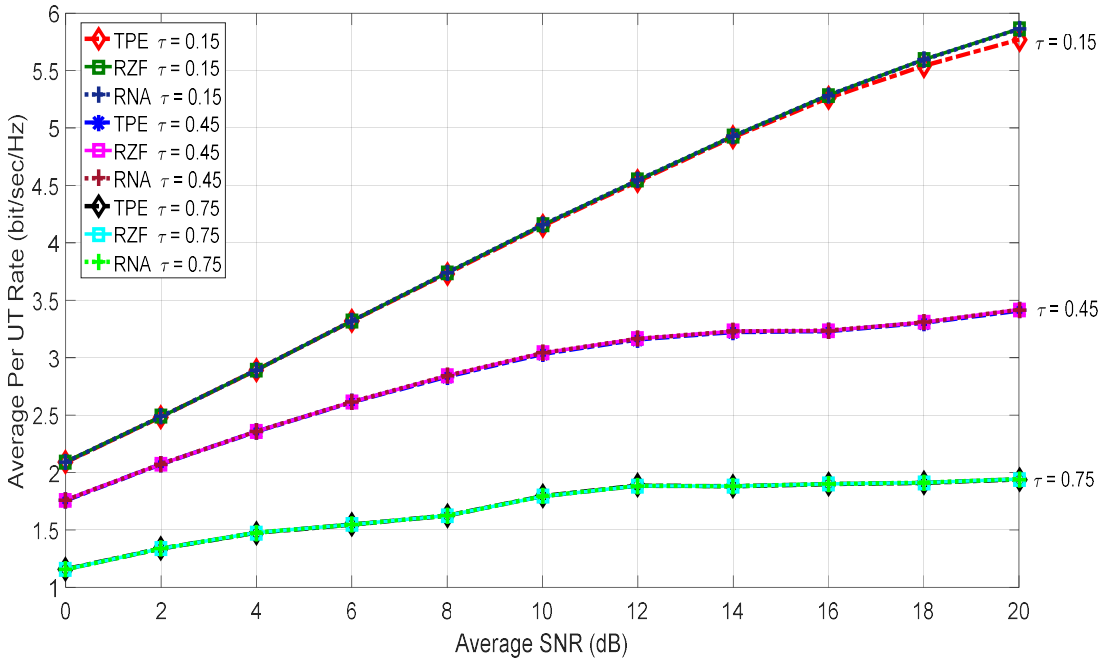

Figure 3.4: Average per UT rate vs. SINR for varying CSI errors at the BS $(\boldsymbol{M = 128, K = 16,\ J = 4})$.

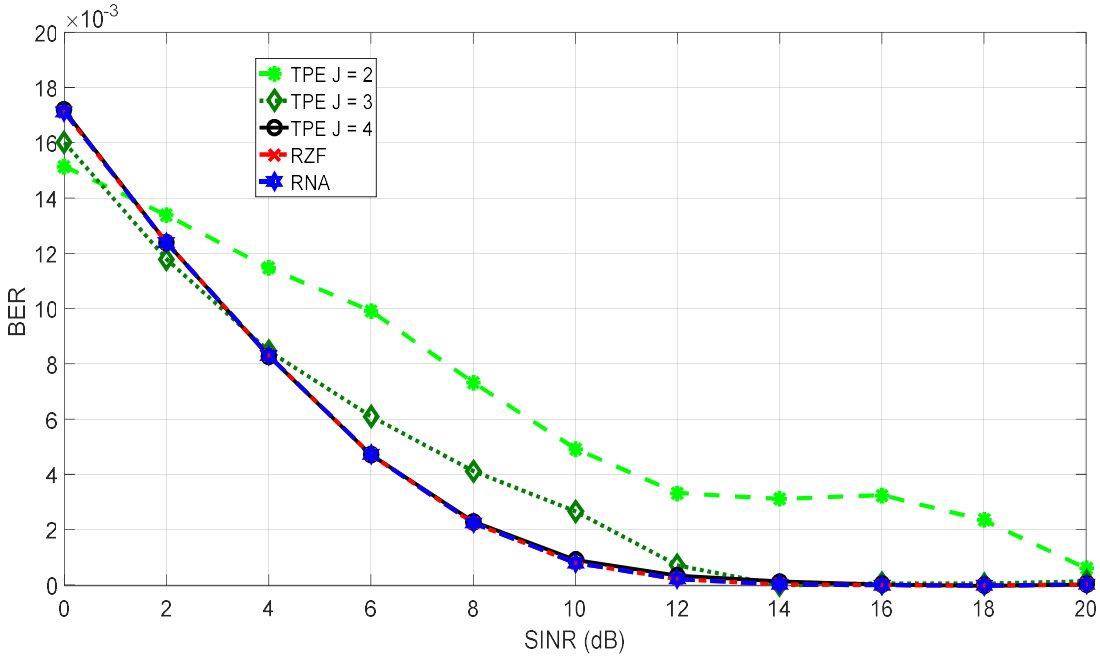

Figure 3.5: BER vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding $(\boldsymbol{M = 128, K = 32, \tau = 0.15})$.

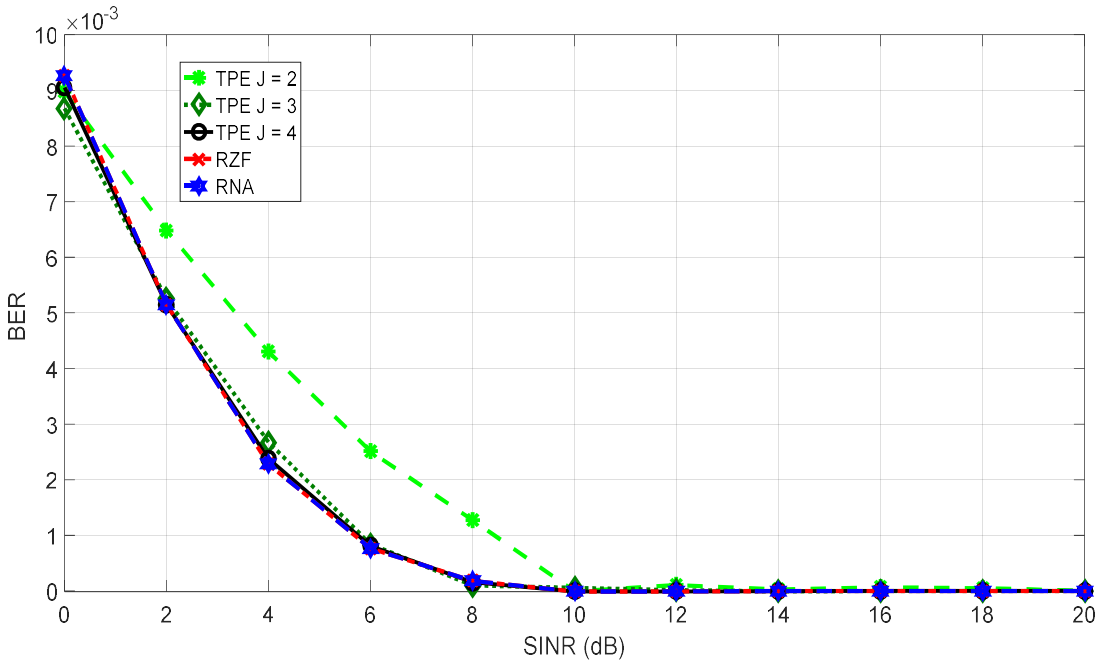

Figure 3.6: BER vs. SINR for different orders of $\boldsymbol{J}$ in TPE precoding $(\boldsymbol{M = 128, K = 16, \tau = 0.15})$.

Figure 3.5 and Figure 3.6 above demonstrate the BER performance comparison in fading channels. From the two figures, we can clearly see that when the number of UTs increases in

massive MIMO systems, the BER execution of all precoding schemes experiences serious performance loss because of the limited number of BS receiving antennas in practical systems, however RNA scheme loss is less than TPE precoding which means its Robust and is the best among them. Then, it is obvious that when J is increasing, BER performance of TPE precoding have improvement in some extend. Yet, the BER performance of RNA precoding is still superior to that of TPE precoding notwithstanding when $J = 4$ and near the RZF. Also, as SINR increases, the performance of the proposed RNA precoding improves faster. Thus, this exploits the massive MIMO diversity and improves the channel performance hence reducing the BER.

3.6 Conclusion Single Cell Precoding

The chapter examines the formulation, performance, analysis, and comparison of the an iterative linear precoder that circumvents matrix inversion called the RNA precoder against the RZF and the TPE precoders within a single cell for a downlink massive MIMO system. The performance of the three precoding schemes in terms of SINR, BER and the data rate for both the perfect and imperfect CSI are studied. These were derived theoretically for each of the precoding schemes under similar assumptions and for the wireless massive MIMO system. From the simulation and the theoretical results, RZF precoding, and RNA precoding have higher data rates and lower BER than the TPE precoder. But when the UTs antennas are halved with the BS antennas remaining the same, TPE precoder data rates and BER nears that of the RZF precoder, and the RNA precoder even without increasing the order J of the polynomial. The performance of TPE precoder also nears that of the other three under poor CSI quality. Thus, RNA precoder have good performance at both lower and large number of BS antennas as compared to UTs antennas than the TPE precoder which only has improved performance at high number of BS antennas. Also, the TPE precoder performance relies on the polynomial power J. which calls for increased polynomial power to improve its performance and thus more hardware and cost of implementation as opposed to the RNA precoder.

CHAPTER 4

Rapid Numerical Algorithms Precoding: Multicell

4.1 Introduction

Inter-cell and intra-cell interference pose a challenge in realization of multicell wireless systems. MIMO wireless system has been earmarked as the enabling technology to enhance quality and capacity in current wireless communication systems [168]. Where we use linear precoding to separate information flags spatially towards their respective UT in the downlink. However, as discussed earlier, the precoding plan in multiuser MIMO needs extremely precise momentary CSI [114] which can be bulky to accomplish since the channel fluctuates rapidly. The quality of CSI received at the BS will enhance or degrade the performance of massive MIMO system. Also, the increased number of antennas for transmission in massive MIMO points to enhanced power radiation and enormous amounts of energy being utilized at the UTs and BSs. This is the consequence of increased RF chains and the processing load due to increase in number of antennas. Complexity of hardware for implementing massive MIMO pose another challenge since it increases exponentially with the number of antennas [193]. Thus, to enhance massive MIMO system output efficiency, the radiated power needs to be regulated [194].

The spectral and energy efficiencies are greatly enhanced if the BS communicates with the UT within the same time-frequency resources. This can be achieved but at a price of enhanced inter-user interference, hence eroding the system performance. However, the channels between UTs and BS becomes approximately orthogonal when massive MIMO is used [18], [41]. It has been said that with massive MIMO, channel approximation errors and thermal noise effectively averages out and simple forms of precoding like Eigen BF and MF tends to be near optimal [170]. Later research have demonstrated that realizable BS antennas demand included linear precoding procedures, for example, MMSE[171]. In line with this requirement a class of RNA and TPE precoders were discussed in [195] for the single cell multi-user case. The multicell TPE-based precoder is discussed in [196] and compared to RZF precoding. In [197] a multicell

MRC-BSC and ZF-BSC precoding to improve system downlink performance, and their influence on the received SINR of users in other cells is presented.

This chapter discusses multicell precoding using RNA-based, RZF and TPE-based precoders with and without BS cooperation. In our work, the proposed precoding schemes with and without BS cooperation are obtained by changing the matrix structure of single-cell RNA, RZF and TPE precoding, so they are both linear precoding and have low computational complexity. We consider the effects and analyse the received SINR of user in other cells. We use a similar analysis method as that employed in [197]. The SINR is used to calculate the capacity per cell for comparison purposes. We also look at energy efficient based on these three precoding schemes for the case of uncoordinated and coordinated BSs and conclude by comparing the performance of these three precoders on overall. The channel model considered includes the imperfect covariance experienced by the BSs.

We begin by modelling the optimal multicell linear precoding and then tailor it to each of the multicell linear precoding schemes. The imperfect covariance channel is modelled before evaluating the SINR and data rate of imperfect channel covariance information is used to analyse the behaviour of the massive MIMO downlink network system over the modelled channels in multicell scenario.

4.2 System Model for Multicell

We presume a multicell massive MIMO system with L cells and K UTs with single antennas that simultaneously transmits signals to M antennas of the BS in the corresponding cell. We obtain the propagation matrix, $\boldsymbol{G} \in \mathbb{C}^{\mathrm{M}\times\mathrm{K}}$, of the corresponding cell through multiplication of an $M \times K$ matrix, [18], [32], [33]. The TDD protocols are presumed to be coordinated across cells and convey pilot signals and data to all cells concurrently. At the mth UT in the jth cell, the received complex baseband signal is $\boldsymbol{y}_{j,m} \in \mathbb{C}^{\mathrm{M}\times\mathrm{K}}$ and computed as

$$\boldsymbol{y}_{j,m} = \sum_{\ell=1}^{L} \boldsymbol{h}_{\ell,j,m}^{H} \boldsymbol{u}_{\ell} + n_{j,m} \tag{4.1}$$

where $\boldsymbol{u}_{\ell} \in \mathbb{C}^{M\times 1}$ describes the ℓth BS transmit signal and $\boldsymbol{h}_{\ell,j.m}^{H} \in \mathbb{C}^{M\times 1}$represents the ℓth BS channel vector towards the mth UT in the jth cell, and $\mathrm{n}_{j,m}\sim\mathcal{CN}(0,\sigma^2)$ represents the Additive White Gaussian Noise (AWGN), with a variance of σ^2 at the receiver's input [196], [106], [198]–[200].

It is assumed that the channel vectors are Rayleigh fading and we model them as

$$\boldsymbol{h}_{\ell,j,m} \sim \mathcal{CN}\left(0, \boldsymbol{\Phi}_{\ell,j,m}\right) \tag{4.2}$$

where $\boldsymbol{\Phi}_{\ell,j,m}$ is the ℓth BS associated matrix of covariance towards the mth UT within the jth cell. The BSs use Gaussian codebooks combined with linear precoding permitting us to write the transmitted signal in the jth cell as

$$\boldsymbol{u}_j = \sum_{k=1}^{K} \boldsymbol{f}_{j,m} x_{j,m} = \boldsymbol{F}_j \boldsymbol{x}_j \tag{4.3}$$

where $\boldsymbol{F}_j = \left[\boldsymbol{f}_{j,1}, \ldots, \boldsymbol{f}_{j,K}\right] \in \mathbb{C}^{M \times K}$ is the precoding matrix and $\boldsymbol{x}_j = \left[x_{j,1}, \ldots, x_{j,K}\right] \in \sim \mathcal{CN}(0, \boldsymbol{I}_K)$is the data vector with information symbols corresponding to all the UTs within the jth cell. The average power for the jth BS to be transmitted is then constrained by $\frac{1}{K} tr\left(\boldsymbol{F}_j{}^* \boldsymbol{F}_j^H\right) = P_j$. Thus, the signal vector received at kth UT is expressed as

$$\boldsymbol{y}_{j,m} = \sum_{\ell=1}^{L} \sum_{k=1}^{K} \boldsymbol{h}_{\ell,j,m}^H \boldsymbol{f}_{\ell,k} s_{\ell,k} + n_{j,m} \tag{4.4}$$

We expand this expression to obtain

$$\boldsymbol{y}_{j,m} = \boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m} s_{j,m} + \sum_{k \neq m} \boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,k} s_{j,k} + \sum_{(\ell,k) \neq (j,m)} \boldsymbol{h}_{\ell,j,m}^H \boldsymbol{f}_{\ell,k} s_{\ell,k} + n_{j,m} \tag{4.5}$$

Channel hardening is a notable feature of massive MIMO systems, which implies that the useful channel $\boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m}$ converges to its average value as $M \to \infty$. The received signal at the mth UT can be broken down further as in [171], [201] to yield

$$\boldsymbol{y}_{j,m} = \mathbb{E}\left[\boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m}\right] s_{j,m} + var\left[\boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m}\right] s_{j,m} + \sum_{(\ell,k) \neq (j,m)} \boldsymbol{h}_{\ell,j,m}^H \boldsymbol{f}_{\ell,k} s_{\ell,k} + n_{j,m} \tag{4.6}$$

In addition, we expect that the fluctuation $var\left[\boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m}\right]$ and the average total interference power $\sum_{(\ell,k) \neq (j,m)} \mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^H \boldsymbol{f}_{\ell,k}\right|^2\right]$ due to simultaneous transmission is perceived together with the channel gain $\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^H \boldsymbol{f}_{j,m}\right|^2\right]$ to different UTs in the same and different cells. We adopt a worst scenario Gaussian noise to cater for interference among users in one cell and the

users from the other cell as well as the channel ambiguity and express the SE of the *m*th UT in *j*th cell by

$$\boldsymbol{r}_{j,m} = log_2\left(1 + \boldsymbol{SINR}_{j,m}\right) \tag{4.7}$$

independent of the knowledge of the transitory values of $\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}$ of its channel [17], [171], [201], [202] and where

$$\boldsymbol{SINR}_{j,m} = \frac{\left|\mathbb{E}\left[\mathbf{h}_{j,j,m}^{H}\mathbf{f}_{j,m}\right]\right|^2}{\sigma^2 + \sum_{\ell=1}^{L}\mathbb{E}\left[\left|\mathbf{h}_{\ell,j,m}^{H}\mathbf{f}_{\ell,k}\right|^2\right] - \left|\mathbb{E}\left[\mathbf{h}_{j,j,m}^{H}\mathbf{f}_{j,m}\right]\right|^2} \tag{4.8}$$

4.2.1 Model of Imperfect Covariance Channel Information

For a given UT, we presume the transmitter to have imperfect information about the instantaneous channel realization $\widehat{\boldsymbol{h}}_{j,j,m}$. We permit the correlation of the pilot sequence with the received signal of a user and leverage on this to approximate the channel corresponding to the *m*th UT within the *j*th cell. This provides the received signal being processed as

$$\boldsymbol{y}_{j,m}^{tr} = \boldsymbol{h}_{j,j,m} + \sum_{\ell \neq j}\boldsymbol{h}_{\ell,j,m} + \frac{1}{\sqrt{\rho_{tr}}}\boldsymbol{n}_{j,m}^{tr} \tag{4.9}$$

where $\boldsymbol{n}_{j,m}^{tr} \sim \mathcal{CN}(0, \boldsymbol{I}_M)$ and $\rho_{tr} > 0$ represents the effective training SINR [171]. We compute the MMSE approximation $\widehat{\boldsymbol{h}}_{j,j,m}$ of $\boldsymbol{h}_{j,j,m}$ by

$$\widehat{\boldsymbol{h}}_{j,j,m} = \boldsymbol{\Phi}_{j,j,m}\,\boldsymbol{S}_{j,m}\boldsymbol{y}_{j,m}^{tr} = \boldsymbol{\Phi}_{j,j,m}\,\boldsymbol{S}_{j,m}\left(\sum_{\ell=1}^{L}\boldsymbol{h}_{\ell,j,m} + \frac{1}{\sqrt{\rho_{tr}}}\boldsymbol{n}_{j,m}^{tr}\right) \tag{4.10}$$

with $\boldsymbol{S}_{j,m} = \left(\frac{1}{\rho_{tr}} + \sum_{\ell=1}^{L}\boldsymbol{\Phi}_{\ell,j,m}\right)^{-1}$ and $\boldsymbol{\Phi}_{j,j,m}$ represents the channel matrix of covariance vector $\boldsymbol{h}_{j,j,m}$. The channels from the *j*th BS to all UTs within the *j*th cell are approximated by

$$\widehat{\boldsymbol{H}}_{j,j} = \left[\widehat{\boldsymbol{h}}_{j,j,1}, \ldots, \widehat{\boldsymbol{h}}_{j,j,K}\right] \in \mathbb{C}^{M\times K} \tag{4.11}$$

Here we define $\boldsymbol{R}_{\ell,j,m} = \boldsymbol{\Phi}_{j,j,m}\,\boldsymbol{S}_{j,m}\,\boldsymbol{\Phi}_{\ell,j,m}$ and point out that $\widehat{\boldsymbol{h}}_{j,j,m} \sim \mathcal{CN}\left(0, \boldsymbol{R}_{j,j,m}\right)$ is independent of the estimation error $\widetilde{\boldsymbol{h}}_{j,j,m} = \widehat{\boldsymbol{h}}_{j,j,m} - \boldsymbol{h}_{j,j,m}$ since MMSE approximation is used.

4.3 Multicell Precoding

4.3.1 RZF Precoding

Based on the notation in [196], we express the RZF precoding matrix corresponding to the BS within the *j*th cell as

$$\mathcal{C}\mathcal{D}_j^{RZF} = \sqrt{\vartheta_j}\left(\sum_{m=1}^{K} \widehat{\boldsymbol{h}}_{j,j,m}\widehat{\boldsymbol{h}}_{j,j,m}^{\boldsymbol{H}} + \boldsymbol{Z}_j\right)^{-1} \widehat{\boldsymbol{h}}_{j,j,m} \tag{4.12}$$

where $\boldsymbol{Z} = \sum_{m=1}^{K}(\boldsymbol{\Phi}_m - \boldsymbol{R}_m) + \frac{1}{\rho_{tr}}\boldsymbol{I}_M$ and we carefully choose the transmit power $\vartheta_j > 0$ so that the power constraint within the cell is satisfied. The precoding matrix coefficient is modified based on norm minimization scheme to enhance the transmission signal SINR performance [203]. The desired precoding matrix can be computed as

$$\boldsymbol{F}_j^{RZF} = \frac{\mathcal{C}\mathcal{D}_j^{RZF}}{\left|\mathcal{C}\mathcal{D}_j^{RZF}\right|} \tag{4.13}$$

4.3.2 TPE Precoding

The TPE concept can be expanded to realize a class of linear precoders with low complexity for the massive MIMO multicell precoding plan. The inverse of matrix A with dimension M can be expressed as a weighted sum of its first *M* powers through the Cayley-Hamilton hypothesis [107], [198]. The suggested TPE precoding matrix corresponding to the *j*th BS with a truncation order of J_j can be expressed as

$$\mathcal{C}\mathcal{D}_j^{TPE} = \sum_{n=0}^{J_j-1} \omega_{n,j}\left(\frac{\sum_{k=1}^{K} \widehat{\boldsymbol{h}}_{j,j,m}\widehat{\boldsymbol{h}}_{j,j,m}^{\boldsymbol{H}}}{K}\right)^{n} \frac{\widehat{\boldsymbol{h}}_{j,j,m}}{\sqrt{K}} \tag{4.14}$$

The BS in *j*th cell uses a total of J_j scalar coefficients given as $\{\omega_{n,j}, j = 0, \dots, J_j - 1\}$ and it controls the precoding matrix energy via the normalization by $\sqrt{K}$. Then, using the norm minimisation, the desired precoding matrix is transformed into the following

$$\boldsymbol{F}_j^{TPE} = \frac{\mathcal{C}\mathcal{D}_j^{TPE}}{\left|\mathcal{C}\mathcal{D}_j^{TPE}\right|} \tag{4.15}$$

For large-scale approximation of SINRs, it is shown that the SINR encountered by the *mth* UT served by the *jth* cell can be calculated by a deterministic term only based on the channel statistics, in the large- (*M, K*) regime. To give a simpler representation of the SINR expression [198], some extra notation is introduced. We set $\boldsymbol{\omega}_j = \left[\omega_{0,j}, \dots, \omega_{(J_j-1),j}\right]$ and let $\boldsymbol{a}_{j,m} \in \mathbb{C}^{J_j \times 1}$ and $\boldsymbol{B}_{\ell,j,m} \in \mathbb{C}^{J_j \times J_j}$ to be computed as

$$\left[\boldsymbol{a}_{j,m}\right]_n = \frac{\boldsymbol{h}_{j,j,m}^H}{\sqrt{K}} \boldsymbol{V}_{n,j} \frac{\widehat{\boldsymbol{h}}_{j,j,m}^H}{\sqrt{K}}, n \in \left[0, J_j - 1\right] \tag{4.16}$$

$$\left[\boldsymbol{B}_{\ell,j,m}\right]_{n,p} = \frac{1}{K} \boldsymbol{h}_{\ell,j,m}^{\boldsymbol{H}} \boldsymbol{V}_{n+p+1,j} \boldsymbol{h}_{\ell,j,m}, p \in \left[0, J_j - 1\right] \tag{4.17}$$

where $\boldsymbol{V}_{n,j} = \left(\frac{\sum_{\ell=1}^{L} \sum_{m=1}^{K} \widehat{\boldsymbol{h}}_{\ell,j,m} \widehat{\boldsymbol{h}}_{\ell,j,m}^H}{K}\right)^n$. Then the *m*th UT within the *j*th cell is associated with the SINR computed as

$$SINR_{j,m}^{TPE} = \frac{\left|\mathbb{E}\left[\boldsymbol{\omega}_j^H \boldsymbol{a}_{j,m}\right]\right|^2}{\frac{\sigma^2}{K} + \sum_{\ell=1}^{L} \mathbb{E}\left[\left|\boldsymbol{\omega}_\ell^H \boldsymbol{B}_{\ell,j,m} \boldsymbol{\omega}_\ell\right|^2\right] - \left|\mathbb{E}\left[\boldsymbol{\omega}_j^H \boldsymbol{a}_{j,m}\right]\right|^2} \tag{4.18}$$

4.3.3 RNA Precoding

In the presence of imperfect CSI, the ideal linear precoding is elusive and requires wide-ranging optimization procedures for the case of perfect CSI [38]. In this way, in fading multicell networks, only heuristic precoding techniques are attainable. RNA is a state-of-the-art heuristic scheme based on MMSE precoding [195], with a simple closed-form precoding expression.

In [195] it is stated that when the value of the polynomial 's power, J_j, is low, the TPE precoder has a poor performance, but an increase in J_j enhances the bit-rates and the performance nears that of RNA precoder. This is achieved at the expense of increasing the hardware needed for

the TPE precoding implementation. With this specified hardware implementation, the RNA precoder will present a computational complexity relative to that of the TPE precoder but with less K^2 computational complexity while giving a better performance. Thus the increase in hardware to enhance performance in TPE precoder is sufficient to minimize the computational complexity of RNA precoder relative to that of the TPE precoder with a better throughput performance [107], [195], [204]. Using [200] notation, the RZF precoding matrix used in the *j*th cell by the BS is

$$\mathcal{C}\mathcal{J}_j^{RZF} = \sqrt{\vartheta_j}\left(\sum_{\ell=1}^{L}\sum_{k=1}^{K}\widehat{\boldsymbol{h}}_{\ell,j,m}\widehat{\boldsymbol{h}}_{\ell,j,m}^{\boldsymbol{H}} + \boldsymbol{Z}_j\right)^{-1}\widehat{\boldsymbol{h}}_{j,j,m} \tag{4.19}$$

where $\boldsymbol{Z} = \sum_{m=1}^{K}(\boldsymbol{\Phi}_m - \boldsymbol{R}_m) + \frac{1}{\rho_{tr}}\boldsymbol{I}_M$ and we carefully choose the transmit power $\vartheta_j > 0$ so that the power constraint within the cell is satisfied. The estimator based on MMSE approximation requires matrix inversion which increases complexity in the order of $\mathcal{O}(M^3)$, where M is the number of BS antennas.

Therefore, we replace the matrix inversion with the rapid numerical algorithm (RNA) method. RNA-based approximation essentially avoids the matrix inversion and replaces it with multiplication and addition instead. Let $\boldsymbol{F}$ be a $M \times M$ nonsingular matrix that we are tasked to compute the inverse. Again, let $\boldsymbol{R}_k$ represent the estimated inverse in the *kth* iteration. The residual matrix represents the divergence of the computed inverse from the real inverse of the matrix $\boldsymbol{F}$. The residual matrix $\boldsymbol{E}_k$ is obtained as

$$\boldsymbol{E}_k = \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k \tag{4.20}$$

This is the residual matrix in the computation of the first inverse $\boldsymbol{R}_k$. From which,

$$\begin{aligned} \boldsymbol{F}\boldsymbol{R}_k &= \boldsymbol{I} - \boldsymbol{E}_k \\ (\boldsymbol{F}\boldsymbol{R}_k)^{-1} &= (\boldsymbol{I} - \boldsymbol{E}_k)^{-1} \\ \boldsymbol{F}^{-1} &= \boldsymbol{R}_k(\boldsymbol{I} - \boldsymbol{E}_k)^{-1} \end{aligned} \tag{4.21}$$

This expression is a power series in $\boldsymbol{E}_k$. Thus,

$$\boldsymbol{F}^{-1} = \boldsymbol{R}_k(\boldsymbol{I} - \boldsymbol{E}_k)^{-1} = \boldsymbol{R}_k\sum_{k=0}^{\infty}(\boldsymbol{E}_k)^k = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{E}_k + \boldsymbol{E}_k^2 + \cdots) \tag{4.22}$$

The first two terms can be used to represent the inverse of an infinite series as in [205], [206]. Limiting the infinite series to the first two terms, we obtain,

$$\boldsymbol{F}^{-1} = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{E}_k) = \boldsymbol{R}_k(\boldsymbol{I} + \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k) = \boldsymbol{R}_k(2\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k) \tag{4.23}$$

This can also be written as,

$$\boldsymbol{R}_k = \boldsymbol{R}_{k-1}(2\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1}) \tag{4.24}$$

Where, $\boldsymbol{R}_k$ represent the inverse in the next iteration. This expression is known as the Schulz iterative method for inverting a matrix [185], [186]. It was pointed out in [187] that the consideration of the initial three terms gives the quickest convergence of the iterative process for finding the inverse. Thus,

$$\begin{aligned}
\boldsymbol{R}_k &= \boldsymbol{R}_{k-1}(\boldsymbol{I} + \boldsymbol{E}_{k-1} + \boldsymbol{E}_{k-1}^2) = \boldsymbol{R}_{k-1}\big(\boldsymbol{I} + \boldsymbol{E}_{k-1}(\boldsymbol{I} + \boldsymbol{E}_{k-1})\big) \\
&= \boldsymbol{R}_{k-1}\Big(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big)\Big) \\
&= \boldsymbol{R}_{k-1}(\boldsymbol{I} + 2\boldsymbol{I} - \boldsymbol{I}\boldsymbol{F}\boldsymbol{R}_{k-1} - 2\boldsymbol{I}\boldsymbol{F}\boldsymbol{R}_{k-1} + (\boldsymbol{F}\boldsymbol{R}_{k-1})^2) \\
&= \boldsymbol{R}_{k-1}\big(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1}(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})\big)
\end{aligned} \tag{4.25}$$

This method was proposed by Amat in [188], and this sequence converges to $\boldsymbol{F}^{-1}$. We let

$$\boldsymbol{\phi}_{j,m}^{RNA} = \left(\sum_{\ell=1}^{L}\sum_{k=1}^{K}\widehat{\boldsymbol{h}}_{\ell,j,m}\widehat{\boldsymbol{h}}_{\ell,j,m}^{\boldsymbol{H}} + \boldsymbol{Z}_j\right)^{-1} \tag{4.26}$$

Applying Cholesky decomposition to $\boldsymbol{\phi}_{j,m}^{RNA}$ [126]

$$\boldsymbol{\phi}_{j,m}^{RNA} = \boldsymbol{L}_{j,m}\boldsymbol{L}_{j,m}^{H} \tag{4.27}$$

where $\boldsymbol{L}_{j,m}$ represents the lower triangular matrix. This implies that

$$\left(\boldsymbol{\phi}_{j,m}^{RNA}\right)^{-1} = \left(\boldsymbol{L}_{j,m}^{H}\right)^{-1}\boldsymbol{L}_{j,m}^{-1} \tag{4.28}$$

Here the computation of the inversion of matrix $\boldsymbol{\phi}_{j,m}^{RNA}$ can be changed into computation of the inversion of matrix $\boldsymbol{L}_{j,m}$. Hence, we compute a lower triangular matrix inverse instead of the symmetric matrix inverse consequently reducing the complexity.

To sustain low computation complexity the matrix $\boldsymbol{L}_{j,m}$ is decomposed further into

$$\boldsymbol{L}_{j,m} = \boldsymbol{B}_{j,m} + \boldsymbol{L}'_{j,m} \tag{4.29}$$

Where $\boldsymbol{B}_{j,m} = diag(\boldsymbol{L}_{j,m})$ and $\boldsymbol{L}'_{j,m} = (\boldsymbol{l}'_1, \boldsymbol{l}'_2, \dots, \boldsymbol{l}'_{K-1}, 0)$ is a matrix based on $\boldsymbol{L}_{j,m}$ with the diagonal elements of $\boldsymbol{L}_{j,m}$ set to zero. From (4.29), we let $\boldsymbol{R}^0_{j,m} = \boldsymbol{B}^{-\mathbf{1}}_{j,m}$ and then the residual matrix can be obtained as

$$\boldsymbol{E}^0_{j,m} = \boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{R}^0_{j,m} \tag{4.30}$$

from which the first iteration can be expressed as

$$\boldsymbol{R}^1_{j,m} = \boldsymbol{R}^0_{j,m}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{R}^0_{j,m}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{R}^0_{j,m}\right)\right) \tag{4.31}$$

and hence the k^{th}iteration is obtained in the same way as follows

$$\boldsymbol{R}^k_{j,m} = \boldsymbol{R}^{k-1}_{j,m}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{R}^{k-1}_{j,m}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{R}^{k-1}_{j,m}\right)\right) \tag{4.32}$$

Once again, the inversion of $\boldsymbol{L}_{j,m}$ can be computed by iterating $K-1$ times. From which $\boldsymbol{L}^{-\mathbf{1}}_{j,m} = \boldsymbol{R}^k_{j,m}$ and the inverse of $\boldsymbol{\phi}^{RNA}_{j,m}$ is computed as per equation (4.28). From which the approximated precoder is expressed as

$$\mathcal{C}\mathcal{J}^{RNA}_j = \sqrt{\vartheta_j}\left(\sum_{\ell=1}^{L}\sum_{m=1}^{K}\boldsymbol{\phi}^{RNA}_{\ell,j,m}\right)^{-1}\widehat{\boldsymbol{h}}_{j,j,m} \tag{4.33}$$

Then using the norm minimization, the desired precoding matrix becomes

$$\boldsymbol{F}^{RNA}_j = \frac{\mathcal{C}\mathcal{J}^{RNA}_j}{\left|\mathcal{C}\mathcal{J}^{RNA}_j\right|} \tag{4.34}$$

4.4 Multicell Linear Precoding with Coordination

Spatial dimensions need to be added to allow the BSs to coordinate beamforming vectors in the case of massive MIMO as the BS have many antennas, this further enhances performance in general. This is realizable because of the duality of uplink-downlink in TDD 's massive MIMO systems. In order to enhance the achievable downlink efficiency, the collaboration between BSs is evaluated in different cells and it is assumed that the *jth* BS will know the precoding of RNA, RZF or TPE and the approximation of channels, $\widehat{\boldsymbol{h}}_{\ell,j,m}$, for each of the

single cell in the network. The linear precoding for multicell with cooperation amongst the BSs is computed as follows [197]

$$y_{j,m} = \left[\boldsymbol{h}_{1,jm}^{H} \ \ldots \ \boldsymbol{h}_{j,j,m}^{H} \ldots \ \boldsymbol{h}_{L,j,m}^{H}\right]$$

$$\begin{bmatrix} \boldsymbol{F}_1 & & & & \\ & \ddots & & & \\ -\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,m}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{F}_1 & -\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,m}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{F}_2 \ldots & \boldsymbol{F}_j \ldots & -\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,m}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{F}_L \\ & & & \ddots & \\ & & & & \boldsymbol{F}_L \end{bmatrix} \begin{bmatrix} S_1 \\ \vdots \\ S_j \\ \vdots \\ S_L \end{bmatrix} \tag{4.35}$$

$$+n_{j,m}$$

These can be rewritten as in (4.5) to yield

$$\begin{aligned} y_{j,m} &= \boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}s_{j,m} + \sum_{k\neq m}\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,k}s_{j,k} \\ &+ \sum_{(\ell,k)\neq(j,m)}\left(\boldsymbol{h}_{\ell,j,m}^{H} - \boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\right)\boldsymbol{f}_{\ell,k}s_{\ell,k} + n_{j,m} \end{aligned} \tag{4.36}$$

From (4.36), the new matrix of the multicell precoding has an extra term $\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell}$ added, which embodies BS cooperation, in comparison to the matrix of the classical multicell precoding in (4.5). By this modification of the classical multicell RNA and TPE precoding matrix structure, we can minimize the inter-cell interference and enhance the achievable downlink sum rate. Although cooperation between BSs will increase the complexity of the system, the base station switch associated with these BSs can effectively acquire the required data. Then (4.36) can be modified by invoking the channel hardening property to yield

$$y_{j,m} = \mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right]s_{j,m} + var\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right]s_{j,m} + \sum_{(\ell,k)\neq(j,m)}\left(\boldsymbol{h}_{\ell,j,m}^{H} - \boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\right)\boldsymbol{f}_{\ell,k}s_{\ell,k} + n_{j,m} \tag{4.37}$$

We apply the identities in (4.38) and (4.39) to the second and third terms on the RHS of (4.37)

$$var\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right] = \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right] - \left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right]\right|^{2} \tag{4.38}$$

$$\sum_{(\ell,k)\neq(j,m)}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] = \sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] - \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right] \tag{4.39}$$

The third term on the RHS of (4.37) is modified using (4.39) to obtain

$$\sum_{(\ell,k)\neq(j,m)}\mathbb{E}\left[\left|\left(\boldsymbol{h}_{\ell,j,m}^{H} - \boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\right)\boldsymbol{f}_{\ell,k}\right|^{2}\right]$$

$$= \sum_{(\ell,k)\neq(j,m)}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] - \sum_{(\ell,k)\neq(j,m)}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,k}\right|^{2}\right]$$

$$= \sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] - \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right] -$$

$$\sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,k}\right|^{2}\right] + \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{j,jm}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{j,m}\right|^{2}\right] \tag{4.40}$$

$$= \sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] - \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right] -$$

$$\sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,k}\right|^{2}\right] + \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right]$$

$$= \sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right] - \sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,k}\right|^{2}\right]$$

Using (4.40) we rewrite (4.8) as follows

$$SINR_{j,m} =$$

$$\frac{\left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right]\right|^{2}}{\sigma^{2}+\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right|^{2}\right]-\left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}\right]\right|^{2}+\sum_{\ell=1}^{L}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,k}\right|^{2}\right]-\boldsymbol{\Upsilon}_{cp}} \tag{4.41}$$

where the term $\boldsymbol{\Upsilon}_{cp}=\sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,k}\right|^{2}\right]$ and captures the cooperation amongst BSs in multicell massive MIMO. For the RZF precoding, we consider the SINR of the *m*th UT in the *j*th cell and use (4.12) and (4.41) to obtain

$$SINR_{j,m}^{RZF}=\frac{\left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RZF}\right]\right|^{2}}{\sigma^{2}+\boldsymbol{\gamma}^{RZF}} \tag{4.42}$$

where

$$\boldsymbol{\gamma}^{RZF}=\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RZF}\right|^{2}\right]-\left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RZF}\right]\right|^{2}+\sum_{\ell=1}^{L}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,m}^{RZF}\right|^{2}\right]-\boldsymbol{\Upsilon}_{cp}^{RZF}$$

and

$$\boldsymbol{\Upsilon}_{cp}^{RZF}=\sum_{\ell,k}\mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,m}^{RZF}\right|^{2}\right]$$

For the TPE precoding, we consider the SINR of the *m*th UT in the *j*th cell and use (4.18) and (4.41) to obtain

$$SINR_{j,m}^{TPE}=\frac{\left|\mathbb{E}\left[\boldsymbol{\omega}_{j}^{H}\boldsymbol{a}_{j,m}\right]\right|^{2}}{\frac{\sigma^{2}}{K}+\boldsymbol{\gamma}^{TPE}} \tag{4.43}$$

where

$$\boldsymbol{\gamma}^{TPE}=\mathbb{E}\left[\left|\boldsymbol{\omega}_{j}^{H}\boldsymbol{a}_{j,m}\right|^{2}\right]-\left|\mathbb{E}\left[\boldsymbol{\omega}_{j}^{H}\boldsymbol{a}_{j,m}\right]\right|^{2}+\sum_{\ell=1}^{L}\mathbb{E}\left[\left|\boldsymbol{\omega}_{\ell}^{H}\boldsymbol{B}_{\ell,j,m}\boldsymbol{\omega}_{\ell}\right|^{2}\right]-\boldsymbol{\Upsilon}_{cp}^{TPE}$$

and

$$\boldsymbol{\Upsilon}_{cp}^{TPE} = \sum_{\ell,k} \mathbb{E}\left[\left|\boldsymbol{\omega}_{\ell}^{H}\boldsymbol{a}_{j,m}\frac{diag\{\boldsymbol{B}_{\ell,j,m}\}}{diag\{\boldsymbol{a}_{j,m}\}}\boldsymbol{\omega}_{\ell}\right|^{2}\right]$$

For the RNA precoding, we consider the SINR of the *m*th UT in the *j*th cell and use (4.41) to obtain

$$SINR_{j,m}^{RNA} = \frac{\left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RNA}\right]\right|^{2}}{\sigma^{2} + \boldsymbol{\gamma}^{RNA}} \tag{4.44}$$

where

$$\boldsymbol{\gamma}^{RNA} = \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RNA}\right|^{2}\right] - \left|\mathbb{E}\left[\boldsymbol{h}_{j,j,m}^{H}\boldsymbol{f}_{j,m}^{RNA}\right]\right|^{2} + \sum_{\ell=1}^{L}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,j,m}^{H}\boldsymbol{f}_{\ell,m}^{RNA}\right|^{2}\right] - \boldsymbol{\Upsilon}_{cp}^{RNA}$$

and

$$\boldsymbol{\Upsilon}_{cp}^{RNA} = \sum_{\ell,k} \mathbb{E}\left[\left|\boldsymbol{h}_{j,j,m}^{H}\frac{diag\{\widehat{\boldsymbol{h}}_{\ell,j,k}\}}{diag\{\widehat{\boldsymbol{h}}_{j,j,m}\}}\boldsymbol{f}_{\ell,m}^{RNA}\right|^{2}\right]$$

The BSs cooperation term in all the precoding schemes works to mitigate inter-cell interference, thus improving the SINR and therefore the sum rate.

4.5 Energy Efficiency

The Energy Efficiency (EE) of the different precoding schemes is evaluated using a realistic Circuit Power (CP) consumption model. Next the EE-SINR trade-off is then evaluated. In [207], [208], the EE is specified as

$$EE = \frac{Throughput[bits/s/cell]}{Power\ consumtion[W/cell]} \tag{4.45}$$

and the computation is performed in bits per joule. We can view this as benefit-cost association where the throughput (quality of service) is pegged on the corresponding costs (power consumption). Realistic EE estimation involves calculating power consumption (PC) based on Effective Transmit Power (ETP) and the CP required to power the cellular network [208].

$$PC = ETP + CP \tag{4.46}$$

The consumption model of CP in a massive MIMO system for the *jth* BS can be computed as [209]–[214]

$$CP_j = P_{FIX,j} + P_{TC,j} + P_{CE,j} + P_{C/D,j} + P_{BH,j} + P_{SP,j} \tag{4.47}$$

where $P_{FIX,j}$ represents a fixed quantity of power. $P_{FIX,j}$ caters for the control signalling power requirement, the backhaul link load-independent power requirement and the baseband processors and the power associated with economic expenses. Also, $P_{TC,j}$ represents the transceiver chains power requirement, $P_{CE,j}$ accounts for channel estimation power requirement, $P_{C/D,j}$ is the power requirement for channel encoding and decoding components, $P_{BH,j}$ accounts for the load-dependent backhaul signalling power requirement, and $P_{SP,j}$ accounts for BS signal processing power requirement.

Then we can express $P_{FIX,j}$ as follows

$$P_{FIX,j} = P_{FIX1,j} + P_{FIX2,j} \tag{4.48}$$

where $P_{FIX1,j}$ is the control signalling power requirement and the backhaul link load-independent and the baseband processors power requirements and $P_{FIX2,j}$ is the power ascribed to economic expenses [215] and expressed as

$$P_{FIX2,j} = \sum_{\mathcal{v}=1}^{V} \mathcal{K}_r(\mathcal{v})\, \mathcal{U}(\mathcal{v}, SE) + (C_0 + \mathcal{K}_c p_{bs}) \tag{4.49}$$

where V is the quantity of traffic classes, $\mathcal{U}(\mathcal{v}, SE)$ is the actual chargeable information throughput corresponding to traffic class $\mathcal{v}$, $\mathcal{K}_r(\mathcal{v})$ is the income generated per unit data of traffic class $\mathcal{v}$, p_{bs} power expended by BSs during data transmission, $\mathcal{K}_c$ is the cost of energy per joule, and C_0 is the additional costs on top of the energy costs. For this work we consider three traffic classes (*V=3*), representing voice traffic (class 1), data traffic with limited volume (class 2), and data traffic with unlimited volume (class 3), are considered in this paper.

The $P_{TC,j}$ for the *jth* cell can be computed as in [210], [211] and expressed as

$$P_{TC,j} = \underbrace{M_j P_{BS,j} + P_{LO,j}}_{BS\ Circuit\ component} + \underbrace{K_j P_{UT,j}}_{UT\ Circuit\ component} \tag{4.50}$$

and the $P_{C/D,j}$ for the *jth* cell is computed as in [212] and expressed as

$$P_{C/D,j} = (P_{COD} + P_{DEC})CT_j \tag{4.51}$$

Then for $P_{BH,j}$ we model it in two parts: the load-independent component and the load-dependent component [213]. The load-independent component is catered for under the $P_{FIX,j}$ while the load-dependent component is computed as

$$P_{BH,j} = \underbrace{P_{BT,j}}_{\textit{Backhaul traffic power}} . CT_j \tag{4.52}$$

Then the power for estimating the channel $P_{CE,j}$ is approximated based on the estimator employed and expressed as

$$P_{CE,j} = \frac{3B}{\tau_c L_{BS}} K_j \begin{cases} M_j\tau_p + M_j^2, & with\ MMSE \\ M_j\tau_p + M_j, & With\ LS \end{cases} \tag{4.53}$$

where B is the bandwidth, L_{BS} is the BS computational efficiency, τ_c is the time of coherence and τ_p is the pilot sequence length [216].

The power consumed by the *jth* BS $P_{SP,j}$, for receive combining and transmit precoding, is dependent on the computational complexity of the precoding methods employed. $P_{SP,j}$ is further decomposed as in [216] to obtain

$$P_{SP,j} = \underbrace{P_{SP-R/T,j}}_{\textit{Reception/ transmission}} + \underbrace{P_{SP-C,j}^{UL}}_{\textit{Computing combining}} + \underbrace{P_{SP-C,j}^{DL}}_{\textit{Computing precoding}} \tag{4.54}$$

where $P_{SP-R/T,j}$ is the total power consumed in UL reception and data stream transmission in DL, $P_{SP-C,j}^{UL}$ is power consumed during combining vector computation and $P_{SP-C,j}^{DL}$ is power consumed during the precoding vector computation at the *j*th BS. The $P_{SP-R/T,j}$ is further decomposed as follows

$$P_{SP-R/T,j} = \frac{3B}{\tau_c L_{BS}} M_j K_j (\tau_u + \tau_d) \tag{4.55}$$

while the $P_{SP-C,j}^{DL}$ is evaluated as

$$P_{SP-C,j}^{DL} = \frac{4B}{\tau_c L_{BS}} M_j K_j \tag{4.56}$$

The $P_{SP-C,j}^{UL}$ power consumption relies on the method of precoding employed.

4.5.1 Receive Combining Computational Complexity

It is presumed that computational complexity is highly impacted by the complex division and multiplication, and in so doing we neglect the addition and subtractions operations [216]. The precoding matrix corresponding to different precoding techniques is thought to highly impact the complexity. We rely on lemmas in appendix B to estimate the complexity.

For RZF precoding, we can express its combining complexity as in [216] and obtain

$$\frac{3K_j^2 M_j}{2} + \frac{3K_j M_j}{2} + \frac{K_j^3 - K_j}{3} + 7K_j \tag{4.57}$$

Starting with equation (4.14), we evaluate the TPE precoder combining complexity using the appendix 0. From which a complexity of $\frac{\left(3K_j^2+K_j\right)M_j}{2}$ is obtained when the multiplication $\sum_{k=1}^{K_l} \widehat{\boldsymbol{h}}_{j,j,k}\widehat{\boldsymbol{h}}_{j,j,k}^{\boldsymbol{H}} = \boldsymbol{H}_j\boldsymbol{H}_j^H$ in (4.14) is considered. The power of the TPE precoder was set to 2 since a J_j of 3 was used, this caters for the power n of the multiplication $\sum_{k=1}^{K_l} \widehat{\boldsymbol{h}}_{j,j,k}\widehat{\boldsymbol{h}}_{j,j,k}^{\boldsymbol{H}} = \boldsymbol{H}_j\boldsymbol{H}_j^H$. Due to this then we obtain a complexity of K_j^3. A cost of K_j divisions per BS is necessitated by the need of precoding vector normalization within the decoding unit. Hence, we obtain the total complexity by the following

$$\sum_{n=0}^{J_j-1}\left(\frac{\left(3K_j^2 + K_j\right)M_j}{2} + K_j^3\right) + 7K_j \tag{4.58}$$

Based on equation (4.33), we evaluate the RNA precoder combining complexity using the Appendix B. If we consider the initial three terms as suggested in [195] we obtain the quickest convergence for the RNA inverse iterative process. We presume that the term $\boldsymbol{Z}_j$ freely obtainable at the *j*th BS. The computational complexity of intercell channel estimates for $\ell \neq$ j used in RNA need to be accounted for. Thus, we express the multiplication complexity as $\sum_{\ell=0}^{L} \frac{\left(3M_j^2+M_j\right)K_\ell}{2}$. Then the computational complexity due to matrix inversion is obtained as

follows. Considering the first three terms *(P = 3)* in a successive iteration as in [195], the inverse can be expressed as

$$\boldsymbol{R}_1 = \boldsymbol{R}_0(\boldsymbol{I} + \boldsymbol{E}_0 + \boldsymbol{E}_0^2)$$

$$\boldsymbol{R}_2 = \boldsymbol{R}_1\left(\boldsymbol{I} + \boldsymbol{E}_0^3 + (\boldsymbol{E}_0^3)(\boldsymbol{E}_0^3)\right) \tag{4.59}$$

$$\boldsymbol{R}_3 = \boldsymbol{R}_2(\boldsymbol{I} + \boldsymbol{E}_0^9 + \boldsymbol{E}_0^{18}) = \boldsymbol{R}_2\left(\boldsymbol{I} + (\boldsymbol{E}_0^6)(\boldsymbol{E}_0^3) + (\boldsymbol{E}_0^9)(\boldsymbol{E}_0^9)\right)$$

and we generally express (4.59) as

$$\boldsymbol{R}_{\hbar} = \boldsymbol{R}_{\hbar-1}\left(\boldsymbol{I} + \left(\boldsymbol{E}_0^{3^{\hbar-1}}\right) + \left(\boldsymbol{E}_0^{3^{\hbar-1}}\right)^2\right) \tag{4.60}$$

For the first iteration $\boldsymbol{R}_1$, we express the total operations count to compute it as

$$\boldsymbol{N}_1 = 2M_j^3 - M_j^2 + M_j^2 + 2M_j^3 + M_j^2 + 2M_j^3 - M_j^2 = 6M_j^3 \tag{4.61}$$

For the first iteration $\boldsymbol{R}_2$, we express the total operations count to compute it as

$$\begin{aligned} N_2 &= 2M_j^3 - M_j^2 + M_j^2 + 2M_j^3 - M_j^2 + M_j^2 + 2M_j^3 - M_j^2 \\ &= 6M_j^3 - M_j^2 \end{aligned} \tag{4.62}$$

For the first iteration $\boldsymbol{R}_3$, we express the total operations count to compute it as

$$\begin{aligned} N_3 &= 2M_j^3 - M_j^2 + M_j^2 + 2M_j^3 - M_j^2 + M_j^2 + 2M_j^3 - M_j^2 \\ &= 6M_j^3 - M_j^2 \end{aligned} \tag{4.63}$$

Hence, in performing $\mathfrak{a}$ iterations, the total operations count for $P = 3$ can be expressed as

$$N_3 = (\mathfrak{a} - 1)\left(6M_j^3 - M_j^2\right) + 6M_j^3 = 6M_j^3\mathfrak{a} - M_j^2(\mathfrak{a} - 1) \tag{4.64}$$

The RNA precoder inversion complexity is computed in equation (4.64). Then we perform a multiplication after inversion yielding a complexity of $M_j\tau_p(\tau_p - K_j)$ and the normalization carried out on the precoding vector in the decoding unit costs M_j divisions per BS. Consequently, we obtain the complexity of the RNA precoder as follows

$$\sum_{\ell=0}^{L} \frac{\left(3M_j^2 + M_j\right)K_\ell}{2} + \left(6M_j^3 \mathfrak{d} - M_j^2(\mathfrak{d} - 1)\right) + 2M_j + M_j\tau_p(\tau_p - K_j) \tag{4.65}$$

It is assumed that each iteration corresponds to the time instance of coherence. The rise in the number of iterations lowers the complexity of the RNA precoding inversion by a factor of $M_j^2(\mathfrak{d} - 1)$.

4.5.2 Receive Combining Power Computation

To obtain the receive combining power for different precoding techniques we rely on the corresponding receive combining computational complexity. Then we consider each of the precoding schemes individually to realize the corresponding receive combining power. First, we look at RZF precoding and compute the receive combining power as follows

$$P_{SP-C,j}^{UL} = \frac{3B}{\tau_c L_{BS}}\left(\frac{3K_j^2 M_j}{2} + \frac{3K_j M_j}{2} + \frac{K_j^3 - K_j}{3} + \frac{7}{3}K_j\right) \tag{4.66}$$

For the TPE precoding, we compute the receive combining power by the following expression

$$P_{SP-C,j}^{UL} = \frac{3B}{\tau_c L_{BS}}\left(\sum_{n=0}^{J_j-1}\left(\frac{\left(3K_j^2 + K_j\right)M_j}{2} + K_j^3\right) + \frac{7}{3}K_j\right) \tag{4.67}$$

Then finally, we look at the RNA precoding and evaluate the receive combining power through the following expression

$$P_{SP-C,j}^{UL} = \frac{3B}{\tau_c L_{BS}}\left(\sum_{\ell=0}^{L} \frac{\left(3M_j^2 + M_j\right)K_\ell}{2} + \left(6M_j^3 \hbar - M_j^2(\hbar - 1)\right) + 2M_j + M_j\tau_p(\tau_p - K_j)\right) \tag{4.68}$$

4.5.3 Throughput and Energy Efficiency

The CP model developed in the previous section is invoked in performing EE-SE evaluation to bring into focus the significance of bandwidth in analysing EE. Thus, we express the EE corresponding to the *j*th cell as follows

$$EE = \frac{TR_j}{ETP_j + CP_j} \tag{4.69}$$

where TR_j represents the capacity for *j*th cell, ETP_j is the ETP for *j*th cell and CP_j is the CP in the *j*th cell. Then we evaluate the TR_j by the following expression

$$TR_j = B \sum_{k=1}^{K_j} \left(\boldsymbol{SINR}_{jk}^{UL} + \boldsymbol{SINR}_{jk}^{DL} \right) \tag{4.70}$$

The ETP_j is computed as per [216] and expressed as

$$ETP_j = \frac{\tau_p}{\tau_c} \sum_{k=1}^{K_j} \frac{1}{\mu_{UE,jk}} p_{jk} + \frac{\tau_u}{\tau_c} \sum_{k=1}^{K_j} \frac{1}{\mu_{UE,jk}} p_{jk} + \frac{1}{\mu_{BS,j}} \frac{\tau_d}{\tau_c} \sum_{k=1}^{K_j} \rho_{jk} \tag{4.71}$$

where τ_u is the uplink time and τ_d is the downlink time, $\mu_{UE,jk}$ is the power amplifier (PA) efficiency of the *k*th UT in the *j*th cell and $\mu_{BS,j}$ is the PA at the *j*th BS.

4.6 Simulation Results and Discussion

The comparison and analysis of RNA-based precoding, TPE precoding and RZF precoding is presented. We consider a multicell massive MIMO system with 16 cells each of which has BS with M antennas varying from $M = 16 \ to \ 128$, with step size of 16 and $K = 6$ user terminals. We choose a polynomial power, J_j, to be 3 for TPE precoding in this particular analysis and comparison. Further we presume CSI quality of $\tau = 0.45$ and the average SNR ranges from 0dB to 20dB.

In this section we choose to vary the reuse frequency because it is the factor that has a major impact on pilot contamination. The effect due to variation in τ will simply follow the observations presented in Chapter 3.

Figure 4.1 contrasts the achievable SE per cell with a reuse factor of 1 between RNA-based precoding, TPE precoding and RZF precoding. We can deduce a number of observations from this figure. Precoding based on RNA has the best achievable SE per cell, followed by RZF precoding and then TPE precoding. It can be observed that the performance of TPE precoding improves to match that of RZF precoding with increase in the number of BS antennas. TPE precoding performance is below the other two since it is merely an approximation of the RZF

precoding. Yet as the number of BS antennas increases, the approximation gets better due to space-time diversity and the hardening of the channels inherent to large antenna numbers. It is evident that the RNA-based precoder performs better than the TPE precoder and the RZF precoder under the same propagation condition. It is clear that, under the same propagation condition, the performance of the RNA-based precoder surpasses that of the TPE precoder and the RZF precoder.

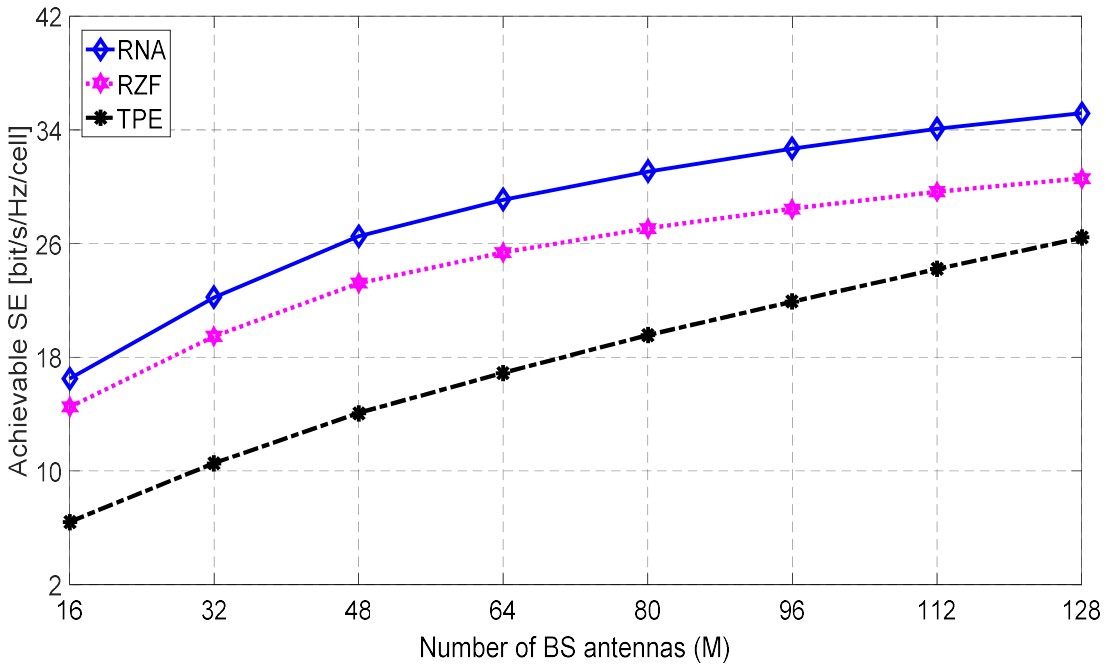

Figure 4.1: Achievable SE per cell vs. number of BS antennas with a reuse factor of $f = 1$.

In the presence of reuse factors 2 and 4, we compare the RNA-based precoding, TPE precoding and RZF precoding as in Figure 4.2 and Figure 4.3. The RNA-based precoding gives the best SE per cell at the reuse factor of 2 as compared to TPE precoding and RZF precoding with TPE precoding yielding the lowest performance. In addition, the SE per cell for RNA-based precoding and the RZF precoding increases with increase in the reuse factor from 1 to 2 while that of TPE precoding decreases. At a reuse factor of 4, the SE per cell from RNA precoding increase but that of TPE precoding and RZF precoding decreases. It clearly shows that the TPE precoding and RZF precoding can be extended to a single cell scenario, but its throughput plummets in a multicell massive MIMO network where reuse factors are invoked to allow efficient use of the limited radio resources.

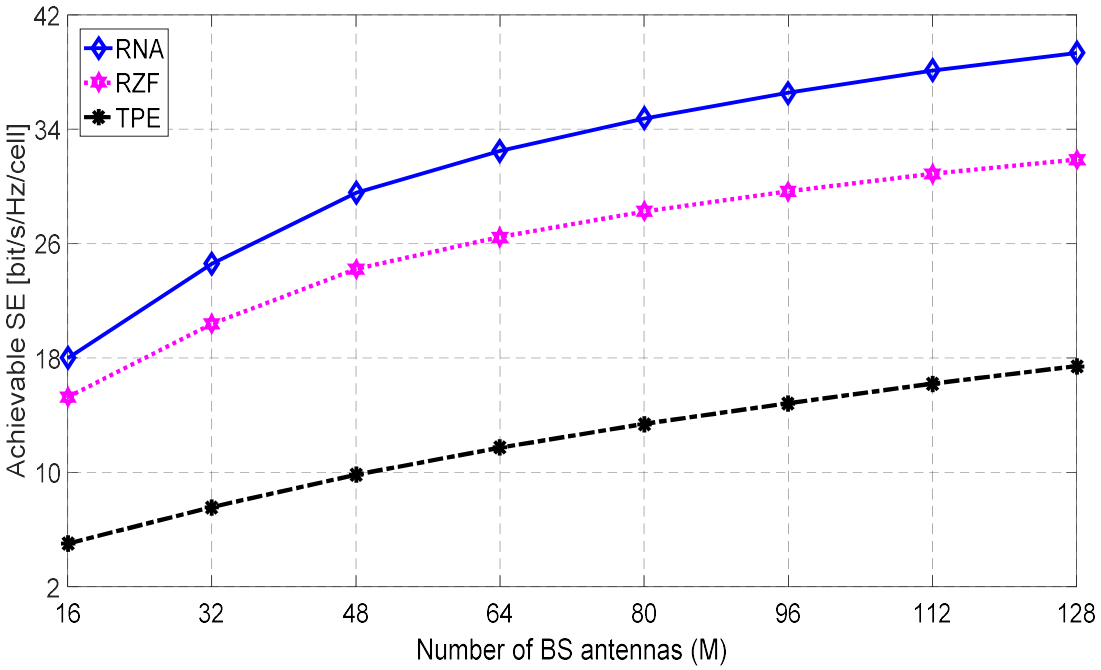

Figure 4.2: Achievable SE per cell vs. number of BS antennas with a reuse factor of $f = 2$.

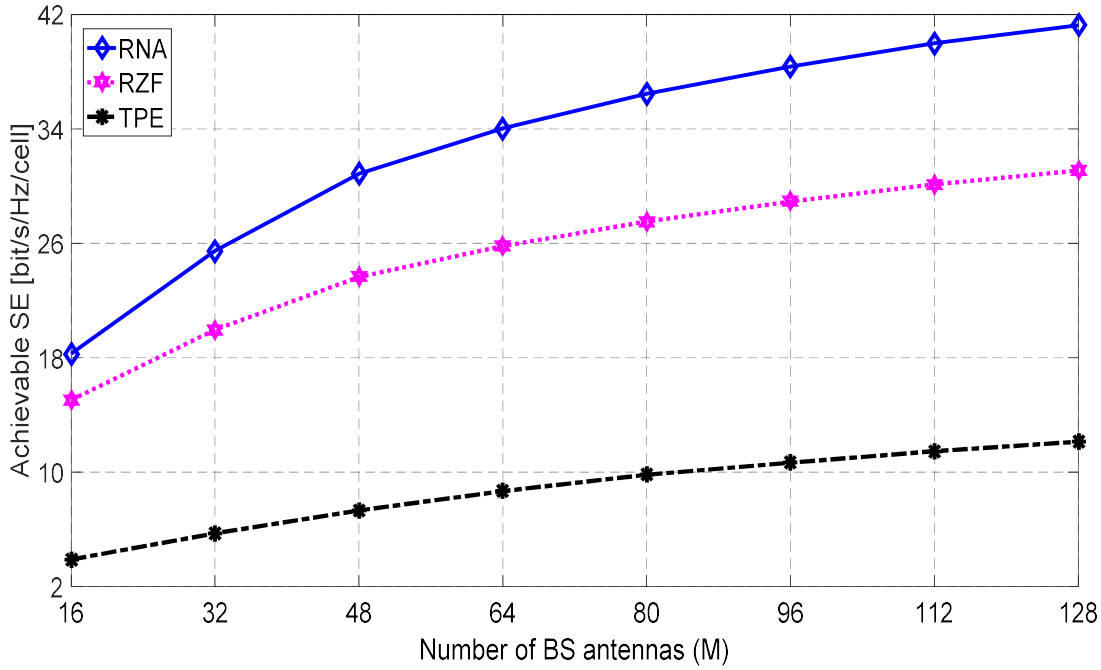

Figure 4.3: Achievable SE per cell vs. number of BS antennas with a reuse factor of $f = 4$.

The rise in SE per cell can be due to the fact that, with the increased number of pilots, the pre-log factor decreases. As the channel estimates get better with reduced pilot interference, it enhances the instantaneous SINRs. Also, there is a reduction in SE for TPE precoding since an increase in approximation efficiency does not outweigh the pre-log factor reduction. It is a fact

as the approximation is evoked only to raise the desired signal and not to minimize interference. Therefore, RNA-based precoding performs well in in presence of a high reuse factor and can easily render itself to multicell massive MIMO systems. The performance of the various precoding schemes for different reuse factors is tabulate in Table 4.1.

Table 4.1: Average down link sum SE for M = 128 and K = 6 for diverse pilot reuse factors *f*. The highest value for each scheme is in bold face. The values are based on Figures 1, 2 and 3.

Scheme	*f = 1*	*f = 2*	*f = 4*
RNA	35.185	39.340	***41.258***
RZF	30.592	***31.880***	31.092
TPE	***26.399***	17.435	12.129

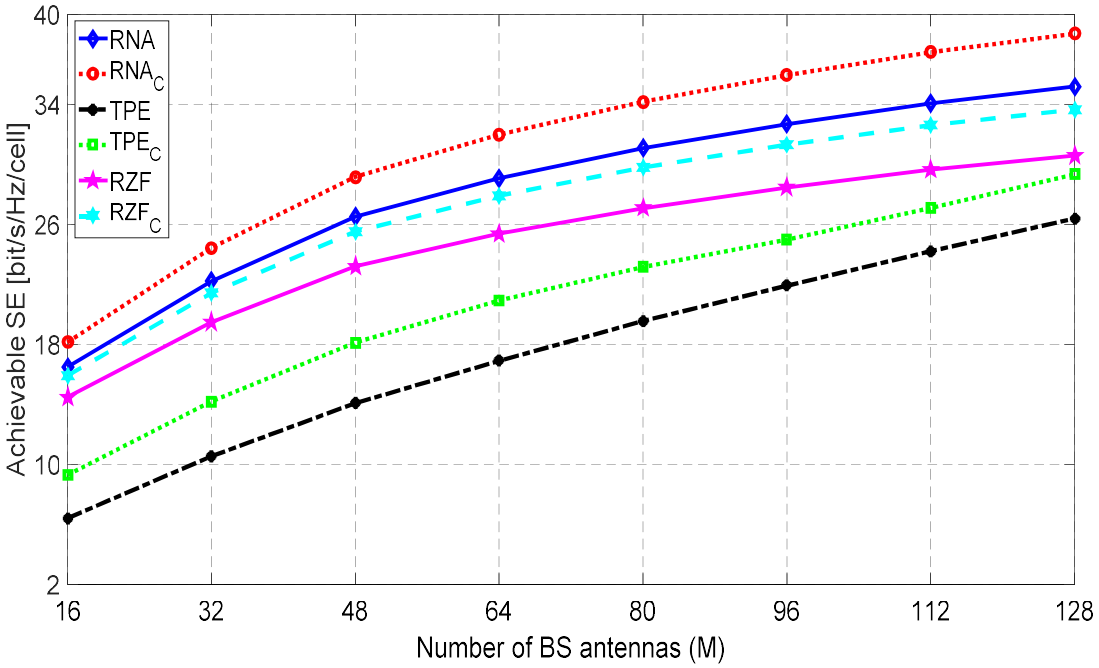

Figure 4.4: Achievable SE per cell vs. number of BS antennas with coordination and a reuse factor of 1.

Figure 4.4 and Figure 4.5 demonstrates the coordination effect in massive MIMO system with multicell. With the selected reuse factor of 1 and 4, it can be observed that the coordination within the multicell massive MIMO network increases the SE per cell for all the precoding schemes under consideration while keeping constant the BS antennas and the UTs. Thus, in multicell massive MIMO with coordination, the pilot contamination is reduced further leading

to general improvement in SE per cell. Although this performance enhancement comes at the expense of increased complexity which merely points to increased hardware requirements.

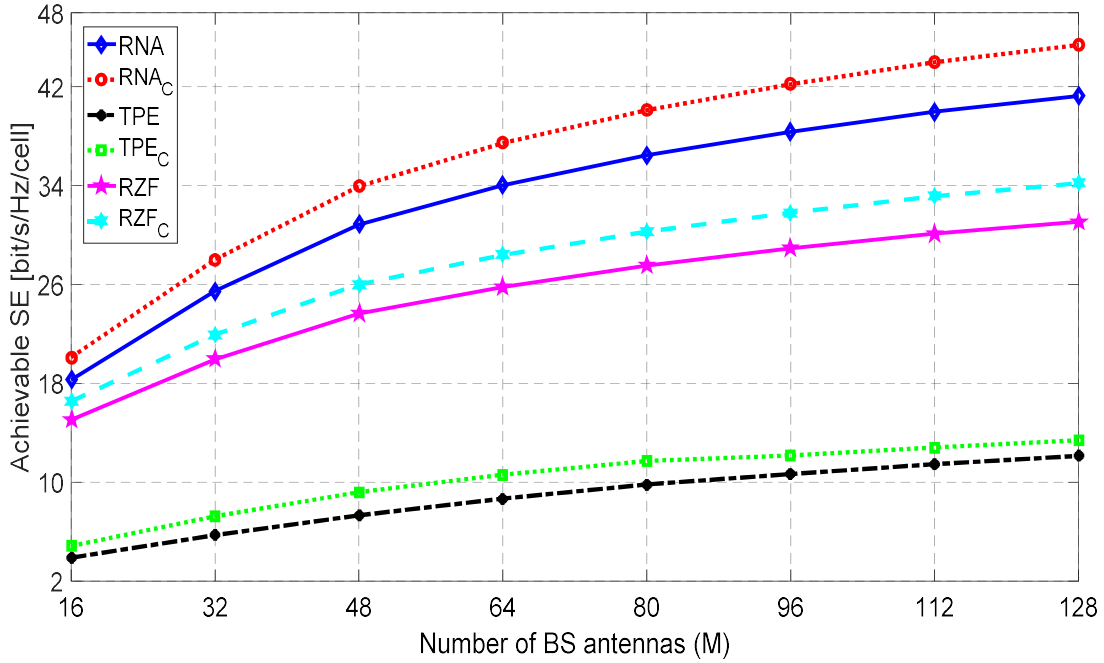

Figure 4.5: Achievable SE per cell vs. number of BS antennas with coordination and a reuse factor of 4.

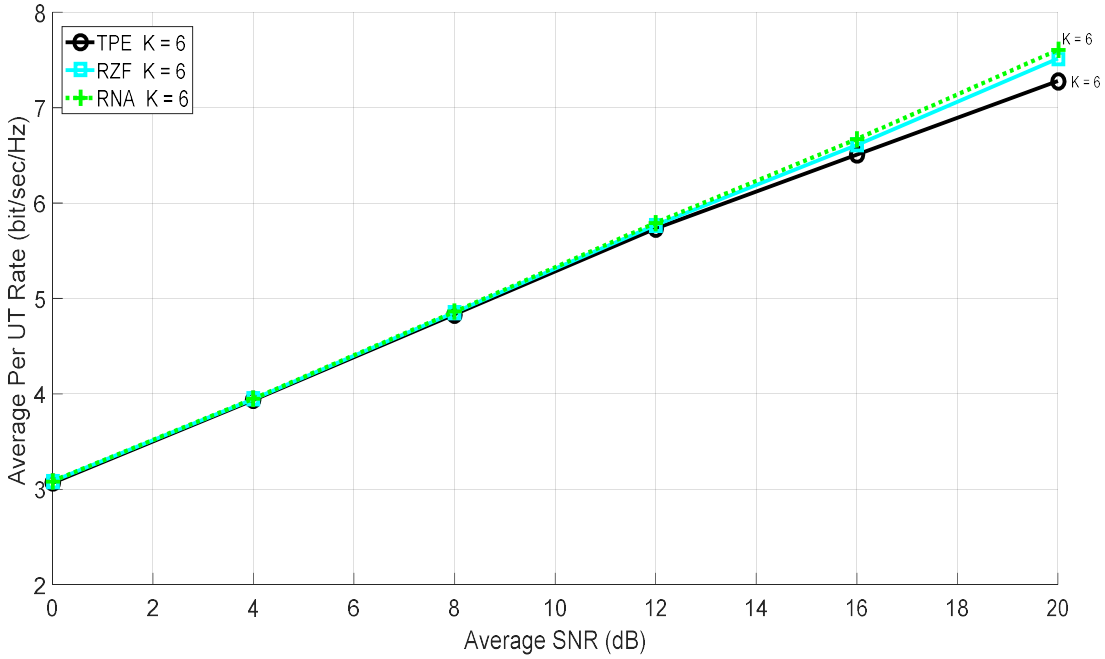

Figure 4.6: Average per UT rate vs. SINR for CSI error of $\tau = 0.45$ at the BS $(M = 128, K = 6, J = 3)$.

The average bit rates are presented in Figure 4.6 above for multicell RZF precoding, RNA precoding and the TPE precoding. The precoding schemes are evaluated under the channel knowledge of $\tau = 0.45$ with $J = 3$. The multicell RNA and RZF precoding have relatively similar performance in terms of the bit rates which surpasses that of the multicell TPE precoder. This again validates the superior performance of the multicell RNA and RZF precoders over multicell TPE precoder particularly at large values of SNR.

Table 4.2: CP model parameters.

Parameter	Value set
Fixed power: P_{FIX}	6.5 W
Power for BS LO: P_{LO}	0.125 W
Power for BS antennas: P_{BS}	0.225 W
Power per UT: P_{UT}	0.125 W
Power for data encoding: P_{COD}	0.02 W/(Gbit/s)
Power for data decoding: P_{DEC}	0.09 W/(Gbit/s)
BS computational efficiency: L_{BS}	775 Gflops/W
Power for backhaul traffic: P_{BT}	0.0225 W/(Gbit/s)

Table 4.3: Cost parameters.

Cost Parameters	Value
$\mathcal{K}_r(1)$	$1.45 \times 10^{-5}\ Rands/bit$
$\mathcal{K}_r(2)$	$7.85 \times 10^{-7}\ Rands/bit$
$\mathcal{K}_r(3)$	$2.75 \times 10^{-8}\ Rands/bit$
$\mathcal{K}_c$	$4.33 \times 10^{-6}\ Rands/J$
C_0	$4.25 \times 10^{-4}\ Rands/s$

The cumulative CP against the number of BS antennas is shown in Figure 4.7 below. The CP was calculated according to the values given in Table 4.2 and Table 4.3. It can be observed that the total CP increases as the number of BS antennas increases. This is anticipated as the rise in BS antennas clearly implies increased circuitry to consume power. But precoding based on RNA has a high consumption of CP compared to RZF precoding and TPE precoding which have the same CP.

Figure 4.8 and Figure 4.9 illustrates the variation of EE with the throughput. It can be observed from the two figures that the EE increases with an increase in the throughput to a certain point then the EE starts to decrease with an increase in the throughput.

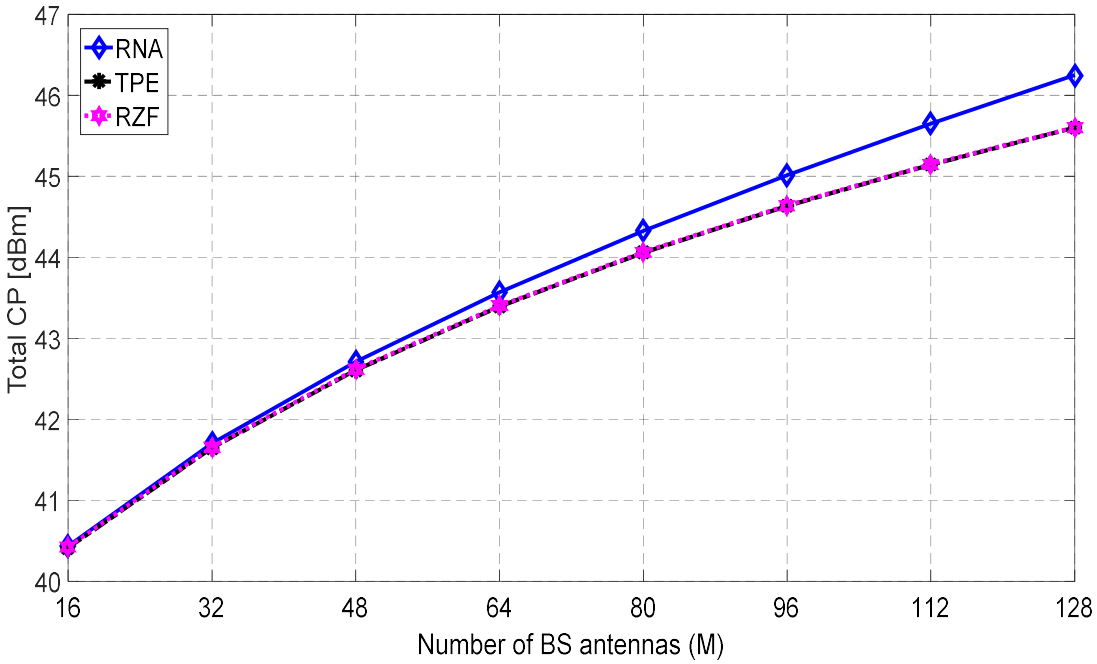

Figure 4.7: Total CP vs. number of BS antennas.

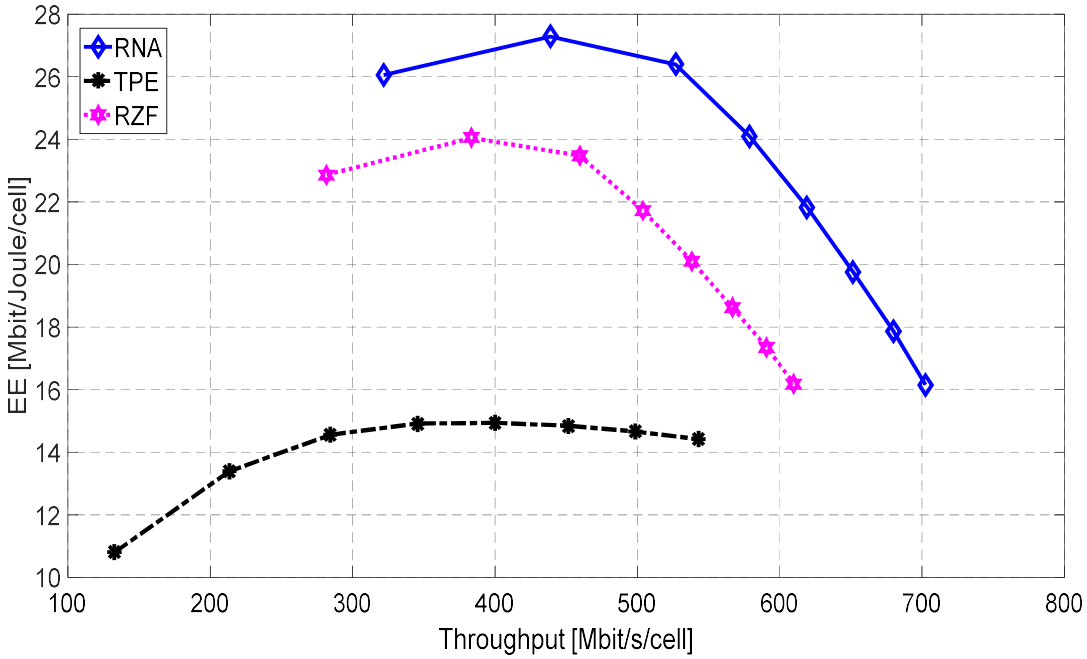

Figure 4.8: EE per cell vs. Capacity per cell for a reuse factor of 1.

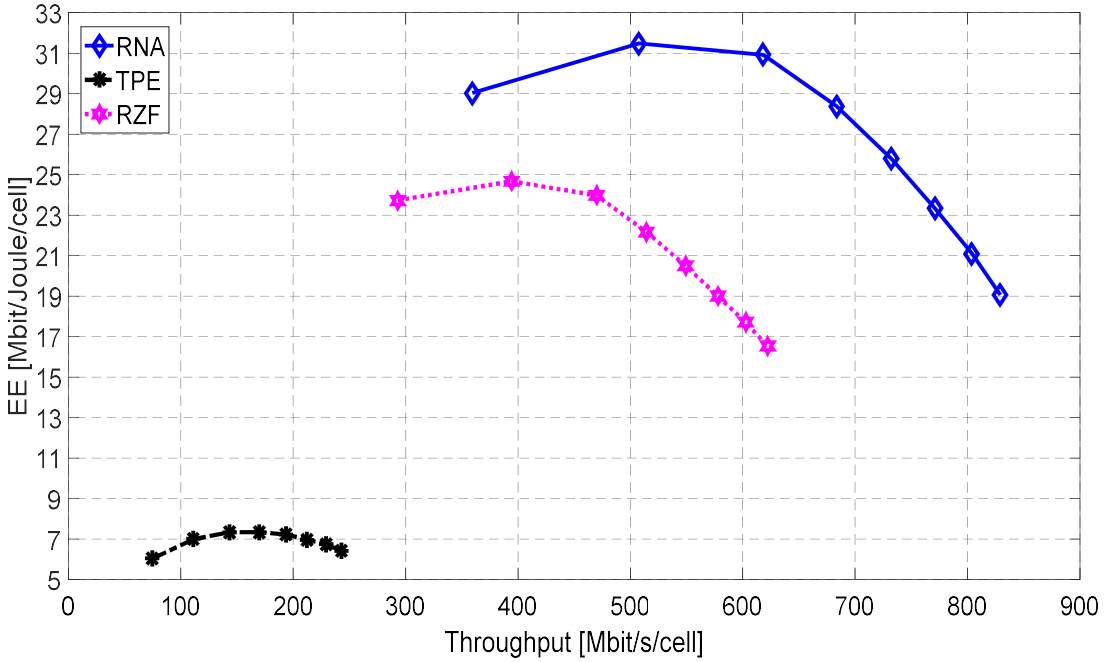

Figure 4.9: EE per cell vs. Capacity per cell for a reuse factor of 4.

The achievable maximum EE serves as a measure of the optimal throughput that will produce maximum EE. This throughput corresponds therefore to the number of BS antennas needed to achieve the optimal EE. It can be observed therefore that the use of all BS antennas may not be optimal. In massive MIMO scheme, this brings in the concept of antenna selection within the BS. This can also be set to be dynamic to compensate for the specific state of fading for the respective receiving points.

Another observation is on the reuse factor. The reuse factor increments lead to an increase in the maximum EE and the corresponding throughput especially for RNA precoder. This is due to the reduction of intercell interference. The RNA precoder has the highest maximum EE followed by the RZF precoder. The TPE precoder has a poor EE and a very low throughput which reduces with the increase in the reuse factor. This limits its use in Massive MIMO as a precoding technique.

4.7 Conclusion Multicell Precoding

The chapter the iterative linear precoder was adopted to multicell massive MIMO with coordination the receive combining computational complexity modelled. We present the performance overview and comparison of the RZF precoder, TPE precoder and RNA-based precoder for a multicell massive MIMO system in the downlink. The performance of the three

precoding schemes for imperfect CSI is analysed in terms of SE and the EE. Under similar assumptions, the SE and EE were theoretically developed for each of the precoding schemes and for the downlink massive MIMO system.

RNA-based precoding posts a higher SE and EE as per the simulation and the theoretical performance followed by RZF precoding, while the TPE precoder has the lowest. Although when the BS antennas are increased, with the reuse factor of 1, the TPE precoder SE and EE approaches that of the RZF precoder. The SE and EE performance of TPE precoder is poor at a higher reuse factor of 4 and cannot be compared to that of the RZF precoder and the RNA-based precoder. Thus, RNA-based precoder and RZF precoder have excellent performance for both lower and higher reuse factor in comparison to TPE precoder which has only better performance with a high number of BS antennas and a lower reuse factor of 1.

TPE and RZF precoders have similar CP consumption which is lower than the CP consumed by the RNA based precoder. While TPE precoder consumes the same CP as the RZF precoding, it has SE and EE that are poor. This points out that in multicell massive MIMO networks, TPE precoder is an inefficient scheme to be realized. The precoder based on RNA has high CP but its efficiency is superior, and this can justify implementing it in multicell massive MIMO networks.

CHAPTER 5

Channel Estimation and Analysis

5.1 Introduction

In chapter 3 and chapter 4 it was assumed that the channel state information is known apriori. This is not practical and channel estimation forms one of the key aspects in mitigating pilot contamination. In this chapter, we seek to propose and analyse channel estimation techniques in view to demonstrate the most probable channel estimation techniques to be used in channel state information approximation in fifth generation and later networks.

Efficient utilization of the constrained amount of accessible spectrum in perspective to exponentially growing interest for throughput has been the focal point in communication and signal processing for a couple of decades. The sporadic rise in technologies has galvanized the once predominantly offline appliances and devices to data generation points through the use of sensors and therefore pushing the demand for throughput higher [217], [218]. The current and future communication systems are being enhanced to cater for this as well as conventional mobile users. One of the key technologies for 5G networks is massive MIMO. Technically massive MIMO puts limitations on channel training and the BS architecture [114]. In any case, the outstanding outcomes on the degrees of freedom realizable in a fixed coherence time puts into consideration identically distributed channel coefficients. The second order information of the channel structure vectors is used to mitigate dimensionality bottleneck [219]–[221].

Even for the simplest techniques employed in beamforming, accurate CSI is needed at the BS to enhance suppression of interference ability in massive MIMO systems. Thus, mitigation of pilot contamination is used to facilitate the approximation of covariance matrix considering one channel vector to a single UT. If we assume particular communication setups having particular channel models, there is a possibility of mitigating pilot contamination provided certain separability conditions are satisfied [222]. From [200] it was argued that the elimination of limits on the UL and DL throughput as a result of pilot contamination can be realized if the covariance matrix are considered to be under certain loose conditions. Realization of this

method dictates the estimation of covariance matrix at the base station (BS) and again they are acquired by virtue of observations which are subjected to pilot contamination.

It was demonstrated in [223] that the throughput of massive MIMO grows boundlessly as antenna number goes to infinity provided there is no linear dependence among covariance matrices of co-channel users. In [224] it was proposed that the coherence interval of channel vectors is less than that of the covariance matrix hence allowing room for accurate covariance matrix estimation. According to [225] covariance matrix corresponding to a particular UT is obtained through computation of estimated channel from sample cross-correlation of two pilot sequences.

The investigation of imperfect statistical data for the UL system is discussed in [226]. The SE is evaluated considering imperfect covariance information. Low complexity covariance matrix is presented in [227] where it is shown that based on this estimated covariance matrix, both UL and DL SEs increases with the increase in the number of antennas. Estimation of covariance matrix from compressed data using unbiased estimator is discussed in [228].

It was presented in [163] that the C-RAN presents minimal expenditure to operate the network with corresponding minimal capital accompanied with enhanced energy and spectral efficiency. The combination of C-RAN and massive MIMO has been envisaged as the leading technology to deliver the various stringent standards involving the uses in 5G networks, [164], [166]. However, this hybrid network suffers from finite capacity in the fronthaul which degrades spectral efficiency and energy efficiency in the C-RAN. To alleviate this, several techniques have been suggested including the large-scale precoding/decoding with compression that exhibits minimized overhead. Another alternative is the partially centralised C-RAN structure in which the RRH cooperate with each other and are inter-connected [163]. For the last case the basic functions like beamforming are performed in the RRH and the BBU carries out the functions like channel estimation [165].

In this work, first we consider massive MIMO system for channel estimation and develop the RNA-based channel estimator. We then combine this with the fast data projection method, to create a semi-blind estimator and compare it with the linear channel estimators. Then we consider the partially centralised C-RAN structure in which the massive MIMO RRH cooperate with each other and are inter-connected. Then apply data compression to reduce the fronthaul traffic in the UL and which is combined with the Givens rotation algorithm for channel estimation at the BBU to allow for greater parallelization.

5.2 System Model for Channel Estimation

This work assumes an L-cell massive MIMO system, each with *BS* carrying *M* antennas and *K* UTs with single antenna. We propose that the TDD protocols are coordinated across cells to relay pilot signals and data to all cells simultaneously. Pilots initially transmitted by UTs to their BS in *j*th cell are the same as those given by $\boldsymbol{\Phi}_K = \left[\boldsymbol{\phi}_{j,1}^T, \boldsymbol{\phi}_{j,2}^T, \ldots, \boldsymbol{\phi}_{j,K}^T\right]$, where $\boldsymbol{\phi}_{j,k}$ represents a pilot used by each *k*th UT in each cell and $\left\|\boldsymbol{\phi}_{j,k}\right\|^2 = 1$. Then, within the *j*th cell, a channel from the *l*th *BS* to *k*th UT is given as $\mathbf{h}_{j,\ell,k} \epsilon \mathbb{C}^M$. The channel vectors are assumed to be Rayleigh fading and modelled as

$$\boldsymbol{h}_{j,\ell,k} \sim \boldsymbol{\mathcal{CN}}\left(\mathbf{0}, \boldsymbol{R}_{j,\ell,k}\right) \tag{5.1}$$

where $\boldsymbol{R}_{\ell,j,k}$ represents a matrix of covariance corresponding to the *k*th UT in the *j*th cell from ℓth BS. If we assume that channels are Rayleigh fading without UT correlation, then $\boldsymbol{R}_{j,\ell,k} = \beta_{j,\ell,k}\boldsymbol{I}_k$. It is suggested from [229] that $\boldsymbol{R}_{j,\ell k}$ will vary slowly over time, compared to $\boldsymbol{h}_{j,\ell,k}$. For this work we assume that $\boldsymbol{R}_{j,\ell,k}$ is constant across the transmission bandwidth and is changing slowly over time. We also assume that the target UT is located at the target cell edge and that the neighbouring cell UTs are also located on the edges of the cell. That ensures maximum inter-cell interference in the uplink for the UT target. Neighbouring cell UTs are on a straight line connecting the target UTs and the neighbouring *BS*s to ensure maximum downlink interference.

5.3 Improved MMSE and Semi-Blind Channel Estimation

5.3.1 Conventional MMSE Channel Estimation

The estimation of the MMSE channel represents an improvement on estimation of the LS channel. In approximating CSI, the MMSE relies on channel statistics. In MMSE channel approximation, the MSE is obtained as the difference between the real channel, $\boldsymbol{h}_{j,\ell,k}$, and the approximate MMSE channel, $\widehat{\boldsymbol{h}}_{j,j,k}^{MMSE}$, is minimized by choosing the appropriate linear approximation and $\widehat{\boldsymbol{h}}_{j,j,k}^{LS}$ which is the approximation of the LS.

$$\widetilde{\boldsymbol{h}}_{j,j,k} = \boldsymbol{h}_{j,j,k} - \widehat{\boldsymbol{h}}_{j,j,k}^{\boldsymbol{MMSE}} \tag{5.2}$$

BS conducts a channel estimate of the minimum mean-squared error (MMSE) that can be written as [226], [230],

$$\widehat{\boldsymbol{h}}_{j,j,k}^{MMSE} = \boldsymbol{R}_{j,j,k}\boldsymbol{\phi}_{j,k}^{-1}\boldsymbol{y}_{j,k}^{p} \tag{5.3}$$

where

$$\boldsymbol{y}_{j,k}^{p} = \boldsymbol{h}_{j,j,k} + \sum_{l=1,l\neq j}^{L} \boldsymbol{h}_{j,\ell,k} + \frac{\mathbf{1}}{\sqrt{\boldsymbol{\rho}^{tr}}}\boldsymbol{N}_{j}^{p}\boldsymbol{\phi}_{j,k}^{\star} \tag{5.4}$$

With the standardized total pilot transmission power ρ^{tr} for each UT and letting $\boldsymbol{\phi}_{j,k} = \mathbb{E}\left[\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^{H}\right]$ which can be exploded into

$$\boldsymbol{\phi}_{j,k} = \sum_{\ell=1}^{L}\sum_{k=1}^{K}\widehat{\boldsymbol{h}}_{j,\ell,k}\widehat{\boldsymbol{h}}_{j,\ell,k}^{H} + \frac{1}{\boldsymbol{\rho}^{tr}}\boldsymbol{I}_{M} = \sum_{\ell=1}^{L}\boldsymbol{R}_{j,\ell,k} + \frac{1}{\boldsymbol{\rho}^{tr}}\boldsymbol{I}_{M} \tag{5.5}$$

5.3.2 Improved MMSE Channel Estimation

The MMSE approximation algorithm evokes inverse operation to solve a linear system of equations. Consequently, the MMSE approximation process suffers from computational complexity and does not render itself to efficient parallelization. The reduction in complexity can be realized if propagation conditions are ideal. However, the existence of pilot contaminations generates a spatial correlated interference degrading approximation process and consequently the spectral efficiency [18], [26], [231].

The estimator utilizing MMSE approximation experience complexity which scales as $\mathcal{O}(K^3)$, where K is the number of UTs per cell. With expansive antenna arrays characterising massive MIMO, the estimation based on MMSE approximation is not viable. We propose the modification in the next section to cater for this shortcoming.

5.3.2.1 SMW-Based MMSE Estimation

In this research we first propose the use of Sherman-Morrison-Woodbury (SMW) algorithm to circumvent the matrix inversion [126]. From equation (5.3) it can be observed that the inversion of $\boldsymbol{\phi}_{j,k}$ is required leading to a high computational complexity since it is a large matrix of size $M \times M$. According to [232] iterative procedures are stable and less susceptible to numerical errors compared to direct procedures, thus highly recommended for manipulating

large matrices. It is proposed that the SMW-based scheme be employed to avoid intricate matrix inversion in estimating the MMSE. The first step in applying this approach is to check if the matrix $\boldsymbol{\phi}_{j,k} = \boldsymbol{\phi}_{j,k}^{SMW}$ is a valid positive definite Hermitian matrix. . If we assume a non-zero, $\boldsymbol{t} \in \mathbb{C}^{K\times 1}$ we can demonstrate that

$$\begin{aligned} \boldsymbol{t}\boldsymbol{\phi}_{j,k}^{SMW}\boldsymbol{t}^{H} &= \boldsymbol{t}\left(\sum_{l=1}^{L}\boldsymbol{R}_{j,\ell,k} + \frac{1}{\boldsymbol{\rho}^{tr}}\boldsymbol{I}_{M}\right)\boldsymbol{t}^{H} \\ &= \boldsymbol{t}\left(\sum_{l=1}^{L}\boldsymbol{R}_{j,\ell,k} + \frac{1}{\boldsymbol{\rho}^{tr}}\boldsymbol{I}_{M}\right)\boldsymbol{t}^{H} > 0 \end{aligned} \tag{5.6}$$

and

$$\left(\boldsymbol{\phi}_{j,k}^{SMW}\right)^{H} = \left(\sum_{l=1}^{L}\boldsymbol{R}_{j,\ell,k} + \frac{1}{\boldsymbol{\rho}^{tr}}\boldsymbol{I}_{M}\right)^{H} = \boldsymbol{\phi}_{j,k}^{SMW} \tag{5.7}$$

Equation (5.7) proofs that $\boldsymbol{\phi}_{j,k}^{SMW}$ is a positive definite Hermitian matrix. When Cholesky decomposition is applied to $\boldsymbol{\phi}_{j,k}^{SMW}$ [126]

$$\boldsymbol{\phi}_{j,k}^{SMW} = \boldsymbol{L}_{j,k}\boldsymbol{L}_{j,k}^{H} \tag{5.8}$$

where $\boldsymbol{L}_{j,k}$ stands for the lower triangular matrix. This means that

$$\left(\boldsymbol{\phi}_{j,k}^{SMW}\right)^{-1} = \left(\boldsymbol{L}_{j,k}^{H}\right)^{-1}\boldsymbol{L}_{j,k}^{-1} \tag{5.9}$$

Here the inverse function of matrix $\boldsymbol{\phi}_{j,k}^{SMW}$ can be modified to matrix inversion function $\boldsymbol{L}_{j,k}$. Hence, we compute a lower triangular matrix inverse instead of the symmetric matrix inverse consequently reducing the complexity. Thus, instead of the symmetric matrix, we calculate a lower triangular matrix inverse thus reducing the complexity. The Sherman-Morrison-Woodbury lemma is introduced [233].

$$\begin{aligned} \boldsymbol{Q} &= \boldsymbol{Z} + \boldsymbol{l}\boldsymbol{e}^{T} \\ \boldsymbol{Q}^{-1} &= (\boldsymbol{Z} + \boldsymbol{l}\boldsymbol{e}^{T})^{-1} \\ &= \boldsymbol{Z}^{-1} + \boldsymbol{Z}^{-1}\boldsymbol{l}(\boldsymbol{I} + \boldsymbol{e}^{T}\boldsymbol{Z}^{-1}\boldsymbol{l})^{-1}\boldsymbol{e}^{T}\boldsymbol{Z}^{-1} \end{aligned} \tag{5.10}$$

With this we can use iterative method to calculate the straight matrix inversion a few times, rather than calculating complex matrix inversion and then streamline the high computational

complexity to a lower computational complexity. The matrix $\boldsymbol{L}_{j,k}$ is further decomposed to preserve low computational complexity

$$\boldsymbol{L}_{j,k} = \boldsymbol{B}_{j,k} + \boldsymbol{L}'_{j,k} \tag{5.11}$$

where $\boldsymbol{B}_{j,k} = diag(\boldsymbol{L}_{j,k})$ and $\boldsymbol{L}'_{j,k} = (\boldsymbol{l}'_1, \boldsymbol{l}'_2, \dots, \boldsymbol{l}'_{K-1}, 0)$ is a $\boldsymbol{L}_{j,k}$ matrix with diagonal elements set to zero. The matrix is further written as

$$\boldsymbol{L}_{j,k} = \boldsymbol{B}_{j,k} + \boldsymbol{l}_{j,k}\boldsymbol{e}_{j,k}^T \tag{5.12}$$

Usage of the Sherman-Morrison-Woodbury lemma yields [233]

$$\boldsymbol{L}_{j,k}^{-1} = \left(\boldsymbol{B}_{j,k} + \boldsymbol{l}_{j,k}\boldsymbol{e}_{j,k}^T\right)^{-1} = \boldsymbol{B}_{j,k}^{-1} + \boldsymbol{B}_{j,k}^{-1}\boldsymbol{l}_{j,k}\left(\boldsymbol{I}_M + \boldsymbol{l}_{j,k}\boldsymbol{B}_{j,k}^{-1}\boldsymbol{e}_{j,k}^T\right)^{-1}\boldsymbol{e}_{j,k}^T\boldsymbol{B}_{j,k}^{-1} \tag{5.13}$$

We set $\boldsymbol{B}_{j,k}^{-1} = \boldsymbol{Z}^{-1}$, and therefore $\boldsymbol{B}_0 = \boldsymbol{B}_{j,k}$ is assumed for the next iteration

$$\boldsymbol{B}_1 = \boldsymbol{B}_0 + \boldsymbol{l}\boldsymbol{e}^T \tag{5.14}$$

Then, we compute the inverse of the 1st iteration, $\boldsymbol{B}_1$ as

$$\boldsymbol{B}_1^{-1} = \boldsymbol{B}_0^{-1} + \boldsymbol{B}_0^{-1}\boldsymbol{l}_{j,k}\left(\boldsymbol{I}_M + \boldsymbol{l}_{j,k}\boldsymbol{B}_0^{-1}\boldsymbol{e}_{j,k}^T\right)^{-1}\boldsymbol{e}_{j,k}^T\boldsymbol{B}_0^{-1} \tag{5.15}$$

and we compute the *ith* iteration inverse as

$$\boldsymbol{B}_i^{-1} = \boldsymbol{B}_{i-1}^{-1} + \boldsymbol{B}_{i-1}^{-1}\boldsymbol{l}_{j,k}\left(\boldsymbol{I}_M + \boldsymbol{l}_{j,k}\boldsymbol{B}_{i-1}^{-1}\boldsymbol{e}_{j,k}^T\right)^{-1}\boldsymbol{e}_{j,k}^T\boldsymbol{B}_{i-1}^{-1} \tag{5.16}$$

Since the matrix $\boldsymbol{B}$ is diagonal matrix, its inversion is simple and efficient. So, the $\boldsymbol{L}_{j,k}$ inversion can be calculated by iterating $N-1$ times. From which $\boldsymbol{L}_{j,k}^{-1} = \boldsymbol{B}_i^{-1}$ and the inverse of $\boldsymbol{\phi}_{j,k}^{SMW}$ as per (5.9) is determined. Then the channel is estimated as

$$\widehat{\boldsymbol{h}}_{j,j,k}^{SMW} = \boldsymbol{R}_{j,j,k}\left(\boldsymbol{\phi}_{j,k}^{SMW}\right)^{-1}\boldsymbol{y}_{j,k}^p \tag{5.17}$$

5.3.2.2 RNA-Based MMSE Estimation

The approximation of the SMW-based MMSE still involves inversion of the lower triangular matrix and therefore we replace it with the Rapid Numerical Algorithm (RNA) method. RNA-based approximation completely avoids the inversion of the matrix, and instead uses multiplication and addition. Let ₣ be a $M \times M$ nonsingular matrix that we are tasked with

calculating the inverse. Let $\boldsymbol{\mathcal{R}}_i$ again represent the inverse estimate in the ith iteration. The residual matrix represents the computed inverse divergence from the actual matrix inverse, ₣. We obtain the residual matrix $\boldsymbol{\mathcal{E}}_i$ as

$$\boldsymbol{\mathcal{E}}_i = \boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_i \tag{5.18}$$

This is the residual matrix of first inverse $\boldsymbol{\mathcal{R}}_i$ in the calculation. Which can be re-written as

$$\begin{aligned} ₣\boldsymbol{\mathcal{R}}_i &= \boldsymbol{I} - \boldsymbol{\mathcal{E}}_i \\ (₣\boldsymbol{\mathcal{R}}_i)^{-1} &= (\boldsymbol{I} - \boldsymbol{\mathcal{E}}_i)^{-1} \\ ₣^{-1} &= \boldsymbol{\mathcal{R}}_i(\boldsymbol{I} - \boldsymbol{\mathcal{E}}_i)^{-1} \end{aligned} \tag{5.19}$$

The expression in $\boldsymbol{\mathcal{E}}_i$ is a power series. Accordingly,

$$₣^{-1} = \boldsymbol{\mathcal{R}}_i(\boldsymbol{I} - \boldsymbol{\mathcal{E}}_i)^{-1} = \boldsymbol{\mathcal{R}}_i \sum_{i=0}^{\infty} (\boldsymbol{\mathcal{E}}_i)^i = \boldsymbol{\mathcal{R}}_i(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_i + \boldsymbol{\mathcal{E}}_i^2 + \cdots) \tag{5.20}$$

The first two terms, as in [205], [206], can be used to represent the inverse of an infinite series. If we Limit the infinite series to the first two terms, we get,

$$₣^{-1} = \boldsymbol{\mathcal{R}}_i(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_i) = \boldsymbol{\mathcal{R}}_i(\boldsymbol{I} + \boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_i) = \boldsymbol{\mathcal{R}}_i(2\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_i) \tag{5.21}$$

Which can be re-written as,

$$\boldsymbol{\mathcal{R}}_i = \boldsymbol{\mathcal{R}}_{i-1}(2\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_{i-1}) \tag{5.22}$$

where, in the next iteration, $\boldsymbol{\mathcal{R}}_i$ represent the inverse. This expression is known as the Schulz iterative process used to invert a matrix [185], [186]. In [187] it was pointed out that the consideration of the initial three terms provides the quickest convergence of the iterative method to find the inverse. Accordingly,

$$\begin{aligned} \boldsymbol{\mathcal{R}}_i &= \boldsymbol{\mathcal{R}}_{i-1}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_{i-1} + \boldsymbol{\mathcal{E}}_{i-1}^2) = \boldsymbol{\mathcal{R}}_{i-1}\big(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_{i-1}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_{i-1})\big) \\ &= \boldsymbol{\mathcal{R}}_{i-1}\Big(\boldsymbol{I} + (\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_{i-1})\big(\boldsymbol{I} + (\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_{i-1})\big)\Big) \\ &= \boldsymbol{\mathcal{R}}_{i-1}(\boldsymbol{I} + 2\boldsymbol{I} - \boldsymbol{I}₣\boldsymbol{\mathcal{R}}_{i-1} - 2\boldsymbol{I}₣\boldsymbol{\mathcal{R}}_{i-1} + (₣\boldsymbol{\mathcal{R}}_{i-1})^2) \\ &= \boldsymbol{\mathcal{R}}_{i-1}\big(3\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_{i-1}(3\boldsymbol{I} - ₣\boldsymbol{\mathcal{R}}_{i-1})\big) \end{aligned} \tag{5.23}$$

Amat suggested this method in [188], and this series is converges to $\mathbb{F}^{-1} = \boldsymbol{L}_{j,k}^{-1}$. From (5.11), we let $\boldsymbol{\mathcal{R}}_0 = \boldsymbol{B}_{j,k}^{-1}$ and then we can obtain the residual matrix as

$$\boldsymbol{\mathcal{E}}_0 = \boldsymbol{I}_M - \boldsymbol{L}_{j,k}\boldsymbol{\mathcal{R}}_0 \tag{5.24}$$

then we express the 1st iteration as

$$\boldsymbol{\mathcal{R}}_1 = \boldsymbol{\mathcal{R}}_0\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,k}\boldsymbol{\mathcal{R}}_0\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,k}\boldsymbol{\mathcal{R}}_0\right)\right) \tag{5.25}$$

and similarly obtain the *ith* iteration as

$$\boldsymbol{\mathcal{R}}_i = \boldsymbol{\mathcal{R}}_{i-1}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{\mathcal{R}}_{i-1}\left(3\boldsymbol{I}_M - \boldsymbol{L}_{j,m}\boldsymbol{\mathcal{R}}_{i-1}\right)\right) \tag{5.26}$$

Setting P as the number of terms in total, we can obtain the inverse by expressing 5.26 as

$$\boldsymbol{R}_i = \boldsymbol{R}_{i-1}\left(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_{i-1} + \boldsymbol{\mathcal{E}}_{i-1}^2 + \cdots + \boldsymbol{\mathcal{E}}_{i-1}^{p-1}\right) \tag{5.27}$$

Hence, we can obtain the inverse for successive iterations through the following expressions

$$\boldsymbol{R}_1 = \boldsymbol{R}_0\left(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_0 + \boldsymbol{\mathcal{E}}_0^2 + \cdots + \boldsymbol{\mathcal{E}}_0^{p-1}\right) \tag{5.28}$$

$$\boldsymbol{R}_2 = \boldsymbol{R}_1\left(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_0^p + \left(\boldsymbol{\mathcal{E}}_0^p\right)^2 + \left(\boldsymbol{\mathcal{E}}_0^p\right)^3 \ldots + \left(\boldsymbol{\mathcal{E}}_0^p\right)^{p-1}\right) \tag{5.29}$$

and generally

$$\boldsymbol{R}_i = \boldsymbol{R}_{i-1}\left(\boldsymbol{I} + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right) + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right)^2 + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right)^3 + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right)^{p-1}\right) \tag{5.30}$$

Performing repeated substitution for $\boldsymbol{R}_{i-1}$ we obtain

$$\begin{aligned} R_i = \boldsymbol{R}_0\left[\boldsymbol{I} + \boldsymbol{\mathcal{E}}_0 + \boldsymbol{\mathcal{E}}_0^2 + \cdots + \boldsymbol{\mathcal{E}}_0^{p-1}\right]\left[\boldsymbol{I} + \boldsymbol{\mathcal{E}}_0^p + \left(\boldsymbol{\mathcal{E}}_0^p\right)^2 + \cdots\right. \\ \left. + \left(\boldsymbol{\mathcal{E}}_0^p\right)^{p-1}\right] \ldots \left[\boldsymbol{I} + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right) + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right)^2\right. \\ \left. + \cdots + \left(\boldsymbol{\mathcal{E}}_0^{p^{i-1}}\right)^{p-1}\right] \end{aligned} \tag{5.31}$$

$$= \boldsymbol{R}_0 \prod_{h=0}^{i-1}\left[\boldsymbol{I} + \left(\boldsymbol{\mathcal{E}}_0^{p^h}\right) + \left(\boldsymbol{\mathcal{E}}_0^{p^h}\right)^2 + \left(\boldsymbol{\mathcal{E}}_0^{p^h}\right)^3 \ldots + \left(\boldsymbol{\mathcal{E}}_0^{p^h}\right)^{p-1}\right]$$

And letting $p = 3$ results in

$$R_i = \boldsymbol{R}_0 \prod_{h=0}^{i-1} \left[\boldsymbol{I} + \left(\boldsymbol{\mathcal{E}}_0^{3^h} \right) + \left(\boldsymbol{\mathcal{E}}_0^{3^h} \right)^2 \right] \tag{5.32}$$

To compute $\boldsymbol{R}_i$, the expression found in 5.32 has considerable significance, in that the inverse calculation in the present iteration makes use of the inverse in the preceding iteration along with two new terms involving the original residual matrices. In computing these terms, we rely on the product of the error matrices previously obtained. Hence, the task of inverting a matrix is reduced into one of multiplying a matrix and the computation of $\boldsymbol{\mathcal{E}}_0^{3^{\mathrm{h}}}$ and $\left(\boldsymbol{\mathcal{E}}_0^{3^{\mathrm{h}}}\right)^2$ in 5.32 requires only a single matrix multiplication as per [195], which makes the iterative process efficiently parallel. Again, the $\boldsymbol{L}_{j,k}$ inversion is realized by $I-1$ iteration times. And we set $\boldsymbol{L}_{j,k}^{-1} = \boldsymbol{\mathcal{R}}_i$ and compute the inversion of $\boldsymbol{\phi}_{j,k}^{RNA} = \boldsymbol{\phi}_{j,k}$ as per equation 5.9. Then we express the approximated RNA channel as in (5.33).

$$\widehat{\boldsymbol{h}}_{j,j,k}^{RNA} = \boldsymbol{R}_{j,j,k} \left(\boldsymbol{\phi}_{j,k}^{RNA} \right)^{-1} \boldsymbol{y}_{j,k}^{p} \tag{5.33}$$

The estimation of the RNA-based channel has reduced computational complexity given by $M^2(I-1)$ compared with conventional estimation of the MMSE channel [234]. But also, according to [235] a method like RNA-based channel estimation that renders itself to parallelization, becomes superior when implemented on machine with more cores.

5.3.2.3 Spectral Efficiency

To process the vector $\boldsymbol{y}_j \in \mathbb{C}^M$ obtained at *j*th BS we rely on $\boldsymbol{v}_{j,k} \in \mathbb{C}^M$ which is the a receive combining vector allocated by the *j*th *BS* to *k*th *UT*. Using MMSE approximation, the channel capacity of the *k*th *UT* in the *j*th *BS* has SE which is lower bounded as [171]

$$\boldsymbol{SE}_{j,k} = \mathbb{E}\{ log_2(1 + \boldsymbol{SINR}_{j,k}) \} \, [bits/s/Hz] \tag{5.34}$$

where the equivalent SINR is computed by the following expression in (5.35) below

$$\boldsymbol{SINR}_{j,m} = \frac{\left| \boldsymbol{v}_{j,k}^H \widehat{\boldsymbol{h}}_{j,j,k} \right|^2}{\boldsymbol{v}_{j,k}^H \left(\sum_{\ell=1}^{L} \sum_{\substack{m=1 \\ (l,m) \neq (j,k)}}^{K} \widehat{\boldsymbol{h}}_{\ell,j,m} \widehat{\boldsymbol{h}}_{\ell,j,m}^H + \boldsymbol{Z}_j \right) \boldsymbol{v}_{j,k}} \tag{5.35}$$

and $\boldsymbol{Z}_j = \sum_{k=1}^{K} \left(\boldsymbol{R}_{\ell,j,k} - \boldsymbol{\phi}_{\ell,j,k} \right) + \frac{1}{\rho^{tr}} \boldsymbol{I}_M$.

This represents the tightest bound capacity that is ever achievable for a massive MIMO system with a receive combination that is linear. Therefore, it calls for the use of MMSE approximation in the BS which means we must have perfect knowledge of the covariance matrices.

5.3.3 Semi-Blind FDPM-based channel estimation

It is a shared understanding that systems anchored on wireless communication technology suffers from radio resources scarcity. To shore up coverage and capacity the radio resources, and specifically the channels of frequency, are reused throughout the wireless communication systems. This points to the fact that a number of UTs use repeated pilot training sequences which leads to pilot contamination and hence degrading the performance of wireless communication systems. This is compounded for the massive MIMO scenario and thus the need to establish estimation techniques that thrive on reduced number of pilots compared to training-based channel estimation methods with precise CSI estimates. Semi-blind channel approximation methods in mitigation of pilot contamination have thus been found to be optimal [25].

The EVD algorithms are leveraged in realizing semi-blind estimators which require reduced number of pilots to determine the problem of ambiguity matrix. The asymptotic orthogonality of UTs is exploited as an alternative to solving the ambiguity matrix by evoking the large numbers law. This is accomplished by SVD method. The SVD-based approximation generally has a better estimate compared to the EVD-based approximation [152], although the two techniques have $\mathcal{O}(\mathrm{M}^3)$ complexity in computation as per the signal dimensions received. For massive MIMO systems characterized by large number of antennas per BS, such schemes are untenable.

Subspace tracking algorithms are proposed to minimize the complexity. To compute the ambiguity matrix through a simplified process of correlation matrix iteration, the authors in [236] propose the Fast Data Projection Method (FDPM). This gives good tracking results accompanied with reduced computational complexity of $\mathcal{O}(MK)$.

The covariance matrix of the received signal can be calculated as set out in [152]

$$\boldsymbol{C}_y = E\{\boldsymbol{y}\boldsymbol{y}^H\} = E\{\boldsymbol{H}\boldsymbol{s}\boldsymbol{s}^H\boldsymbol{H}^H + \boldsymbol{z}\boldsymbol{z}^H\} = \boldsymbol{H}\boldsymbol{H}^H + \boldsymbol{I}_M \tag{5.36}$$

Decomposing $\mathbf{C}_y$ using SVD technique results in

$$\boldsymbol{C}_y = [\boldsymbol{U}_s\, \boldsymbol{U}_n] \wedge [\boldsymbol{U}_s\, \boldsymbol{Y}_n] \tag{5.37}$$

where the subspace of the signal is given by $\boldsymbol{U}_s \in \mathbb{C}^{M\times K}$ and the subspace of the noise is given by $\boldsymbol{U}_n \in \mathbb{C}^{M\times(M-K)}$. The scalar multiplicative ambiguity matrix, $\boldsymbol{A} \in \mathbb{C}^{K\times K}$, is leveraged to compute the channel matrix, $\boldsymbol{H}$, from the signal subspace, $\boldsymbol{U}_s$, as per [152], thus

$$\widehat{\boldsymbol{H}} = \boldsymbol{U}_s \boldsymbol{A} \tag{5.38}$$

Then we compute the ambiguity matrix using a short training sequence as in (5.4) and obtain

$$\boldsymbol{A} = (\boldsymbol{U}_s)^H \widehat{\boldsymbol{H}}^{RNA} \tag{5.39}$$

where the $\widehat{\boldsymbol{H}}^{RNA}$ is the initial channel estimate from the estimation of the RNA channel in (5.33). Then we calculate the estimate for channels as

$$\widehat{\boldsymbol{H}} = \boldsymbol{U}_s (\boldsymbol{U}_s)^H \widehat{\boldsymbol{H}}^{RNA} \tag{5.40}$$

Table 5.1: FDPM sub-space tracking algorithm.

for $n = 1, 2, \dots, N_{data}$

$\boldsymbol{r}_n = \boldsymbol{W}_{n-1}^H \boldsymbol{y}_n$

$\boldsymbol{T}_n = \boldsymbol{W}_{n-1} \pm \frac{\mu}{\|\boldsymbol{y}_n\|^2} \boldsymbol{y}_n * \boldsymbol{r}_{n-1}^H$

$\boldsymbol{a}_n = \boldsymbol{r}_n - \|\boldsymbol{r}_n\| [1\ 0 \dots 0]^T$

$\boldsymbol{Z}_n = \boldsymbol{T}_n - \frac{2}{\|\boldsymbol{a}_n\|^2} [\boldsymbol{T}_n * \boldsymbol{a}_n] * \boldsymbol{a}_n^H$

$\boldsymbol{W}_n = normalise\{\boldsymbol{Z}_n\}$

end

Based on its rapid convergence combined with a good tracking efficiency, the subspace tracking algorithm known as the FDPM in [236] was espoused. The algorithm is summed up in Table 5.1.

The subspace of the signal corresponding to the *n*th sample, $\boldsymbol{W}_n \in \mathbb{C}^{M\times K}$, is tracked as in Table 5.1 above. The effect of the old data is controlled by a forgetting factor $0 < \beta < 1$ and expressed as $\mu = 1 - \beta$. N_{data} denotes the duration of signals transmitted without the pilots. Then we express the approximated $\boldsymbol{U}_s$ through $\boldsymbol{W}_{N_{data}}$ which is the tracked signal subspace as

$$\boldsymbol{U}_s = \boldsymbol{W}_{N_{data}} \tag{5.41}$$

Then the approximate channel matrix is determined as

$$\widehat{\boldsymbol{H}}^{FDPM} = \boldsymbol{W}_{N_{data}}\left(\boldsymbol{W}_{N_{data}}\right)^{H}\widehat{\boldsymbol{H}}^{\boldsymbol{RNA}} \tag{5.42}$$

We estimate the initial value $\boldsymbol{W}_0$ of $\boldsymbol{W}$ by

$$\boldsymbol{W}_0 = \begin{bmatrix} \boldsymbol{I}_K \\ \boldsymbol{0}_{(M-K)\times K} \end{bmatrix} \tag{5.43}$$

5.3.4 Simulation Results and Discussion

We start off by analysing the Normalized Mean Square Error (NMSE) and M performance parameters for the previously discussed channel estimation techniques. We assume an uplink massive MIMO system and evaluate the trade-offs amongst these parameters in lieu of the channel estimation techniques. By computing the NMSE as

$$NMSE = \frac{E\{trace(\widetilde{\boldsymbol{h}}_{j,j,k}\widetilde{\boldsymbol{h}}_{j,j,k}^{H})\}}{trace(\boldsymbol{R}_{j,j,k})} \tag{5.44}$$

in view of the NMSE and the respective corresponding M, the performance of the four different channel estimation techniques is analysed. First, the different values of reuse factor, f and BS antennas, M are evoked to evaluate the performance of conventional MMSE, improved MMSE and the FDPM channel estimation techniques.

We assume a multicell massive MIMO scenario to compare these channel estimation techniques. The analysis is performed for $M = 32\ to\ 176$, with a step of 16 and $K = 16$ massive MIMO system. We use three reuse factor variations ($f = [1, 2, 4]$) and evaluate the performance of each of the channel estimation techniques.

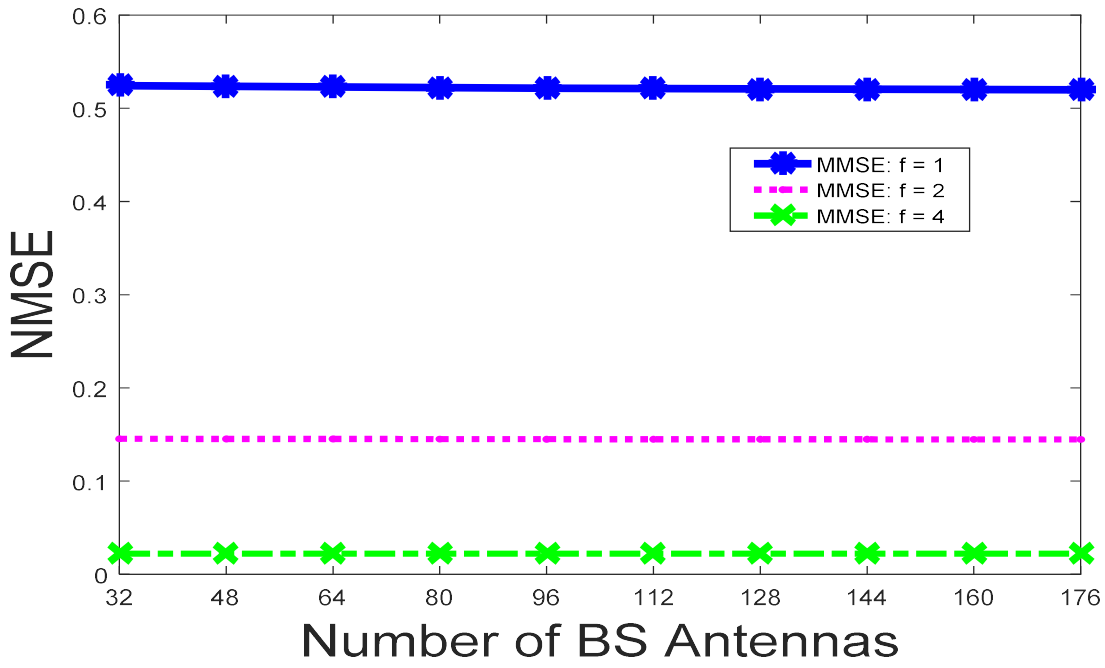

Figure 5.1: Achievable NMSE vs. number of BS antennas for conventional MMSE estimator ($\boldsymbol{f = [1, 2, 4]}$).

For a multicell massive MIMO, Figure 5.1 contrasts the possible NMSE vs. number of BS antennas for conventional MMSE estimator with varying reuse factor f. A number of observations can be deduced on the basis of this figure. The performance of the conventional MMSE estimator improves as the number of BS antennas increases and the same is observed when the reuse factor increases. This is represented by a decrease in NMSE as the number of BS antennas, M, and the reuse factor, f, increases as in Table 5.2. At an f of 4 the lowest NMSE is obtained due to further reduction in pilot contamination. Figure 5.2, Figure 5.3 and Figure 5.4 indicate a similar pattern for SMW-based MMSE, RNA-based MMSE and the FDPM-based semi-blind estimators with NMSE decreasing with increasing number of BS antennas and reuse factor f, as in Table 5.3, Table 5.4, and Table 5.5. Again, the values in Table 5.3, Table 5.4, and Table 5.5 are the same, since SMW and RNA estimators present efficient approximation of the MMSE yielding similar performance though with reduced or complete avoidance of matrix inversion.

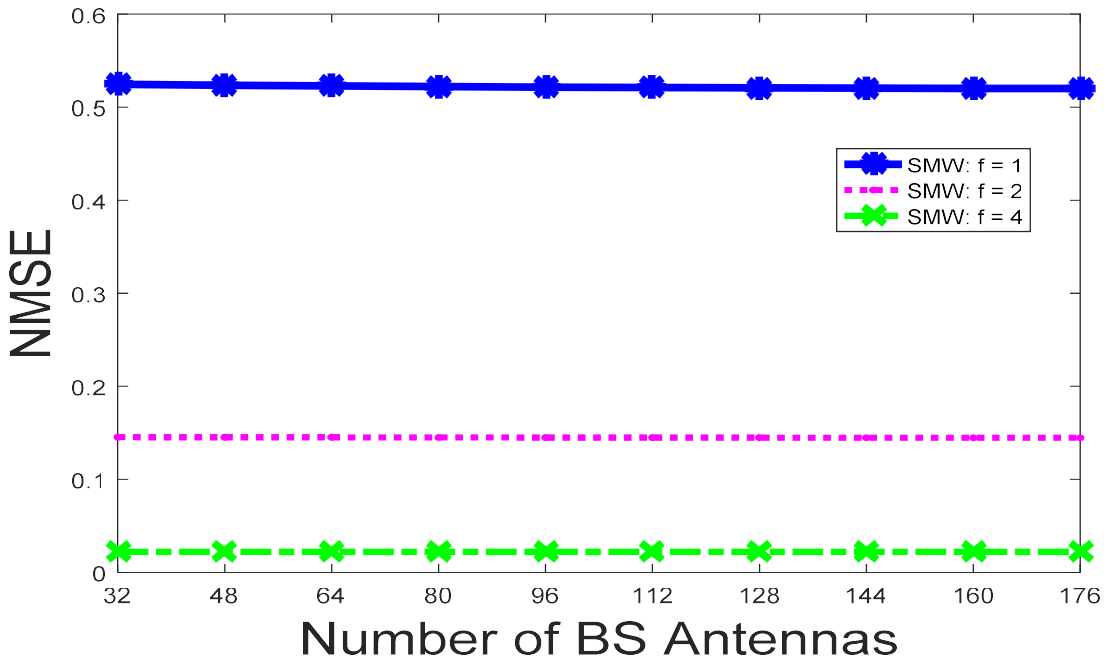

Figure 5.2: Achievable NMSE vs. number of BS antennas for SMW-based MMSE estimator ($\boldsymbol{f} = [\mathbf{1}, \mathbf{2}, \mathbf{4}]$).

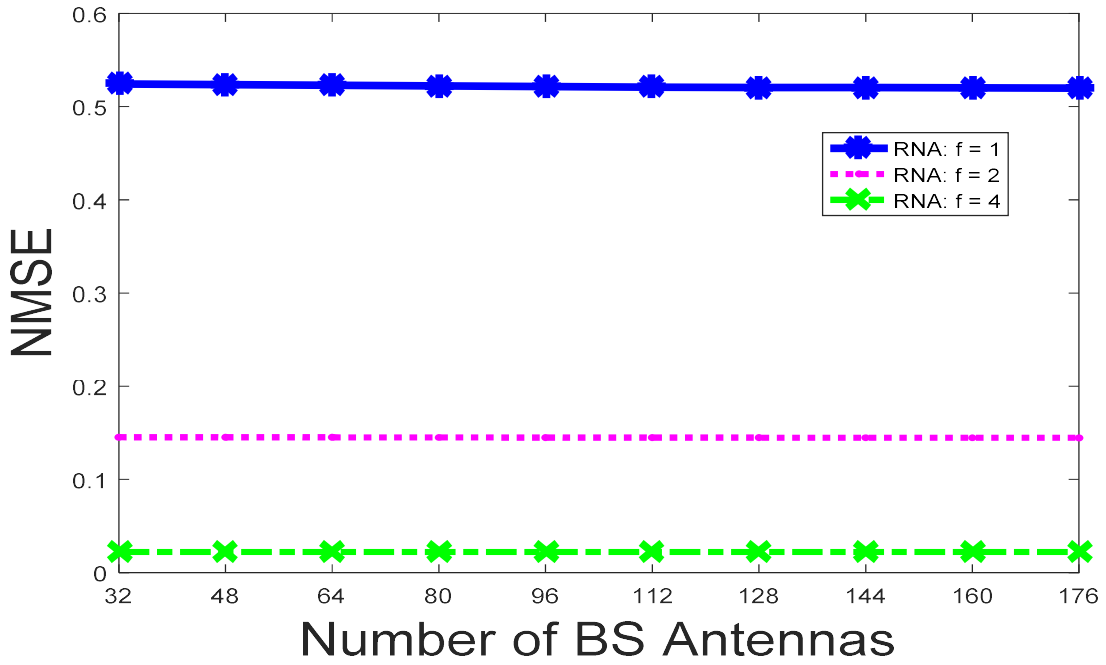

Figure 5.3: Achievable NMSE vs. number of BS antennas for RNA-based MMSE estimator ($\boldsymbol{f} = [\mathbf{1}, \mathbf{2}, \mathbf{4}]$).

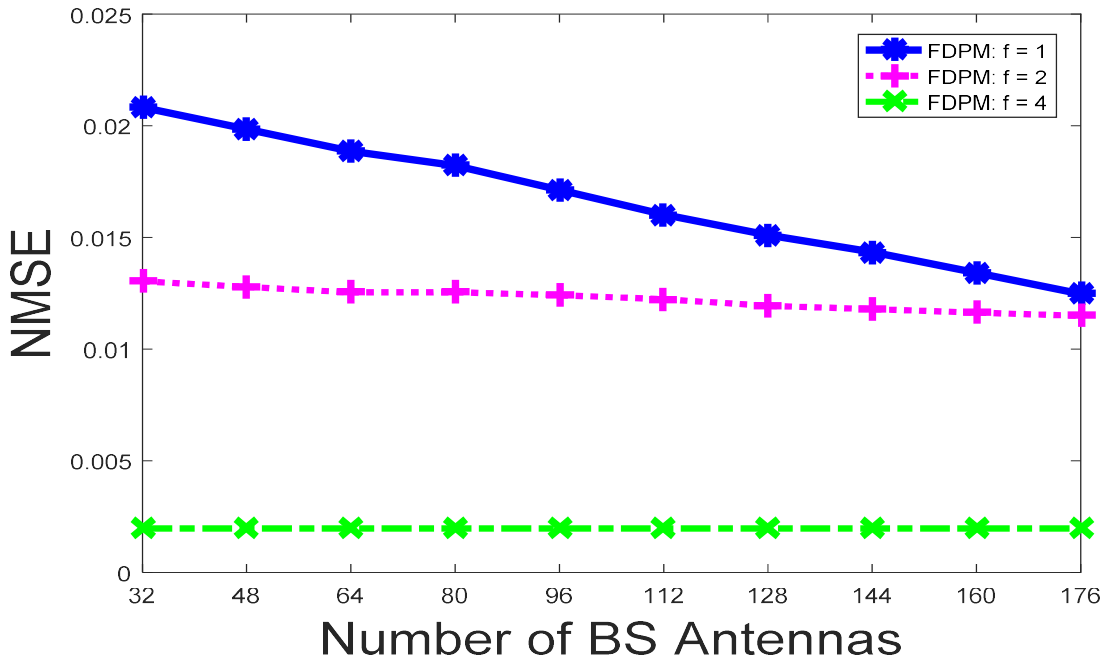

Figure 5.4: Achievable NMSE vs. number of BS antennas for FDPM-based semi-blind estimator ($f = [1, 2, 4]$).

Table 5.2: Performance of conventional MMSE estimator.

M	***f***		
	1	**2**	**4**
32	0.524419	0.145363	0.0221061
112	0.521059	0.144879	0.0221025
176	0.519950	0.144688	0.0221011

Table 5.3: Performance of RNA estimator.

M	***f***		
	1	**2**	**4**
32	0.524419	0.145363	0.0221061
112	0.521059	0.144879	0.0221025
176	0.519950	0.144688	0.0221011

Table 5.4: Performance of SMW estimator.

M	***f***		
	1	**2**	**4**
32	0.524419	0.145363	0.0221061
112	0.521059	0.144879	0.0221025
176	0.519950	0.144688	0.0221011

Table 5.5: Performance of FDPM estimator.

M	***f***		
	1	**2**	**4**
32	0.0208206	0.0130384	0.00196498
112	0.0160260	0.0122331	0.00196466
176	0.0124696	0.0114883	0.00196453

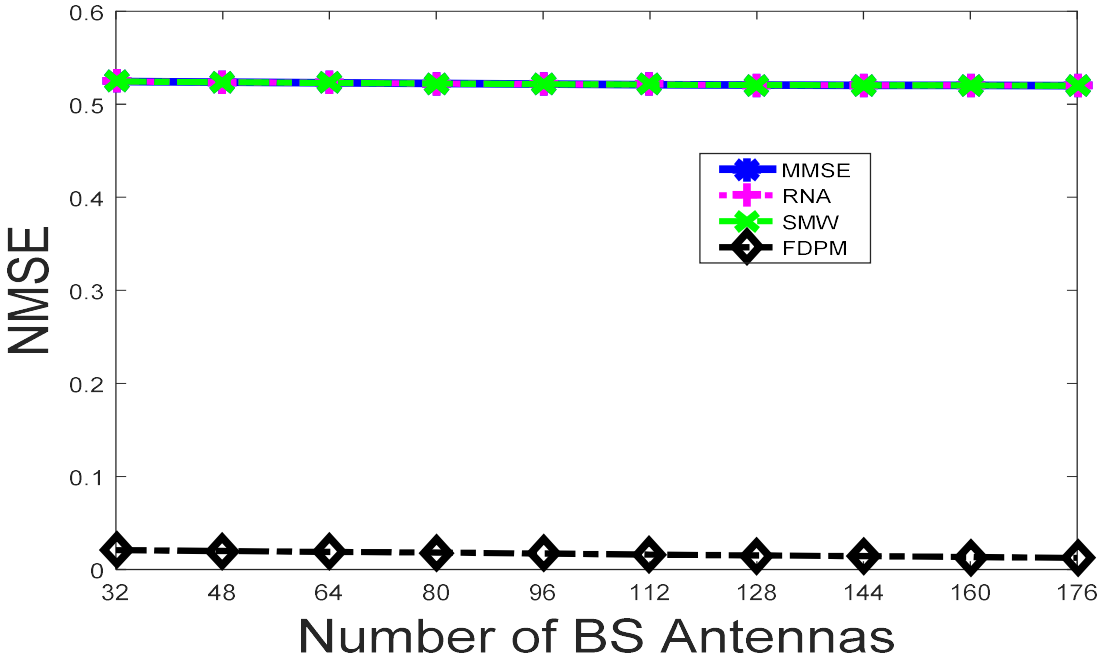

Figure 5.5: Comparison of achievable NMSE vs. number of BS antennas for all the estimators ($\boldsymbol{f = 1}$).

Figure 5.5 and Figure 5.6 provide a comparison of all four channel estimation techniques performance. Figure 5.5 suggests that the semi-blind estimation technique based on the FDPM holds the lowest at NMSE at $f = 1$. However, with growing numbers of BS antennas, the conventional MMSE, SMW-based MMSE and RNA-based MMSE all have a similar NMSE. This points to the fact that SMW-based MMSE and RNA-based estimation techniques gives unparalleled approximation of the conventional MMSE technique and hence worth techniques to circumvent matrix inversion. From Figure 5.6, it is clear that the conventional MMSE, SMW-based MMSE and RNA-based MMSE techniques are still inferior, at $f = 4$, to FDPM-based semi-blind estimation technique. But in overall the NMSE is lower at $f = 4$ than it is at $f = 1$.

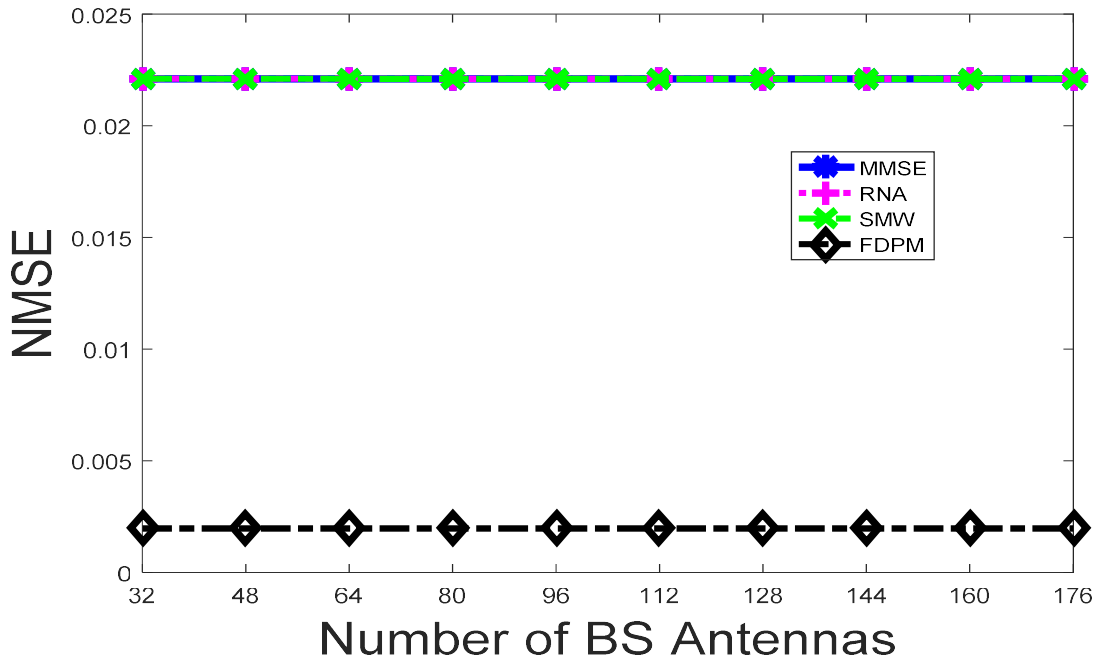

Figure 5.6: Comparison of achievable NMSE vs. number of BS antennas for all the estimators ($f = 4$).

Thus, we can conclude that the use of the RNA-based method in MMSE enhances the channel estimation functionality and makes the estimation process efficiently parallel, which is a very important aspect of high computation systems used in massive MIMO networks at the BS. We have also shown that RNA-based MMSE method circumvents the inverse matrix and instead uses matrix multiplication and addition that can be easily implemented in hardware. And it has an equally good performance comparable to the conventional estimator of MMSE channels. The estimation of semi-blind channels is a better estimation approach with a lower complexity and higher efficiency than traditional methods of estimating channels. The approximation of channel parameters requires a reduced number of pilots; normally the pilots used in the initial channel estimation method based on RNA. It thus further mitigates contamination by the pilot.

5.4 Channel Estimation in MPC-RAN

To minimize C-RAN fronthaul capacity and associated time latency we use partially centralized C-RAN with interconnected RRHs according to [163]. The cooperative processing at the BBU is envisaged to suppress the inter-RRH interference by use of the CSI both from the RRHs and the wireless fronthaul [237]. Then we apply data compression to reduce the fronthaul traffic in the UL for CSI computation. The compressed data is then put on the

fronthaul for transmission to BBU. The received compressed CSI data is then processed using the method described in [238] to compute the estimated CSI. The initial computed CSI is exported to Givens Data Projection Method (GDPM) to create a semi-blind channel estimation process for better CSI estimation and high parallelization to speed up the CSI computation in the BBU.

5.4.1 System Model

We assume an L-RRH massive MIMO system, each with M antennas and K single-antenna UTs. We propose to organize the TDD protocols around the RRHs to simultaneously transmit pilot signals and data to all RRHs and BBUs. Pilots initially transmitted by UTs to their BS in *j*th are the same as those given by $\boldsymbol{\psi}_K = [\boldsymbol{\varphi}_{j,1}^T, \boldsymbol{\varphi}_{j,2}^T, \ldots, \boldsymbol{\varphi}_{j,K}^T]$, where $\boldsymbol{\varphi}_{j,k}$ represents a pilot used by each *k*th UT in each cell and $\|\boldsymbol{\varphi}_{j,k}\|^2 = 1$. Then, within the *j*th RRH, a channel from the *k*th UT is given by $\boldsymbol{h}_{j,k} \epsilon \mathbb{C}^M$. We assume Rayleigh fading channels vectors and model them as

$$\boldsymbol{h}_{j,k} \sim \mathcal{CN}(0, \boldsymbol{R}_{j,k}) \tag{5.45}$$

where $\boldsymbol{R}_{j,k}$ represents a matrix of covariance corresponding to the *k*th UT in the *j*th RRH. If we assume that channels are Rayleigh fading without UT correlation, then $\boldsymbol{R}_{j,k} = \beta_{j,k} \boldsymbol{I}_M$. It is suggested from [229], that $\boldsymbol{R}_{j,k}$ will vary slowly over time, compared to $\boldsymbol{h}_{j,k}$. For this work we assume that $\boldsymbol{R}_{j,k}$ is constant across the bandwidth of transmission and change slowly over time. Hence the training sequences received $\boldsymbol{Y}_j \epsilon \mathbb{C}^M$, is computed as

$$\boldsymbol{Y}_j = \boldsymbol{H}_{j,k} \boldsymbol{\psi}_k + \boldsymbol{Z}_j \tag{5.46}$$

where $\boldsymbol{\psi}_k \epsilon \mathbb{C}^K$ is the pilot matrix representing total transmitted sequences by K UTs and $\boldsymbol{Z}_j \epsilon \mathbb{C}^M$ represents the AWGN noise matrix.

5.4.2 Compressed Data CSI Model

In practice, we need to estimate the matrices of covariances based on the pilot samples received at the RRH. We set out to investigate the BBU's approximation of the covariance information needed and the impact these estimates would have. Calculating the MMSE estimation of $\boldsymbol{h}_{j,k}$ at the *j*th RRH from (5.40) requires the knowledge of $\boldsymbol{R}_{j,k} = \mathbb{E}[\boldsymbol{h}_{j,k} \boldsymbol{h}_{j,k}^H]$ and $\boldsymbol{\phi}_{j,k} =$

$\mathbb{E}\left[\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^{H}\right]$. Considering that these are $M \times M$ (quite large) matrices, it may be necessary to regularize estimates [239], [240].

Because using MPC-RAN results in high-dimensional data, we require huge communication and storage resources to compute these covariance matrices. There is a critical need for huge bandwidth and power resources [238] to deliver to the BBU the CSI information from the RRHs. To alleviate this shortcoming, we presume a partially centralised C-RAN system that has interconnected and cooperating massive MIMO RRHs [163]. As noted earlier, this makes fronthaul traffic largely dependent on UT data rates and not on the number of antennas. Then we use compressed data to estimate the matrix of covariance. We adopt the via-Q approach to estimate the matrices of covariance [226] but use the compressed data estimate as in [238].

5.4.2.1 Approximation of $\boldsymbol{\phi}_{j,k}$

It is presumed that the pilot signal $\boldsymbol{y}_{j,k}^{p}$ will arrive at the *j*th *RRH* over N_{ϕ} blocks of coherence. We denote these N_{ϕ} observations by $\boldsymbol{y}_{j,k}^{p}[1], \ldots, \boldsymbol{y}_{j,k}^{p}[N_{\phi}]$. Then formulate the sample the sample observation as in [226]

$$\boldsymbol{y}_{j,k}^{(Sample)} = \frac{1}{N_{\phi}} \sum_{n=1}^{N_{\phi}} \boldsymbol{y}_{j,k}^{p}[n] \tag{5.47}$$

We consider an antenna with index *k,* and equation (5.47) almost surely (a. s.) tends to true $\boldsymbol{y}_{j,k}^{p}$ as $N_{\phi} \to \infty$.

$$\frac{1}{N_{\phi}} \sum_{n=1}^{N_{\phi}} \left[\boldsymbol{y}_{j,k}^{p}[n]\right]_{m,k} \overset{a.s.}{\longrightarrow} \left[\boldsymbol{y}_{j,m}^{p}\right]_{m,k} \tag{5.48}$$

This is a derivative of the law of large numbers and the ergodicity of channels. To find a good approximation to $\boldsymbol{y}_{j,k}^{p}$ we need only a few observations as the standard deviation for a sample $\boldsymbol{y}_{j,k}^{p}$ decays as $1/\sqrt{N_{\phi}}$. The elements in $\mathbf{y}_{j,k}^{(Sample)}$ will tend to the corresponding elements in $\boldsymbol{y}_{j,k}^{p}$, individually. Then we follow the method set out in [238] for mitigating this problem.

Weighted sampling matrices $\{\boldsymbol{S}_{j,k}\}_{k=1}^{K} \in \mathbb{C}^{M \times Z}$ are used to compress data via $\boldsymbol{S}_{j,k}^{T}\boldsymbol{y}_{j,k}^{(Sample)}$ and the data is projected back to the original space through $\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{T}\boldsymbol{y}_{j,k}^{(Sample)}$. The obtained data is then used in covariance matrix approximation. At least M-Z elements are removed from the

kth vector by the weighted sampling matrix $\boldsymbol{S}_{j,k}$, the remaining ones are kept as they may be most informative. If the probabilities of sampling are carefully designed, the unbiased estimator $\widehat{\boldsymbol{\phi}}_{j,k}$ would perform accurately in relation to the spectral norm of the matrix $\left\|\widehat{\boldsymbol{\phi}}_{j,k}-\boldsymbol{\phi}_{j,k}\right\|_2$ [228], [241], [242].

The weighted sampling evoked is strong enough to explore the most appropriate entries to minimize the approximation error of $\left\|\widehat{\boldsymbol{\phi}}_{j,k}-\boldsymbol{\phi}_{j,k}\right\|_2$. We start by setting the required variables and then proceeding with the approximation process.

We presume $\mathbf{y}_{j,k}^{(Sample)} \epsilon \mathbb{C}^{M\times K}$ and set $\alpha\epsilon[0,1]$ as our regularization factor. The information obtained from the uplink is then compressed as follows with $l=[1,2,\dots,M]$

$$\boldsymbol{v}_{j,k}=\left\|\boldsymbol{y}_{j,k}^{(Sample)}\right\|_1=\sum_{l=1}^{M}\left|\boldsymbol{y}_{j,l,k}^{(Sample)}\right| \tag{5.49}$$

and

$$\boldsymbol{\omega}_{j,k}=\left\|\boldsymbol{y}_{j,k}^{(Sample)}\right\|_2^2=\sum_{l=1}^{M}\left(\boldsymbol{y}_{j,l,k}^{(Sample)}\right)^2 \tag{5.50}$$

Out of all the M rows, we sample Z rows of $\mathbf{y}_{j,k}^{(Sample)}$ to compact the matrix. We set $z\in[Z]$, and pick $t_{z,k}\in[M]$ to compute T as

$$\boldsymbol{p}_{j,l,k}\equiv\mathbb{P}\left(t_{z,k}=l\right)=\alpha\frac{\left|\boldsymbol{y}_{j,l,k}^{(Sample)}\right|}{\boldsymbol{v}_{j,k}}+(1-\alpha)\frac{\left(\boldsymbol{y}_{j,l,k}^{(Sample)}\right)^2}{\boldsymbol{\omega}_{j,k}} \tag{5.51}$$

and set

$$\boldsymbol{x}_{j,l,k}=\boldsymbol{y}_{j,t_{z,k},k}^{(Sample)} \tag{5.52}$$

Then the compressed data $\mathbf{X}$, the indices $\mathbf{T},\boldsymbol{V},\boldsymbol{W}\ and\ \alpha$ used in sampling are propagated from the RRH to the BBU and used to create the unbiased covariance matrix estimator from the compressed data as

$$\boldsymbol{p}_{j,t_{z,k},k}=\alpha\frac{\left|\boldsymbol{x}_{j,z,k}^{(Sample)}\right|}{\boldsymbol{v}_{j,k}}+(1-\alpha)\frac{\left(\boldsymbol{x}_{j,z,k}^{(Sample)}\right)^2}{\boldsymbol{\omega}_{j,k}} \tag{5.53}$$

and

$$\boldsymbol{s}_{j,t_{z,k},k} = \frac{1}{\sqrt{Z\boldsymbol{p}_{j,t_{z,k},k}}} \tag{5.54}$$

Because of imperfection in knowledge of matrix correlation, we perform robust approximation by experimentally optimizing the parameter α. With recent computing improvements, it is possible to manipulate vectors with a length of $O(M)$ in memory. Therefore, compression of data by weighted sampling would require a single pass from the RRH to the BBU when transferring data to memory. This lets the algorithm to be efficient in streaming data and is therefore suitable for use in MPC-RAN systems.

This gives unbiased estimator represented by $\{\boldsymbol{S}_{j,k}\}_{k=1}^{K}$ and $\left\{\boldsymbol{S}_{j,k}^{T}\boldsymbol{y}_{j,k}^{(Sample)}\right\}_{k=1}^{K}$. We presume that $\boldsymbol{y}_{j,k}^{(Sample)} \in \mathbb{C}^{M\times K}$ and the sampling window is set to be $2 \leq Z < M$. Z entries are sampled for every $\mathbf{y}_{j,k}^{(Sample)} \epsilon\, \mathbb{C}^{M}$. The probabilities of sampling are expressed by $\{p_{j,l,k}\}_{l=1}^{M}$ and the matrix of sampling is expressed as $\boldsymbol{S}_{j,k}\epsilon\, \mathbb{C}^{M\times Z}$. The unbiased estimator corresponding to the target covariance matrix $\boldsymbol{\phi}_{j,k} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^{H} = \frac{1}{K}\boldsymbol{Y}^{p}(\boldsymbol{Y}^{p})^{H}$ is recovered as follows

$$\widehat{\boldsymbol{\phi}}_{j,k}^{Compressed} = \widehat{\boldsymbol{\phi}}_{j,k}^{1} - \widehat{\boldsymbol{\phi}}_{j,k}^{2} \tag{5.55}$$

with $\mathbb{E}\left[\widehat{\boldsymbol{\phi}}_{j,k}^{Compressed}\right] = \boldsymbol{\phi}_{j,k}$

$$\widehat{\boldsymbol{\phi}}_{j,k}^{1} = \frac{Z}{KZ-K}\sum_{k=1}^{K}\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{(Sample)}\left(\boldsymbol{y}_{j,k}^{(Sample)}\right)^{H}\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H} \tag{5.56}$$

$$\widehat{\boldsymbol{\phi}}_{j,k}^{2} = \frac{Z}{KZ-K}\sum_{k=1}^{K}\mathbb{D}\left(\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{(Sample)}\left(\boldsymbol{y}_{j,k}^{(Sample)}\right)^{H}\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\right)\mathbb{D}(\boldsymbol{b}_{j,k}) \tag{5.57}$$

with $\boldsymbol{b}_{j,k} = \frac{1}{1+(Z-1)p_{j,k}}$. For each $\boldsymbol{b}_{j,k}$, a maximum of Z entries need to be computed because a given $\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^{H}\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}$ has a maximum of Z none zero elements on the diagonal.

5.4.2.2 Approximation of $\mathbf{R}_{j,k}$

We adopt the same method used for $\boldsymbol{\phi}_{j,k}$ in approximating $\boldsymbol{R}_{j,k} \in \mathbb{C}^{M\times M}$. We aim to minimize the interference from other UTs as we acquire the $\boldsymbol{h}_{j,k}$ observations. It has been pointed out

from [226], [243] that the UT uses a set of unique orthogonal pilots to carry out a training process in order to learn $\boldsymbol{R}_{j,k}$. We assume that the *j*th RRH has N_R observations of the noisy $\boldsymbol{h}_{j,k}$ which sets the foundation for building the approximate matrix of covariance $\widehat{\boldsymbol{R}}_{j,k}$.

It would basically mean more data transmission over the fronthaul from RRH to BBU and higher computations to that effect. The via-Q method is used as outlined in [226] to evaluate the matrix of covariance $\widehat{\boldsymbol{R}}_{j,k}$. This allows the calculation of $\widehat{\boldsymbol{\phi}}_{j,-k}^{(Sample)} = \widehat{\boldsymbol{\phi}}_{j,k} - \boldsymbol{h}_{j,k}\boldsymbol{h}_{j,k}^{H}$ combining all interfering UTs. We presume $\widehat{\boldsymbol{\phi}}_{j,k}^{(Sample)} = \widehat{\boldsymbol{\phi}}_{j,k}^{1}$. It is therefore possible to calculate the covariance matrix $\widehat{\boldsymbol{R}}_{j,k}^{(Sample)}$ as

$$\widehat{\boldsymbol{R}}_{j,k}^{(Sample)} = \widehat{\boldsymbol{\phi}}_{j,k}^{(Sample)} - \widehat{\boldsymbol{\phi}}_{j,-k}^{(Sample)} \tag{5.58}$$

The approximate matrix of covariance $\widehat{\boldsymbol{R}}_{j,k}^{Compressed}$ is computed as

$$\widehat{\boldsymbol{R}}_{j,k}^{Compressed} = \beta\widehat{\boldsymbol{R}}_{j,k}^{(Sample)} - (1-\beta)\widehat{\boldsymbol{R}}_{j,k}^{(Sample)} \tag{5.59}$$

where $\beta \in [0,1]$ is the regularizing parameter used in approximation of $\widehat{\boldsymbol{R}}_{j,k}$.

5.4.2.3 Channel Estimate Approximation

We compute the MMSE estimate corresponding to $\widehat{\boldsymbol{h}}_{j,k}$ based on presumed correct matrices of covariance $\widehat{\boldsymbol{R}}_{j,k}^{Compressed}$ and $\widehat{\boldsymbol{\phi}}_{j,k}^{Compressed}$ as

$$\widehat{\boldsymbol{h}}_{j,k}^{Compressed} = \boldsymbol{W}_{j,k}\boldsymbol{y}_{j,k}^{p} \tag{5.60}$$

where $\boldsymbol{W}_{j,k} = \widehat{\boldsymbol{R}}_{j,k}^{Compressed}\left(\boldsymbol{\phi}_{j,k}^{Compressed}\right)^{-1}$.

5.4.2.4 Spectral Efficiency Estimate Approximation

The presence of imperfect covariance information negatively impacts the quantification of SE hence a lower capacity bound, independent of MMSE estimates is needed. From [3], the *k*th UT in the *j*th cell has a lower bounded channel capacity given by

$$\boldsymbol{SE}_{j,k} = log_2\left(1 + \boldsymbol{SINR}_{j,k}\right)\ [bit/s/Hz] \tag{5.61}$$

where

$$\boldsymbol{SINR}_{j,k} = \frac{\left|\mathbb{E}\left[\boldsymbol{h}_{j,k}^{H}\boldsymbol{f}_{j,k}\right]\right|^{2}}{\sigma^{2} + \mathbb{E}\left[\left|\boldsymbol{h}_{j,k}^{H}\boldsymbol{f}_{j,k}\right|^{2}\right] - \left|\mathbb{E}\left[\boldsymbol{h}_{j,k}^{H}\boldsymbol{f}_{j,k}\right]\right|^{2} + \sum_{\ell,m}\mathbb{E}\left[\left|\boldsymbol{h}_{\ell,k}^{H}\boldsymbol{f}_{\ell,m}\right|^{2}\right] - \boldsymbol{\Upsilon}_{cp}} \quad (5.62)$$

with the channel realization expectation assumed. We assumed partially centralized C-RAN with interconnected RRHs and so in section 4.4 we use (4.41) to capture the cooperation between RHHs. It is evident here that the power bound is independent of the channel estimation approach used and the combination scheme.

5.4.2.5 Simulation Results and Discussion

This section provides the NMSE, SE, and M comparison and evaluation of the RNA-MMSE channel estimation technique in section 5.3, via-Q method in [226] and the compressed channel data channel estimation in MPC-RAN. This comparison is performed for M ranging from $32 \ to \ 176$, with $K = 16$ MPC-RAN system and the SNR ranges from 0dB to 20dB. Then, we average the SE and NMSE over this SNR range to obtain the average SE and NMSE.

Figure 5.7 compares the achievable SE per RRH in MPC-RAN between MR and the RNA-MMSE precoding techniques with a reuse factor of 1 for RNA-MMSE, the via-Q and the compressed data channel estimation techniques. A number of observations can be deduced on the basis of Figure 5.7. The RNA-MMSE based SE has the best attainable SE per RRH as predicted from previous analyses, but it can be deduced that the SE calculated using the compressed data channel approximation is inferior to the RNA-MMSE channel approximation and the via-Q channel approximation for both MR and RNA-MMSE. This is expected since the compressed data channel approximates the RNA-MMSE channel. Also, there is enhanced SE performance both in MR and RNA-MMSE based precoding with increasing number of RRH antennas due to the channel hardening phenomenon.

As per Figure 5.8, the SE obtain from MR and RNA precoders employing RNA-MMSE, via-Q and compressed data channel approximations is enhanced when the reuse factor is increased to 4. The increase in SE per RRH is due to the fact that, with the increased number of pilots, the pre-log factor is minimized. As the channel estimates get better with reduced pilot contamination, this further leads to improved instantaneous SINRs.

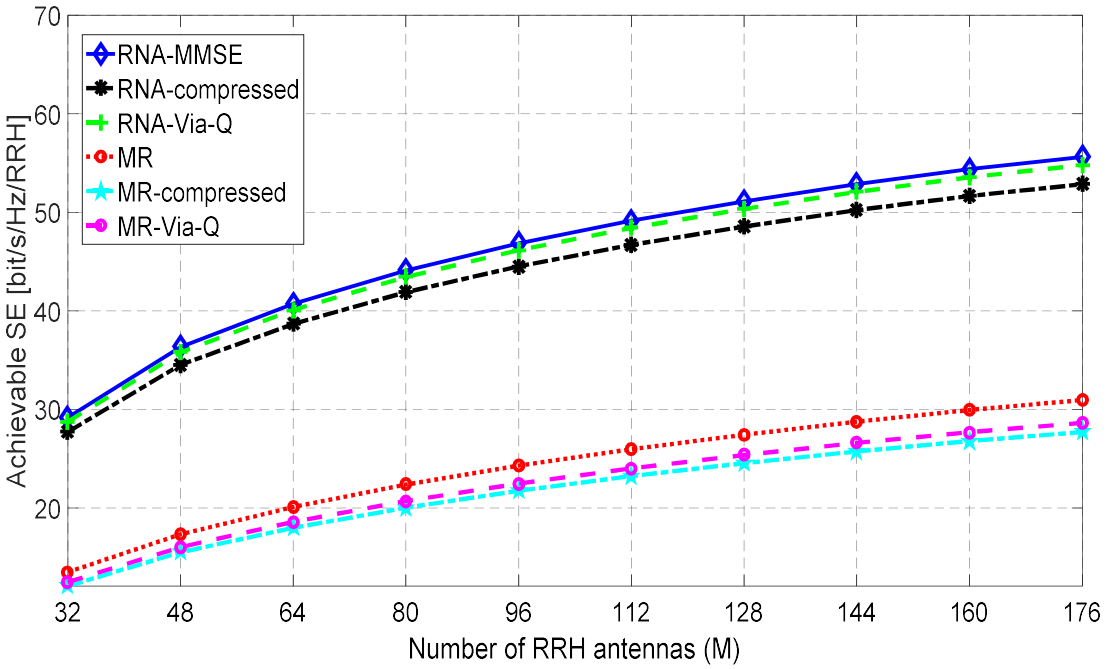

Figure 5.7: Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 1.

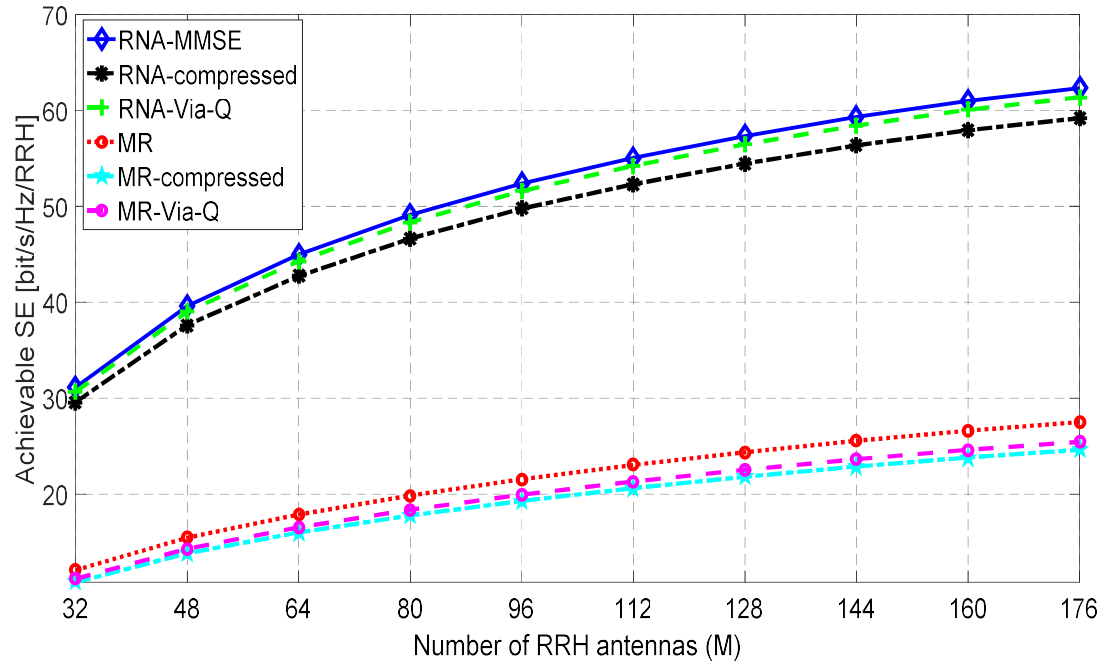

Figure 5.8: Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 4.

The trade-offs between NMSE and the number of RRH antennas is shown in Figure 5.9, for a reuse factor of 1. The channel hardening phenomenon realized from increased RRH antennas enhances the channel approximation leading to reduced NMSE. Again, the approximation of the RNA-MMSE channel and the via-Q channel have lower NMSE in comparison to the

compressed data channel approximation since the compressed data channel approximates the RNA-MMSE channel. But the compressed data channel NMSE nears that of the RNA-MMSE channel with increasing M because the channel approximation is enhanced as the number of RRH antennas increase.

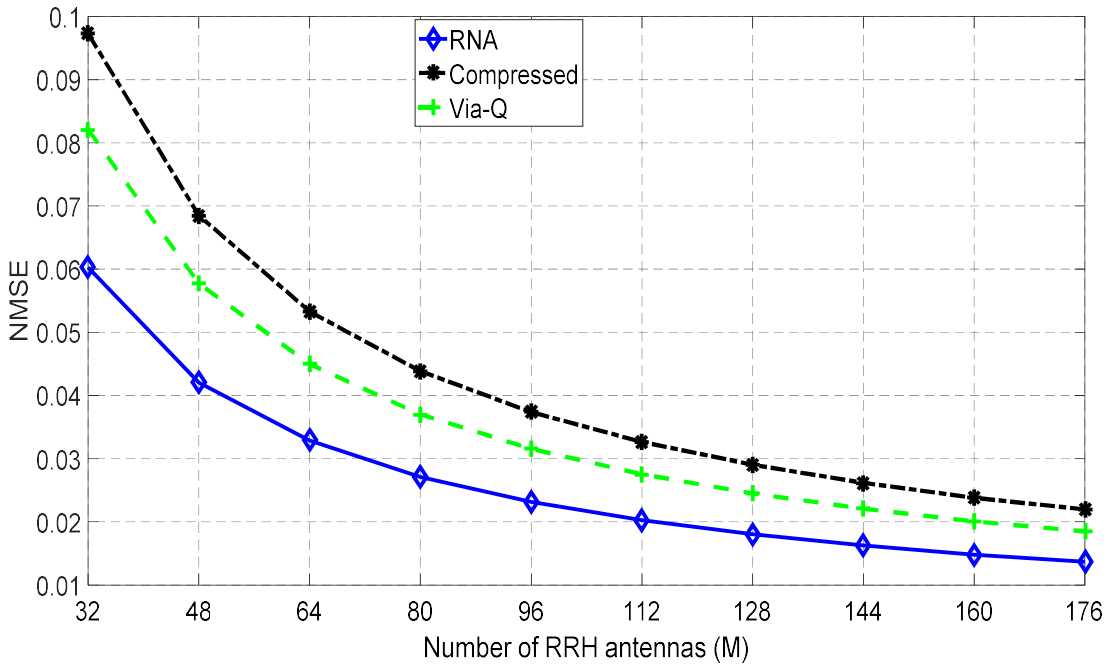

Figure 5.9: The normalized MSE vs. number of RRH antennas with a reuse factor of 1.

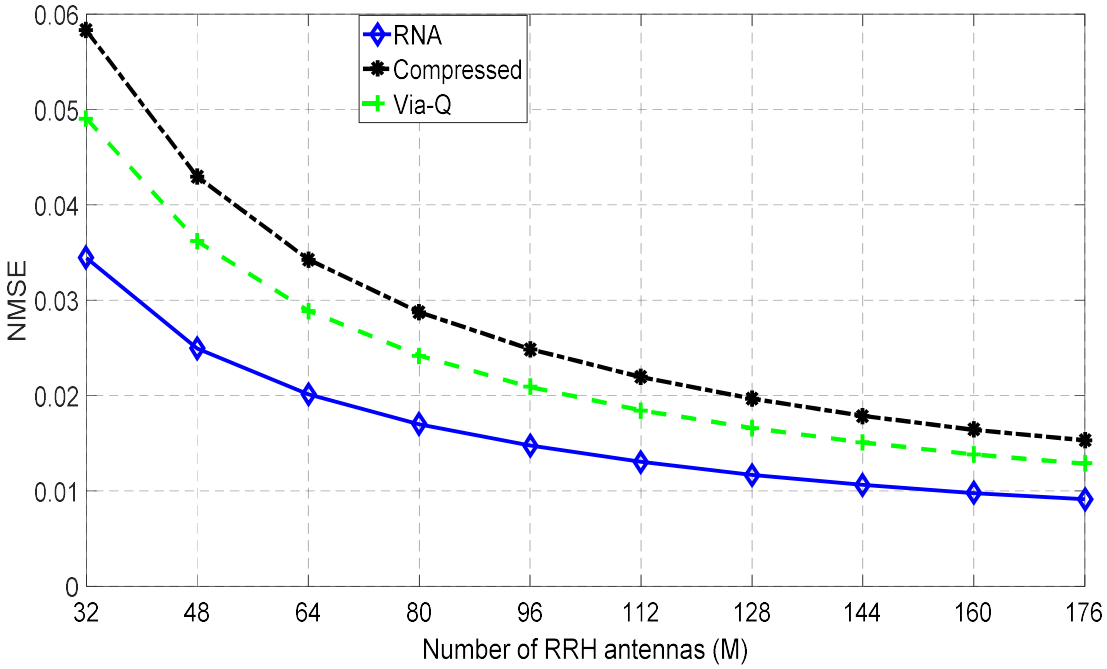

Figure 5.10: The normalized MSE vs. number of RRH antennas with a reuse factor of 4.

A reuse factor of 4 is used in Figure 5.10, which shows a reduction in the NMSE when compared to Figure 5.9 with a reuse factor of 1. The increase in reuse factor from 1 to 4 minimizes pilot contamination and in turn enhances the channel approximation leading to reduced NMSE for all RNA-MMSE, the via-Q and compressed data channel estimation techniques.

5.4.3 GDPM Based Semi-Blind Model for Channel Estimation

We mentioned earlier that computation of the covariance matrices, needs huge communication and storage resources since the use of MPC-RAN results in high-dimensional data. But it is also worth noting that the computation of the covariance matrices requires enormous computation time. But we reckon that channel estimation is done at the BBU. The computational resource here at the BBU is great with availability of multicore processing and with properly thought out parallelization, the computation time can be reduced and hence speed up the channel estimation process.

Here we rely on the initial approximation of the compressed data channel to avail the initial input which the subspace estimation methods leverage on to approximate the matrix of covariance for channel estimation. DPM and FDPM are among the commonly used subspace methods [244]. The FDPM relies on Householder orthonormalization while the Gram-Schmidt method is used to perform orthonormalization process in DPM. Consequently, the computation complexity of FDPM is $\mathcal{O}(MK)$ while that of DPM is $\mathcal{O}(\mathrm{MK}^2)$**.** None of these two, however, provides effective parallelization, and are therefore not ideal for operating in multicore systems that require efficiently parallel algorithms. Thus, in order to obtain an efficient parallelization which is highly needed in MPC-RAN BBUs with enormous computing resources, Givens rotation is used in orthonormalization process despite the fact that it carries $\mathcal{O}(\mathrm{MK}^2)$ computational complexity.

Lower complexity or higher complexity is just but a product of the flop-count and might not point to the fact that an algorithm is more efficient than the other. This becomes a motivating factor when multicore systems are involved in computing as is the case of BBUs in MPC-RAN. Because an efficiently parallel algorithm might become more superior even though with higher complexity [235]. This forms our basis of choosing Givens orthonormalization for the MPC-RAN system channel estimation.

Givens Rotations renders itself to calculations where it is of importance to selectively zero particular elements [245]. Each rotation can only impact two rows of the given matrix, we can

interchange the order of rotations that impact different rows, hence facilitating the application of sets of rotation in parallel [245]. This is the reason we said Givens transformation lends itself to efficient parallelization. The Givens transformation also come in handy when updating of a matrix is needed after a row is added or a column is deleted. This is equivalent to addition of RRH antenna as a result of evolution of the communicating RRH antennas for a UT in motion and column deletion is when a UT drops out of network for whatever reasons.

Again from [152], we compute the covariance matrix for the received signal as follows

$$\boldsymbol{\phi}_y = E\{\boldsymbol{y}\boldsymbol{y}^H\} = \boldsymbol{H}\boldsymbol{H}^H + \boldsymbol{I}_M \tag{5.63}$$

Then using SVD $\boldsymbol{\phi}_y$ is decomposed to obtain

$$\boldsymbol{\phi}_y = [\boldsymbol{U}_s\, \boldsymbol{U}_n]\Lambda[\boldsymbol{U}_s\, \boldsymbol{Y}_n] \tag{5.64}$$

where $\boldsymbol{U}_s \in \mathbb{C}^{M\times K}$ is subspace of the signal and $\boldsymbol{U}_n \in \mathbb{C}^{M\times(M-K)}$ is subspace of the noise. From [152] we use the scalar multiplicative ambiguity matrix $\boldsymbol{B} \in \mathbb{C}^{K\times K}$ to compute the channel matrix $\boldsymbol{H}$ from $\boldsymbol{U}_s$, hence;

$$\widehat{\boldsymbol{H}} = \boldsymbol{U}_s\boldsymbol{B} \tag{5.65}$$

To obtain the ambiguity matrix we leverage the short training sequence in (5.4) and compute as

$$\boldsymbol{B} = (\boldsymbol{U}_s)^H \widehat{\boldsymbol{H}}^{Compressed} \tag{5.66}$$

Where the $\widehat{\boldsymbol{H}}^{Compressed}$ is the initial channel estimate obtained from the compressed channel estimation in (5.60). From which we compute the channel estimate as

$$\widehat{\boldsymbol{H}} = \boldsymbol{U}_s(\boldsymbol{U}_s)^H \widehat{\boldsymbol{H}}^{Compressed} \tag{5.67}$$

DPM subspace tracking algorithm was adopted as in [246] with Givens rotation used in orthonormalization process instead of the Gram-Schmidt orthonormalization. Though this does not reduce the complexity of the algorithm, it renders the algorithm to efficient parallelization which is an important aspect in MPC-RAN system.

5.4.3.1 Serial Givens Data Projection Method (SGDPM)

We compute the multiplication of a matrix by exploiting the simple structure of the Givens rotation matrix. We assume a matrix $\boldsymbol{A} \in \mathbb{C}^{M\times K}$, and parameters c and s. Supposing $\boldsymbol{G} \in \mathbb{C}^{M\times M}$, the updated matrix $\widetilde{\boldsymbol{A}} = \boldsymbol{GA}$ will only affect two rows, $[i, k]$ thus

$$\widetilde{\boldsymbol{A}}_{[i,k]} = \boldsymbol{G}_{[i,k]}\boldsymbol{A}_{[[i,k]]} \tag{5.68}$$

with $\boldsymbol{G}_{[i,k]} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$. This is then expressed as an algorithm in Table 5.6 and we refer to it as the Serial Givens Data Projection Method (SGDPM) in this work.

Table 5.6: SGDPM sub-space tracking algorithm.

for $k = 1, 2, \dots, K$
$\boldsymbol{A}_k = \widetilde{\boldsymbol{A}}_{k-1}^H \boldsymbol{y}_k$
$\boldsymbol{T}_k^1 = \boldsymbol{A}_{\boldsymbol{i-1},k}$
$\boldsymbol{T}_k^2 = \boldsymbol{A}_{\boldsymbol{i},k}$
$\widetilde{\boldsymbol{A}}_{\boldsymbol{i-1},k} = c\boldsymbol{T}_k^1 - s\boldsymbol{T}_k^2$
$\widetilde{\boldsymbol{A}}_{\boldsymbol{i},k} = s\boldsymbol{T}_k^1 + c\boldsymbol{T}_k^2$
end

Table 5.7: SGDPM sub-space tracking algorithm parameters.

if $\boldsymbol{T}_k^2 = 0$
$c = 1$
$s = 0$
else if $|\boldsymbol{T}_k^2| > |\boldsymbol{T}_k^1|$
$t = \boldsymbol{T}_k^1/\boldsymbol{T}_k^2$
$s = \frac{1}{\sqrt{1+t^2}}$
$c = st$
else
$t = \boldsymbol{T}_k^2/\boldsymbol{T}_k^1$
$c = \frac{1}{\sqrt{1+t^2}}$
$s = ct$
end

The parameters c and s are computed as in the algorithm in Table 5.7. From the Table 5.6, $\widetilde{\boldsymbol{A}}_k \in \mathbb{C}^{M\times K}$ is tracked signal subspace corresponding to the kth sample. The length of signals received without pilots is denoted by N. The estimated $\boldsymbol{U}_s$ is obtained from the tracked signal subspace $\widetilde{\boldsymbol{A}}_K$ and expressed as

$$\boldsymbol{U_s} = \widetilde{\boldsymbol{A}}_K \tag{5.69}$$

Hence, the estimated channel matrix is computed as

$$\widehat{\boldsymbol{H}}^{GDPM} = \widetilde{\boldsymbol{A}}_K \left(\widetilde{\boldsymbol{A}}_K\right)^H \widehat{\boldsymbol{H}}^{Compressed} \tag{5.70}$$

5.4.3.2 Parallel Givens Data Projection Method (PGDPM)

The algorithm adapted in section 5.4.3.1 is the classical Givens rotation. This method can be improved on by the column-wise Givens rotation where the annihilation of multiple elements of a column within the input matrix can be realized. This modification has the advantage of less multiplications than the classical Givens rotation implementation [247], [248]. It also has the capability of exploiting both coarse- and fine-grained parallelism.

First, we condition the input matrix $\boldsymbol{Y} \epsilon \mathbb{C}^{M\times K}$, through multiplication with an initial matrix $\widetilde{\boldsymbol{A}}_{\boldsymbol{0}} = \begin{bmatrix} \boldsymbol{I}_K \\ \boldsymbol{0}_{(M-K)\times K} \end{bmatrix}$ to obtain a matrix $\boldsymbol{A} \epsilon \mathbb{C}^{M\times M}$. We presume $\boldsymbol{G} \epsilon \mathbb{C}^{M\times M}$, to obtain the updated matrix $\widetilde{\boldsymbol{A}} = \boldsymbol{GA}$ and rewrite it as

$$\widetilde{\boldsymbol{A}}_{[i,k]} = \boldsymbol{G}_{[i,k]} \boldsymbol{A}_{[i,k]} \tag{5.71}$$

and

$$\boldsymbol{G}_{[i,k]} = diag\left(\boldsymbol{I}_{i-2}, \widetilde{\boldsymbol{G}}_{[i,k]}, \boldsymbol{I}_{M-i}\right) \tag{5.72}$$

with $\widetilde{\boldsymbol{G}}_{[i,k]} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$. We presume an element *(m, 1)*, located in the 1st column of the *m*th raw and apply a single Givens transformation which allows the rewriting of (5.71) as

$$\widetilde{\boldsymbol{A}} = \boldsymbol{G}_{[M,1]} \boldsymbol{A} = \begin{bmatrix} R^1 \\ 0 \end{bmatrix} \tag{5.73}$$

where R^k epitomises an upper triangular matrix after undergoing k-updates. Multiple Givens transformations can be applied in unison to eliminate multiple elements in a column of an input matrix [248]. Thus, we can eliminate 2-elements in the 1st column of the input matrix by extension of equation (5.73), hence

$$\boldsymbol{G}_{[M-1,1]}\boldsymbol{G}_{[M,1]}\boldsymbol{A} = \begin{bmatrix} R^2 \\ 0 \end{bmatrix} \tag{5.74}$$

With a slight modification on (5.74) $M-1$ elements can be eliminated from the 1st column of an input matrix, hence

$$\boldsymbol{G}_{[2,1]}\boldsymbol{G}_{[3,1]} \dots \boldsymbol{G}_{[M-1,1]}\boldsymbol{G}_{[M,1]}\boldsymbol{A} = \begin{bmatrix} R^{M-1} \\ 0 \end{bmatrix} \tag{5.75}$$

Then equation (5.75) can be extended to eliminate *M-1* elements in 1st column and $M-2$ elements in 2nd column. Hence, the total elements eliminated are $(M-1) + (M-2)$ and the resultant matrix will be $R^{(M-1)+(M-2)}$.

$$\begin{aligned} &\left(\boldsymbol{G}_{[3,2]}\boldsymbol{G}_{[4,2]} \dots \boldsymbol{G}_{[M-1,2]}\boldsymbol{G}_{[M,2]}\right)\left(\boldsymbol{G}_{[2,1]}\boldsymbol{G}_{[3,1]} \dots \boldsymbol{G}_{[M-1,1]}\boldsymbol{G}_{[M,1]}\right)\boldsymbol{A} \\ &= \begin{bmatrix} R^{(M-1)+(M-2)} \\ 0 \end{bmatrix} \end{aligned} \tag{5.76}$$

And in general, we can eliminate all the elements in the respective columns and hence we rewrite equation (5.76) as

$$\begin{aligned} &\left(\boldsymbol{G}_{[M,M-1]}\right)\left(\boldsymbol{G}_{[M-1,M-2]}\boldsymbol{G}_{[M,M-2]}\right)\cdots\left(\boldsymbol{G}_{[3,2]}\boldsymbol{G}_{[4,2]} \dots \boldsymbol{G}_{[M-1,2]}\boldsymbol{G}_{[M,2]}\right) \\ &\left(\boldsymbol{G}_{[2,1]}\boldsymbol{G}_{[3,1]} \dots \boldsymbol{G}_{[M-1,1]}\boldsymbol{G}_{[M,1]}\right)\boldsymbol{A} = \begin{bmatrix} R^{(M-1)+(M-2)+\cdots+(M-(M-2))+(M-(M-1))} \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} R^{(M-1)+(M-2)+\cdots+2+1} \\ 0 \end{bmatrix} \end{aligned} \tag{5.77}$$

and Givens rotation is written in general for as in equation (5.77).

Column-wise Givens rotation transforms a single column per iteration while the generalized Givens rotation operates column-wise and row-wise concurrently to triangulate a matrix of $M \times M$ dimension in a single iteration. This is illustrated in Figure 5.12.

The number of theoretical iterations required to carry out Givens transformation-based computations to upper triangulate a matrix of size $M \times M$ for classical Givens rotation is $\frac{M(M-1)}{2}$ and for column-wise Givens rotation is $M-1$ while the generalised Givens rotation yields a 1. Then we realize the generalized Givens algorithm as illustrated in Table 5.8. This algorithm forms the basis for the sub-space tracking of the Parallel GDPM (PGDPM), and hence the semi-blind estimator based on PGDPM.

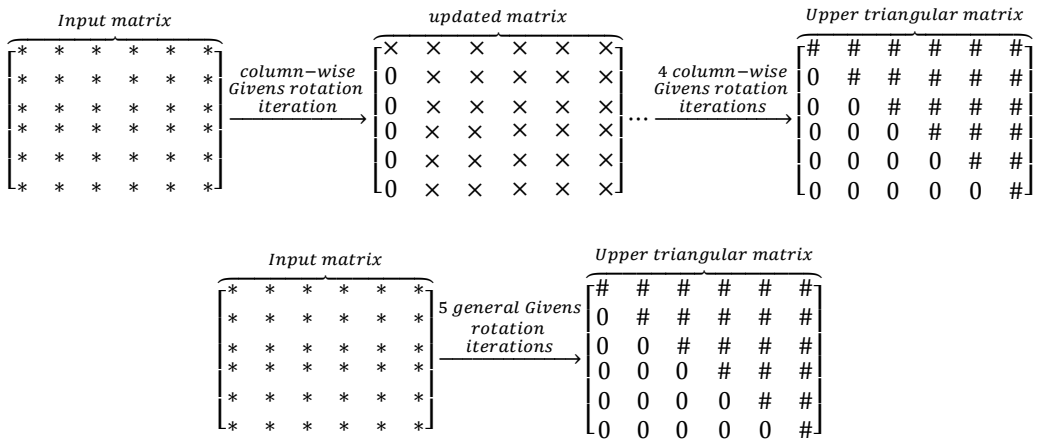

Figure 5.11: The column-wise Givens rotation and general Givens rotation.

Table 5.8: PGDPM sub-space tracking algorithm.

```
for j = 1, 2, ..., L
    for m = 1, 2, ..., M
        for i = 1, 2, ..., M
            A_j = Ã_0^H y_j
            T(m) = ||A_j(:, m)||_2
            c = A_j(i − 1, m)/T(m)
            s = A_j(i, m)/T(m)
            Ã_{j,1} = G_1 A_j
            Ã_{j,2} = G_2 A_j
                    ⋮
            Ã_{j,M−1} = G_{M−1} A_j
            Ã_{j,M} = G_M A_j
        end
    end
end
```

where $\widetilde{\boldsymbol{A}}_{j,1}, \ldots, \widetilde{\boldsymbol{A}}_{j,M}$ represents the update of row 1 through to row M in the *m*th column. It is important to note that *j* represents the particular RRH we are currently operating in, *m* represents the columns of the matrix $\boldsymbol{A}_j$ and *i* represents the rows of the matrix $\boldsymbol{A}_j$. From Table 5.8 updating of the first row up to the last can be executed concurrently hence allowing parallelization in updating of rows in a column. But again, out of loop iteration can be executed concurrently meaning that the updating of row in different columns can be parallelized. Furthermore, the computation of the generalized Givens rotation for different RRHs in MPC-RAN system can be parallelized as per the algorithm in Table 5.8. This then makes the PGDPM an efficient algorithm to be employed at the BBU for the channel estimation process. To further

enhance parallelization, we adopt the method in Appendix C to carry out matrix-matrix multiplication. This has the advantage that it does not need synchronization of cores in BBU.

5.4.3.3 Simulation Results and Discussion

We proceed to perform a comparison and analysis of the NMSE, reuse fact f and M for the RNA-MMSE channel estimation technique, FDPM-based semi-blind channel estimation technique and PGDPM-based semi-blind channel estimation technique in MPC-RAN. We compare these channel estimation techniques in MPC-RAN. This comparison is carried out for M varying from $32 \ to \ 176$, with a step of 16 and $K = 16$ MPC-RAN system and for f varying as $[1, 2, 4]$. In this work the PGDPM and GDPM are used interchangeably and refer to one and the same thing.

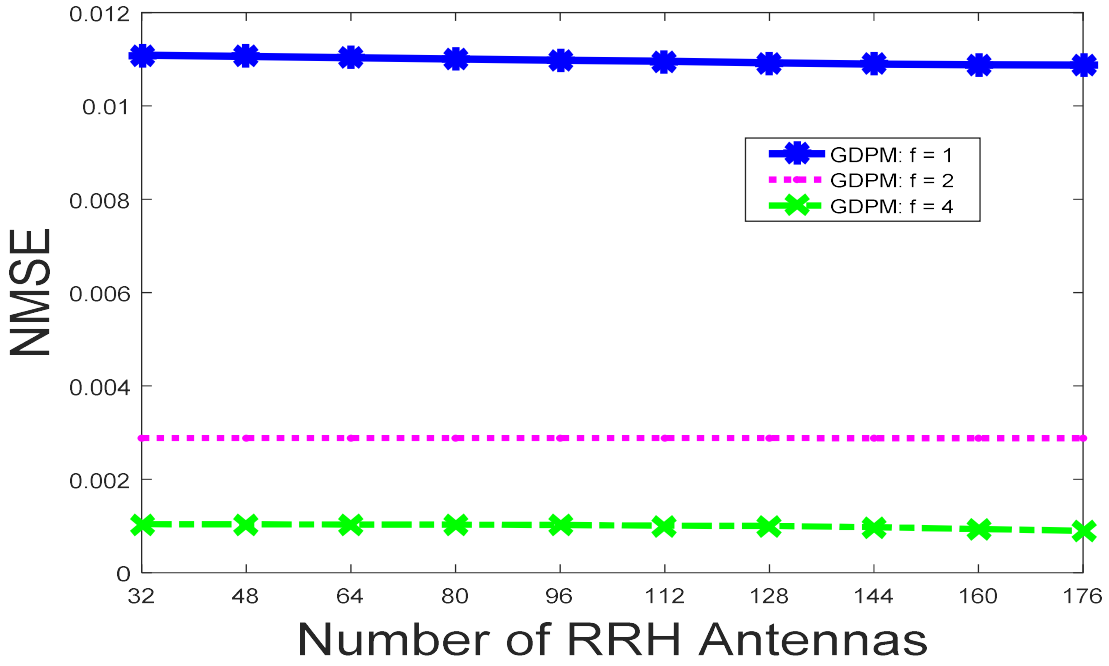

Figure 5.12: Achievable NMSE vs. number of RRH antennas for PGDPM-based semi-blind estimator ($\boldsymbol{f = [1, 2, 4]}$).

Table 5.9: Performance of PGDPM estimator.

M	***f***		
	1	**2**	**4**
16	0.0110891	0.00288340	0.00103900
96	0.0109548	0.00288293	0.00100669
160	0.0108728	0.00288274	0.000895608

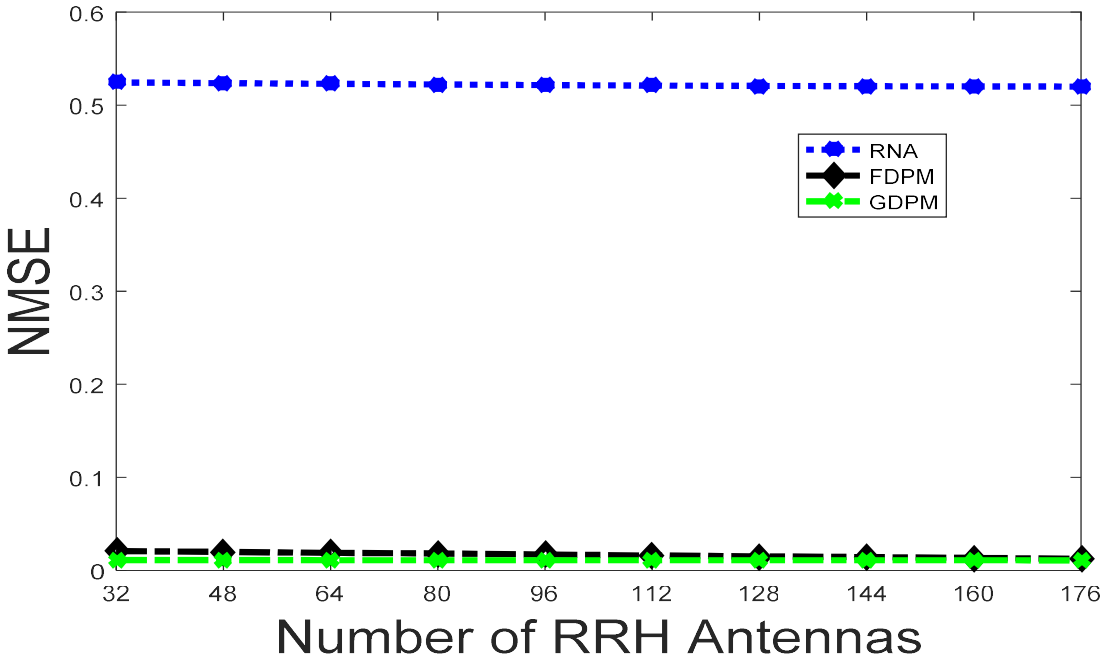

Figure 5.13: Comparison of achievable NMSE vs. number of RRH antennas for RNA, FDPM and PGDPM estimators ($\boldsymbol{f = 1}$).

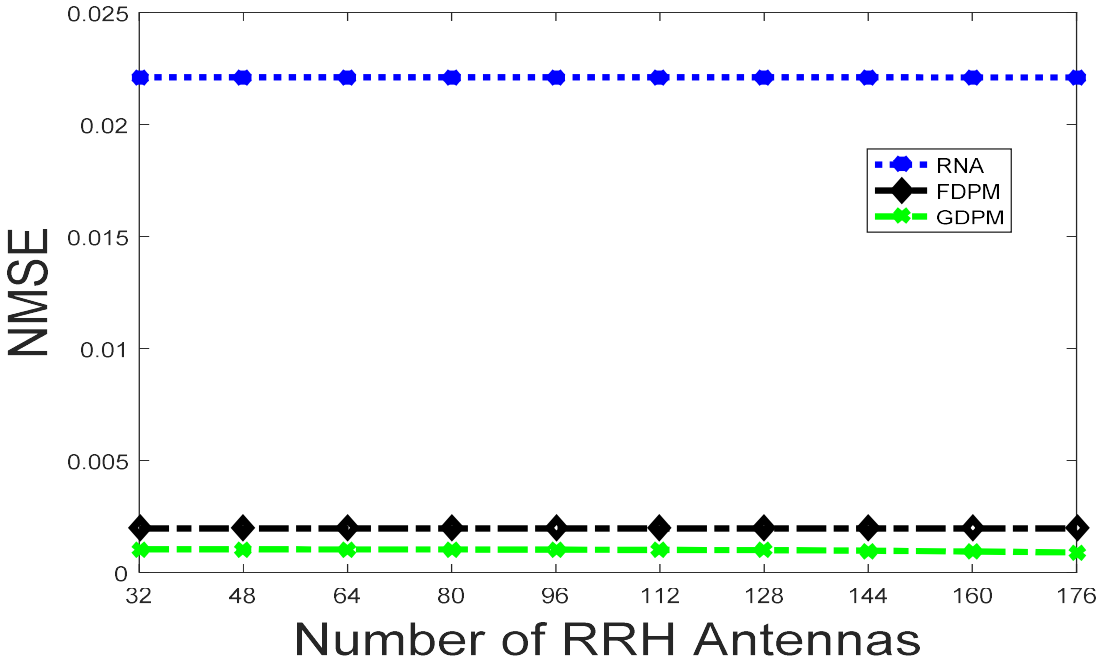

Figure 5.14: Comparison of achievable NMSE vs. number of RRH antennas for RNA, FDPM and PGDPM estimators ($f = 4$).

Figure 5.12 contrasts the achievable NMSE vs. number of RRH antennas for PGDPM-based semi-blind estimator in MPC-RAN. From this figure, a number of observations can be deduced.

The performance of PGDPM-based semi-blind estimator improves with the increase in the number of RRH antennas as well as increase in the reuse factor. This is depicted by a reduction in the NMSE when the number of RRH antennas increase. It is also obvious that the NMSE decreases with increase in the reuse factor Table 5.9. This points to the fact that the increase in M reduces the NMSE at all values of f because of channel hardening inherent in MPC-RAN. At an f of 4 the increase in BS antennas still impacts on the NMSE reducing it further.

Figure 5.13 and Figure 5.14 presents a comparison on the performance of RNA-based, FDPM-based and PGDPM-based channel estimation techniques. Figure 5.13 shows that PGDPM-based semi-blind estimation technique has the lowest NMSE at $f = 1$ followed by FDPM and then RNA. The FDPM closely follows the PGDPM. From Figure 5.14, it's clear that at $f = 4$ PGDPM-based semi-blind estimation technique still outperforms the conventional FDPM and RNA-based MMSE techniques. But in overall we observe that the NMSE is lower at $f = 4$ than it is at $f = 1$.

In overall it can be observed that the graphs at $f = 1$ have steeper gradient than that at $f = 4$. That is the graph looks flattened at a higher reuse factor. This is attributed to the fact that at a higher reuse factor the pilot contamination is minimized improving the network condition which in turn reduces the NMSE and thus the impact of increase in the RRH antenna has minimal effect on the NMSE hence the observed flattening.

Next, we provide the comparison and analysis of NMSE, SNR and M for the RNA-MMSE channel estimation technique, the channel estimation using compressed channel data, FDPM and PGDPM in MPC-RAN. This comparison is carried out for M varying from $32\ to\ 176$, with a step of 16 and $K = 16$ MPC-RAN system and the SNR varies from 0dB to 20dB in steps of 2dB. In order to get the average NMSE we take the NMSE over the RRH antenna range at particular value of f for each SNR between 0dB to 20dB in 2dB steps. Next, we average this NMSE for all SNR considered at a given f over the range of RRH antennas. This then yields the NMSE over a given RRH range for a specified f and plotted as follows.

Figure 5.15 shows the normalized MSE (NMSE) against the number of RRH antennas for RNA-MMSE, compressed data, FDPM and PGDPM channel estimation techniques for a reuse factor 1. As the number of RRH antennas increase the NMSE decreases since the channel estimation improves due channel hardening phenomenon. Again, the PGDPM channel estimation and the FDPM channel estimation have less NMSE compared to the RNA-MMSE and compressed data channel estimation because the semi-blind channel estimation techniques are superior to linear channel estimation techniques. But as M increases the RNA-MMSE and

compressed data channel estimation NMSE nears that of the FDPM and PGDPM channel estimation since the approximation improves with the increase in the number of antennas due to channel hardening phenomenon.

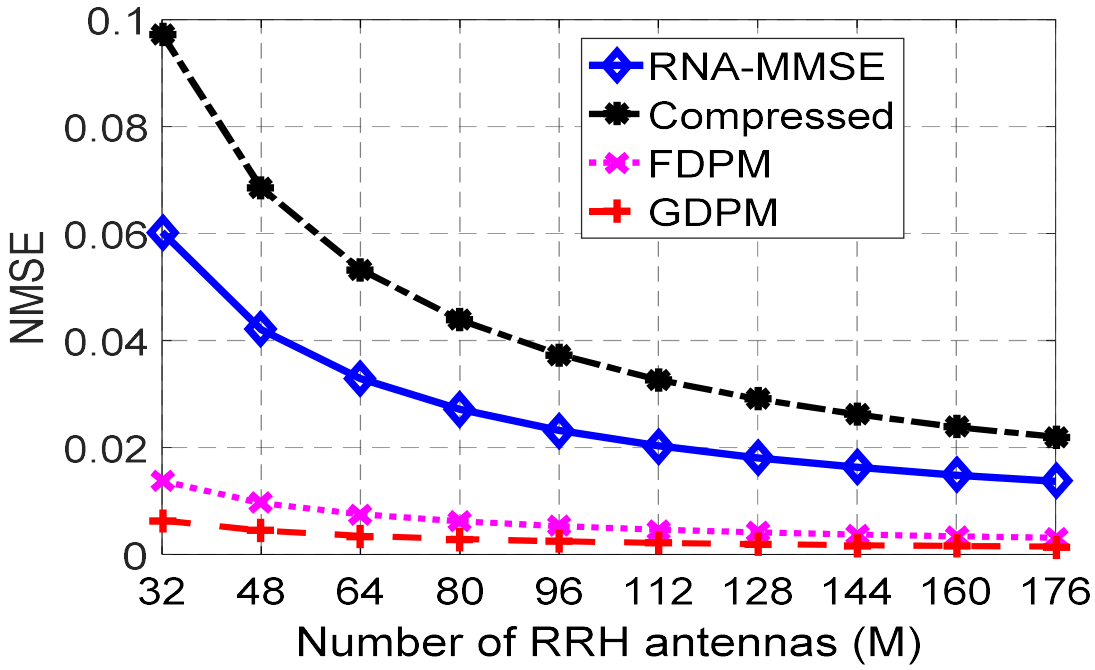

Figure 5.15: The normalized MSE vs. number of RRH antennas with a reuse factor of 1.

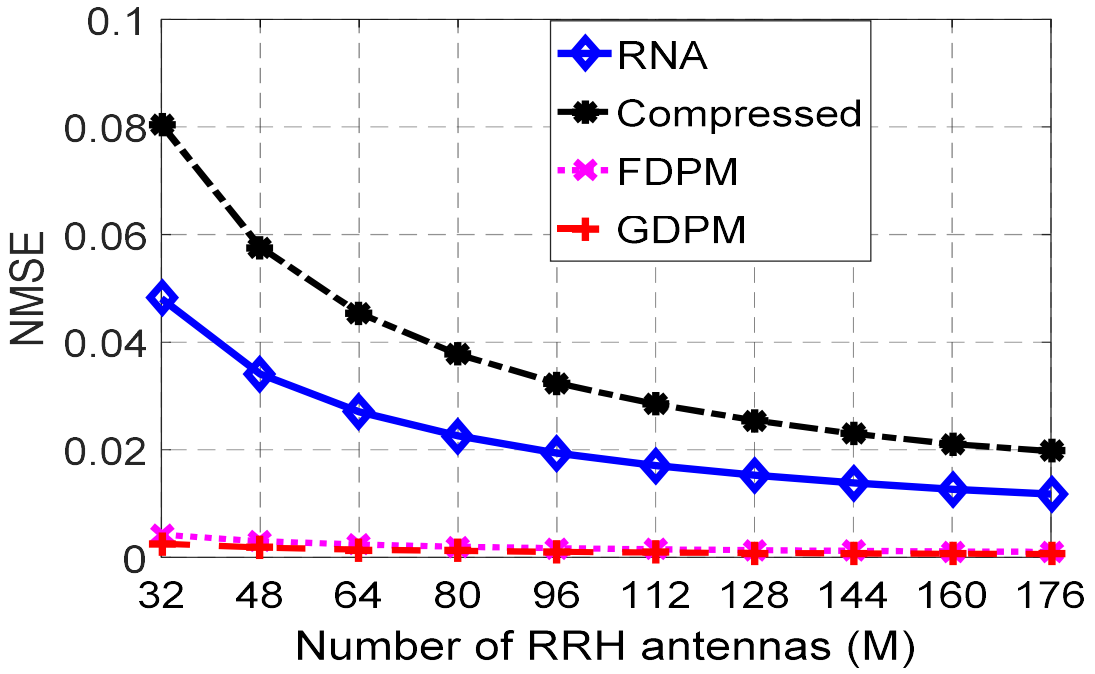

Figure 5.16: The normalized MSE vs. number of RRH antennas with a reuse factor of 2.

In Figure 5.16 and Figure 5.17, the reuse factor is set to2 and 4 respectively and the NMSE is reduced as compared to the case when the reuse factor is set to 1 and also a reuse factor of 2 has a higher NMSE than the reuse factor of 4. This can be attributed to the fact that as the reuse factor increases the pilot contamination reduces and this enhances the channel estimation process leading to a reduction in NMSE for the RNA-MMSE, compressed data, FDPM and PGDPM channel estimation techniques.

Again, it can be noted that the FDPM has a slightly higher NMSE than the PGDPM pointing to the fact that though PGDPM has a high complexity its performance is superior to that of FDPM. Thus, with parallelization we can lower the complexity in PGDPM and thus exploit its superior performance in channel estimation.

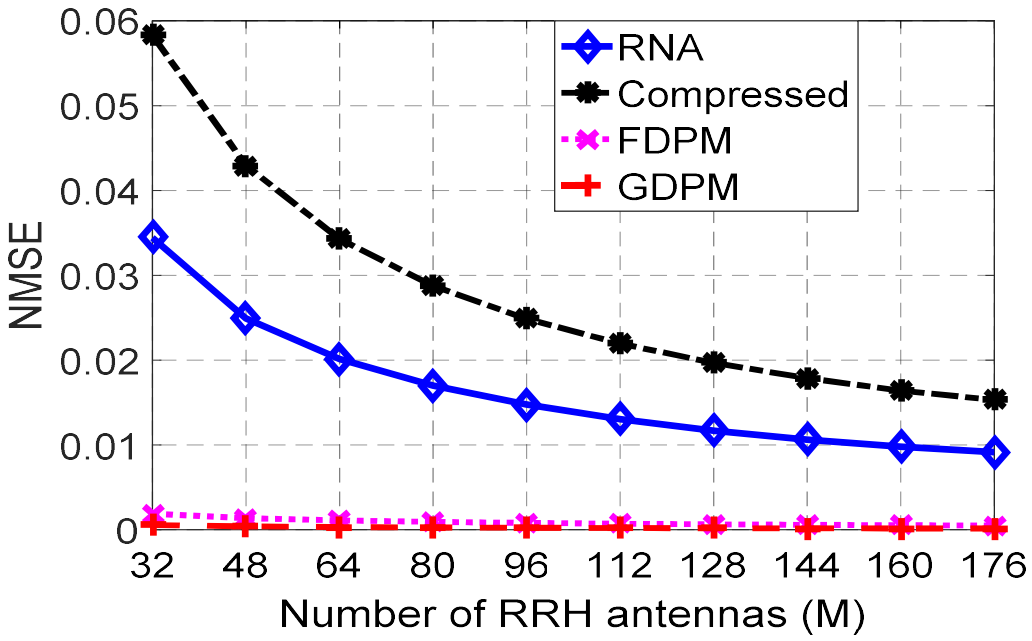

Figure 5.17: The normalized MSE vs. number of RRH antennas with a reuse factor of 4.

5.5 Conclusion on Channel Estimation

In this chapter the performance analysis and comparison of the channel estimation for multicell massive MIMO system and MPC-RAN system are formulated and evaluated. The performance of the various channel estimation schemes in terms of SE and the NMSE is studied. The SE and NMSE were derived theoretically for each of the channel estimation schemes under similar assumptions and for the wireless massive MIMO system and MPC-RAN system. From the simulation and the theoretical results, RNA-based precoding has higher SE than the MR. The

NMSE for the semi-blind channel estimation techniques is lower than that of the linear channel estimation as expected from the literature. Data compression can be leveraged in MPC-RAN system to reduce the fronthaul and use PGDPM with efficient parallelization to exploit computing power at BBU to perform channel estimation.

CHAPTER 6

Conclusion and Recommendations

6.1 Introduction

Massive MIMO communication systems and its applications stands out as one of the key technologies that are and will continue satisfying the quest for high network capacity spectral efficiency, energy efficiency, security, and robustness. The massive MIMO Partially Centralized C-RAN (MPC-RAN) is envisaged as a game changer in 5G networks and beyond. However, massive MIMO is inherently susceptible to pilot contamination and MPC-RAN suffers from finite fronthaul capacity due to centralized BBU though with enormous computing resources. Thus, efficient precoding methods combined with optimal channel estimation techniques are required to mitigate these problems. Highly parallelizable techniques are required to be implemented at the BBU to exploit the enormous computing resources.

6.2 Conclusions

In this section we discuss the conclusions drawn from chapter 3, chapter 4 and chapter 5 in this study as follows:

In **chapter 3** the use of RNA-based iterative linear precoding for the single cell massive MIMO system has been presented. A class of rapid numerical algorithms are used to circumvent the matrix inversion inherent in the generation of the precoding matrix. This was realized by formulating the linear iterative RNA-based precoding. Then, a criterion to stop the iterative method was formulated followed by the formulation of the initial value of the linear iterative RNA-based precoder. The complex analysis of the linear iterative RNA-based precoder was done and the performance of this precoder was validated against the conventional linear precoders. These algorithms conditions the precoder matrix generation process to endear itself to high parallelization which when effectively implemented lowers its complexity considerably.

In **chapter 4** the RNA-based precoder was adapted for the multicell massive MIMO system and coordination between the corresponding cell BS introduced. Then the energy efficient of

the multicell linear iterative RNA-based precoder was computed and the receive combining computational complexity modelled. Then the modelled multicell RNA-based precoder was validated against the conventional multicell linear precoders.

Then channel estimation scenario was considered in **chapter 5**. First, we modelled an improved MMSE channel estimator for multicell massive MIMO system. Then this was combined with the FDPM sub-space channel estimator to realize a semi-blind channel estimator. Then these models were validated and compared to each other in terms of their performance. Then the MPC-RAN was considered next, and a compressed data channel estimation model was realized but followed the via-Q method to perform channel estimation. This was done to compress the channel estimation pilot data to ease the fronthaul capacity constraint. This was validated against conventional channel estimators.

Then we sought to exploit the intense computational resources and enhance channel estimation in MPC-RAN by modelling a sub-space channel estimator, PGDPM, which is highly parallelizable estimator based on Givens transformations. This was then combined with the compressed data channel estimator to realize a semi-blind channel estimator for MPC-RAN. Then the performance of this semi-blind model was validated.

The performance and analysis of the RNA-based precoder, for single-cell and multicell massive MIMO as well as the channel estimation models was evaluated in MATLAB for the downlink massive MIMO system. By comparison to other conventional precoders the performance of RNA-based precoder was shown to be optimal and with low complexity due to high parallelization ability. It was found to perform equally well in comparison to other precoders in terms of SINR, SE and EE. The compressed data channel estimator was found to perform well in relation to the conventional channel estimators in terms of SE and the NMSE. The PGDPM performance was good, and the resultant semi-blind channel estimator performed very well also.

6.3 Recommendations

There are several recommendations we put forward which can emanate from this study.

- This RNA precoding scheme need to be combined with nonlinear precoding scheme and to come up with ways to optimize the hybrid precoding scheme and evaluate its performance. Also, the RNA precoding scheme can be extended to space-time coding.
- Implementation of the iterative linear RNA-based precoder to multicore computing systems to test the parallelization aspect and how efficient it is can be undertaken.

- The semi-blind channel estimator modelled for the MPC-RAN can be implemented to test the compressed data channel estimation in a real-time system and also the parallelization efficiency of the estimator tested over multicore computing system to verify the speeding up of channel estimation process and then compare it to the less complex semi-blind estimators.
- The performance of iterative linear RNA based precoder should be investigated for the multicell scenario with lossy links.

Bibliography

[1] I. Khan, "A Robust Signal Detection Scheme for 5G Massive," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9597–9604, 2018, doi: 10.1109/TVT.2018.2858922.

[2] A. Abboud, "Interference Mitigation in 5G Mobile Networks," University of Limoges, Limoges, 2017.

[3] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge: Cambridge University Press, 2016.

[4] H. Arshad, A. Habib, A. Mustafa, B. Khan, and S. Zahid, "Pairing and Scheduling for Large Array MIMO Using Regularized Channel Inversion Receivers Over Nakagami-m Fading," *Wirel. Pers. Commun.*, vol. 96, no. 4, pp. 6397–6424, 2017, doi: 10.1007/s11277-017-4484-9.

[5] F. A. P. De Figueiredo, F. A. C. M. Cardoso, I. Moerman, and G. Fraidenraich, "Channel Estimation for Massive MIMO TDD Systems Assuming Pilot Contamination and Frequency Selective Fading," *IEEE Access*, vol. 5, no. September, pp. 17733–17741, 2017, doi: 10.1109/ACCESS.2017.2749602.

[6] O. A. Saraereh, I. Khan, B. M. Lee, and A. Tahat, "Efficient pilot decontamination schemes in 5G massive MIMO systems," *Electron.*, vol. 8, no. 1, 2019, doi: 10.3390/electronics8010055.

[7] S. Ali, Z. Chen, and F. Yin, "Design of orthogonal uplink pilot sequences for TDD massive MIMO under pilot contamination," *J. Commun.*, vol. 12, no. 1, pp. 40–48, 2017, doi: 10.12720/jcm.12.1.40-48.

[8] E. Björnson, "Pilot Contamination in a Nutshell," *Massive MIMO*, 2017. https://ma-mimo.ellintech.se/2017/01/14/pilot-contamination-in-a-nutshell/ (accessed Aug. 18, 2019).

[9] A. S. Al-Hubaishi, N. K. Noordin, A. Sali, S. Subramaniam, and A. M. Mansoor, "An efficient pilot assignment scheme for addressing pilot contamination in multicell massive mimo systems," *Electron.*, vol. 8, no. 4, 2019, doi: 10.3390/electronics8040372.

[10] A. Salh, L. Audah, N. S. M. Shah, and S. A. Hamzah, "Analyze of pilot reuse with achievable sum rate for massive MIMO cellular uplink," *ARPN J. Eng. Appl. Sci.*, vol.

12, no. 2, pp. 570–578, 2017.

[11] A. S. Al-Hubaishi, N. K. Noordin, A. Sali, S. Subramaniam, and A. M. Mansoor, “An efficient pilot assignment scheme for addressing pilot contamination in multicell massive mimo systems,” *Electron.*, vol. 8, no. 4, pp. 1–19, 2019, doi: 10.3390/electronics8040372.

[12] A. M. Girgis, B. Abdelhamid, and S. Elramly, “Improved Pilot Sequences Allocation inMassive MIMO Systems,” 2017. https://bscw.5g-mmmagic.eu/pub/bscw.cgi/187743 (accessed Feb. 21, 2020).

[13] J. Mietzner and R. Schober, “Multiple-antenna techniques for wireless communications-a comprehensive literature survey,” *Commun. ...*, vol. 11, no. 2, pp. 87–105, 2009, doi: 10.1109/SURV.2009.090207.

[14] M. AlKhaled and E. Alsusa, “Impact of pilot sequence contamination in massive MIMO systems,” *IET Commun.*, vol. 11, no. 13, pp. 2005–2011, 2017, doi: 10.1049/iet-com.2017.0062.

[15] Cisco, “Cisco Visual Networking Index: Global MobileDataTraffic Forecast Update, 2015–2020,” CIsco Systems, San Jose, CA, 2016. Accessed: Oct. 29, 2016. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf.

[16] E. Björnson, “Radio Resource Management in Massive MIMO Communication Systems,” 2016. http://www.commsys.isy.liu.se/en/research/projects/CENIIT-Radio-Resource-Management (accessed Oct. 29, 2016).

[17] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” *IEEE Trans. Wirel. Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010, doi: 10.1109/TWC.2010.092810.091092.

[18] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. L. Marzetta, and F. Tufvesson, “Scaling up MIMO : Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013, doi: 10.1109/MSP.2011.2178495.

[19] C. T. Neil, M. Shafi, P. J. Smith, and P. A. Dmochowski, “On the impact of antenna topologies for massive MIMO systems,” in *IEEE International Conference on Communications*, 2015, vol. 2, pp. 2030–2035, doi: 10.1109/ICC.2015.7248624.

[20] X. Artiga, J. Perruisseau-Carrier, and A. I. Perez-Neira, “Antenna array configurations for massive MIMO outdoor base stations,” in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2014, no. 4, pp. 281–284, doi:

10.1109/SAM.2014.6882396.

[21] A. Muller, A. Kammoun, E. Bjornson, and M. Debbah, “Efficient linear precoding for massive MIMO systems using truncated polynomial expansion,” in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2014, no. 2, pp. 273–276, doi: 10.1109/SAM.2014.6882394.

[22] A. L. Moustakas, H. U. Baranger, L. Balents, A. M. Sengupta, and S. H. Simon, “Communication Through a Diffusive Medium: Coherence and Capacity,” *Science (80-.).*, vol. 287, no. January, pp. 287–290, 2000, doi: 10.1126/science.287.5451.287.

[23] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, 2013, doi: 10.1109/TCOMM.2013.020413.110848.

[24] P. I. Tebe, Y. Kuang, A. U. E. Ampoma, and K. A. Opare, “Mitigating pilot contamination in massive MIMO using cell size reduction,” *IEICE Trans. Commun.*, vol. E101B, no. 5, pp. 1280–1290, 2018, doi: 10.1587/transcom.2017EBP3189.

[25] H. Quoc Ngo and E. G. Larsson, “EVD-based Channel Estimations for Multicell Multiuser MIMO with Very Large Antenna Arrays,” in *IEEE International Conference on Acoustics, Speed and Signal Processing*, 2012, pp. 3249–3252, Accessed: Oct. 29, 2016. [Online]. Available: http://liu.diva-portal.org/smash/record.jsf?pid=diva2:480042.

[26] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, “Blind pilot decontamination,” *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014, doi: 10.1109/JSTSP.2014.2310053.

[27] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, “Multiple-antenna channel hardening and its implications for rate feedback and scheduling,” *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, 2004, doi: 10.1109/TIT.2004.833345.

[28] H. Q. Ngo and E. G. Larsson, “No Downlink Pilots Are Needed in TDD Massive MIMO,” *IEEE Trans. Wirel. Commun.*, vol. 16, no. 5, pp. 2921–2935, 2017, doi: 10.1109/TWC.2017.2672540.

[29] Z. Chen and E. Bjoernson, “Can We Rely on Channel Hardening in Cell-Free Massive MIMO?,” in *2017 IEEE Globecom Workshops, GC Wkshps 2017 - Proceedings*, 2018, vol. 2018-Janua, pp. 1–6, doi: 10.1109/GLOCOMW.2017.8269162.

[30] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, “A Comprehensive Survey of Pilot Contamination in Massive MIMO-5G System,” *IEEE Commun. Surv.*

Tutorials, vol. 18, no. 2, pp. 905–923, 2016, doi: 10.1109/COMST.2015.2504379.

[31] N. Mehrotra and A. K. Chaubey, "Pilot Contamination Effect in Massive MIMO and Analysis of Mitigation Techniques," *Int. J. New Technol. Res.*, vol. 3, no. 2, pp. 19–23, 2017.

[32] K. Zheng, S. Ou, and X. Yin, "Massive MIMO Channel Models : A Survey," *Int. J. Antennas Propag.*, vol. 2014, pp. 1–8, 2014.

[33] J. Li and Y. Zhao, "Channel Characterization and Modeling for Large- Scale Antenna Systems," in *International Symposium on Communications and Information Technologies (ISCIT)*, 2014, no. 1, pp. 559–563.

[34] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, 2014, doi: 10.1109/TIT.2014.2354403.

[35] E. Bjornson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *GLOBECOM - IEEE Global Telecommunications Conference*, 2012, pp. 4775–4780, doi: 10.1109/GLOCOM.2012.6503874.

[36] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in Massive MIMO systems subject to phase noise," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 2, pp. 711–723, 2014, doi: 10.1109/TWC.2014.2359018.

[37] F. Kaltenberger, J. Haiyong, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *in Proc. Future Network and Mobile Summit*, 2010, pp. 1–10.

[38] E. Björnson, "Optimal Resource Allocation in Coordinated Multi-Cell Systems," *Found. Trends® Commun. Inf. Theory*, vol. 9, no. 2–3, pp. 113–381, 2013, doi: 10.1561/0100000069.

[39] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, 2000, doi: 10.1109/26.837052.

[40] R. Janaswamy, "Effect of element mutual coupling on the capacity of fixed length linear arrays," *IEEE Antennas Wirel. Propag. Lett.*, vol. 1, no. 1, pp. 157–160, 2002, doi: 10.1109/LAWP.2002.807570.

[41] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for Next

Generation Wireless Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014, doi: 10.1109/MCOM.2014.6736761.

[42] A. Puglielli *et al.*, "A scalable massive MIMO array architecture based on common modules," in *IEEE International Conference on Communication Workshop*, 2015, pp. 1310–1315, doi: 10.1109/ICCW.2015.7247359.

[43] X. Ge, R. Zi, H. Wang, J. Zhang, and M. Jo, "Multi-User Massive MIMO Communication Systems Based on Irregular Antenna Arrays," *IEEE Trans. Wirel. Commun.*, vol. 15, no. 8, pp. 5287–5301, 2016, doi: 10.1109/TWC.2016.2555911.

[44] K. Technologies, "Examining the Challenges in Implementing and Testing Massive MIMO for 5G," 2016. http://literature.cdn.keysight.com/litweb/pdf/5992-1448EN.pdf (accessed Oct. 30, 2016).

[45] X. Li, E. Bjornson, S. Zhou, and J. Wang, "Massive MIMO with multi-antenna users: When are additional user antennas beneficial?," in *International Conference on Telecommunications*, 2016, vol. 1, no. 1, pp. 1–6, doi: 10.1109/ICT.2016.7500452.

[46] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *European Conference on Antennas and Propagation*, 2012, pp. 433–437, doi: 10.1109/EuCAP.2012.6206345.

[47] X. Gao, F. Tufvesson, and O. Edfors, "Massive MIMO channels - Measurements and models," in *IEEE-Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 280–284, doi: 10.1109/ACSSC.2013.6810277.

[48] D. Fei, R. He, B. Ai, B. Zhang, K. Guan, and Z. Zhong, "Massive MIMO Channel Measurements and Analysis at 3.33 GHz," in *International Conference on Communications and Networking in China (ChinaCom)*, 2015, pp. 194–198.

[49] T. Sidabras and M. Salman, "Analysis of Channel Measurements Using a Very Large Antenna Array," Lund University, 2015.

[50] À. O. Martínez, E. De Carvalho, and J. Ø. Nielsen, "Towards Very Large Aperture Massive MIMO : a measurement based study," in *Globecom Workshops*, 2014, pp. 281–286.

[51] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel measurements for large antenna arrays," in *International Symposium on Wireless Communication Systems (ISWCS)*, 2012, pp. 811–815, doi: 10.1109/ISWCS.2012.6328480.

[52] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance

evaluation based on measured propagation data," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 7, pp. 3899–3911, 2015, [Online]. Available: http://arxiv.org/abs/1403.3376.

[53] C.-X. Wang, S. Wu, L. Bai, X. You, J. Wang, and C.-L. I, "Recent advances and future challenges for massive MIMO channel measurements and models," *Sci. China Inf. Sci.*, vol. 59, no. 2, pp. 1–16, 2016, doi: 10.1007/s11432-015-5517-1.

[54] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, 2013, doi: 10.1109/TCOMM.2013.012913.110827.

[55] J. Zhang, X. Yuan, and L. Ping, "Hermitian precoding for distributed MIMO systems with individual channel state information," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 241–250, 2013, doi: 10.1109/JSAC.2013.130212.

[56] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*, vol. 1. New Jersey: now Publishers Inc., 2004.

[57] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications Blending*, vol. 1. Cambridge: Cambridge University Press, 2011.

[58] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot Beam Pattern Design for Channel Estimation in Massive MIMO Systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 787–801, 2014.

[59] R. Couillet, M. Debbah, and J. W. Silverstein, "A deterministic equivalent for the analysis of non-Gaussian correlated MIMO multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3493–3514, 2011, doi: 10.1109/TIT.2012.2218571.

[60] N. Costa and S. Haykin, *Multiple-input multiple-output channel models: theory and practice*. New York: John Wiley & Sons Inc., 2010.

[61] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wirel. Commun.*, vol. 5, no. 1, pp. 90–99, 2006, doi: 10.1109/TWC.2006.1576533.

[62] X. Gao, F. Tufvesson, O. Edfors, and F. Rusek, "Measured propagation characteristics for very-large MIMO at 2.6 GHz," in *Asilomar Conference on Signals, Systems and Computers*, 2012, pp. 295–299, doi: 10.1109/ACSSC.2012.6489010.

[63] S. Wu, C. X. Wang, E. H. M. Aggoune, and M. M. Alwakeel, "A novel Kronecker-based stochastic model for massive MIMO channels," in *IEEE International Conference on Communications in China*, 2015, pp. 1–6, doi: 10.1109/ICCChina.2015.7448642.

[64] C. K. Wen, S. Jin, and K. K. Wong, "On the sum-rate of multiuser MIMO uplink channels with jointly-correlated Rician fading," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2883–2895, 2011, doi: 10.1109/TCOMM.2011.081111.100585.

[65] W. Weichselberger, "Spatial Structure of Multiple Antenna Radio Channels: A signal Processing View Point," Technische Universität Wien, 2003.

[66] E. Bonek Ove, "Experimental validation of analytical MIMO channel models," *e i Elektrotechnik und Informationstechnik*, vol. 122, no. 6, pp. 196–205, 2005.

[67] S. Wu, "Massive MIMO Channel Modelling for 5G Wireless Communication Systems," Heriot-Watt University, 2015.

[68] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO Wireless Channels: Capacity, Optimal Signaling, and Asymptotics," *IEEE Trans. Inf. THEORY*, vol. 51, no. 6, pp. 2058–2072, 2005.

[69] A. M. Sayeed, "Deconstructing Multiantenna Fading Channels," *IEEE Trans. SIGNAL Process.*, vol. 50, no. 10, pp. 2563–2579, 2002, doi: 10.1109/TSP.2002.803324.

[70] H. Ozcelik, N. Czink, and E. Bonek, "What Makes a Good MIMO Channel Model?," in *Proc. VTC'05*, 2005, pp. 156–160.

[71] X. Yin and X. Cheng, "Geometry-based Stochastic Channel Modeling," in *Propagation Channel Characterization, Parameter Estimation, and Modeling for Wireless Communications*, IEEE, 2016, pp. 77–105.

[72] X. Gao, "Massive MIMO in real propagation environments," Lund University, 2016.

[73] S. Wu, C. X. Wang, H. Haas, E. H. M. Aggoune, M. M. Alwakeel, and B. Ai, "A Non-Stationary Wideband Channel Model for Massive MIMO Communication Systems," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 3, pp. 1434–1446, 2015, doi: 10.1109/TWC.2014.2366153.

[74] C. F. Lopez, C.-X. Wang, and R. Feng, "A novel 2D non-stationary wideband massive MIMO channel model," in *IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, 2016, pp. 1–6, [Online]. Available: http://arxiv.org/abs/1611.00617.

[75] A. L. Swindlehurst and T. Kailath, "Passive direction-of-arrival and range estimation for near-field sources," in *Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, 1988, pp. 123–128, doi: 10.1109/SPECT.1988.206176.

[76] A. Kuchar, J. P. Rossi, and E. Bonek, "Directional macro-cell channel characterization

from urban measurements," *IEEE Trans. Antennas Propag.*, vol. 48, no. 2, pp. 137–146, 2000, doi: 10.1109/8.833062.

[77] L. Liu *et al.*, "The COST 2100 MIMO channel model," *IEEE Wirel. Commun.*, vol. 19, no. 6, pp. 92–99, 2012, doi: 10.1109/MWC.2012.6393523.

[78] X. Gao, J. Flordelis, G. Dahman, F. Tufvesson, and O. Edfors, "Massive MIMO Channel Modeling-Extension of the COST 2100 Model," in *Joint NEWCOM/COST Workshop on Wireless Communications*, 2015, pp. 1–8.

[79] S. Wu, C. X. Wang, E. H. M. Aggoune, M. M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, 2014, doi: 10.1109/JSAC.2014.2328131.

[80] V. Nurmela *et al.*, "METIS Channel Models," Turin, ICT-317669-METIS/D1.4, 2015.

[81] S. Wu, C. Wang, Y. Yang, W. Wang, X. Gao, and R. Samsung, "Performance Comparison of Massive MIMO Channel Models," in *IEEE/CIC International Conference on Communications in China (ICCC)*, 2016, pp. 1–6.

[82] P. Kyösti *et al.*, "WINNER II Channel Models," Munich, IST-4-027756 WINNER II D1.1.2 V1.2, 2007. Accessed: Nov. 18, 2016. [Online]. Available: http://www.cept.org/files/1050/documents/winner2 - final report.pdf.

[83] A. A. M. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987, doi: 10.1109/JSAC.1987.1146527.

[84] L. Bai, C. X. Wang, S. Wu, H. Wang, and Y. Yang, "A 3-D wideband multi-confocal ellipsoid model for wireless MIMO communication channels," in *IEEE International Conference on Communications,* 2016, pp. 1–6, doi: 10.1109/ICC.2016.7510852.

[85] E. G. Larsson and E. Björnson, "Physical layer security in massive MIMO - Linköping University," 2019. https://liu.se/en/research/physical-layer-security-in-massive-mimo (accessed Nov. 01, 2019).

[86] Y. Liang, H. V. Poor, and S. Shamai, "Information Theoretic Security," *Found. Trends Commun. Inf. Theory*, vol. 5, no. 4–5, pp. 355–580, 2009.

[87] R. Scober, "Massive MIMO Systems," 2016. http://www.idc.lnt.de/en/forschung/massive-mimo-systems/ (accessed Aug. 22, 2016).

[88] H. Wen, *Physical layer approaches for securing wireless communication systems*, no. 9781461465096. New York: Springer, 2013.

[89] S. Goel and R. Negi, "Guaranteeing Secrecy Using Artificial Noise," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 6, pp. 2180–2189, 2008.

[90] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. B. Collings, "Secrecy Sum-Rates for Multi-User MIMO Regularized Channel Inversion Precoding," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3472–3482, 2012.

[91] D. Kapetanović, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 21–27, 2015, doi: 10.1109/MCOM.2015.7120012.

[92] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. K. Wong, and X. Gao, "A Survey of Physical Layer Security Techniques for 5G Wireless Networks and Challenges Ahead," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 679–695, 2018, doi: 10.1109/JSAC.2018.2825560.

[93] Y. O. Basciftci, C. E. Koksal, and A. Ashikhmin, "Physical-Layer Security in TDD Massive MIMO," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7359–7380, 2018, doi: 10.1109/TIT.2018.2855058.

[94] R. F. Schaefer, G. Amarasuriya, and H. V. Poor, "Physical layer security in massive MIMO systems," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, 2017.

[95] X. Zhou and M. R. McKay, "Secure Transmission With Artificial Noise Over Fading Channels: Achievable Rate and Optimal Power Allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, 2010.

[96] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in mimo wiretap channels with imperfect csi," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, 2011, doi: 10.1109/TSP.2010.2078810.

[97] Q. Li and W. K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, 2011, doi: 10.1109/TSP.2011.2146775.

[98] X. Zhou, B. Maham, and A. Hjørungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 3, pp. 903–907, 2012, doi: 10.1109/TWC.2012.020712.111298.

[99] D. Tse and P. Vishwanath, *Fundamentals of Wireless Communications 12*. Cambridge: Cambridge University, 2005.

[100] H. Jafarkhani, *Space-Time Coding*. California: Cambridge University, 2005.

[101] M. Renzo, H. Haas, and P. Grant, "Spatial modulation for multiple-antenna wireless systems: a survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, 2011, doi: 10.1109/MCOM.2011.6094024.

[102] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, 2006, doi: 10.1109/TIT.2006.880064.

[103] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003, doi: 10.1109/TIT.2003.817421.

[104] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, 2003, doi: 10.1109/TIT.2003.814483.

[105] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Massive MU-MIMO Downlink TDD Systems with Linear Precoding and Downlink Pilots," in *Fifty-first Annual Allerton Conference, Control, and Computing*, 2013, vol. 51, pp. 293–298.

[106] X. Li, E. Björnson, E. G. Larsson, S. Zhou, and J. Wang, "A Multi-cell MMSE Detector for Massive MIMO Systems and New Large System Analysis," in *2015 IEEE Global Communications Conference*, 2015, no. SEPTEMBER, pp. 1–6.

[107] A. Müller, A. Kammoun, E. Björnson, and M. Debbah, "Linear Precoding Based on Polynomial Expansion : Reducing Complexity in Massive MIMO," *EURASIP J. Wirel. Commun. Netw.*, vol. 2016, no. 63, pp. 1–18, 2016, doi: 10.1109/JSTSP.2014.2322582.

[108] V. P. Selvan, M. S. Iqbal, and H. S. Al-Raweshidy, "Performance Analysis of Linear Precoding Schemes for Very Large Multi-user MIMO Downlink System," in *International Conference on Innovative Computing Technology (INTECH),* 2014, pp. 219–224.

[109] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 861–875, 2014, doi: 10.1109/JSTSP.2014.2322582.

[110] P. Patcharamaneepakorn, S. Armour, and A. Doufexi, "On the equivalence between SLNR and MMSE precoding schemes with single-antenna receivers," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1034–1037, 2012, doi: 10.1109/LCOMM.2012.050912.120329.

[111] M. Joham, W. Utschick, and J. a. Nossek, “Linear transmit processing in MIMO communications systems,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, 2005, doi: 10.1109/TSP.2005.850331.

[112] P. Liu, M. Q. Wu, C. X. Xu, and F. Zheng, “Multi-user MIMO linear precoding schemes in OFDM systems,” in *Proceedings - 2010 3rd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2010*, 2010, vol. 9, pp. 528–531, doi: 10.1109/ICCSIT.2010.5564646.

[113] T. Parfait, Y. Kuang, and K. Jerry, “Performance analysis and comparison of ZF and MRT based downlink massive MIMO systems,” in *International Conference on Ubiquitous and Future Networks (ICUFN)*, 2014, pp. 383–388, doi: 10.1109/ICUFN.2014.6876818.

[114] D. Gesbert, M. Kountouris, R. W. Heath, C. B. Chae, and T. Sälzer, “Shifting the MIMO Paradigm,” *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, 2007, doi: 10.1109/MSP.2007.904815.

[115] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, “MIMO precoding for decentralized receivers,” in *IEEE International Symposium on Information Theory*, 2002, p. 496, doi: 10.1109/isit.2002.1023768.

[116] R. F. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, “Space-time transmission using Tomlinson-Harashima precoding,” in *ITG Conference on Source and Channel coding*, 2002, pp. 139–147, doi: 10.1049/iet-opt.2014.0152.

[117] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication - Part II: Perturbation,” *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, 2005, doi: 10.1109/TCOMM.2004.841997.

[118] C. B. Peel, B. M. Hochwald, and a L. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization,” *Technology*, vol. 53, no. 1, pp. 1–22, 2005, doi: 10.1109/TCOMM.2004.840638.

[119] J. Stillwell, *Elements of number theory*. New York: Springer-Verlag, 2003.

[120] C. Masouros, M. Sellathurai, and T. Ratnarajah, “Computationally efficient vector perturbation precoding using thresholded optimization,” *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1880–1890, 2013, doi: 10.1109/TCOMM.2013.022713.120632.

[121] J. Maurer, J. Jaldén, D. Seethaler, and G. Matz, “Vector perturbation precoding for

receivers with limited dynamic range," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, no. 2, pp. 2709–2712, doi: 10.1109/ICASSP.2009.4960182.

[122] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wirel. Commun. Lett.*, vol. 3, no. 6, pp. 653–656, 2014, doi: 10.1109/LWC.2014.2363831.

[123] L. Kong, S. Han, and C. Yang, "Wideband hybrid precoder for massive MIMO systems," in *IEEE Global Conference on Signal and Information Processing*, 2016, pp. 305–309, doi: 10.1109/GlobalSIP.2015.7418206.

[124] S. Zarei, W. Gerstacker, and R. Schober, "Low-Complexity Hybrid Linear / Tomlinson-Harashima Precoding for Downlink Large-scale Mu-mimo Systems," in *IEEE Globecom Workshops*, 2016, pp. 1–7.

[125] Y. Xu, W. Zou, and L. Du, "A fast and low-complexity matrix inversion scheme based on CSM method for massive MIMO systems," *EURASIP J. Wirel. Commun. Netw.*, vol. 2016, no. 1, p. 251, 2016, doi: 10.1186/s13638-016-0749-3.

[126] A. A. Rontogiannis, V. Kekatos, and K. Berberidis, "A square-root adaptive V-BLAST algorithm for fast time-varying MIMO channels," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 3135–3139, 2006, doi: 10.1109/ICC.2006.255287.

[127] B. Y. J. Sherman and W. J. Morrison, "Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix," *Ann. Math. Stat.*, vol. 21, no. 1, pp. 124–127, 1950.

[128] A. Müller, A. Kammoun, E. Björnson, and M. Debbah, "Linear Precoding Based on Polynomial Expansion: Reducing Complexity in Massive MIMO," no. October, 2013, doi: 10.1186/s13638-016-0546-z.

[129] A. Müller, A. Kammoun, E. Björnson, and M. Debbah, "Linear Precoding Based on Truncated Polynomial Expansion - Part I: Large-Scale Single-Cell Systems," *arXiv: 1310.1806v2 [Cs.IT]* , Oct. 2013, Accessed: Oct. 12, 2017. [Online]. Available: http://arxiv.org/abs/1310.1806.

[130] J. G. Andrews *et al.*, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014, doi: 10.1109/JSAC.2014.2328098.

[131] J. Jose, A. Ashikhminl, T. L. Marzettal, and S. Vishwanath, "Pilot Contamination Problem in Multi-Cell TDD Systems," in *IEEE International Symposium Information Theory*, 2009, pp. 2184–2188.

[132] Q. Cheng, “Novel Channel Estimation Methods under Pilot Contamination in Massive MIMO,” Macquarie University, 2016.

[133] D. Neumann, M. Joham, and W. Utschick, “Channel Estimation in Massive MIMO Systems,” 2015. http://arxiv.org/abs/1503.08691 (accessed Nov. 23, 2016).

[134] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, “Training Signal Design for Correlated Massive MIMO Channel Estimation,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 6499–6503.

[135] K. Li, X. Song, M. Omair Ahmad, and M. N. S. Swamy, “An improved multicell MMSE channel estimation in a massive MIMO system,” *Int. J. Antennas Propag.*, vol. 2014, no. 2014, pp. 1–9, 2014, doi: 10.1155/2014/387436.

[136] A. Zaib, M. Masood, A. Ali, W. Xu, and T. Y. Al-Naffouri, “Distributed Channel Estimation and Pilot Contamination Analysis for Massive MIMO-OFDM Systems,” *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4607–4621, 2015, doi: 10.1109/TCOMM.2016.2593924.

[137] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, “Low-complexity channel estimation in large-scale MIMO using polynomial expansion,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, 2013, pp. 1157–1162, doi: 10.1109/PIMRC.2013.6666313.

[138] D. kong, D. Qu, K. Luo, and T. Jiang, “Channel Estimation under Staggered Frame Structure for Uplink Massive MIMO System,” *IEEE Trans. Wirel. Commun.*, vol. 15, no. 2, pp. 1469–1479, 2016, doi: 10.1109/TWC.2015.2490666.

[139] C. K. Wen, S. Jin, K. K. Wong, J. C. Chen, and P. Ting, “Channel Estimation for Massive MIMO Using Gaussian-Mixture Bayesian Learning,” *IEEE Trans. Wirel. Commun.*, vol. 14, no. 3, pp. 1356–1368, 2015, doi: 10.1109/TWC.2014.2365813.

[140] H. Al-Salihi and M. R. Nakhai, “Bayesian Compressed Sensing-based Channel Estimation for Massive MIMO Systems,” in *European Conference on Networks and Communications (EUCNC 2016)*, 2016.

[141] D. C. Araújo, A. L. F. De Almeida, J. Axnäs, and J. C. M. Mota, “Channel estimation for millimeter-wave Very-Large MIMO systems,” in *European Signal Processing Conference (EUSIPCO)*, 2014, no. 1, pp. 81–85.

[142] C. Qi, Y. Huang, S. Jin, and L. Wu, “Sparse Channel Estimation Based on Compressed Sensing for Massive MIMO Systems,” in *IEEE ICC 2015-Signal Processing for Communications Symposium*, 2015, pp. 4558–4563.

[143] Y. Nan, X. Sun, and L. Zhang, "Joint channel estimation algorithm via weighted Homotopy for massive MIMO OFDM system," *Digit. Signal Process.*, vol. 50, no. 2016, pp. 34–42, 2016, doi: 10.1016/j.dsp.2015.11.010.

[144] J. Jinga and W. Ni, "A Low Complexity Channel Estimation Algorithm for Massive MIMO System," *Int. J. Grid Distrib. Comput.*, vol. 7, no. 4, pp. 81–92, 2014.

[145] H. Wang, W. Du, and L. Xu, "A New Sparse Adaptive Channel Estimation Method Based on Compressive Sensing for FBMC/OQAM Transmission Network," *Sensors*, vol. 16, no. 7, p. 966, 2016, doi: 10.3390/s16070966.

[146] M. Guo, G. Xie, J. Gao, and Y. Liu, "Enhanced EVD based channel estimation and pilot decontamination for Massive MIMO networks," *J. China Univ. Posts Telecommun.*, vol. 22, no. 6, pp. 72–77, 2015, doi: 10.1016/S1005-8885(15)60697-5.

[147] A. L. Sarker, S. M. Ieee, M. H. Lee, and S. M. Ieee, "A fast channel estimation and the reduction of pilot contamination problem for massive MIMO based on a diagonal Jacket matrix," in *International Workshop on Fiber Optics in Access Network (FOAN)*, 2013, pp. 26–30.

[148] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Blind Pilot Decontamination," in *17th International ITG Workshop on Smart Antennas (WSA)*, 2013, pp. 1–6.

[149] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Analysis of blind pilot decontamination," *Asilomar Conference on Signals, Systems and Computers*. IEEE, California, pp. 1016–1020, 2013, doi: 10.1109/ACSSC.2013.6810444.

[150] L. Cottatellucci, R. R. Müller, and M. Vehkaperä, "Analysis of pilot decontamination based on power control," in *IEEE Vehicular Technology Conference*, 2013, pp. 1550–2252, doi: 10.1109/VTCSpring.2013.6691891.

[151] D. Neumann, A. Gr, M. Joham, and W. Utschick, "Pilot Coordination for Large-Scale Multi-Cell TDD Systems," in *International ITG Workshop on Smart Antennas (WSA),* 2014, pp. 1–6, [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6776879&sortType%3Dasc_p_Sequence%26filter%3DAND(p_IS_Number%3A6776878).

[152] A. Hu, T. Lv, and Y. Lu, "Subspace-based semi-blind channel estimation for large-scale multi-cell multiuser MIMO systems," in *IEEE Vehicular Technology Conference*, 2013, pp. 1–6, doi: 10.1109/VTCSpring.2013.6692772.

[153] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, 1995, [Online]. Available:

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=365290.

[154] F. Xu, Y. Xiao, and D. Wang, “Adaptive semi-blind channel estimation for massive mimo systems,” in *International Conference on Signal Processing*, 2014, pp. 1698–1702.

[155] D. Hu, L. He, and X. Wang, “Semi-Blind Pilot Decontamination for Massive MIMO Systems,” *IEEE Trans. Wirel. Commun.*, vol. 15, no. 1, pp. 525–536, 2016, doi: 10.1109/TWC.2015.2475745.

[156] G. Sangirov, Y. Fu, M. Rakib Uddin, and J. Sangirov, “An enhanced semi-blind channel estimation for MIMO-OFDM systems,” *Wirel. Networks*, vol. 22, no. 6, pp. 2101–2110, 2016, doi: 10.1007/s11276-016-1201-7.

[157] T. Lv, S. Yang, and H. Gao, “Semi-Blind Channel Estimation Relying on Optimum Pilots Designed for Multi-Cell Large-Scale MIMO Systems,” *IEEE Access*, vol. 4, pp. 1190–1204, 2016, doi: 10.1109/ACCESS.2016.2543300.

[158] H. Al-salihi and M. R. Nakhai, “An Enhanced Whitening Rotation Semi-Blind Channel,” in *International Conference on Telecommunications (ICT)*, 2016, pp. 1–6.

[159] T. Salman, “Cloud RAN : Basics , Advances and Challenges,” 2016. .

[160] A. Sibley, “Cloud Radio Access Network (C-RAN) Market to be worth US$1,870.8 mn by 2025,” 2020. https://www.hashtap.com/@anna.sibley/cloud-radio-access-network-c-ran-market-to-be-worth-us-1-870-8-mn-by-2025-bQYMA1bL4gJm (accessed Feb. 20, 2020).

[161] S. Parsaeefard, V. Jumba, A. Dalili, M. Derakhshani, and T. Le-Ngoc, “User Association in Cloud RANs with Massive MIMO,” *IEEE Trans. Cloud Comput.*, vol. 2018, no. August, pp. 0–13, 2018, doi: 10.1109/TCC.2018.2867224.

[162] Crcw. News, “What is C-RAN?,” 2020. https://www.rcrwireless.com/20151222/featured/what-is-c-ran-tag4 (accessed Feb. 18, 2020).

[163] M. Peng, C. Wang, V. Lau, and H. V. Poor, “Fronthaul-constrained cloud radio access networks: Insights and challenges,” *IEEE Wirel. Commun.*, vol. 22, no. 2, pp. 152–160, 2015, doi: 10.1109/MWC.2015.7096298.

[164] J. Francis and G. Fettweis, “Power Allocation for Massive MIMO-based , Fronthaul-constrained Cloud RAN Systems,” *eprint arXiv:1810.04529*, vol. arXiv:1810, pp. 1–10, 2018.

[165] J. Park, D. M. Kim, E. De Carvalho, and C. N. Manch, “Hybrid Precoding for Massive MIMO Systems in Cloud RAN Architecture with Capacity-Limited Fronthauls,” *arXiv.org*, vol. arXiv:1709, pp. 1–13, 2017.

[166] J. Francis and G. Fettweis, “Energy Efficiency Maximization in Massive MIMO-aided, Fronthaul-constrained C-RAN,” *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2019-Septe, pp. 1–6, 2019, doi: 10.1109/PIMRC.2019.8904159.

[167] S. Park, H. Lee, C. B. Chae, and S. Bahk, “Massive MIMO operation in partially centralized cloud radio access networks,” *Comput. Networks*, vol. 115, no. 2017, pp. 54–64, 2017, doi: 10.1016/j.comnet.2017.01.013.

[168] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, “Recent Trend of Multiuser MIMO in LTE-Advanced,” *IEEE Commun. Mag.*, vol. 51, no. March, pp. 127–135, 2013, [Online]. Available: http://dblp.uni-trier.de/db/journals/cm/cm51.html#LimYCLS13.

[169] P. Piantanida and P. Duhamel, “Dirty-paper coding without channel information at the transmitter and imperfect estimation at the receiver,” in *IEEE International Conference on Communications*, 2007, pp. 5406–5411, doi: 10.1109/ICC.2007.895.

[170] J. Hoydis, S. Ten Brink, and M. Debbah, “Comparison of Linear Precoding Schemes for the Massive MIMO Downlink,” in *IEEE ICC 2012*, 2012, pp. 2135–2139.

[171] J. Hoydis, S. Ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013, doi: 10.1109/JSAC.2013.130205.

[172] A. Bjorck, *Numerical Methods in Matrix Computations*. Cham Heidelberg: Springer International Publishing Switzerland, 2015.

[173] D. Zhu, B. Li, and P. Liang, “On the matrix inversion approximation based on neumann series in massive MIMO systems,” *IEEE Int. Conf. Commun.*, vol. 2015-Septe, pp. 1763–1769, 2015, doi: 10.1109/ICC.2015.7248580.

[174] X. Gao, L. Dai, Y. Ma, and Z. Wang, “Low-complexity near-optimal signal detection for uplink large-scale MIMO systems,” *Electron. Lett.*, vol. 50, no. 18, pp. 1326–1328, 2014, doi: 10.1049/el.2014.0713.

[175] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, “Conjugate gradient-based soft-output detection and precoding in massive MIMO systems,” in *IEEE Global Communications Conference*, 2014, no. 1, pp. 3696–3701, doi: 10.1109/GLOCOM.2014.7037382.

[176] T. Xie, Q. Han, H. Xu, Z. Qi, and W. Shen, “A Low-Complexity Linear Precoding

Scheme Basedon SOR Method for Massive MIMO Systems," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5, doi: 10.1109/VTCSpring.2015.7145618.

[177] T. Xie, L. Dai, X. Gao, X. Dai, and Y. Zhao, "Low-Complexity SSOR-Based Precoding for Massive MIMO Systems," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 744–747, 2016, doi: 10.1109/LCOMM.2016.2525807.

[178] X. Gao, L. Dai, J. Zhang, S. Han, and I. Chih-Lin, "Capacity-approaching linear precoding with low-complexity for large-scale MIMO systems," in *IEEE International Conference on Communications*, 2015, vol. 2015-Septe, pp. 1577–1582, doi: 10.1109/ICC.2015.7248549.

[179] E. V. Krishnamurthy and S. K. Sen, *Numerical Algorithms: Computations In Science & Engineering*. New Delhi: Affiliated East-West Press, 2008.

[180] F. Soleymani, "A rapid numerical algorithm to compute matrix inversion," *Int. J. Math. Math. Sci.*, vol. 2012, pp. 1–12, 2012, doi: 10.1155/2012/134653.

[181] H. Prabhu, J. Rodrigues, O. Edfors, and F. Rusek, "Approximative matrix inverse computations for very-large MIMO and applications to linear pre-coding systems," in *IEEE Wireless Communications and Networking Conference, WCNC*, 2013, pp. 2710–2715, doi: 10.1109/WCNC.2013.6554990.

[182] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large System Analysis of Linear Precoding in MISO Broadcast Channels Under Limited Feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, 2012, doi: 10.1109/TIT.2012.2191700.

[183] C. Shepard *et al.*, "Argos: Practical Many-Antenna Base Stations," in *18th annual international conference on Mobile computing and networking*, 2012, pp. 53–64, Accessed: Oct. 10, 2017. [Online]. Available: http://argos.rice.edu/pubs/Shepard-MobiCom12.pdf.

[184] G. M. A. Sessler and F. K. Jondral, "Low complexity polynomial expansion multiuser detector for CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1379–1391, 2005, doi: 10.1109/TVT.2005.851322.

[185] H. B. Li, T. Z. Huang, Y. Zhang, X. P. Liu, and T. X. Gu, "Chebyshev-type methods and preconditioning techniques," *Appl. Math. Comput.*, vol. 218, no. 2, pp. 260–270, 2011, doi: 10.1016/j.amc.2011.05.036.

[186] A. Ben-Israel, "An iterative method for computing the generalized inverse of a matrix," *Math. Comput.*, vol. 19, no. 91, pp. 452–455, 1965, doi: 10.1007/s11075-014-9913-1.

[187] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*. New York: Dover Publications, INC., 1994.

[188] S. Amat, S. Busquier, and J. M. Gutiã Errez, "Geometric constructions of iterative functions to solve nonlinear equations," *J. Comput. Appl. Math.*, vol. 157, no. 1, pp. 197–205, 2003, doi: 10.1016/S0377-0427(03)00420-5.

[189] S. Zarei, W. Gerstacker, R. R. Müller, and R. Schober, "Low-complexity linear precoding for downlink large-scale MIMO systems," in *International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2013, pp. 1119–1124, Accessed: Oct. 10, 2017. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6666306.

[190] L. Annell and M. Törngren, *Nordic transputer applications : proceedings of the 1st and 2nd Nordic Transputer Seminars*. Oxford: IOS Press, 1992.

[191] U. Madhow, *Fundamentals of Digital Communication*. Cambridge: Cambridge University Press, 2008.

[192] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the Distribution of SINR for the MMSE MIMO Receiver and Performance Analysis," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 271–286, 2006.

[193] P. Suthisopapan, A. Meesomboon, K. Kasai, and V. Imtawil, "Ultra low complexity soft output detector for non-binary LDPC coded large MIMO systems," *International Symposium on Turbo Codes and Iterative Information Processing, ISTC*. Gothenburg, pp. 230–234, 2012, doi: 10.1109/ISTC.2012.6325233.

[194] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Netw.*, vol. 28, no. 6, pp. 6–11, 2014, doi: 10.1109/MNET.2014.6963798.

[195] E. Mukubwa, O. A. Sokoya, and D. S. Ilcev, "Comparison and Analysis of Massive MIMO Linear Precoding Schemes in the Downlink," in *IEEE Africon 2017*, 2017, pp. 187–191, [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8095479&tag=1.

[196] A. MÜLLER, "Random Matrix Analysis of Future Multi Cell MU-MIMO Networks," SUPÉLEC and École doctorale, 2005.

[197] H. Zhi and Y. Hu, "Novel Multi-cell Precoding Schemes for TDD Massive," *Wirel. Pers. Commun.*, vol. 97, no. 4, pp. 6111–6129, 2017, doi: 10.1007/s11277-017-4829-4.

[198] A. Kammoun, M. Axel, and E. Björnson, “Low-complexity linear precoding for multi-cell massive MIMO systems,” in *22nd European Signal Processing Conference (EUSIPCO)*, 2014, no. 1, pp. 2150–2154, [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6952770.

[199] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-Free Massive MIMO Versus Small Cells,” *IEEE Trans. Wirel. Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017, doi: 10.1109/TWC.2017.2655515.

[200] E. Bjornson, J. Hoydis, and L. Sanguinetti, “Massive MIMO Has Unlimited Capacity,” *IEEE Trans. Wirel. Commun.*, vol. 17, no. 1, pp. 574–590, 2018, doi: 10.1109/TWC.2017.2768423.

[201] J. Jose *et al.*, “Pilot Contamination and Precoding in Multi-Cell TDD Systems,” *IEEE Trans. Wirel. Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.

[202] M. Medard, “The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel,” *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, 2000.

[203] H. Tu, Y. Hu, Y. Xu, and F. Li, “An improved precoding of Approximative Matrix Inverse Computations based on norm minimization algorithm in massive MIMO system,” in *Proceedings of 2015 IEEE International Conference on Progress in Informatics and Computing, PIC 2015*, 2016, pp. 414–418, doi: 10.1109/PIC.2015.7489880.

[204] E. Mukubwa and O. Sokoya, “Efficient and low-complexity matrix inversion scheme for massive MIMO systems using rapid numerical algorithms,” *to IET J. Eng.*, 2018.

[205] V. Pan and J. Reif, “Efficient Parallel Solution of Linear Systems.,” *Conf. Proc. Annu. ACM Symp. Theory Comput.*, pp. 143–152, 1985, doi: 10.1145/22145.22161.

[206] V. Pan, “Complexity of Algorithms for Linear Systems of Equations,” *Comput. Algorithms Solving Linear Algebr. Equations*, pp. 27–56, 1991, doi: 10.1007/978-3-642-76717-3_2.

[207] C. Isheden, Z. Chong, E. Jorswieck, S. Member, and G. Fettweis, “Framework for Link-Level Energy Efficiency Optimization with Informed Transmitter,” *IEEE Trans. Wirel. Commun.*, vol. 11, no. 8, pp. 2946–2957, 2012.

[208] A. Zappone and E. Jorswieck, “Energy Efficiency in Wireless Networks via Fractional Programming Theory,” *Found. Trends Commun. Inf. Theory*, vol. 11, no. 3, pp. 185–399, 2015, doi: 10.1561/0100000088.

[209] G. A. et Al, “EARTH Project D2.3 - Energy efficiency analysis of the reference systems, areas of improvements and target breakdown,” *Energy Aware Radio Netw. Technol. (EARTH), INFSO-ICT-247733, ver. 2.0. [Online]. Available http//www.ict-earth.eu/*, 2012.

[210] V. K. Sachan, S. a. Imam, and M. T. Beg, “Energy-Efficiency of MIMO and Cooperative MIMO Techniques in Sensor Networks,” *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, 2004, doi: 10.1109/ICCCI.2012.6158868.

[211] R. V. R. Kumar and J. Gurugubelli, “How green the LTE technology can be?,” in *2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, Wireless VITAE 2011*, 2011, pp. 1–5, doi: 10.1109/WIRELESSVITAE.2011.5940909.

[212] A. Mezghani and J. A. Nossek, “Power efficiency in communication systems from a circuit perspective,” in *Proceedings - IEEE International Symposium on Circuits and Systems*, 2011, pp. 1896–1899, doi: 10.1109/ISCAS.2011.5937958.

[213] S. Tombaz, A. Västberg, and J. Zander, “Energy- and cost-efficient ultra-high-capacity wireless access,” *IEEE Wirel. Commun.*, vol. 18, no. 5, pp. 18–24, 2011, doi: 10.1109/MWC.2011.6056688.

[214] H. Yang and T. L. Marzetta, “Total energy efficiency of cellular large scale antenna system multiple access mobile networks,” in *2013 IEEE Online Conference on Green Communications, OnlineGreenComm 2013*, 2013, pp. 27–32, doi: 10.1109/OnlineGreenCom.2013.6731024.

[215] P. Patcharamaneepakorn *et al.*, “Spectral, Energy, and Economic Efficiency of 5G Multicell Massive MIMO Systems with Generalized Spatial Modulation,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9715–9731, 2016, doi: 10.1109/TVT.2016.2526628.

[216] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, vol. 1, no. 1. now Publishers Inc., 2017.

[217] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, “Millimeter wave vehicular communications: A survey,” *Found. Trends Netw.*, vol. 10, no. 1, pp. 1–116, 2016, doi: 10.1561/1300000054.

[218] J. Choi, V. Va, N. González-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, “Millimeter-Wave Vehicular Communication to Support Massive Automotive Sensing,” *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, 2016, doi: 10.1109/MCOM.2016.1600071CM.

[219] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving 'massive MIMO' spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 9, pp. 3226–3239, 2012, doi: 10.1109/TWC.2012.070912.111383.

[220] D. Neumann, M. Joham, and W. Utschick, "On MSE Based Receiver Design for Massive MIMO," 2017, pp. 1–6.

[221] D. Neumann, A. Grundinger, M. Joham, and W. Utschick, "Rate-balancing in massive MIMO using statistical precoding," *IEEE Work. Signal Process. Adv. Wirel. Commun. SPAWC*, vol. 2015-Augus, no. 1, pp. 226–230, 2015, doi: 10.1109/SPAWC.2015.7227033.

[222] S. Haghighatshoar and G. Caire, "Massive MIMO pilot decontamination and channel interpolation via wideband sparse channel estimation," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 12, pp. 8316–8332, 2017, doi: 10.1109/TWC.2017.2760825.

[223] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Pilot contamination is not a fundamental asymptotic limitation in massive MIMO," in *IEEE International Conference on Communications*, 2017, doi: 10.1109/ICC.2017.7996674.

[224] D. Neumann, M. Joham, and W. Utschick, "Covariance Matrix Estimation in Massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, 2018, doi: 10.1109/LSP.2018.2827323.

[225] K. Upadhya and S. A. Vorobyov, "Covariance Matrix Estimation for Massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 4, pp. 546–550, 2018, doi: 10.1109/LSP.2018.2827323.

[226] E. Bjornson *et al.*, "Massive MIMO with imperfect channel covariance information," in *50th Asilomar Conference on Signals, Systems and Computers*, 2016.

[227] X. Zhang, S. Wu, S. Yu, and J. Liu, "Spectral efficiency of massive MIMO networks with pilot contamination and channel covariance matrix estimation," *IET Commun.*, vol. 13, no. 1, pp. 59–65, 2018, doi: 10.1049/iet-com.2018.5032.

[228] F. Pourkamali-Anaraki, "Estimation of the sample covariance matrix from compressive measurements," *IET Signal Process.*, pp. 1089–1095, 2016, doi: 10.1049/iet-spr.2016.0169.

[229] I. Viering, H. Hofstetter, and W. Utschick, "Spatial Long-term Variantions in Urban, Rurla and Indoor Enviroments," in *in the 5th Meeting of COST273*, 2002.

[230] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ,USA: Prentice-Hall, Inc., 1993.

[231] 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, no. 2, pp. 264–273, 2013, doi: 10.1109/JSAC.2013.130214.

[232] D. K. Faddeev and V. N. Faddeva, *Computational Methods of Linear Algebra*. San Francisco: W.H. Freeman, 1963.

[233] W. W. . Hager, "Updating the Inverse of a Matrix," *Soc. Ind. Appl. Math. Stable*, vol. 31, no. 2, pp. 221–239, 1989.

[234] E. Mukubwa and O. A. Sokoya, "Efficient and low-complexity matrix inversion scheme for massive multiple-input multiple-output systems using rapid numerical algorithms," *J. Eng.*, vol. 2019, no. 10, pp. 7244–7249, 2019, doi: 10.1049/joe.2018.5015.

[235] W. Ford, *Numerical Linear Algebra with Applications*, 4th ed. London: Elsevier Inc., 2015.

[236] X. G. Doukopoulos, G. V Moustakides, and S. Member, "Fast and Stable Subspace Tracking," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1452–1465, 2008.

[237] S. H. Park, O. Simeone, O. Sahin, and S. Shamai, "Robust layered transmission and compression for distributed uplink reception in cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 204–216, 2014, doi: 10.1109/TVT.2013.2271642.

[238] X. Chen, M. R. Lyu, and I. King, "Toward Efficient and Accurate Covariance Matrix Estimation on Compressed Data," in *Proceedings of the 34th International Conference on Machine Learning*, 2017, vol. 70, pp. 767–776, [Online]. Available: http://proceedings.mlr.press/v70/chen17g.html.

[239] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivar. Anal.*, vol. 88, no. 2, pp. 365–411, 2004, doi: 10.1016/S0047-259X(03)00096-4.

[240] N. Shariati, E. Bjornson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 815–830, 2014, doi: 10.1109/JSTSP.2014.2316063.

[241] A. Gittens, "The spectral norm error of the naive Nystrom extension," *arXiv Prepr. arXiv1110.5305*, pp. 1–9, 2011, [Online]. Available: http://arxiv.org/abs/1110.5305.

[242] D. Achlioptas, Z. Karnin, and E. Liberty, “Near-Optimal Entrywise Sampling for Data Matrices,” in *Neural Information Processing Systems*, 2013, pp. 1565–1573, [Online]. Available: http://arxiv.org/abs/1311.4643.

[243] 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, no. 2, pp. 264–273, 2013.

[244] X. G. Doukopoulos and G. V Moustakides, “The fast Data Projection Method for stable subspace tracking,” in *13th European Signal Processing Conference*, 2005, pp. 1–4.

[245] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, Maryland: The Johns Hopkins University Pres, 2013.

[246] J. F. Yang and M. Kaveh, “Adaptive Eigensubspace Algorithms for Direction or Frequency Estimation and Tracking,” *IEEE Trans. Acoust.*, vol. 36, no. 2, pp. 241–251, 1988, doi: 10.1109/29.1516.

[247] F. Merchant, A. Chattopadhyay, G. Garga, S. K. Nandy, R. Narayan, and N. Gopalan, “Efficient QR decomposition using low complexity column-wise givens rotation (CGR),” in *Proceedings of the IEEE International Conference on VLSI Design*, 2014, pp. 258–263, doi: 10.1109/VLSID.2014.51.

[248] F. Merchant *et al.*, “Efficient Realization of Givens Rotation through Algorithm-Architecture Co-design for Acceleration of QR Factorization,” *ArXiv*, vol. 1803.05320, pp. 1–12, 2018, doi: 10.1109/TPDS.2018.2803820.

[249] A. Krishnamoorthy and D. Menon, “Matrix Inversion Using Cholesky Decomposition,” in *2013 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, 2011, no. 3, pp. 10–12, [Online]. Available: http://arxiv.org/abs/1111.4144.

[250] C. Ingemarsson and O. Gustafsson, “On fixed-point implementation of symmetric matrix inversion,” in *2015 European Conference on Circuit Theory and Design, ECCTD 2015*, 2015, pp. 1–4, doi: 10.1109/ECCTD.2015.7300068.

[251] S. Boyd and L. Vandenberghe, *Convex optimization theory*, 7th ed. Cambridge University Press, 2009.

Appendix A

Appendix A.1: Proof of Theorem 3.2

From (3.27) and if we let $\boldsymbol{E}_k = \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k$ then,

$$\begin{aligned} \boldsymbol{E}_k &= \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k \\ &= \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1}(3\boldsymbol{I} - 3\boldsymbol{F}\boldsymbol{R}_{k-1} + (\boldsymbol{F}\boldsymbol{R}_{k-1})^2) = (\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_{k-1})^3 = (\boldsymbol{E}_{k-1})^3 \end{aligned} \tag{A-1}$$

Since

$$\|\boldsymbol{E}_0\| = \|\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_0\| < 1$$

So

$$\|\boldsymbol{E}_k\| \leq \|\boldsymbol{E}_{k-1}\| \leq \cdots \|\boldsymbol{E}_0\|^{3^k} \to 0$$

when $k \to \infty$. Namely, $\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k \to 0$, when $k \to \infty$, that is, $\lim_{k\to\infty} \boldsymbol{R}_k = \boldsymbol{F}^{-1}$.

If we let

$$\boldsymbol{z}_k = \boldsymbol{F}^{-1} - \boldsymbol{R}_k$$

then

$$\boldsymbol{F}\boldsymbol{z}_k = \boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k = \boldsymbol{E}_k$$

From (A-1), we can obtain

$$(\boldsymbol{F}\boldsymbol{z}_k)(\boldsymbol{F}\boldsymbol{z}_k)^2 = \boldsymbol{E}_k^3 = \boldsymbol{E}_{k+1}$$

By $\boldsymbol{F}\boldsymbol{z}_{k+1} = \boldsymbol{E}_{k+1}$, we have that,

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k(\boldsymbol{F}\boldsymbol{z}_k)^2$$

which follows immediately that,

$$\|\boldsymbol{z}_{k+1}\| \leq \|\boldsymbol{z}_k(\boldsymbol{F}\boldsymbol{z}_k)^2\| \leq \|\boldsymbol{F}\|^2\|\boldsymbol{z}_k\|^3$$

From this, it is a proved that the iterative formula (3.27) is convergent and $\boldsymbol{R}_k$ at least converges cubically to $\boldsymbol{F}^{-1}$.

Appendix A.2: Proof of Theorem 3.3

To begin with, we have that $\boldsymbol{F}\boldsymbol{R}_0 = \boldsymbol{R}_0\boldsymbol{F}$ and starting from (3.27) we have that

$$\boldsymbol{F}\boldsymbol{R}_1 = \boldsymbol{F}\boldsymbol{R}_0\big(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_0(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_0)\big) = \boldsymbol{R}_0\boldsymbol{F}\big(3\boldsymbol{I} - \boldsymbol{R}_0\boldsymbol{F}(3\boldsymbol{I} - \boldsymbol{R}_0\boldsymbol{F})\big)$$

$$= \boldsymbol{R}_0\big(3\boldsymbol{F} - \boldsymbol{F}\boldsymbol{R}_0\boldsymbol{F}(3\boldsymbol{F} - \boldsymbol{F}\boldsymbol{R}_0\boldsymbol{F})\big) = \boldsymbol{R}_0\big(3\boldsymbol{I} - \boldsymbol{R}_0\boldsymbol{F}(3\boldsymbol{I} - \boldsymbol{R}_0\boldsymbol{F})\big)\boldsymbol{F}$$

$$= \boldsymbol{R}_1\boldsymbol{F}$$

This is the case when $k = 1$, and Theorem 3.3 is valid. For the case when $k \gg 1$ we evoke proof by mathematical induction that $\boldsymbol{F}\boldsymbol{R}_k = \boldsymbol{R}_k\boldsymbol{F}$ and thus

$$\boldsymbol{F}\boldsymbol{R}_{k+1} = \boldsymbol{F}\boldsymbol{R}_k\big(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k(3\boldsymbol{I} - \boldsymbol{F}\boldsymbol{R}_k)\big) = \boldsymbol{R}_k\boldsymbol{F}\big(3\boldsymbol{I} - \boldsymbol{R}_k\boldsymbol{F}(3\boldsymbol{I} - \boldsymbol{R}_k\boldsymbol{F})\big)$$

$$= \boldsymbol{R}_k\big(3\boldsymbol{F} - \boldsymbol{R}_k\boldsymbol{F}(3\boldsymbol{F} - \boldsymbol{F}\boldsymbol{R}_k\boldsymbol{F})\big) = \boldsymbol{R}_k\big(3\boldsymbol{I} - \boldsymbol{R}_k\boldsymbol{F}(3\boldsymbol{I} - \boldsymbol{R}_k\boldsymbol{F})\big)\boldsymbol{F}$$

$$= \boldsymbol{R}_{k+1}\boldsymbol{F}$$

Thus $\boldsymbol{F}\boldsymbol{R}_{k+1} = \boldsymbol{R}_{k+1}\boldsymbol{F}$ and the proof is complete.

Appendix A.3: Initial Inverse Formulation Proof

We can calculate the initial inverse as

$$\boldsymbol{R}_0 = \beta\boldsymbol{L} \tag{A-2}$$

and

$$\beta = \frac{1}{\|\boldsymbol{L}^{-1}\|^2} \tag{A-3}$$

Using the infinity norm, we have that

$$\frac{\|\boldsymbol{R}_0\|}{\|\boldsymbol{L}^{-1}\|} = \frac{\|\beta\boldsymbol{L}\|}{\|\boldsymbol{L}^{-1}\|} = \frac{\|\boldsymbol{L}\|}{\|\boldsymbol{L}\|_\infty^2\|\boldsymbol{L}^{-1}\|} \tag{A-4}$$

If we evoke this equation, then

$$\frac{\|\boldsymbol{R}_0\|_\infty}{\|\boldsymbol{L}^{-1}\|_\infty} = \frac{\|\boldsymbol{L}\|_\infty}{\|\boldsymbol{L}\|_\infty^2\|\boldsymbol{L}^{-1}\|_\infty} \tag{A-5}$$

From this we can write the condition number of **L** with respect to matrix norm as

$$K(\boldsymbol{L}) = \|\boldsymbol{L}^{-1}\|\|\boldsymbol{L}\| \tag{A-6}$$

This gives an indication whether slight change in matrix $\boldsymbol{L}$ occasions a similar change in the calculation of the inverse. For a well-conditioned matrix, we expect a small condition number and otherwise if ill-conditioned. Also, the condition number is normally larger than or equal to 1 for any given matrix.

$$\|\boldsymbol{L}\|\|\boldsymbol{L}^{-1}\| = \|\boldsymbol{I}\| = 1 \tag{A-7}$$

If substitute (A-6) in (A-5), we get

$$\frac{\|\boldsymbol{R}_0\|_\infty}{\|\boldsymbol{L}^{-1}\|_\infty} = \frac{1}{K_\infty(\boldsymbol{L})} \tag{A-8}$$

where $\mathrm{K}_\infty(\boldsymbol{L})$ represents the condition number of $\boldsymbol{L}$ with reference to the infinity norm. it can be seen that $0 = \mathscr{b}' \leq 1$ and (3.55) is satisfied. Thus, the inverse formulation based on infinity norm is justified.

Appendix A.4: Strassen's Algorithm

The Strassen's algorithm reduces the number of sub-calls to matrix-multiplication to 7, using just a bit of algebra. This brings the work down to $\mathcal{O}\left(\mathrm{K}^{\log_2 7}\right)$. If we assume matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, to multiply theses matrices to get a matrix **C**, we can follow the procedure below.

We partition matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ into four square sub-matrices, each sub-matrix having a dimension of $\frac{K}{2} \times \frac{K}{2}$.

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}$$

Where

$$\boldsymbol{C}_{11} = \boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21}$$

$$\boldsymbol{C}_{12} = \boldsymbol{A}_{11}\boldsymbol{B}_{12} + \boldsymbol{A}_{12}\boldsymbol{B}_{22}$$

$$\boldsymbol{C}_{21} = \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{A}_{22}\boldsymbol{B}_{21}$$

$$\boldsymbol{C_{22} = A_{21}B_{12} + A_{22}B_{22}}$$

Strassen observed that:

$$\boldsymbol{P_1 = A_{11}(B_{12} + B_{22})}$$

$$\boldsymbol{P_2 = (A_{11} + A_{12})B_{22}}$$

$$\boldsymbol{P_3 = (A_{21} + A_{22})B_{11}}$$

$$\boldsymbol{P_4 = A_{22}(B_{21} + B_{11})}$$

$$\boldsymbol{P_5 = (A_{11} + A_{22})(B_{11} + B_{22})}$$

$$\boldsymbol{P_6 = (A_{12} - A_{22})(B_{21} + B_{22})}$$

$$\boldsymbol{P_7 = (A_{11} - A_{21})(B_{11} + B_{12})}$$

Essentially, every one of the above components can be assessed utilizing precisely single matrix multiplication. But then, since every one of the $\boldsymbol{P_k}$'s grows by the distributive property of matrix multiplication, they capture extra data. Likewise, significant, is that these matrices $\boldsymbol{P_k}$ might be calculated independent of each other, i.e. this is the point the parallelization of our algorithm happens.

It can be verified that

$$\mathbf{C}_{11} = \boldsymbol{P_1 + P_4 - P_5 + P_7}$$

$$\mathbf{C}_{12} = \boldsymbol{P_3 + P_5}$$

$$\mathbf{C}_{21} = \boldsymbol{P_2 + P_4}$$

$$\mathbf{C}_{22} = \boldsymbol{P_1 - P_2 + P_3 + P_6}$$

Appendix B

Appendix B.1: Matrix Analysis

Computational Complexity of Matrix Operations:

Elementary linear operations of algebra assume identical characteristic form and in any hardware configuration render themselves efficient for execution. The computational complexity, however, poses a challenge when large matrices have to be processed after each millisecond. The definite complexity of the processing of a matrix categorically relies on the realization of hardware, together with the bit width and type of data. We provide first-order approximations by looking at the amount of complex multiplications and divisions that are necessary, while the complexity resulting from additions / subtractions is dismissed as these tasks are far less demanding to be realized in hardware.

Lemma A.1. Suppose we have $\boldsymbol{H} \in \mathbb{C}^{K_1 \times K_2}$ and $\boldsymbol{G} \in \mathbb{C}^{K_2 \times K_3}$ matrices. Then we need complex multiplications $K_1 K_2 K_3$ for matrix-matrix, **HG**, multiplication. Nevertheless, in multiplying $\boldsymbol{H}\boldsymbol{H}^H$ we require a total of $\frac{K_1^2+K_1}{2} K_3$complex multiplications when applying Hermitian symmetry.

Proof. **HG** has $K_1 K_3$ elements in total and we require K_2 multiplications to compute each and every element (we take elements in a given row of **H** and multiply it by elements in a given column of **G**). If we have that $\boldsymbol{G} = \boldsymbol{H}^H$, Hermitian symmetry is invoked to allow a total of $\frac{K_1^2+K_1}{2}$ elements to be computed , which represents the leading diagonal elements and half of the other-diagonal elements.

The $\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^H$ decomposition is invoked when the inverse of a matrix is multiplied with another matrix [249] to obtain hardware realization that is efficient with regard to computation and utilization of memory. **L** is a lower triangular matrix with ones on the main diagonal and **D** is a diagonal matrix whose elements are from the main diagonal.

Lemma A.2. Suppose a positive semi-definite matrix $\boldsymbol{H} \in \mathbb{C}^{K_1 \times K_1}$ which is Hermitian and a matrix $\boldsymbol{G} \in \mathbb{C}^{K_1 \times K_2}$. We decompose **H** through $\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^H$ at a cost of $\frac{K_1^3+K_1}{3}$ complex multiplications. If **H** decomposition using $\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^H$ is known, then $\boldsymbol{H}^{-1}\boldsymbol{G}$ can be computed at a cost of $K_1^2 K_2$ complex multiplications and K_1 complex divisions.

Proof. Study of efficient $\boldsymbol{LDL}^H$ decomposition algorithms is presented in [249], [250], and multiplication quantities are given in Table I of [250]. $\boldsymbol{H}^{-1}\boldsymbol{G}$ is calculated by using an evaluation of K_2 linear equation systems. With the knowledge of $\boldsymbol{LDL}^H$ decomposition, we invoke the forward-backward substitution to resolve the linear equation systems [251], this will come at a cost of K_1^2 multiplications per system. Also, we need K_1 divisions to compute $\boldsymbol{D}^{-1}$.

Appendix C

Appendix C.1: Parallel Computation of Covariance Matrix

In this section we follow the method suggested in [238] to enable high parallelization and with no multiplication repetition and no inter-core synchronization. We assume that all rows and columns begin from 1, and our channel matrix $\boldsymbol{h}_{j,k}^{(Sample)}$ and the UL received matrix $\boldsymbol{y}_{j,k}^{(Sample)}$ are our input matrices. Then we set the sliding window to be $\boldsymbol{V}^{p,q}$ with superscript p, q referring to the top left corner of the window P, Q within $\mathbf{h}_{j,k}^{(Sam\quad)}$ and $\mathbf{y}_{j,k}^{(Sample)}$. We let $\boldsymbol{A}^{p,q}$ be the column stack of $\boldsymbol{V}^{p,q}$ and we

$$\boldsymbol{V}^{p,q}(j,k) \equiv \boldsymbol{A}^{p,q}(p(j-1)+k) \tag{C-1}$$

And the results anticipated is generally expressed as

$$C = \sum_{p=1}^{M-P}\sum_{q=1}^{K-Q} \boldsymbol{C}^{p,q} = \sum_{p=1}^{M-P}\sum_{q=1}^{K-Q} \boldsymbol{A}^{p,q}(\boldsymbol{A}^{p,q})^{H} \tag{C-2}$$

To achieve efficient parallelization, we let $\mathbf{A}_{\mathrm{r_1 c_1}}$ and $\mathbf{A}_{\mathrm{r_2 c_2}}$ to represent two elements of the input matrix A. We define the distance between these two elements as

$$(\Delta r, \Delta c) = (r_2 - r_1, c_2 - c_1) \tag{C-3}$$

The elements relative position within the matrix influence the combinations. Given an element in one row then the second element will assume any position defined by $Q-1$ different places. From which we realize number of different combinations totalling to $(P-1)(Q-1)$. Unique combination (UC) set is then specified as follows

$$UC = \begin{Bmatrix} 0 \le \Delta r \le P-1 \\ 0 \le \Delta c \le Q-1 \end{Bmatrix} \cup \begin{Bmatrix} -(P-1) \le \Delta r \le -1 \\ 1 \le \Delta c \le Q-1 \end{Bmatrix} \tag{C-4}$$

Thus, the combinations in total will be

$$|UC| = PQ + (P-1)(Q-1) \tag{C-5}$$

It can be observed that Δc remains positive throughout implying that for a given product the second element is always appears on right of the first element. While Δr can either carry a negative or a positive value.

Combination centric is offered as the solution in parallel algorithm where we allocate work to various tasks at combination granularity. To optimize the number of multiplications we calculate each product only once. This gives the unique multiplications for parallel algorithm as,

$$UM = \sum_{\Delta r, \Delta c} \mu(\Delta r, \Delta c) \tag{C-6}$$

Then we express the number of indices a combination $(\Delta \mathrm{r}, \Delta \mathrm{c})$ can write to as

$$\eta(\Delta r, \Delta c) = (P - |\Delta r|)(Q - |\Delta c|) \tag{C-7}$$

Appendix C.2: Computational Complexity

It is common understanding that covariance matrices are arbitrary quantities that need to be computed every other time when the channel information need to be updated. This happens at every coherence time; thus, we are tasked to compute expansive dimensional matrix product several hundred times in a second. The number of arithmetic operations needed for the matrix product scales as a square in the matrix rank, making this operation intractable for large-scale networks. To try and reduce this implementation complexity, we employ the compression process to work on reduced matrix size but still achieve comparable estimation of covariance matrix. Then next we introduce the parallelization method mentioned above to try and bring complexity of implementation of the estimation down.

Assume we have matrices $\boldsymbol{H} \in \mathbb{C}^{K_1 \times K_2}$ and $\boldsymbol{G} \in \mathbb{C}^{K_2 \times K_3}$. Then for matrix-matrix multiplication $\boldsymbol{HG}$ we need $K_1 K_2 K_3$ complex multiplications. To multiply $\boldsymbol{H}\boldsymbol{H}^H$ we only need $\frac{K_1^2 + K_1}{2} K_3$complex multiplications when Hermitian symmetry is applied. Since there is a total of $K_1 K_3$ elements in $\boldsymbol{HG}$ and the calculation of every element needs K_2 multiplications (the $\boldsymbol{H}$ elements in a given row are multiplied by the $\boldsymbol{G}$ elements in a corresponding column). For the case when $\boldsymbol{G} = \boldsymbol{H}^H$, we use Hermitian symmetry to set the number of elements computed to $\frac{K_1^2 + K_1}{2}$, which is a representation of the leading diagonal and half of the elements from the other-diagonal.

Based on this and noting that our input matrices have dimensions of $M \times K$, then the complexity of estimating the covariance matrices $\boldsymbol{R}_{j,k}$ and $\boldsymbol{y}_{j,k}$ can easily be seen to be

$$\frac{\left(3M_j^2+M_j\right)K_\ell}{2}ll \tag{C-8}$$

When the parallelization technique is employed this then comes to

$$UM = \sum_{\Delta r,\Delta c\in UC} \mu(\Delta r,\Delta c) = UM_1 + UM_2 \tag{C-9}$$

$$\begin{aligned} UM_1 &= \sum_{\Delta r=-(P-1)}^{0} \sum_{\Delta c=0}^{Q-1} (M-\Delta r)\,(K-\Delta c) \\ &= \sum_{\Delta r=-(P-1)}^{0} (M-\Delta r) \sum_{\Delta c=0}^{Q-1} (K-\Delta c) \\ &= \frac{P}{2}(2M-P+1)\frac{Q}{2}(2K-Q+1) \end{aligned} \tag{C-10}$$

and

$$\begin{aligned} UM_2 &= \sum_{\Delta r=1}^{P-1} \sum_{\Delta c=1}^{Q-1} (M-\Delta r)\,(K-\Delta c) \\ &= \frac{P-1}{2}(2M-P)\frac{Q-1}{2}(2K-Q) \end{aligned} \tag{C-11}$$

Then we have that

$$\begin{aligned} UM &= \frac{P}{2}(2M-P+1)\frac{Q}{2}(2K-Q+1) \\ &+ \frac{P-1}{2}(2M-P)\frac{Q-1}{2}(2K-Q) \end{aligned} \tag{C-12}$$

Approximating *UM* conservatively, we have that

$$UM \leq \widehat{UM} = 2\frac{P}{2}\left(2M - (P-1)\right)\frac{Q}{2}\left(2K - (Q-1)\right) \quad \text{(C-13)}$$

It can be clearly seen that the number of operations scales as $\mathcal{O}(M^2)$ for the normal covariance estimation. For the parallel method, the number of operations scales as $\mathcal{O}(M)$ which is lower than the normal covariance estimation.