

LOW COMPLEXITY CHANNEL ESTIMATION FOR
SISO, MIMO AND MASSIVE MIMO OFDM
WIRELESS SYSTEMS

BY

ALAM ZAIB

A Dissertation Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

1963 ١٣٨٣
In Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In

ELECTRICAL ENGINEERING

DECEMBER 2015

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This dissertation, written by **ALAM ZAIB** under the direction of his dissertation adviser and approved by his dissertation committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING**.

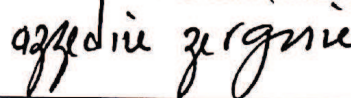
Dissertation Committee



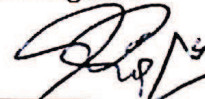
Dr. Tareq Y. Al-Naffouri (Adviser)



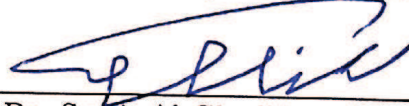
Dr. Merouane Debbah (Member)



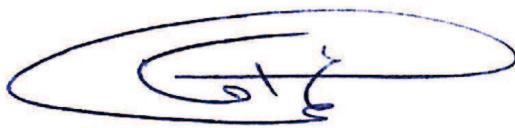
Dr. Azzedine Zerguine
(Member)



Dr. Ali H. Muqaibel (Member)



Dr. Samir Al-Ghadhban (Member)



Dr. Ali A. Al-Shaikhi
Department Chairman

Dr. Salam A. Zummo
Dean of Graduate Studies

28/1/16
Date



© Alam Zaib
2015

*Dedicated to my loving
parents and family*

ACKNOWLEDGMENTS

First of all, I would like to thank all mighty Allah (SWT) for His countless blessing on me at every sphere of life. May He help us to follow Islam in its true spirit according to the teachings of His prophet Muhammad (PBUH).

Secondly, I would like to express deepest gratitude to my advisor, Dr. Tareq Al-Naffouri for his continuous support and guidance. He is one of the leading researcher, teacher and a great mentor whom I have learnt a lot. I also appreciate him for giving many research directions, often more than I could handle, and enough freedom to work on various topics. It has been a valuable and enriching experience working under his supervision.

I am grateful to my dissertation committee members, Dr. Mérouane Debbah, Dr. Azzedine Zerguine, Dr. Ali Muqaibel and Dr. Samir Al-Ghadhban for their careful review and constructive feedback on my work. Specially, I am grateful to Dr. Mérouane for travelling from France to attend my defense.

I take this opportunity to express gratitude to KFUPM and its knowledgable faculty members for stimulating my research interests in signal processing and communications through various courses and projects. I also acknowledge the support provided by the Deanship of Scientific Research at KFUPM for partially funding this work through project RG –1314.

I would also like to thank my friends and colleagues at KFUPM: Naveed Iqbal, Adeel Sabir, Waqas Afzal, Hussain Ali, Ali Qureshi and Asjad Ameen (to name only a few) for their indirect support, creating conducive work environment and making my stay at KFUPM fully enjoyable and memorable. I also acknowledge the time and efforts of my friends at KUST: Mudassir Masood and Anum Ali, for reviewing papers, extending help and sharing their ideas related to this work during my visits to KAUST.

Last but not least, I owe the greatest debt to my father for imparting us quality education (both at school and college level) despite difficult financial situations, to my mother for her irredeemable love and prayers, to my wife and kids for their patience and to entire family for their generous help and support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
NOMENCLATURE	xiv
ABSTRACT (ENGLISH)	xx
ABSTRACT (ARABIC)	xxii
CHAPTER 1. INTRODUCTION	1
1.1 Overview	1
1.2 Characteristics of the Wireless Channel	5
1.3 Channel Estimation Techniques	8
1.3.1 Previous Works	9
1.4 Aims and Scope	12
1.5 Thesis Contributions and Layout	14
CHAPTER 2. LOW-COMPLEXITY CHANNEL INTERPOLA-	
TION FOR SISO-OFDM SYSTEMS	17
2.1 System Model	18
2.1.1 Linear Interpolation	22
2.1.2 Polar Linear Interpolation	23

2.1.3	Adaptive Polar Linear Interpolation	24
2.1.4	LMMSE Interpolation and Variants	25
2.2	The Proposed Method	28
2.2.1	Further Reducing the Complexity	33
2.3	Simulation Results	33
2.4	Concluding Remarks	38
CHAPTER 3. BLIND AND SEMI-BLIND ML DETECTION FOR		
MIMO OFDM SYSTEMS		44
3.1	Motivation	45
3.2	Problem Formulation	46
3.3	Joint ML/MAP solution	51
3.3.1	Recursive Derivation of Bound	54
3.4	Low-Complexity Blind Algorithm	57
3.4.1	Reducing Complexity by Avoiding \mathbf{P}_i	57
3.4.2	Reducing Complexity by Carrier Reordering	58
3.5	Complexity Reduction by Reliable Carriers	60
3.5.1	Measuring the Reliability	61
3.6	Simulation Results	64
CHAPTER 4. CHANNEL ESTIMATION FOR MASSIVE MIMO		
OFDM SYSTEMS		71
4.1	System Model	72
4.1.1	Channel Model	73
4.1.2	Signal Model	77
4.2	LMMSE and LS based Channel Estimation	78
4.2.1	The Localized LMMSE (L-LMMSE) Estimation	79
4.2.2	The Optimal LMMSE (O-LMMSE) Solution	80
4.2.3	Estimation Using Least Squares	81
4.3	The Proposed Distributed LMMSE Estimation	83
4.3.1	Complexity Analysis	93

4.3.2	Linkage Between Localized and Centralized Solutions . . .	95
4.4	Data-Aided D-LMMSE Estimation	96
4.4.1	Reliable Carriers Selection	97
4.4.2	Revisiting the Estimation Step	99
4.5	Simulation Results	100
4.6	Concluding Remarks	104
 CHAPTER 5. EFFECT OF PILOT CONTAMINATION ON		
CHANNEL ESTIMATION		108
5.1	Pilot Contamination and Implications	109
5.2	Mathematical Preliminaries	112
5.2.1	Point Processes	112
5.2.2	Poisson Point Process (PPP)	113
5.3	Modified Network Model	114
5.4	Interference Characterization	115
5.5	Effect of PC on MSE Performance	117
5.6	Simulation Results	122
 CHAPTER 6. CONCLUSIONS AND FUTURE RECOMMENDA-		
TIONS		128
6.1	Achievements of the Work	128
6.2	Summary of Contributions	129
6.3	Future Recommendations	130
 APPENDIX A.		132
A.1	Proof of Lemma 5.1	132
A.2	Proof of Theorem 5.2	133
A.3	Proof of Theorem 5.4	134
 REFERENCES		136
 VITAE		148

LIST OF TABLES

2.1	Complexity-Performance Trade-off	37
4.1	Computational Complexity	94
4.2	Simulation Parameters	102
5.1	Summary of MSE expressions	122

LIST OF FIGURES

2.1	Simplified block diagram of OFDM system	19
2.2	Different types of pilot arrangements used in OFDM systems. . .	21
2.3	CFR coefficients in an OFDM symbol of $N = 64$ sub-carriers and $K = 16$ equi-spaced pilots	23
2.4	Normalized correlation as function of l for $N=64$ and $L=16$. . .	31
2.5	Pictorial representation of interpolation process.	34
2.6	Effect of interpolation depth on (a) MSE and (b) BER performance. Number of required multiplications per CFR coefficient is propor- tional to d	39
2.7	Effect of channel PDP mismatch for (a) General PDP profile and (b) Exponentially decaying profile. The parameters are $N=128$ and $L=8$ and $K=16$ with QPSK symbols.	40
2.8	Effect of (a) Channel delay spread and (b) Number of pilots, on MSE performance using $N=256$, 16-QAM symbols at 20dB SNR.	41
2.9	(Top) MSE performance comparison of the proposed method with various algorithms. The parameters are $N=256$, $K=32$ with QPSK symbols. (Bottom) Snapshot of the magnitude of CFR at each subcarrier. Red circles represent pilot locations where the CFR is known.	42

2.10	(Top) MSE performance comparison of the proposed method with various algorithms. The parameters are $N=256$, $K=32$ with QPSK symbols. (Bottom) Snapshot of the magnitude of CFR at each subcarrier. Red circles represent pilot locations where the CFR is known.	43
3.1	(a) Alamouti coded OFDM system (b) Frame structure of OFDM data blocks over two consecutive time instants.	47
3.2	Correlation between partial vectors \mathbf{a}_1 and \mathbf{a}_i for $N=16$ and $L=4$	59
3.3	Reliability of data carriers $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ in decoding them to the nearest neighbour constellation point \mathcal{X}	63
3.4	BER Performance of blind algorithm over Rayleigh fading channel with $N = 16$ and $L = 4$	67
3.5	BER performance of semi-blind algorithm over Rayleigh fading channel with $N = 32$ and $L = 4$	68
3.6	(a) Computational complexity and (b) BER performance of the proposed algorithm with various degrees of reliability measurements using BPSK symbols with $N=16$ and $L=4$	69
3.7	Complexity of the proposed semi-blind algorithm without (solid lines) and with reliable carriers (dashed lines) for different modulations with $N=16$ and $L=4$	70
4.1	(a) Multi-cell massive MIMO system layout (b) An example of $M \times G$ UPA structure with antenna indexing.	74
4.2	The working principle of D-LMMSE Algorithm	85
4.3	Details of (a) Information diffusion process and (b) Information sharing process.	87
4.4	The working principle of DAD-LMMSE Algorithm	98
4.5	Number of iterations (D) required for convergence of the proposed distributed algorithm	105

4.6	Effect of spatial correlation on convergence of the proposed distributed algorithm.	106
4.7	MSE performance comparison of different algorithms.	106
4.8	MSE performance comparison of the proposed data-aided algorithm with pilot-based techniques.	107
4.9	Average runtime of various algorithms.	107
5.1	Effect of pilot contamination due to reuse of pilots, on the uplink and downlink data transmissions.	111
5.2	Realization of interferes distributed according to PPP of $\lambda=0.3$, $\gamma_o=2m$ and $\gamma_m=5m$ with BS at the origin.	116
5.3	Mean and variance of interference as a function of: (a) Intensity λ , with $b = 2$ and (b) Pathloss exponent b , with $\lambda = 1$	125
5.4	MSE performance in presence of AWGN and PC with parameters $\lambda=0.1$ and $b=2$	126
5.5	Effect of interfering user intensity λ and pathloss exponent b on MSE performance at SNR = 10 dB.	127

Nomenclature

Abbreviations

3GPP	3rd Generation Partnership Project
AoA	Angle of Arrival
AoD	Angle of Departure
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
bps	Bits Per Second
BS	Base Station
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CP	Cyclic Prefix
CSI	Channel State Information
DAB	Digital Audio Broadcast
DFT	Discrete Fourier Transform
DVB	Digital Video Broadcast
EVD	Eigen Value Decomposition
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
ISI	Inter-Symbol-Interference
LMMSE	Linear Minimum Mean Square Error
LMS	Least Mean Square

LS	Least Squares
LTE	Long Term Evolution
M2M	Machine-to-Machine
MAP	Maximum A Posteriori
MIMO	Multiple-Input-Multiple-Output
ML	Maximum-Likelihood
MMSE	Minimum-Mean-Square-Error
MPCs	Multipath Components
MRT	Maximum Ratio Transmission
MU-MIMO	Multi-User MIMO
OFDM	Orthogonal Frequency Division Multiplexing
PDF	Probability Density Function
PDP	Power Delay Profile
PPP	Poisson Point Process
RLS	Recursive Least Squares
RMS	Root Mean Square
SDMA	Space Division Multiple Access
SIMO	Single Input Multiple Output
SINR	Signal to Interference Noise Ratio
SISO	Single-Input-Single-Output
SNR	Signal to Noise Ratio
SS	Spherically Symmetric

STBC	Space Time Block Coding
SU-MIMO	Single-User MIMO
SVD	Singular Value Decomposition
TDD	Time Division Duplex
UPA	Uniform Planar Array
WLAN	Wireless Local Area Networks
WLS	Weighted Least Squares
ZF	Zero Forcing

Notations and Symbols

$(.)^*$	Conjugate operation
$(.)^T$	Transpose operation
$(.)^H$	Conjugate Transpose (or Hermitian) operation
\mathbf{x}	Represents a column vector
$\hat{\mathbf{x}}$	Represents the estimate of \mathbf{x}
\mathbf{X}	Represents a matrix
\mathcal{X}	Represents a column vector in frequency domain
\mathbf{F}	(Unitary) FFT matrix
\mathbf{I}_N	An $N \times N$ identity matrix
$\mathbf{0}_{M \times N}$	An $M \times N$ matrix with all elements equal to zero
\mathbf{x}_i	Represents i -th column of matrix \mathbf{X}
$x(i)$	Represents i -th entry of vector \mathbf{x}
$\mathbf{x}_{(i)}$	Partial vector formed by entries of N -dimensional vector \mathbf{x} up to the index $i < N$
$\mathbf{x}(\mathcal{P})$	A vector formed by selecting entries of \mathbf{x} indexed by \mathcal{P}
$[\mathbf{X}]_{i,j}$ or $x_{i,j}$	Denotes (i, j) -th entry of matrix \mathbf{X}
\mathbf{X}_{ij}	Denotes (i, j) -th block entry of a block matrix
$\mathbf{X}_{(i)}$	Partial matrix formed by rows of $N \times M$ -dimensional matrix \mathbf{X} up to the index $i < N$
$\mathbf{X}(\mathcal{P})$	A matrix formed by selecting rows of \mathbf{X} indexed by \mathcal{P}
ν	Wavelength (in meters)

B_c	Coherence bandwidth of channel
T_c	Coherence time of channel
f_c	Carrier frequency
F_d	(Normalized) Doppler frequency
O	Origin of 2D plane OR the order of complexity
Θ	Number of Monte Carlo trials
\otimes	Kronecker product
ϕ	Mean horizontal AoD (in radians)
σ	Standard deviation of horizontal AoD
θ	Mean vertical AoD (in radians)
ξ	Standard deviation of vertical AoD
λ	(Interfering) user's intensity
b	Pathloss exponent
Ω	Set of constellation alphabets
$ \Omega $	Cardinality of set Ω
Ψ	Denotes Poisson Point Process (PPP)
$N(\mathcal{B})$	Number of poisson points in the subset \mathcal{B}
$\mathcal{A}(\mathcal{B})$	Lebesgue measure of the subset \mathcal{B}
\mathbb{R}^d	d -dimensional space of real numbers
\mathbb{C}^d	d -dimensional space of complex numbers
$f_X(x)$	PDF of random variable X
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian distribution with mean μ and variance σ^2

$\text{trace}(\mathbf{X})$	Trace operator, equals sum of all diagonal entries of \mathbf{X}
$\mathbb{E}\{.\}$	Statistical expectation operator
$\text{diag}(\mathbf{x})$	A diagonal matrix with entries of \mathbf{x} along the diagonal
$\max(.)$	Max function that returns maximum value of its arguments
$\ \mathbf{x}\ $	The ℓ_2 norm of vector \mathbf{x}
$\ \mathbf{x}\ _{\mathbf{A}}^2$	The weighted squared norm of \mathbf{x} defined as $\ \mathbf{x}\ _{\mathbf{A}}^2 \triangleq \mathbf{x}^H \mathbf{A} \mathbf{x}$
$\langle \hat{\mathcal{X}}(k) \rangle$	Hard decoding i.e., mapping of $\hat{\mathcal{X}}(k)$ to nearest constellation point $\mathcal{X}(k) \in \Omega$

THESIS ABSTRACT

NAME: Alam Zaib
TITLE OF STUDY: Low Complexity Channel Estimation for SISO, MIMO
and Massive MIMO OFDM Wireless Systems
MAJOR FIELD: Electrical Engineering
DATE OF DEGREE: December 2015

In OFDM based wireless communication systems, whether employing single or multiple antennas, channel state information has to be estimated accurately and that too within a fraction of time, making channel estimation very crucial. Against various state-of-the art on channel estimation, this thesis presents several low complexity channel estimation techniques for SISO, MIMO and massive MIMO OFDM systems by exploiting the structure and some of the constraints of communication problem.

We first present a reduced complexity optimal interpolation technique for SISO-OFDM systems based on MMSE criteria. By utilizing the structure of channel frequency correlation, it is shown that if pilots are placed appropriately across OFDM subcarriers, the matrix inversion in conventional MMSE estimation can be com-

pletely avoided with no loss in performance. Next, we present a blind ML algorithm for joint channel estimation and data detection for MIMO-OFDM systems with Alamouti coding where the complexity is reduced by again utilizing the correlation structure and the finite alphabet property of symbols. A semi-blind algorithm is also introduced which has much lower complexity than the blind algorithm but at the cost of few training symbols.

As for the massive MIMO systems, the complexity is of primary concern because with increased number of base station antennas (BS), the number of unknown channel parameters also grow large. Unlike the optimal MMSE approach, which is prohibitively complex, we present a distributed MMSE algorithm whose complexity is linear in the number of BS antennas while at the same time achieves near-optimal performance by sharing the information locally in a large antenna array. A data-aided version of distributed algorithm is also presented to minimize the pilot overhead in massive MIMO. Finally, we investigate the effect of pilot contamination (i.e., interference due to reuse of pilots) on MSE performance of various algorithms. We use stochastic geometry to derive closed-form expressions for channel MSE under both noise and pilot contamination regime, which are validated by simulations. Our results indicate severe implications of pilot contamination on channel estimation performance.

ملخص الرسالة

الاسم: عالم زيب

عنوان الرسالة: طريقة تقدير سهلة لقنوات SISO ، MIMO وقنوات MIMO الضخمة اللاسلكية لتقنية OFDM

التخصص: هندسة كهربائية

تاريخ المناقشة: ديسمبر 2015

في نظم الاتصالات اللاسلكية OFDM ، سواء بتوظيف هوائي واحد أو هوائيات متعددة، لابد من معرفة حالة القناة بكل دقة ويجب أن يكون في خلال جزء ضئيل جدا من الوقت، مما يجعل تقدير القناة محمة حاسمة. بالرغم من وجود طرق حديثة لتقدير القنوات، فإن هذه الأطروحة تقدم عدة طرق سهلة لتقدير قنوات SISO، MIMO وقنوات MIMO الضخمة، تحت تقنيات OFDM، وذلك من خلال استغلال بنية وبعض القيود من مسائل الاتصالات.

نقدم أولا تقنيات تضمنين مثالية سهلة لأنظمة SISO-OFDM بناءً على معيار MMSE. وذلك من خلال الاستفادة من بنية ارتباط ترددات القناة، فقد لوحظ أن بيانات الفحص للقناة إذا وضحت بعناية عبر الحوامل الجزئية في تقنية OFDM، فإن إيجاد مقلوب المصفوفة في معيار MMSE ليس له ضرورة، ولن يؤثر على دقة التقدير. وبعد ذلك، قمنا بتصميم خوارزمية ML التي لا تحتاج لبيانات فحص للقناة، الذي يستخدم لتقدير القناة وكشف المعلومات لنظم MIMO-OFDM مع استخدام ترميز Alamouti ، حيث تم خفض التعقيد مرة أخرى باستخدام بنية الارتباط ورموز أبجدية مخصصة ومحدودة. لقد تم أيضاً تطوير خوارزمية تحتوي جزئياً على بيانات فحص، والتي تتميز بتعقيد أقل بكثير من الخوارزمية التي لا تعتمد على أي بيانات فحص للقناة ، ولكن على حساب بعض رموز للتدريب.

أما بالنسبة للأنظمة MIMO الضخمة، فإن التعقيد هو شاغل رئيسي لأنه مع زيادة عدد الهوائيات في محطة القاعدة (BS)، فإن عدد القنوات الغير معروفة تنمو أيضاً بشكل كبير. وخلافاً لمعيار MMSE الأمثل، والذي لا يمكن تطبيقه لتعقيده، فإننا نقترح خوارزمية MMSE موزعة تتميز بقلّة التعقيد وهي دالة خطة تعتمد على عدد الهوائيات في الـ BS. في الوقت نفسه، يحقق الأداء الأمثل من خلال تبادل المعلومات محليا في مجموعة كبيرة للهوائيات. وقد عرضنا أيضاً خوارزميات تعتمد على البيانات وزعت أيضاً للحد من النفقات العامة الرائدة في MIMO الضخمة. وأخيراً، فإننا تحققنا من تأثير تلوث بيانات الفحص (التلوث الطارئ على بيانات الفحص) على أداء MSE من خوارزميات مختلفة. تم أيضاً استخدام الهندسة العشوائية لاشتقاق قاعدة لقناة MSE في إطار كل من الضوضاء والتلوث الطارئ على بيانات الفحص، والتي يتم التصديق عليها من قبل المحاكاة. نتائجنا تشير إلى آثار خطيرة ناتجة عن تلوث بيانات الفحص على أداء تقدير القناة.

CHAPTER 1

INTRODUCTION

This Chapter gives a general overview of advances in wireless communications. Some of the important characteristics of wireless channel and state-of-the-art on channel estimation in OFDM-based systems are also presented. Towards the end, we highlight various key research challenges, motivations and scope of our work, major contributions and layout of the thesis.

1.1 Overview

Transmitting information from one place to another without having any physical medium was not plausible till the advent of wireless communication about a century ago. Since then, the wireless communications has advanced significantly and has played pivotal role in shaping our society. In the recent decades, there is even more rapid development of wireless mobile communications. The ever increasing demand for higher data rates, anytime and anywhere connectivity, security and robustness has led to evolution of four generations of mobile communications (1G

to 4G); with 4G currently being deployed around the world.

Over the years, the two prime technologies for sustaining the higher data-rates and spectral efficiency for wireless communications are: Multiple-Input-Multiple-Output (MIMO) and Orthogonal Frequency Division Multiplexing (OFDM). The use of multiple antennas whether at the transmitter or the receiver or both, can substantially increase data throughput and the reliability of a radio link [1, 2]. Multi-antenna wireless systems offer additional degrees of freedom provided by the spatial dimension, which can be exploited to either simultaneously transmit independent data-streams (spatial multiplexing) thereby increasing the data-rate, or multiplicative transmission of single data stream (spatial diversity) to increase the system reliability [3]. In cellular scenario when the base station (BS) serves a multitude of terminals over same time-frequency resources, these spatially separated streams can be used to transmit (or receive) the data to (or from) each terminal. This technique, commonly known as multi-user MIMO (MU-MIMO) or space-division-multiple-access (SDMA), can significantly increase the throughput gains of MIMO wireless systems. However, the gains with MIMO are achieved at increased processing complexity and hardware costs. Hence, it is necessary to adopt modulation techniques with simplified signal processing at transmitter and receiver for precoding and equalization respectively.

OFDM emerged as a promising modulation scheme to achieve these objectives [4]. In OFDM the data is transmitted on orthogonal subcarriers each experiencing a flat-fading channel conditions and can be processed individually. OFDM also

efficiently utilizes the bandwidth, a scarce and expensive resource, by allowing subcarriers to overlap and offers low complexity modulation and demodulation structures by efficient Discrete Fourier Transform (DFT) based signal processing. The combination of MIMO technology with OFDM, called MIMO-OFDM, has enabled high speed data transmission and broadband multimedia services over wireless links that we enjoy today. MIMO-OFDM has been studied extensively and is becoming a mature technology. It has also been incorporated into many existing wireless standards and products such as Wireless Local Area Networks (WLAN) standards (IEEE 802.11 a/b/g) [5], WiFi standard IEEE 802.11n [6] , WiMAX standard IEEE 802.16e [7], digital audio broadcast (DAB), digital video broadcast (DVB) and also adopted for existing 4G cellular communication systems (3GPP-LTE Advanced) [8, 9].

The unprecedented usage of smart phones, tablets, super-phones etc., equipped with data-intensive applications like video streaming, graphics heavy social media interfaces and real time navigation services, has called for revolutionary changes for the next generation wireless systems. Data communication speeds of 10Gbps are expected that must be provided by 10-fold enhanced spectral efficiency and 1000-fold greater system capacity [10]. The Wireless World Research Forum (WWRF) predicts that with addition of emerging machine-to-machine (M2M) communications, 7 trillion wireless devices will serve 7 billion people by 2017. This means the number of network-connected devices to become 1000 times the world population [11]. Such substantial growth in capacity can only be envis-

aged when paralleled by similar improvements in energy efficiency. Addressing these multi-fold challenges requires a revolutionary technology; one that could deliver an economically sustainable capacity and performance improvements, better coverage and superior user experience than the existing wireless systems.

More recently in 2010, in a seminal paper by Marzetta [12], it was established that installing excessively large number of antennas at the BS, in the order of a few hundred, can achieve huge gains in spectral and energy efficiencies. Such systems are commonly known as massive MIMO or large-scale MIMO systems [13, 14, 15]. Massive MIMO overcome several limitations of the traditional MIMO systems such as security, robustness and throughput scalability. The current state of the art in cellular technology (LTE Release 10) allows for 8 antenna ports at the BS and an equal number of antenna ports at the terminal [8]. As such, they cannot fully achieve the gains offered by MU-MIMO and rely mostly on single-user-MIMO (SU-MIMO). Thus making a clean break from traditional MIMO, the use of very large antenna arrays at BS has the potential to achieve full benefits of MIMO on much larger scale and where the bulk of processing complexity is handled at the BS. It has been demonstrated that massive MIMO systems hold great promises of boosting system throughput by 10 times or more by simultaneously serving tens of users in the same time-frequency resource. This has become possible by fully exploiting the MU-MIMO through excess degrees of freedom available at the BS array. They also have the potential to increase energy efficiency because of their ability to precisely focus their transmission energy towards intended users

through the use of a large number of small active antennas with very low power. Moreover, in massive MIMO the effects of fast fading are averaged out and intra-cell interference almost vanishes. In light of these advantages, the massive MIMO is expected to play increasingly important role in development of future broadband wireless systems and envisioned as one of the enabling technologies for 5G [16].

To fully realize the potentials of aforementioned technologies the knowledge of channel state information (CSI) is indispensable. For improved system performance, it is essential that CSI is available at both transmitter and the receiver. The knowledge of CSI is utilized for coherent detection of signals at the receiver. On transmitter side, CSI is crucial to design effective precoding schemes for minimizing inter-user interference. However, the perfect knowledge of CSI is not available in real life, therefore it has to be estimated. This thesis is concerned with efficient and low complexity channel estimation algorithms for single antenna (i.e., SISO) as well as multiple antenna (i.e., MIMO and massive MIMO) OFDM wireless systems.

1.2 Characteristics of the Wireless Channel

The wireless channel is highly dynamic and unpredictable as opposed to the typically static and predictable wired channel and therefore limits the performance of communication systems. This makes the channel estimation an essential part of the receiver design. In wireless medium, the propagation of a radio wave is governed by reflection, diffraction, scattering and relative motion of objects in

the environment, which distort the amplitude, phase and frequency of the received signal. The most important characteristics of wireless channel and systems considered in this thesis are,

- **Fading:** Fading is the main characteristic of wireless channel which describes the random fluctuations of received signal strength over time and frequency. Fading can be classified as large-scale and small-scale fading. The former is caused by shadowing due to large obstacles such as buildings, trees or walls etc., while the later is caused by multipath propagation and time-varying nature of the channel.

Due to scattering, the transmitted signal arrives at the receiver via several paths each having its own attenuation and phase-shift. These multipath components (MPCs) add up constructively or destructively depending on their path-lengths, giving rise to small-scale fading. If there is no direct path, the small-scale fading is assumed to follow zero mean complex Gaussian distribution, called Rayleigh fading [17].

The properties of MPCs may vary due to mobility of the transmitter, receiver and or scatterers giving rise to Doppler shift in the frequency of MPCs. The multipath and Doppler effect causes the signal to spread in time and frequency respectively and manifest themselves as frequency-selective and time-selective behaviour of wireless channel. The small-fading effects are typically modeled by a tapped-delay line filter with time-varying coefficients.

- **Path loss:** The attenuation of received signal power as a function of the dis-

tance between transmitter and receiver is usually determined by path loss. Calculation of path loss depends largely on the type of environment (e.g., rural or urban) and the radio propagation conditions (e.g., precipitation) [18]. In free space the pathloss exponent is 2 meaning that signal strength decreases with inverse square of the distance, while in urban environment it is found to be between 2.5 and 6. The effects of path-loss are especially important in multi-cell and multi-user scenario, where the users are located at different positions in the cells.

- **Channel coherence:** As described above, the variation of small-scale fading with time and frequency give rise to frequency and time selectivity of wireless channel. Both of these effects are quantified by channel coherence bandwidth, B_c and channel coherence time, T_c , defined respectively as the bandwidth or the time over which the channel is approximately constant. They can be well approximated as $B_c \approx c/T_{rms}$ and $T_c \approx c/(4vf_c)$ respectively, where c is the speed of light, T_{rms} is the Root Mean Square (RMS) of path lengths, v is the velocity of receiver/and or transmitter and f_c is the carrier frequency.

Channel coherence has important implications in channel estimation as it places a fundamental limit on the duration and the bandwidth, called coherence interval, over which the channel estimates are valid.

1.3 Channel Estimation Techniques

Channel estimation techniques can be broadly categorized into pilot-based, blind, semi-blind and data-aided. These approaches are summarized as follows.

- **Pilot-based techniques:** The pilot-based approach is the most common and widely used approach in which the training sequences (called pilots), known a priori at the receiver, are transmitted along with data symbols to obtain reliable channel estimates. Estimation performance generally improves with increasing the number of pilots but the transmission efficiencies are reduced due to required overhead of training symbols. So, usually, there is a trade-off.
- **Blind techniques:** In blind methods no pilots are transmitted and the channel is estimated using the statistical properties of the received symbols. This reduces the incurred overhead of pilots, however, often a large number of data symbols are required to extract statistical properties. Furthermore, their complexity is higher and performance is usually worse than pilot-based techniques.
- **Semi-blind:** The pilot-based and blind are the two extreme cases in the sense that the former only uses pilots while the latter doesn't make use of pilots. The semi-blind techniques are the hybrid of pilot-based and blind techniques that utilize statistical properties of data symbols and require fewer pilots. These techniques are suitable in slowly time-varying channels where initial estimates can be obtained from pilots and the channel can be tracked subsequently from received data symbols.

- **Data-aided techniques:** Data-aided techniques are perhaps the most sensible approach, where the data is first decoded from pilot-based channel estimates and then used to enhance the estimation performance. Both decoding and estimation process are interdependent and form the basis of many iterative and joint estimation-detection algorithms.

1.3.1 Previous Works

The problem of channel estimation for wireless systems has been well investigated and a number of techniques have been proposed in the literature. Here, we present a brief overview of some state-of-the-art methods on channel estimation for OFDM wireless systems and also discuss their limitations that forms the basis for this work.

Pilot-based techniques for OFDM systems mostly rely on estimating the channel frequency response(CFR) at pilot sub-carriers using Least-Squares (LS) and the remaining CFR coefficients via some form of interpolation [19]. Some interpolation approaches are based on minimum mean square error estimation (MMSE), which are optimal while others are non-MMSE based ranging from simple techniques e.g., [20, 21, 22, 23, 24, 25] to more sophisticated ones such as [26, 27]. Although sophisticated techniques may have better performance, simple linear interpolation techniques are preferred due to computational and implementation advantages. The MMSE type methods exploit frequency correlation of the channel to yield optimal performance but require prohibitively higher complexity due

to large matrix inversion and large matrix-vector products [28]. The sub-optimal methods to reduce the complexity of MMSE estimators have also been proposed [29, 30, 31, 32, 33] but have certain drawbacks e.g., heavy computation of singular value decomposition (SVD). Thus there is a natural desire to look for other alternatives which give better trade off between complexity and performance.

Blind techniques for channel estimation can be grouped into maximum-likelihood (ML) based which are optimal, and non-ML based approaches. The later includes subspace-based methods [34, 35], second-order-statistics [36], cholesky factorization [37] or iterative methods [38]. These methods either suffer from slow convergence, higher computational costs or assume channel to be stationary over several OFDM blocks. These drawbacks make ML based approaches e.g., [39, 40] more attractive due to their fast convergence. Usually suboptimal techniques are employed to reduce computational cost by restricting the search space of exhaustive ML search. These suboptimal techniques, however, are applicable to specific constant modulus constellations [41, 42]. Recently in [43] and [44], the authors have proposed a low-complexity blind ML method for general constellations for single-input-multiple-output (SIMO) and single-input-single-output (SISO) systems respectively, which gives motivation for extending this work to multi-antenna systems.

Channel estimation is critical as well as more challenging in massive MIMO systems. Having a very large number of antennas means that a significant number of channel coefficients need to be estimated – far more than that could be

handled by traditional pilot-based MIMO channel estimation techniques (see [45] and references therein). In this regard, the Bayesian MMSE estimator provides an optimal estimate in the presence of additive white Gaussian noise (AWGN) by exploiting antenna spatial correlations. However, the direct generalization of MMSE estimator to massive MIMO has some drawbacks. In particular, it suffers from huge complexity due to matrix inversion of very large dimensionality, making it impractical. Some ways to reduce the complexity of MMSE estimators in massive MIMO have also been proposed e.g., [46, 47, 48, 49, 50, 51, 52]. It is important to note that most of these methods make assumptions that are not always realistic. For example, many methods deal with flat fading channels only while others assume that the channels are sparse. This gives motivation to investigate low complexity channel estimation approaches suited to correlated Rayleigh fading channels in massive MIMO-OFDM systems.

In a multi-cell environment, allocation of orthogonal pilot sequences for all users cannot be guaranteed due to finite coherence time and limited bandwidth [12, 53]. Therefore, the pilots must be reused across cells. Consequently, when the BS in a given cell performs channel estimation via uplink training, the channel estimates will be severely distorted (contaminated) by pilots of neighboring cell users. The impact of pilot contamination is far greater than AWGN and is one of the limiting factors in performance of the massive MIMO systems as demonstrated in [54, 55]. Some ways to reduce the pilot contamination have recently been proposed. The line of work in [48, 56] shows that the impact of pilot contam-

ination can be reduced by allocating pilot resources by base station cooperation that require sharing of second-order statistics. In [57], it is shown that pilot contamination is reduced if adjacent cells are unsynchronized in the sense that one sends data when the other one is sending pilots, and vice versa. An effective way to deal with pilot contamination is to use blind and semi-blind approaches as they do not rely much on pilots [58, 59]. Despite several works on pilot contamination, only few have analysed its impact on channel estimation performance [47] with known user's locations. As such, these works cannot analytically answer how the randomness of user's locations would effect channel estimation performance under pilot contamination. Moreover, these techniques also ignore the effects of strong spatial antenna correlations that is more likely to exist in massive MIMO when the BS antennas are co-located.

1.4 Aims and Scope

Despite several state-of-the-art techniques on channel estimation, there are many research challenges that need to be addressed. One of the major challenges, that is tackled in this thesis is: how to reduce the complexity of the optimum channel estimation algorithms without losing their performance? In certain scenarios, the optimum solutions might be prohibitively complex to implement directly. Hence, in practical implementation of channel estimators, the designer may be willing to compromise a bit on performance to lower the complexity.

The approach used in this thesis to reducing the complexity of channel es-

timators is motivated by the fact that most of communication problems exhibit inherently rich structure that can be exploited to reduce the complexity. More specifically, the structure of Fast Fourier Transform (FFT) matrices induced by OFDM, the Finite alphabet property of symbols (since they belong to certain constellation), the structure of channel correlation in time, frequency and space will be utilized in SISO and MIMO OFDM systems resulting in low complexity solutions attaining optimal or near optimal performance.

The issue of complexity is more pronounced in massive MIMO systems utilizing several hundreds of antennas at the BS. This raises serious questions on complexity of the optimal MMSE estimator, which grows according to the cube of the number of antennas, if implemented centrally. Moreover, all received (thousands of) signals in massive MIMO cannot be processed efficiently at one central processor. These limitations motivate the use of algorithms which are distributed, computationally efficient and require very little inter-processor communication.

Moreover, with increasing number of users, a large number transmission resources are needed for estimating channel coefficients. This will greatly limit the bandwidth and power resources that can be allocated for the transmission of information. Data-aided techniques can significantly reduce the pilot overhead and improve the spectral efficiency of the system. In this work, we aim to develop a low complexity data-aided distributed algorithm with a simple equalizer structure, such as zero-forcing (ZF). In massive MIMO, the rapidly increasing number of users will also quickly exhaust the available supply of orthogonal pilot sequences.

The necessary reuse of pilots from one cell to another results in negative consequences i.e, pilot contamination. Pilot contamination severely distorts the channel estimates, which not only affects the decoding process but also has implications on down-link beamforming. We aim to quantify the effect of pilot contamination on channel estimation performance using a realistic network model. Specifically, the approach used in our work is inspired by stochastic geometry based analysis where the users are assumed to be distributed randomly according to a poisson point process (PPP).

1.5 Thesis Contributions and Layout

The main contributions of the thesis can be summarized into three parts corresponding to SISO, MIMO and massive MIMO OFDM systems as follows:

Part-I (SISO)

- A low-complexity and spectrally efficient channel interpolation method is developed based on MMSE criteria. The proposed channel estimator exploits the correlation structure of CFR matrix to reduce the complexity. In fact, it is shown that if pilots are distributed appropriately across OFDM subcarriers, then the proposed algorithm requires no matrix inversion, thus substantially reducing the computational complexity.

Part-II (MIMO)

- A low-complexity blind algorithm is developed for joint channel estimation and data detection for Alamouti-coded OFDM wireless systems with two

transmit and single receive antenna. The blind approach, again utilizes the structure of FFT matrices to significantly reduce the complexity. The proposed algorithm is applicable for both constant modulus and non-constant modulus constellations.

- A semi-blind version of blind ML algorithm is developed which utilizes few pilots to further reduce the complexity and subsequently uses reliable data to improve the channel estimation. Simulation results reveal that significant reduction in complexity can be achieved in comparison with the blind algorithm.

Part-III (Massive MIMO)

- A distributed MMSE algorithm is developed for estimation of generally correlated fading channel in massive MIMO-OFDM systems. The distributed approach involves local estimation of CIRs at each array element followed by sharing of the estimates through collaboration among array elements. The distributed approach outperforms the centralized solution in terms of communication, memory requirements and computational complexity while at the same time attains almost the same performance as the centralized (optimal) solution.
- A novel data-aided distributed MMSE channel estimation algorithm is developed for massive MIMO-OFDM systems that can significantly reduce the pilot overhead and/or enhance the channel estimation performance. The data-aided approach increases the number of measurements without adding

more pilots by carefully picking the most reliable data-carriers after the hard-decision decoding process.

- The impact of pilot contamination on MSE performance of different algorithms is quantified by using tools from stochastic geometry. Specifically, the analytical expressions for MSE are derived in presence of AWGN and pilot contamination resulting from interfering users in a multi-cell system. The derived expressions are validated by numerical simulations.

Rest of the thesis is organized as follows: Chapter 2 is concerned with pilot based channel estimation for SISO OFDM systems where a low complexity MMSE-based algorithm is presented. Chapter 3 is focused on blind and semi-blind algorithms for MIMO OFDM systems. Chapter 4 is devoted to various channel estimation techniques (e.g., LS or MMSE) for massive MIMO OFDM systems. Specifically, distributed MMSE algorithms for estimation of correlated Rayleigh fading channels based on pilots and data-aided techniques will be presented. The impact of pilot contamination in massive MIMO is studied in Chapter 5 where the MSE performance of various algorithms is derived under both AWGN and pilot contamination effects. The Chapter concludes with intensive simulation results to validate the theoretical analysis. Finally, the general conclusions and future directions of research are discussed in Chapter 6.

CHAPTER 2

LOW-COMPLEXITY CHANNEL INTERPOLATION FOR SISO-OFDM SYSTEMS

Incorporating OFDM technology in wireless systems greatly simplifies the receiver structure as it converts a frequency selective channel into a number of parallel flat-fading channels each requiring a single-tap equalizer. Thus all that is required is the accurate estimate of each channel frequency response (CFR) coefficient to perform coherent detection at the receiver. This Chapter deals with estimating the CFR with the aid of pilots with a comb type pilot arrangement i.e., the equispaced pilots are interleaved with the data in an OFDM symbol (see Fig. 2.2). Channel estimation is accomplished by first estimating the CFR coefficients at pilot locations using LS and then interpolated at other sub-carriers using some form of interpolation. This Chapter briefly describes various interpolation techniques

keeping an eye on their estimation performance and complexity trade-offs. Then, an optimal LMMSE type channel interpolator is presented that is demonstrated to be superior to existing interpolation schemes both in terms of computations and MSE performance.

The key idea behind our approach is the exploitation of the correlation structure of CFR matrix to reduce the computational complexity. It is shown that if the pilots are placed appropriately across OFDM subcarriers, then the matrix inversion in LMMSE interpolation can be completely avoided. Further reduction in complexity can be traded-off with performance by restricting the interpolation depth of the algorithm.

2.1 System Model

Consider SISO-OFDM system as depicted in Fig. 2.1. Assume that OFDM system has N subcarriers and let \mathbf{x} represent N -dimensional information symbols drawn from certain constellation (e.g., M-QAM) so that after IFFT operation the time-domain OFDM symbol can be written as:

$$\mathbf{x} = \mathbf{F}^H \boldsymbol{\mathcal{X}} \quad (2.1)$$

where \mathbf{F} is a unitary FFT matrix whose (l, k) th entry is defined as $f_{l,k} = N^{-1/2} e^{-j2\pi lk/N}$, $l, k = 0, 1, 2, \dots, N - 1$ for an N -dimensional Fourier Transform. The multi-path Rayleigh fading channel is modeled by

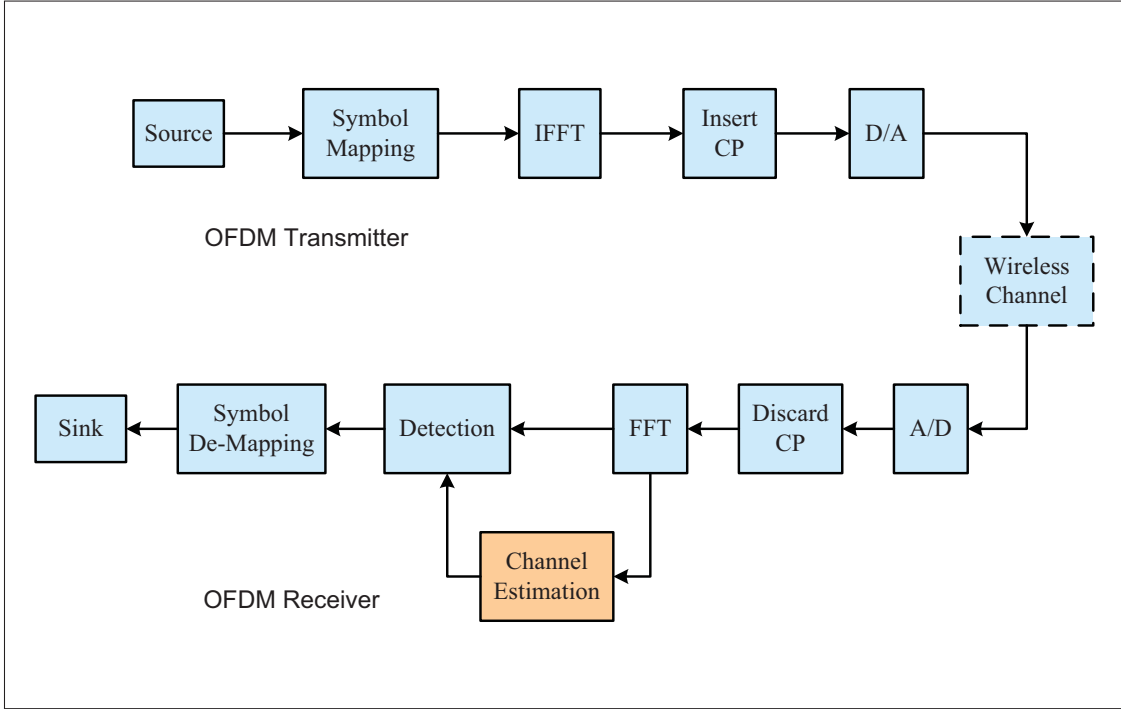


Figure 2.1: Simplified block diagram of OFDM system

a Gaussian L -tap channel impulse response (CIR) vector defined by $\mathbf{h} \triangleq [h(0), h(1), \dots, h(L-1)]^T \in \mathbb{C}^{L \times 1}$, where $h(l) \in \mathbb{C}$ represents l -th tap complex channel gain. In OFDM, the time-domain symbol \mathbf{x} is transmitted after inserting a cyclic prefix (CP) of length at least $L-1$ to avoid inter-symbol-interference (ISI). The received OFDM symbol after discarding the CP and FFT operation can be described as

$$\mathbf{y} = \text{diag}(\mathbf{x})\mathbf{H} + \mathbf{w} \quad (2.2)$$

where \mathbf{y} is the received OFDM symbol in the frequency domain, $\text{diag}(\mathbf{x})$ is the diagonal matrix of transmitted information symbols, \mathbf{H} is CFR vector and \mathbf{w} is complex AWGN with zero mean and covariance matrix $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_N$.

In pilot-aided channel estimation schemes some subcarriers are reserved for

pilots or training symbols, which are known to the receiver. Henceforth, these subcarriers will be called pilot subcarriers. Let \mathcal{P} and \mathcal{D} represent the set of pilot and data indices with cardinalities $|\mathcal{P}| = K$ and $|\mathcal{D}| = N - K$ respectively. The choice of set \mathcal{P} , or which subcarriers should be reserved for pilots, is also crucial. In slow varying channels i.e., the channels that almost remain constant during several OFDM symbols, the block type pilot arrangement as shown in Fig. 2.2(a) is usually considered. The block type pilot arrangement is used in WLAN standard IEEE 802.11a [5]. For fast varying channels, pilots have to be inserted in each OFDM symbol as shown in Fig. 2.2(b). This type of pilot arrangement is known as comb type pilot arrangement, which is used in WLAN standard IEEE 802.11g. The other standards, like multi-user WiMAX standard IEEE 802.16a uses hybrid pilots arrangement as depicted in Fig. 2.2(c). In this thesis, we will use comb-type pilot arrangement which is more appropriate for fast varying channels that may change according to symbol-by-symbol basis. Further, the equi-spaced pilots will be considered which results in optimal performance in the presence of AWGN [60, 61].

For the set of K pilot indices, represented by set \mathcal{P} , the system equation (2.2) can be written as

$$\mathbf{Y}(\mathcal{P}) = \text{diag}(\mathbf{X}(\mathcal{P}))\mathbf{H}(\mathcal{P}) + \mathbf{W}(\mathcal{P}) = \mathbf{A}\mathbf{H}(\mathcal{P}) + \mathbf{W}(\mathcal{P}) \quad (2.3)$$

where $\mathbf{A} \triangleq \text{diag}(\mathbf{X}(\mathcal{P}))$ and $\mathbf{Y}(\mathcal{P})$, $\mathbf{H}(\mathcal{P})$ and $\mathbf{W}(\mathcal{P})$ are formed by selecting the entries of \mathbf{Y} , \mathbf{H} and \mathbf{W} indexed by \mathcal{P} .

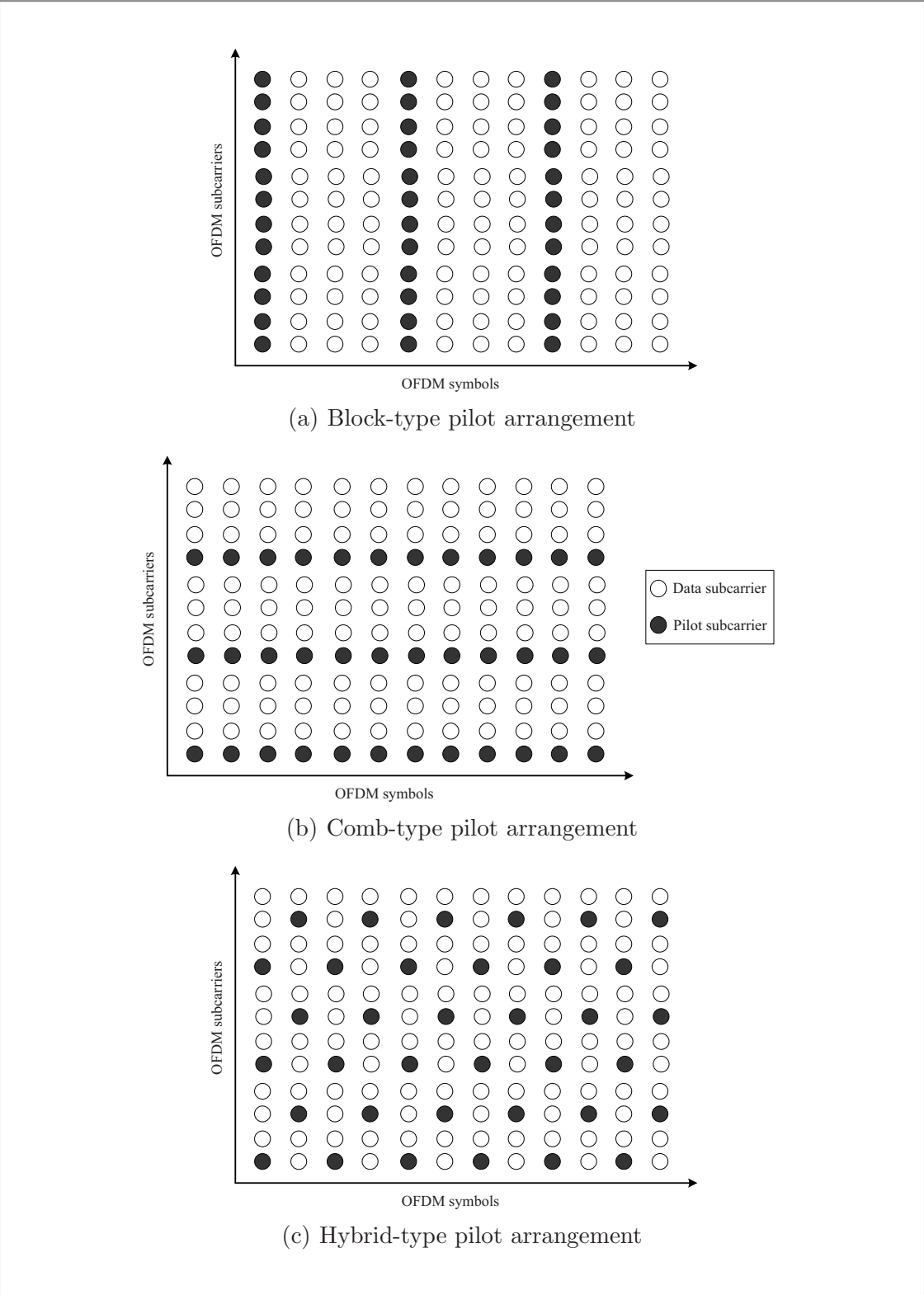


Figure 2.2: Different types of pilot arrangements used in OFDM systems.

The CFR coefficients at pilot sub-carriers can be obtained using LS or LMMSE estimation. We use the LS approach due to its low complexity and simplicity. The LS solution of (2.3) is given by

$$\begin{aligned}\widehat{\mathbf{H}}^{LS}(\mathcal{P}) &= (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}(\mathcal{P}) \\ &= \mathbf{A}^{-1} \mathbf{y}(\mathcal{P})\end{aligned}\tag{2.4}$$

In other words, owing to the diagonal structure of data matrix \mathbf{A} , we have

$$\widehat{\mathbf{H}}^{LS}(k) = \frac{\mathbf{y}(k)}{\mathbf{x}(k)}, \quad k \in \mathcal{P}\tag{2.5}$$

Once the CFR coefficients at pilot sub-carriers are determined, the remaining channel coefficients at data sub-carriers are estimated using various interpolation techniques. Fig. 2.3 shows CFR coefficients for $N = 64$ sub-carriers in OFDM symbol with equi-spaced pilots positioned at $\mathcal{P} = \{2, 6, 10, \dots, 62\}$ and data carriers with indices $\mathcal{D} = \{1, 3, 4, 5, 7, 8, \dots, 63, 64\}$. Referring to Fig. 2.3, we summarize different interpolation techniques.

2.1.1 Linear Interpolation

Linear interpolation (LI) is the simplest of all interpolation techniques where CFR coefficients are obtained by piece-wise linear approximation. Consider any two known complex CFR channel coefficients $\widehat{\mathbf{H}}(k)$ and $\widehat{\mathbf{H}}(k+1)$, where $k \in \mathcal{P}$. Without loss of generality we assume that there are M unknown channel

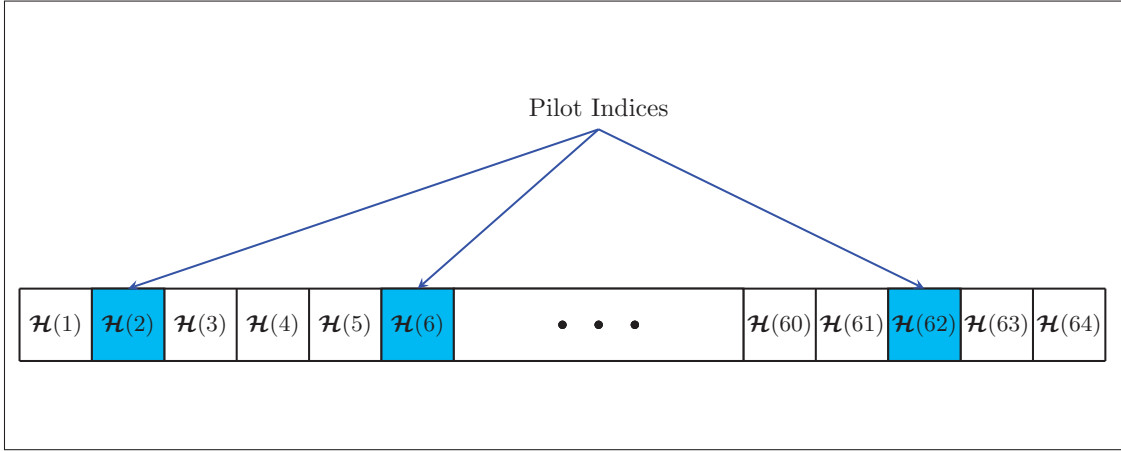


Figure 2.3: CFR coefficients in an OFDM symbol of $N = 64$ sub-carriers and $K = 16$ equi-spaced pilots

coefficients between two consecutive pilots with indices k and $k + 1$. Then the LI of these unknown coefficients is given by:

$$\hat{\mathbf{H}}^{LI}(n) = \hat{\mathbf{H}}(k) + \frac{n+1}{M+1} \left(\hat{\mathbf{H}}(k+1) - \hat{\mathbf{H}}(k) \right), \quad (2.6)$$

where, $n = 0, 1, 2, \dots, M - 1 \in \mathcal{D}$.

2.1.2 Polar Linear Interpolation

Polar Linear interpolation (PLI) exploits both magnitude and phase information of LS estimates at pilot locations. In PLI, first each CFR coefficient is separated into amplitude and phase. The linear interpolation is then performed separately on the amplitudes and the phases to estimate CFR coefficients at data carriers. Again consider two consecutive known CFR coefficients at pilot indices k and $k+1$

expressed in polar form as:

$$\begin{aligned}\widehat{\mathcal{H}}(k) &= \hat{a}(k)e^{j\hat{\phi}(k)} \\ \widehat{\mathcal{H}}(k+1) &= \hat{a}(k+1)e^{j\hat{\phi}(k+1)}\end{aligned}\tag{2.7}$$

Linear interpolation of amplitudes and phases of M CFR coefficients between consecutive pilots is then given by [22]

$$\begin{aligned}\hat{a}^{PLI}(n) &= \hat{a}(k) + \frac{n+1}{M+1}(\hat{a}(k+1) - \hat{a}(k)), \\ \hat{\phi}^{PLI}(n) &= \hat{\phi}(k) + \frac{n+1}{M+1}(\hat{\phi}(k+1) - \hat{\phi}^{LS}(k)),\end{aligned}\tag{2.8}$$

where, $n = 0, 1, 2, \dots, M-1 \in \mathcal{D}$. Just like LI, PLI is also very simple and easy to implement for real systems. When the variation in CFR phases is small, PLI is expected to yield better performance than LI. Essentially, PLI aims to track the true trajectory from $\widehat{\mathcal{H}}(k)$ to $\widehat{\mathcal{H}}(k+1)$, $k \in \mathcal{P}$ in complex plane as opposed to LI which follows the straight path. The major problem in PLI is that when the variation in CFR phases is large (typically larger than π), then the PLI cannot detect changes in direction of polar route (hence does not follow the true trajectory). The performance of PLI is even worse than LI in such situations.

2.1.3 Adaptive Polar Linear Interpolation

Build around PLI, Adaptive polar linear interpolation (APLI) is a new geometric based algorithm recently proposed in [25] which alleviates the weaknesses of PLI

by involving more pilots (instead of just 2) and introducing additional steps tailored to forcing the PLI to follow the true trajectory. It is observed that APLI outperforms both LI and PLI provided that the CFR is smooth enough and the CFR coefficients are highly correlated. This happens when $L \ll K < N$ which requires increasing the pilot density, thus compromising the system throughput. In short, APLI, like LI and PLI is not appropriate for high frequency selective channels. Moreover, neither of these simple interpolation techniques is capable of achieving the optimal performance, which makes the optimal LMMSE based approaches more attractive. However, they do serve as a benchmark for computational complexity.

2.1.4 LMMSE Interpolation and Variants

The simple approaches described earlier are not optimal in the sense of minimizing the MSE as they do not exploit the channel correlations and noise statistics. The LMMSE based interpolation is widely used on OFDM channel estimation and it is optimum in terms of MSE in the presence of AWGN. Beginning with the system model in (2.3), the LMMSE estimator of CFR is obtained by minimizing the MSE criteria

$$\hat{\mathcal{H}}^{LMMSE} = \underbrace{\operatorname{argmin}}_{\mathcal{H}} \mathbb{E} \left\{ \|\mathcal{H} - \hat{\mathcal{H}}\|^2 \right\} \quad (2.9)$$

The solution is given by [62]:

$$\begin{aligned}\widehat{\mathbf{H}}^{LMMSE} &= \mathbf{R}_{\mathcal{H}\mathcal{Y}(\mathcal{P})} \mathbf{R}_{\mathcal{Y}(\mathcal{P})\mathcal{Y}(\mathcal{P})}^{-1} \mathcal{Y}_{\mathcal{P}} \\ &= \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \mathbf{A}^H \left[\sigma_w^2 \mathbf{I}_K + \mathbf{A} \mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})} \mathbf{A}^H \right]^{-1} \mathcal{Y}(\mathcal{P})\end{aligned}\quad (2.10)$$

where $\mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})}$ is the cross correlation matrix between channel coefficients at all sub-carriers and coefficients at the pilot sub-carriers while $\mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})}$ is the auto-correlation matrix of channel coefficients at pilot sub-carriers. By simple algebraic manipulations, (2.10) can be re-written as,

$$\widehat{\mathbf{H}}^{LMMSE} = \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \left[\mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})} + \sigma_w^2 (\mathbf{A}^H \mathbf{A})^{-1} \right]^{-1} \widehat{\mathbf{H}}^{LS}(\mathcal{P}) \quad (2.11)$$

The above formulation in (2.11) suggests that LMMSE interpolates/smoothes the LS estimates over all OFDM sub-carriers. If the noise is AWGN, it is also the optimal interpolator. However, due to the dependency of LMMSE on transmitted data, the non-trivial matrix inversion and large matrix products would be required for each estimation process. The dependence of LMMSE on data can be removed by assuming that transmitted signals use the same signal constellation so that the expression $(\mathbf{A}^H \mathbf{A})^{-1}$ can be replaced by taking its expected value as follows

$$\sigma_w^2 \mathbb{E} \left\{ (\mathbf{A}^H \mathbf{A})^{-1} \right\} = \frac{\beta}{\rho} \mathbf{I}_K \quad (2.12)$$

where, $\beta = \mathbb{E}\{|\mathcal{X}(k)|^2\}\mathbb{E}\{1/|\mathcal{X}(k)|^2\}$, $\mathcal{X}(k)$ is the constellation point and ρ is the SNR. For BPSK, QPSK constellations, $\beta = 1$ and for 16-QAM constellation $\beta = 17/9$. Substitution of (2.12) into (2.11) yields:

$$\hat{\mathcal{H}}^{LMMSE} = \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \left[\mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})} + \frac{\beta}{\rho} \mathbf{I}_K \right]^{-1} \hat{\mathcal{H}}^{LS}(\mathcal{P}) \quad (2.13)$$

The resulting minimum MSE is given by

$$\text{MSE} = \text{trace} \left(\mathbf{R}_{\mathcal{H}\mathcal{H}} - \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \left[\mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})} + \frac{\beta}{\rho} \mathbf{I}_K \right]^{-1} \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})}^H \right) \quad (2.14)$$

Although (2.13) is simple, it still requires multiplications of the order $O(K^3)$ due to matrix inversion. Different variants of LMMSE have been proposed to reduce the complexity of LMMSE.

Subspace Methods

The subspace methods are based on singular value decomposition (SVD) [30, 31, 63, 64]. Since the CIR length is generally much smaller than the total number of subcarriers, the SVD of autocorrelation and cross correlation matrices in LMMSE estimator results in only few significant singular values corresponding to CIR coefficients. Therefore, retaining only few singular values (say $r \ll N$) can result in significant reduction in complexity. However the computation of SVD of auto/cross correlation matrices is by itself very complex, requiring multiplications of the order $O(K^3)$. This makes SVD based approaches impractical for real time

implementations.

Approximate LMMSE Algorithm

The original LMMSE involves matrix inversion of size $K \times K$ and multiplication of two huge matrices of sizes $N \times K$ and $K \times 1$. Both of these operations require intensive computations in practical systems where N is large, e.g. in LTE $N = 1200$ and $K = 200$. To overcome this, Approximate LMMSE algorithm (ALMMSE) was proposed in [32] for WiMAX which was later adopted for LTE [33]. The idea is to split the whole band into three portions; 1 to $M/2$, $M/2 + 1$ to $N - M/2$ and $N - M/2 + 1$ to N , where M is predefined. At the edges of the bands (i.e. the first and the last), the LMMSE follows the expression (2.11) with K replaced by $M/6$, while at the middle band and for each subcarrier k , only the middle element of computed LMMSE estimate of size M is used i.e., the $(M/2)^{th}$ element. Although the number of computations is reduced due to this divide-and-conquer type strategy, it has two main drawbacks. The first is that $M \times M$ matrix operations are repeated approximately as the number of sub-carriers over different overlapping frequency bands. Secondly, many redundant computations for adjacent sub-carriers are also carried out which are later discarded.

2.2 The Proposed Method

We propose a structure-based approach that exploits the structure of auto/cross correlation matrices of CFR to reduce the number of computations. Since it is

well known that the number of pilot spacing is inversely proportional to channel delay spread [65], we will make a reasonable assumption that the number of pilots K , is an integer multiple of channel length, i.e., $K = qL$ for some integer $q \geq 1$. This gives the pilot spacing of $N/(qL)$ which is inversely proportional to channel length. Based on this, the CFR correlation matrix can be easily computed as follows

$$\mathbf{R}_{\mathcal{H}\mathcal{H}} = \mathbb{E} \{ \mathcal{H}\mathcal{H}^H \} = \mathbb{E} \{ \underline{\mathbf{F}} \mathbf{h} \mathbf{h}^H \underline{\mathbf{F}}^H \} = \underline{\mathbf{F}} \mathbb{E} \{ \mathbf{h} \mathbf{h}^H \} \underline{\mathbf{F}}^H = \underline{\mathbf{F}} \mathbf{R}_{\mathbf{h}} \underline{\mathbf{F}}^H \quad (2.15)$$

where $\underline{\mathbf{F}}$ is the partial FFT matrix consisting of first L columns of \mathbf{F} and $\mathbf{R}_{\mathbf{h}}$ is the autocorrelation matrix of CIR vector \mathbf{h} . Observe that the auto/cross correlation matrices in (2.13) are the subsets of the full correlation matrix given in (2.15). The design of the proposed estimator is based on the assumption that CIR coefficients are uncorrelated, so that $\mathbf{R}_{\mathbf{h}} = \sigma_h^2 \mathbf{I}_L$ where σ_h^2 represents the channel variance that will be assumed to be unity. This corresponds to the channel exhibiting a uniform PDP. Based on this assumption, (2.15) simplifies to

$$\mathbf{R}_{\mathcal{H}\mathcal{H}} = \underline{\mathbf{F}} \underline{\mathbf{F}}^H \quad (2.16)$$

By using the definition of $\underline{\mathbf{F}}$ in (2.16), the correlation between any two CFR coefficients can be analytically computed and in fact it can be easily shown that

the magnitude of (i, l) -th element of $\mathbf{R}_{\mathcal{H}\mathcal{H}}$ is:

$$|[\mathbf{R}_{\mathcal{H}\mathcal{H}}]_{i,l}| = |\mathbf{a}_i \mathbf{a}_l^H| = \begin{cases} L & \text{if } i = l \\ \frac{1}{L} \left| \frac{\sin(\pi(i-l)L/N)}{\sin(\pi(i-l)/N)} \right| & \text{if } i \neq l \end{cases} \quad (2.17)$$

where \mathbf{a}_i represents the i^{th} row of \mathbf{F} . In Fig. 2.4 we plot this magnitude correlation function for $i = 1$, $N = 64$ and $L = 16$ as a function of l i.e., the first row of $\mathbf{R}_{\mathcal{H}\mathcal{H}}$.

From Fig. 2.4, it is clear that correlation between any two coefficients which are N/L (or its integer multiples) apart is zero. So, if we place the pilots at those very positions then the autocorrelation matrix $\mathbf{R}_{\mathcal{H}(\mathcal{P})\mathcal{H}(\mathcal{P})}$ would essentially become an identity matrix. Hence the matrix inversion in LMMSE estimation process would be trivial, thanks to the special structure of FFT matrix. In practical OFDM systems where N is large, this can result in significant computational advantages. Exploiting of above facts is carried out in the proposed solutions as described below for different choices of parameter q (or equivalently K).

The Case When $q = 1$

Evidently when $q = 1$, the pilot spacing would be N/L , so that one would end up with $K = L$ pilots. Further, from (2.13), due to orthogonality of CFR coefficients at pilot positions, we get

$$\hat{\mathcal{H}} = \frac{\mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \hat{\mathcal{H}}^{LS}(\mathcal{P})}{(1 + \beta/\rho)} \quad (2.18)$$

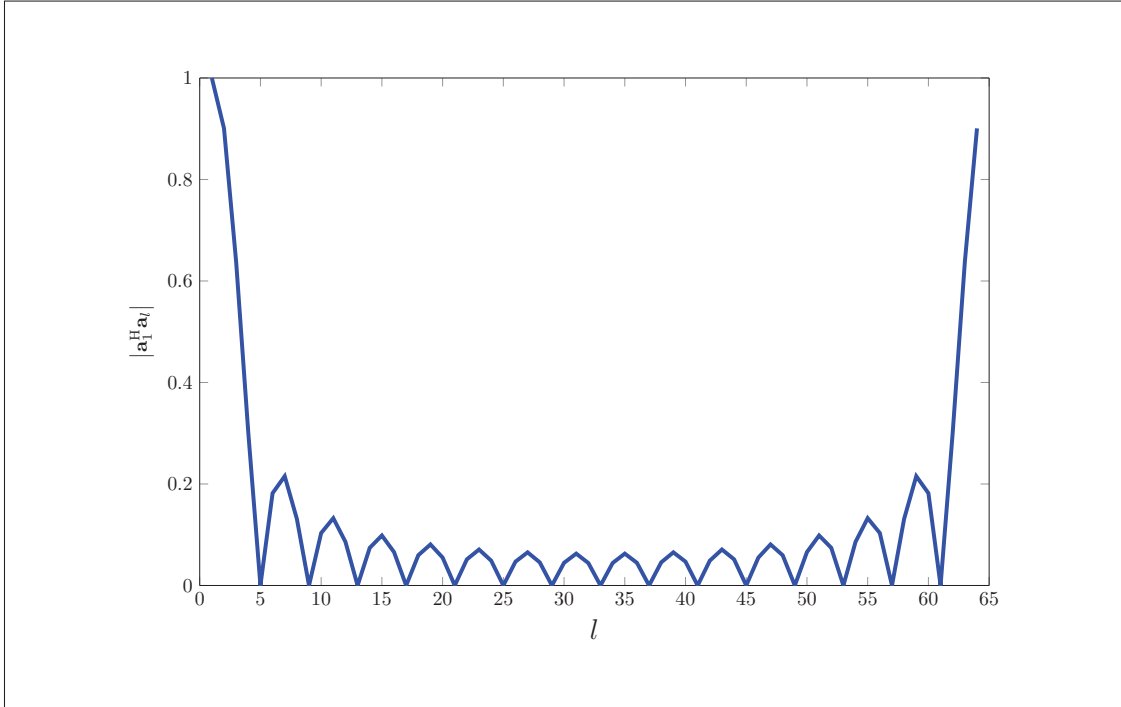


Figure 2.4: Normalized correlation as function of l for $N=64$ and $L=16$

where matrix $\mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})}$ that will be referred to as interpolation matrix, can be easily read-off from $\mathbf{R}_{\mathcal{H}\mathcal{H}}$. This choice of K renders our method spectrally efficient as only few pilots (i.e., equal to the number of unknown channel coefficients) are required to achieve the optimal performance. The resulting minimum MSE from (2.14) is given by

$$\text{MSE} = \text{trace} \left(\mathbf{R}_{\mathcal{H}\mathcal{H}} - \frac{\mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})} \mathbf{R}_{\mathcal{H}\mathcal{H}(\mathcal{P})}^H}{1 + \beta/\rho} \right) \quad (2.19)$$

The Case When $q > 1$

Despite the fact that utilizing less number of pilots is spectrally more efficient, in many practical systems (e.g., LTE) the number of pilots are considered to be much larger than the length of CIR, in order to obtain fairly accurate channel

estimates. Let the number of pilots be $K = qL$ and for $q > 1$, the pilot spacing would be $N/(qL)$. In this case the consecutive pilot subcarriers would not be orthogonal, since the orthogonality is only guaranteed when the pilot spacing is an integer multiple of N/L . However, one can always construct q disjoint sets of pilot indices such that the pilot spacing in each set is exactly N/L and thus the carriers within each set are orthogonal. For example, if $N = 64$, $L = 8$ and $q = 2$, the pilot spacing is 4 which gives the set of pilot indices \mathcal{P} as depicted in Fig. 2.3. This set can be further decomposed into two disjoint sets, $\mathcal{P}^1 = \{2, 10, \dots, 58\}$ and $\mathcal{P}^2 = \{6, 14, \dots, 62\}$, each fulfilling the orthogonality condition. Hence, (2.18) can be applied to each set of pilot indices to obtain q estimates of CFR coefficients corresponding to each set. The final CFR estimate is then obtained by combining the individual CFR estimates by means of simple averaging as given in the following equations,

$$\begin{aligned} \widehat{\mathbf{h}} &= \frac{1}{q} \sum_{i=1}^q \widehat{\mathbf{h}}^{(i)} \\ &= \frac{1}{q(1 + \beta/\rho)} \sum_{i=1}^q \mathbf{R}_{\mathcal{H}\mathcal{H}_{\mathcal{P}}}^{(i)} \widehat{\mathbf{h}}_{\mathcal{P}}^{i(LS)} \end{aligned} \quad (2.20)$$

where, $\widehat{\mathbf{h}}_{\mathcal{P}}^{i(LS)}$ is the LS estimate of CFR at the i th set of pilot indices, \mathcal{P}^i . Note that (2.20) is more general than (2.18) and applicable for $q = 1$. Also observe that the interpolation matrices (i.e., $\mathbf{R}_{\mathcal{H}\mathcal{H}_{\mathcal{P}}}^{(i)}$) associated with different sets of indices are related to each other, such that given for the i -th set, the rest can be obtained by simple shift operations. This alleviates the requirement to compute them

individually for all the sets, thus reducing the computational complexity.

2.2.1 Further Reducing the Complexity

Despite avoiding matrix inversion, (2.18) and (2.20) still require large matrix-vector product of size $N \times K$ i.e., the computation of each CFR coefficient requires K complex multiplications. This is because all the LS channel estimates are involved in the interpolation process. The fact that the correlations among CFR coefficients decrease exponentially (see Fig. 2.4), we can take only few LS estimates in the interpolation process. This allows us to further reduce the complexity of the proposed algorithm. Hence, by defining a parameter $d \leq K$, the interpolation depth parameter, as the number of LS estimates centered around a CFR coefficient, the number of multiplications can be reduced from K to d . As the depth increases the performance improves but at the cost of increasing the complexity. Fig. 2.5 describes the pictorial representation of interpolation process for $d = 1$ and $d = 2$ respectively. Note that the computational complexity of the proposed algorithm is comparable to the benchmark (i.e., simple linear interpolation) schemes for $d = 2$.

2.3 Simulation Results

For simulations we consider BPSK, QPSK and QAM modulation schemes with FFT size N varying from 64 to 1024 with pilots assumed to be uniformly distributed in an OFDM symbol. The channel is assumed to vary independently

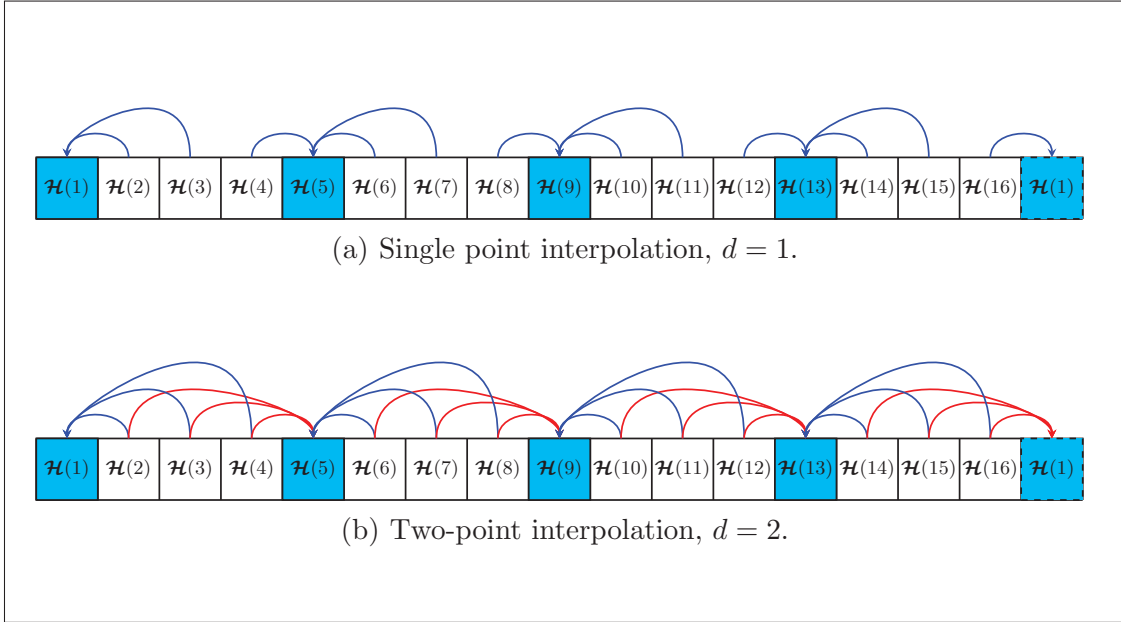


Figure 2.5: Pictorial representation of interpolation process.

from one symbol to another but assumed to be stationary within an OFDM symbol duration. The MSE performance is evaluated based on following criteria,

$$\text{NMSE} = \frac{1}{\Theta} \sum_{i=1}^{\Theta} \frac{\|\mathcal{H}^i - \hat{\mathcal{H}}^i\|^2}{\|\mathcal{H}^i\|^2} \quad (2.21)$$

where, \mathcal{H}^i and $\hat{\mathcal{H}}^i$ are true and estimated CFR in the i th trial respectively, and Θ is the total number of trials. We use $\Theta = 500$ in our simulations.

Impact of Interpolation Depth

Fig. 2.6(a) shows the MSE performance of the proposed algorithm with various levels of interpolation depth determined by parameter d . It is clear that the MSE decreases with increasing the interpolation depth and vice versa. The case $d = 1$ corresponds to the lowest complexity with just one multiplication per CFR

coefficient. The BER vs SNR in Fig. 2.6(b) shows a similar behaviour.

Impact of PDP Mismatch

Since the proposed algorithm was designed based on uniform PDP, it is important to investigate the performance under non-uniform channel PDP. Fig. 2.7(a) compares the performance when channel correlation matrix has more general form $[\mathbf{R}_{\mathbf{h}}]_{i,j} = a^{|i-j|}$, where a is the correlation parameter such that $0 \leq a \leq 1$. Thus $a = 0$ corresponds to uniform PDP while $a = 1$ corresponds to the perfect correlation. The results in Fig. 2.7(a) show that there is only marginal loss in performance for different choices of a . In Fig. 2.7(b), we evaluate the proposed estimator for the exponential PDP i.e., $\mathbb{E}\{|h(k)|^2\} = e^{-0.1k}, k = 0, 1, \dots, L - 1$. Again, the result indicates that the proposed estimator is robust against PDP mismatch.

The Impact of Channel Spreading and Pilots

The frequency selectivity of the channel increases with increasing the channel delay spread L . First, we compare the MSE performance of the proposed algorithm by fixing the number of pilots K and varying the channel length L . The simulation results at SNR of 20dB in Fig. 2.8(a) show that the MSE performance deteriorates with increasing the CIR length, which is typical of LMMSE estimation. Also it is obvious that the proposed method almost attains LMMSE performance and is robust against frequency selectivity.

Next we fix the CIR length to 8 and vary the number of pilots at SNR of

20dB. Obviously, the MSE performance improves by increasing the number of pilot measurements as depicted in Fig. 2.8(b). Again observe that there is a close match between the proposed algorithm and LMMSE estimator.

Comparison With Other Approaches

We compare the performance of proposed algorithm with various channel interpolation techniques such as LI, PLI and DFT based interpolation. In all algorithms CFR estimates are computed using LS. The results are plotted for two different values of CIR lengths i.e., 8 and 32 in Fig. 2.9 and Fig. 2.10, respectively. It can be seen that the proposed algorithm attains LMMSE performance at all SNR values and that too with very low complexity. The MSE of simple interpolation techniques reaches an error floor at higher SNRs. Their performance is generally better over low frequency selective channels (i.e., when L is small), where the CFR is rather smooth enough as shown in Fig.2.9(b). However when the channel exhibits high frequency selective behaviour as depicted in Fig.2.10(b), simple interpolation techniques suffer from huge performance loss as indicated in Fig.2.10(a). Further, the DFT-based interpolation attains LMMSE performance only at higher SNR and that too at comparatively higher complexity (because the FFT has to be done twice).

Complexity Vs Performance

The complexity and performance comparison of various algorithm is summarized in Table 2.1. Evidently, the conventional MMSE based techniques or those based

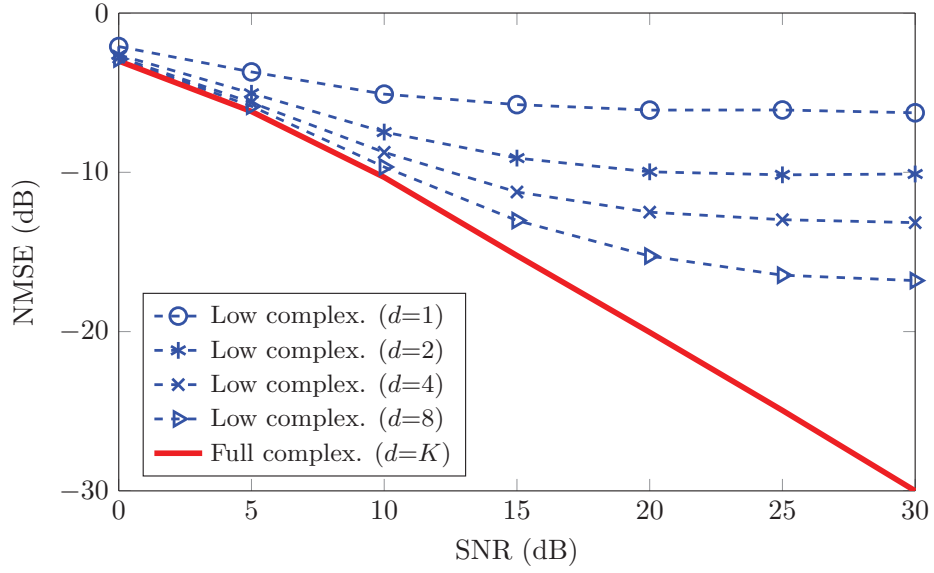
on SVD yield optimal MSE performance but are too complex to be realized. The simple interpolation techniques have the advantage of low complexity but at the cost of performance. More advanced interpolation techniques yield better performance but at the cost of higher complexity than simple interpolation schemes. The proposed algorithm has the best trade-off between complexity and performance as it does not require any matrix inversion nor any channel statistics.

Table 2.1: Complexity-Performance Trade-off

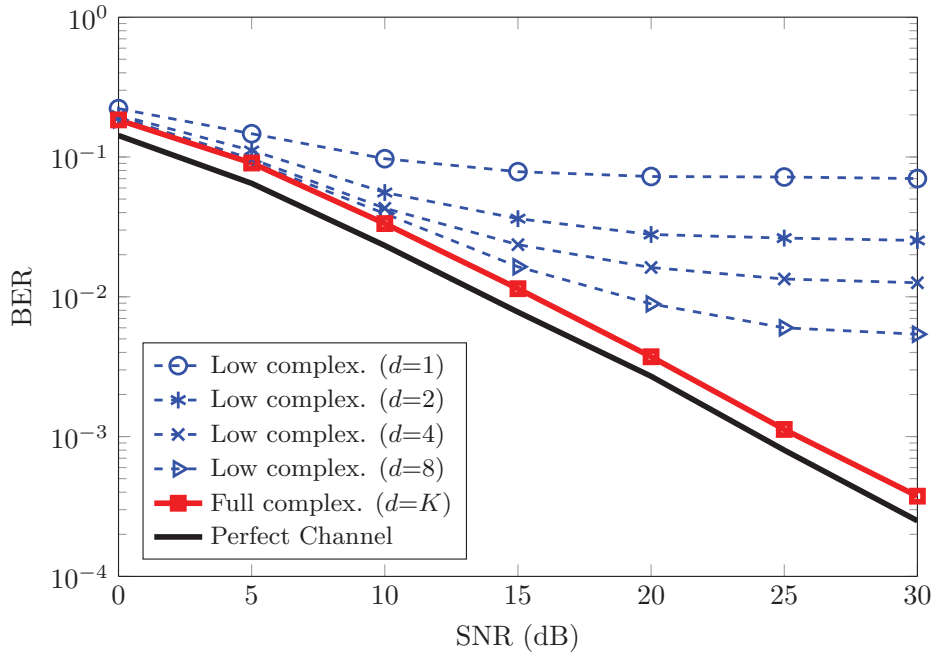
CE methods	Need Matrix Inversion	Complexity	MSE performance
LS in TD	No	Low	Not Optimal
LMMSE (conventional)	Yes	Very high	Optimal
SVD based	No	Very high	Optimal
DFT/IDFT based	No	High	Near Optimal
Simple Interpolation schemes	No	Very low	Not Optimal
Sophisticated Interpolation schemes	No	High	Sub-Optimal
Proposed	No	Low	Near Optimal

2.4 Concluding Remarks

In this Chapter we proposed a simple yet a novel strategy to reduce the complexity of LMMSE based interpolation for channel estimation in OFDM systems. The key to achieving the objective was to use the inherent structure of channel correlation in the frequency domain. Moreover, the correlations required for interpolation process could be computed off-line, since they are designed on the basis of uniform PDP. The simulation results demonstrate that if pilots are placed appropriately i.e., $q(N/L)$ apart, then the matrix inversion in LMMSE can be avoided without losing MSE performance. Further, the computations due to large matrix-vector products in LMMSE can be traded off with the performance.

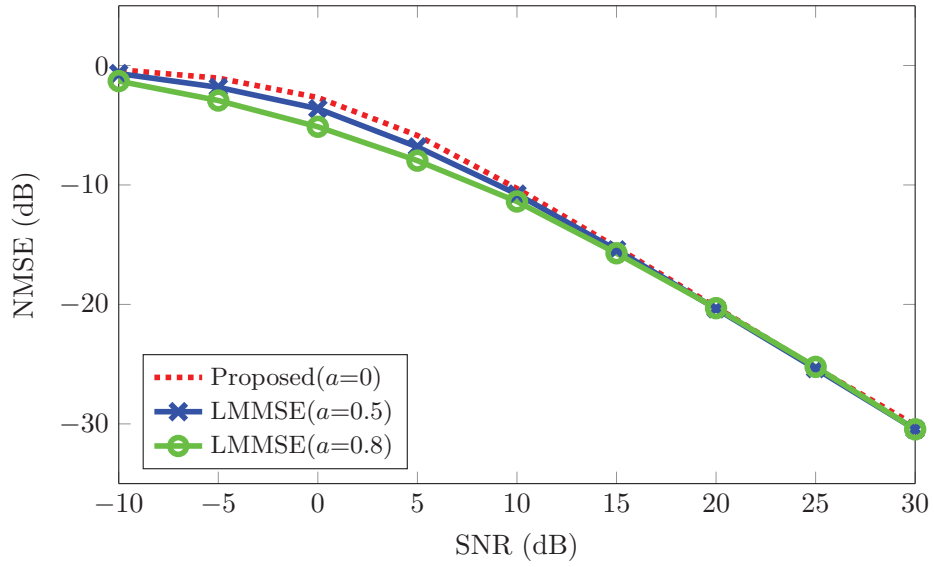


(a)

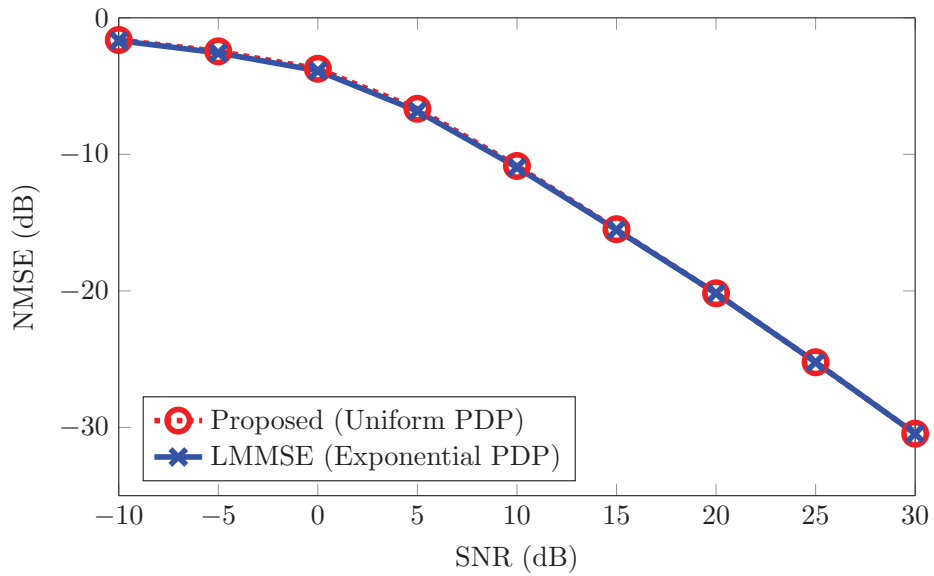


(b)

Figure 2.6: Effect of interpolation depth on (a) MSE and (b) BER performance. Number of required multiplications per CFR coefficient is proportional to d .

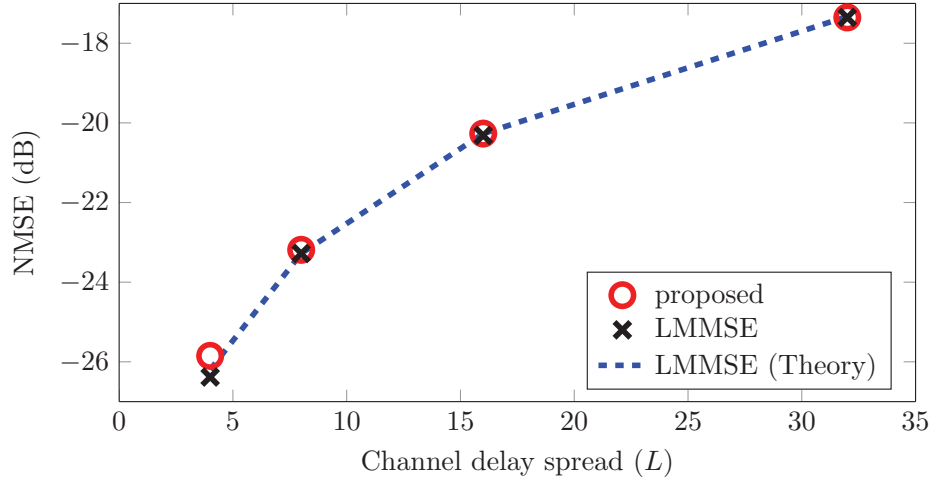


(a)

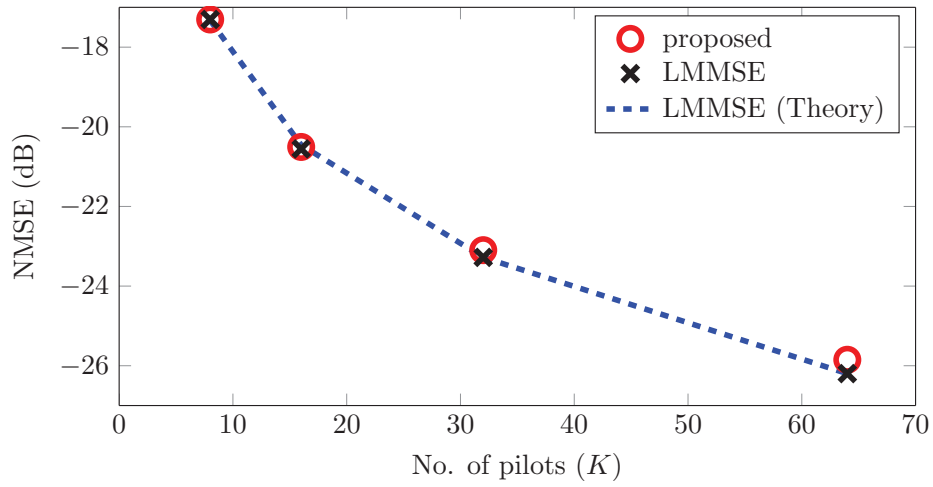


(b)

Figure 2.7: Effect of channel PDP mismatch for (a) General PDP profile and (b) Exponentially decaying profile. The parameters are $N=128$ and $L=8$ and $K=16$ with QPSK symbols.

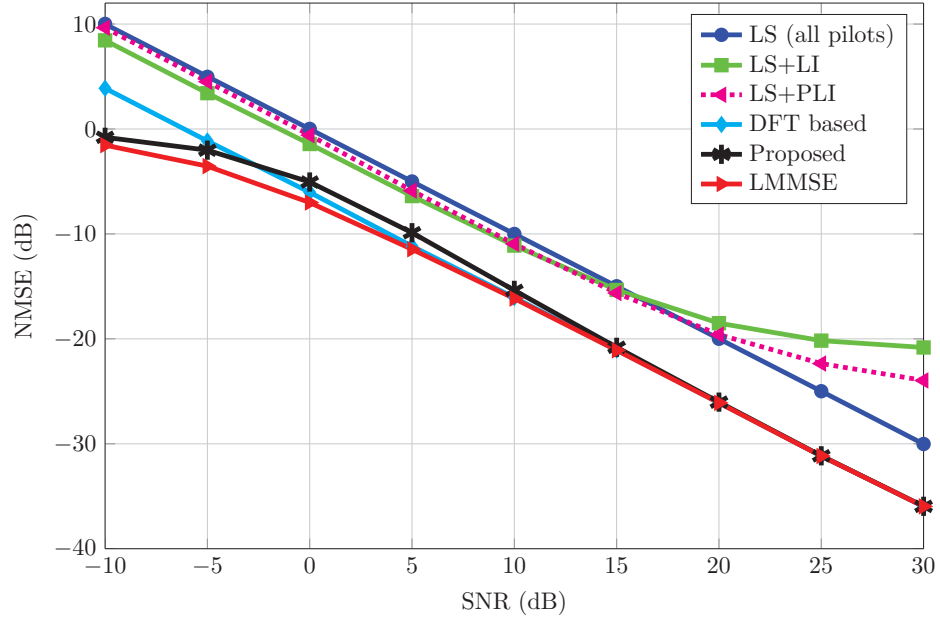


(a) $K = 32$

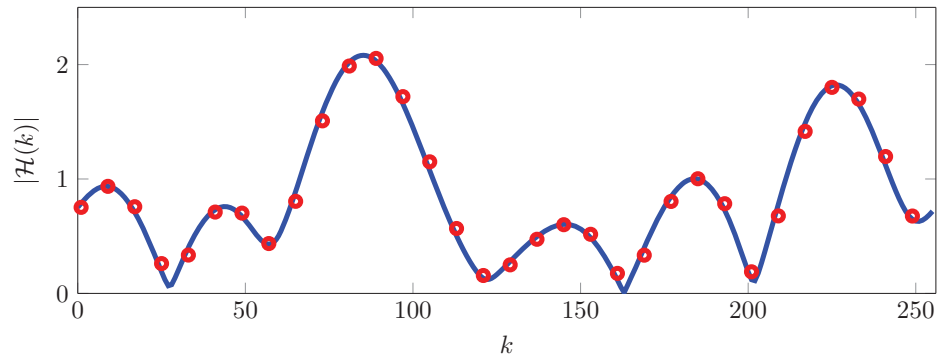


(b) $L = 8$

Figure 2.8: Effect of (a) Channel delay spread and (b) Number of pilots, on MSE performance using $N=256$, 16-QAM symbols at 20dB SNR.

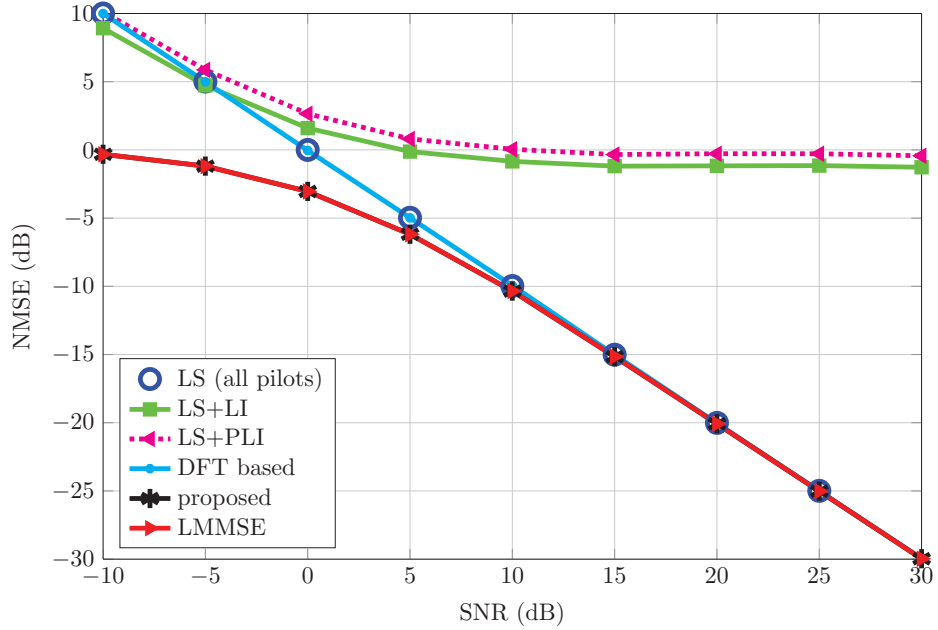


(a) $L = 8$

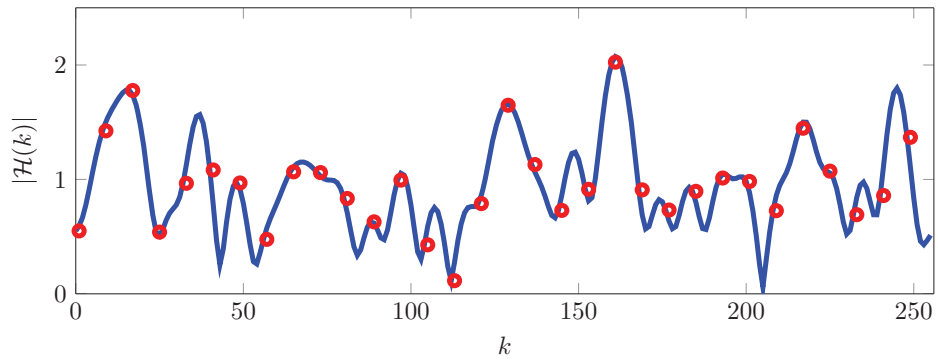


(b) $L = 8$

Figure 2.9: (Top) MSE performance comparison of the proposed method with various algorithms. The parameters are $N=256$, $K=32$ with QPSK symbols. (Bottom) Snapshot of the magnitude of CFR at each subcarrier. Red circles represent pilot locations where the CFR is known.



(a) $L = 32$



(b) $L = 32$

Figure 2.10: (Top) MSE performance comparison of the proposed method with various algorithms. The parameters are $N=256$, $K=32$ with QPSK symbols. (Bottom) Snapshot of the magnitude of CFR at each subcarrier. Red circles represent pilot locations where the CFR is known.

CHAPTER 3

BLIND AND SEMI-BLIND ML DETECTION FOR MIMO OFDM SYSTEMS

The current standards use pilot symbols to estimate the channel thus sacrificing bandwidth which otherwise would have been available for data transmission. In high mobility wireless systems, the channels may even change so rapidly that this approach will become infeasible. Blind or semi-blind detection over the time-varying wireless channels has shown to enhance the system performance considerably [66, 67]. Unlike the pilot-based techniques as discussed in the previous Chapter, this Chapter focuses on blind and semi-blind methods for channel estimation that can significantly improve the spectral efficiency of overall system.

Specifically, we investigate the joint ML data detection and channel estimation problem for Alamouti space-time-block-coded (STBC) OFDM systems. However,

the joint ML estimation and data detection is generally considered a hard combinatorial optimization problem. We propose an efficient low-complexity algorithm based on branch-estimate-bound strategy that renders exact joint ML solution. However, the computational complexity of blind algorithm becomes critical at low signal-to-noise-ratio (SNR) regime as the number of OFDM carriers and constellation size are increased. To overcome this problem a semi-blind algorithm is proposed based on subcarrier reordering according to their reliability. The proposed algorithm can reliably track the wireless Rayleigh fading channel without requiring any channel statistics. Simulation results presented against the perfect coherent detection demonstrate the effectiveness of the proposed technique.

3.1 Motivation

The use of multiple antennas at transmitter offers many advantages over single antenna systems including multiplexing gain and diversity gain [1]. Of several diversity schemes available in the literature, the major motivation for using Alamouti scheme [68] with two transmit and one receive antenna is that it is the optimum in both the capacity and the diversity. Alamouti coding achieves full spatial diversity at full transmission rate for any signal (real or complex) constellation and offers very simple receiver structures. However, to decouple the signals at the receiver side via simple decoding, the Alamouti scheme requires the channel between each transmit-receive antenna to be constant over two consecutive OFDM symbols. Moreover, when dealing with frequency selective channels, Alamouti scheme has

to be implemented over the block level.

The proposed research in blind estimation is motivated by recent works [43] and [44], where the authors have developed a low-complexity blind ML method for general constellations for SIMO and SISO systems, respectively. Specifically, we extend the work in [44] to Alamouti block-coded OFDM systems with two transmit antennas. Parallelizing the results and discussions therein, we first derive the exact blind ML algorithm and then reduce its complexity using different methods. Then a semi-blind algorithm is proposed by assuming that few training symbols are available.

3.2 Problem Formulation

Consider a single user OFDM system with two-transmit and one-receive antenna as shown in Fig. 3.1(a). The frequency selective channels from two transmit antennas to the receive antenna are modelled as finite impulse response (FIR) filters. We assume that both channels are independent Rayleigh-fading channels having maximum length L and CP length is at least $L-1$ to avoid ISI.

Let \mathcal{X} represent information symbols and that OFDM system has N sub-carriers so that after IFFT operation the time-domain information symbols can be written in vector form as:

$$\mathbf{x} = \mathbf{F}^H \mathcal{X} \quad (3.1)$$

where \mathbf{F} is unitary FFT matrix defined as $[\mathbf{F}]_{l,k} = N^{-1/2} e^{-j2\pi lk/N}$. Let the n^{th} symbol of k^{th} transmitted block from antenna i ($= 1$ or 2) be denoted by

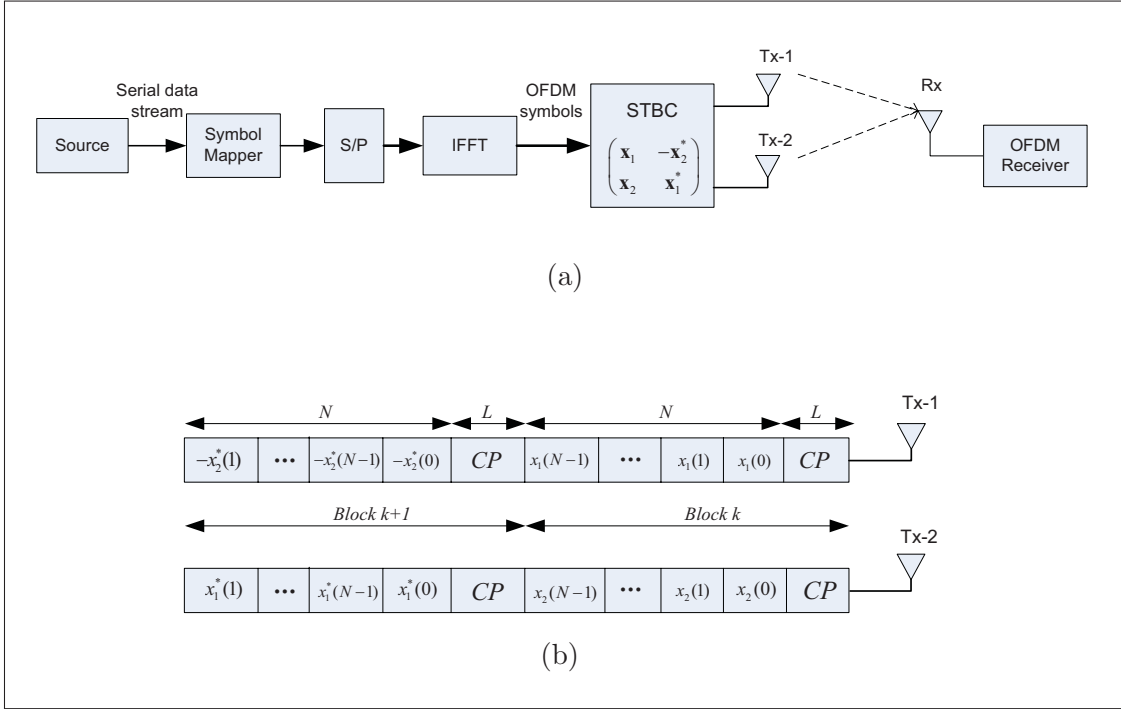


Figure 3.1: (a) Alamouti coded OFDM system (b) Frame structure of OFDM data blocks over two consecutive time instants.

$x_i^{(k)}(n), n = 0, 1, \dots, N - 1$. At times $k = 0, 2, 4, \dots$ pair of blocks $x_1^{(k)}(n)$ and $x_2^{(k)}(n)$ are generated according to the Alamouti STBC coding rule, which is defined as [68, 69]

$$\begin{aligned} x_1^{(k+1)}(n) &= -x_2^{*(k)}((n)_N) \\ x_2^{(k+1)}(n) &= x_1^{*(k)}((n)_N) \end{aligned} \quad (3.2)$$

where, $(\cdot)_N$ is modulo N operation and $(\cdot)^*$ denotes the complex conjugation operation. Each antenna transmits a data block of length N according to the above STBC scheme after appending the CP. Adding CP eliminates inter-block interference and converts linear convolution into circular convolution. The structure

of data-blocks over consecutive time instants is depicted in Fig. 3.1(b). In the presence of additive white Gaussian noise (AWGN), the received data blocks over two consecutive time instants after discarding the CPs can be written as:

$$\mathbf{y}^{(j)} = \sqrt{\rho} \mathbf{H}_1 \mathbf{x}_1^{(j)} + \sqrt{\rho} \mathbf{H}_2 \mathbf{x}_2^{(j)} + \mathbf{w}^{(j)}, j = k, k + 1 \quad (3.3)$$

where ρ is the SNR, \mathbf{H}_1 and \mathbf{H}_2 are circular channel matrices from two transmit antennas to the receive antenna and \mathbf{w} is circular symmetric AWGN with *pdf*: $\mathbf{w} \sim CN(\mathbf{0}, \mathbf{I})$. In (3.3), we also assumed that channel is static over two consecutive OFDM blocks at time instants k and $k + 1$. Specifically, the structure of two circular channel matrices is:

$$\mathbf{H}_i = \begin{pmatrix} h_i(0) & 0 & \cdots & h_i(L-1) & \cdots & h_i(1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_i(L-2) & \cdots & h_i(0) & 0 & \cdots & h_i(L-1) \\ h_i(L-1) & h_i(L-2) & \cdots & h_i(0) & 0 & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & h_i(L-1) & h_i(L-2) & \cdots & h_i(0) \end{pmatrix}$$

and where,

$$\mathbf{h}_i = \left[h_i(0) \ h_i(1) \ \cdots \ h_i(L-1) \right]^T \quad (3.4)$$

represents the impulse response sequence of i^{th} channel matrix. At the receiver side, the frequency domain received symbols after FFT operations are obtained

as:

$$\mathbf{y}^{(j)} = \sqrt{\rho} \Lambda_1 \mathbf{x}_1^{(j)} + \sqrt{\rho} \Lambda_2 \mathbf{x}_2^{(j)} + \mathbf{w}^{(j)}, j = k, k + 1 \quad (3.5)$$

where $\mathbf{x}_i^{(j)} = \mathbf{F}\mathbf{x}_i^{(j)}$, $\Lambda_i = \mathbf{F}\mathbf{H}_i\mathbf{F}^H$ are diagonal matrices whose entries are N -point DFT of \mathbf{h}_i after zero-padding and $\mathbf{w}^{(j)} = \mathbf{F}\mathbf{w}^{(j)}$. Expanding (3.5) and using DFT properties we get:

$$\begin{aligned} \mathbf{y}^{(k)} &= \sqrt{\rho} \Lambda_1 \mathbf{x}_1^{(k)} + \sqrt{\rho} \Lambda_2 \mathbf{x}_2^{(k)} + \mathbf{w}^{(k)}, \\ \mathbf{y}^{(k+1)} &= \sqrt{\rho} \Lambda_1 \mathbf{x}_1^{(k+1)} + \sqrt{\rho} \Lambda_2 \mathbf{x}_2^{(k+1)} + \mathbf{w}^{(k+1)} \end{aligned} \quad (3.6)$$

By stacking the received data symbols over consecutive intervals in one column and so as the DFT channel coefficients, (3.6) can be written in matrix-vector notation as

$$\begin{bmatrix} \mathbf{y}^{(k)} \\ \mathbf{y}^{(k+1)} \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} \text{diag}(\mathbf{x}_1^{(k)}) & \text{diag}(\mathbf{x}_2^{(k)}) \\ -\text{diag}(\mathbf{x}_2^{*(k)}) & \text{diag}(\mathbf{x}_1^{*(k)}) \end{bmatrix} \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{w}^{(k)} \\ \mathbf{w}^{(k+1)} \end{bmatrix} \quad (3.7)$$

where $\mathcal{H}_i = \text{diag}(\Lambda_i) = \mathbf{F} \begin{bmatrix} \mathbf{h}_i \\ \mathbf{0} \end{bmatrix}$. Let $\underline{\mathbf{F}}$ consists of first L columns of \mathbf{F} , then

$$\mathcal{H}_i = \underline{\mathbf{F}}\mathbf{h}_i \quad \text{and} \quad \mathbf{h}_i = \underline{\mathbf{F}}^H \mathcal{H}_i \quad (3.8)$$

which allows us to rewrite (3.7) as:

$$\underbrace{\begin{bmatrix} \mathbf{y}^{(k)} \\ \mathbf{y}^{(k+1)} \end{bmatrix}}_{\mathbf{y}} = \sqrt{\rho} \underbrace{\begin{bmatrix} \text{diag}(\boldsymbol{\mathcal{X}}_1^{(k)}) \underline{\mathbf{F}} & \text{diag}(\boldsymbol{\mathcal{X}}_2^{(k)}) \underline{\mathbf{F}} \\ -\text{diag}(\boldsymbol{\mathcal{X}}_2^{*(k)}) \underline{\mathbf{F}} & \text{diag}(\boldsymbol{\mathcal{X}}_1^{*(k)}) \underline{\mathbf{F}} \end{bmatrix}}_{\mathbf{X}_a} \underbrace{\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}}_{\mathbf{h}} + \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{W}}^{(k)} \\ \boldsymbol{\mathcal{W}}^{(k+1)} \end{bmatrix}}_{\boldsymbol{\mathcal{W}}} \quad (3.9)$$

or even more compactly as:

$$\mathbf{y} = \sqrt{\rho} \mathbf{X}_a \mathbf{h} + \boldsymbol{\mathcal{W}} \quad (3.10)$$

where \mathbf{y} and $\boldsymbol{\mathcal{W}}$ are observation and noise vectors each of size $2N \times 1$, \mathbf{X}_a is of $2N \times 2L$ data matrix, which we shall refer to as *Alamouti matrix*, and \mathbf{h} is $2L \times 1$ dimensional composite channel vector. The above model can be easily transformed to SISO-OFDM system of [44] by replacing \mathbf{X}_a with $N \times N$ square matrix $\text{diag}(\boldsymbol{\mathcal{X}})$ consisting of N data symbols on its diagonal as follows,

$$\mathbf{y} = \sqrt{\rho} \text{diag}(\boldsymbol{\mathcal{X}}) \underline{\mathbf{F}} \mathbf{h} + \boldsymbol{\mathcal{W}} \quad (3.11)$$

where \mathbf{y} and $\boldsymbol{\mathcal{W}}$ are N -dimensional received OFDM symbol and noise vector respectively while \mathbf{h} is the length- L SISO channel vector. In either case, the task of receiver is to jointly estimate the channel \mathbf{h} and the data vector $\boldsymbol{\mathcal{X}}$ given only the received data symbol \mathbf{y} .

3.3 Joint ML/MAP solution

Considering the data model in (3.10), the joint ML channel estimation and data detection problem reduces to minimizing the following objective function,

$$J_{ML} = \underset{\mathbf{h}, \boldsymbol{\mathcal{X}} \in \Omega^{2N}}{\operatorname{argmin}} \left\{ \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho} \mathbf{X}_a \mathbf{h}\|^2 \right\} \quad (3.12)$$

where Ω^{2N} denotes all possible $2N$ -dimensional signal vectors. As seen from (3.12), the joint ML problem is a combinatorial problem involving $|\Omega|^{2N}$ hypothesis tests and it is almost impossible to solve it exactly for sufficiently large Ω and N . For instance, if $N = 16$ and 4-QAM constellation is used, the exhaustive search would require to examine $4^{(32)} \approx 1.84 \times 10^{19}$ hypothesis for each coherence time of two OFDM blocks.

To solve it efficiently, we propose the following strategy. We start by decomposing the original cost function as,

$$J_{ML} = \min_{\mathbf{h}, \boldsymbol{\mathcal{X}} \in \Omega^{2N}} \left\{ \underbrace{\|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho} \mathbf{X}_{a(i)} \mathbf{h}\|^2}_{M_{\mathcal{X}_{(i)}}} + \sum_{j=i+1}^N \|\boldsymbol{\mathcal{Y}}_{(j)} - \sqrt{\rho} \mathbf{X}_{a(j)} \mathbf{h}\|^2 \right\} \quad (3.13)$$

and define,

$$M_{\mathcal{X}_{(i)}} = \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho} \mathbf{X}_{a(i)} \mathbf{h}\|^2 \quad (3.14)$$

as the partial joint ML metric for $\boldsymbol{\mathcal{X}}$ up to the index i , i.e., $\boldsymbol{\mathcal{X}}_{(i)}$ and where

$$\mathbf{X}_{a(i)} = \begin{bmatrix} \text{diag}(\boldsymbol{\mathcal{X}}_{1(i)}^{(k)}) \underline{\mathbf{F}}_{(i)} & \text{diag}(\boldsymbol{\mathcal{X}}_{2(i)}^{(k)}) \underline{\mathbf{F}}_{(i)} \\ -\text{diag}(\boldsymbol{\mathcal{X}}_{2(i)}^{*(k)}) \underline{\mathbf{F}}_{(i)} & \text{diag}(\boldsymbol{\mathcal{X}}_{1(i)}^{*(k)}) \underline{\mathbf{F}}_{(i)} \end{bmatrix}$$

is a partial Alamouti-matrix of dimension $2i \times 2L$ corresponding to $\boldsymbol{\mathcal{X}}_{(i)}$, $\mathbf{X}_{a(j)}$ is $2 \times 2L$ matrix corresponding to $\boldsymbol{\mathcal{X}}_{(i)}$, which is the same as $\mathbf{X}_{a(j)}$ with all $\boldsymbol{\mathcal{X}}_{(j)}$ replaced by $\boldsymbol{\mathcal{X}}_{(i)}$, $\boldsymbol{\mathcal{Y}}_{(i)} = \left[\left(\boldsymbol{\mathcal{Y}}_{(i)}^{(k)} \right)^T \left(\boldsymbol{\mathcal{Y}}_{(i)}^{(k+1)} \right)^T \right]^T$ is the partial data vector of dimension $2i \times 1$ and the partial matrix $\underline{\mathbf{F}}_{(i)}$ consists of first i rows of $\underline{\mathbf{F}}$. It should be noted that partial Alamouti-matrix $\mathbf{X}_{a(i)}$ is the function of first i data points while $\mathbf{X}_{a(i)}$ is only a function of i^{th} data point. Obviously, the solution that minimizes this partial joint ML metric would not be globally optimal, but we have the following Lemma.

Lemma 3.1 *Let R represent the optimal value of the objective function in (3.12).*

If $M_{\mathcal{X}_{(i)}} > R$, then $\boldsymbol{\mathcal{X}}_{(i)}$ cannot be the ML solution $\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{ML}$ of (3.12). In other words, for any estimate $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ to correspond to the ML solution, we should have $M_{\mathcal{X}_{(i)}} < R$.

Proof. This Lemma was proved in [44] for SISO case, we simply extend it here to the multi-antenna case. ▮

The above Lemma suggests that if the optimal value of the objective function (3.12) R , can be estimated then we can adopt the following tree search procedure for joint estimation and detection: At each subcarrier i , make a guess of new value of $\boldsymbol{\mathcal{X}}_{(i)} = \left[\mathcal{X}_1(i) \quad \mathcal{X}_2(i) \right]^T$ and use that along with previous estimates to construct

$\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ and $\hat{\mathbf{X}}_{a(i)}$. Then estimate \mathbf{h} to minimize the associated cost function,

$$M_{\hat{\mathcal{X}}_{(i)}} = \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ \left\| \boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho} \hat{\mathbf{X}}_{a(i)} \mathbf{h} \right\|^2 \right\} \quad (3.15)$$

and calculate the resulting metric $M_{\hat{\mathcal{X}}_{(i)}}$. If $M_{\hat{\mathcal{X}}_{(i)}} < R$, then proceed to the next subcarrier $i + 1$, otherwise backtrack and change the guess of $\boldsymbol{\mathcal{X}}(j)$ for some $j \leq i$. We call this approach as the branch-estimate-and-bound strategy, which reduces the search space of exhaustive ML search to those (partial) sequences that satisfy the given constraint $M_{\hat{\mathcal{X}}_{(i)}} < R$. This approach however doesn't work for $i \leq L$ as $\mathbf{X}_{a(i)}$ will be full rank for any choice of $\boldsymbol{\mathcal{X}}_{(i)}$ and therefore \mathbf{h} with $2L$ degrees of freedom can always be chosen by Least Squares (LS) to yield the trivial (i.e., zero) value for $M_{\hat{\mathcal{X}}_{(i)}}$. To obtain a non-trivial value of $M_{\hat{\mathcal{X}}_{(i)}}$, we have to use L pilots, but it would defeat our original motive of *blind* estimation.

To overcome this problem, we adopt a weighted regularized LS and instead of minimizing the ML objective function, J_{ML} , we minimize the maximum *a posteriori* (MAP) objective function

$$J_{MAP} = \underset{\mathbf{h}, \boldsymbol{\mathcal{X}} \in \Omega^{2N}}{\operatorname{argmin}} \left\{ \|\mathbf{h}\|_{\mathbf{R}_h}^2 + \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho} \mathbf{X}_a \mathbf{h}\|^2 \right\} \quad (3.16)$$

where \mathbf{R}_h is the block diagonal autocorrelation matrix of the composite channel vector \mathbf{h} i.e. $\mathbf{R}_h = \mathbb{E} \{ \mathbf{h} \mathbf{h}^H \}$. The objective function in (3.16) can also be

decomposed as

$$J_{MAP} = \underset{\mathbf{h}, \mathcal{X} \in \Omega^{2N}}{\operatorname{argmin}} \left\{ \underbrace{\|\mathbf{h}\|_{\mathbf{R}_h}^2 + \|\mathcal{Y}_{(i)} - \sqrt{\rho} \mathbf{X}_{a(i)} \mathbf{h}\|^2}_{M_{\mathcal{X}_{(i)}}} + \sum_{j=i+1}^N \|\mathcal{Y}_{(j)} - \sqrt{\rho} \mathbf{X}_{a(j)} \mathbf{h}\|^2 \right\} \quad (3.17)$$

So, if we have the guess $\hat{\mathcal{X}}_{(i-1)}$, then the partial metric for \mathcal{X} up to index $i-1$ can be written as

$$M_{\hat{\mathcal{X}}_{(i-1)}} = \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ \|\mathbf{h}\|_{\mathbf{R}_h}^2 + \|\mathcal{Y}_{(i-1)} - \sqrt{\rho} \hat{\mathbf{X}}_{a(i-1)} \mathbf{h}\|^2 \right\} \quad (3.18)$$

whose optimum value $\hat{\mathbf{h}}$ and the minimum cost can be computed [62].

3.3.1 Recursive Derivation of Bound

For our blind search strategy, the calculation of the metric or bound $M_{\hat{\mathcal{X}}_{(i)}}$ is needed at each tree node for comparison with the optimal value of objective function, R . This bound can be derived recursively by simply expressing $M_{\hat{\mathcal{X}}_{(i)}}$ in terms of new observation and an additional regressor $\hat{\mathbf{X}}_a(i)$ as follows:

$$\begin{aligned} M_{\hat{\mathcal{X}}_{(i)}} &= \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ \|\mathbf{h}\|_{\mathbf{R}_h}^2 + \|\mathcal{Y}_{(i)} - \sqrt{\rho} \hat{\mathbf{X}}_{a(i)} \mathbf{h}\|^2 \right\} \\ &= \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ \|\mathbf{h}\|_{\mathbf{R}_h}^2 + \left\| \begin{bmatrix} \mathcal{Y}_{(i-1)} \\ \mathcal{Y}_{(i)} \end{bmatrix} - \sqrt{\rho} \begin{bmatrix} \hat{\mathbf{X}}_{a(i-1)} \\ \hat{\mathbf{X}}_{a(i)} \end{bmatrix} \mathbf{h} \right\|^2 \right\} \end{aligned} \quad (3.19)$$

By invoking the block version of recursive least squares (RLS) algorithm to the cost function in (3.19) with the data vector of size 2×1 and the regressor matrix of dimension $2 \times 2L$ we get [62],

$$M_{\hat{\chi}_{(i)}} = M_{\hat{\chi}_{(i-1)}} + \mathbf{e}_i^H \mathbf{\Gamma}_i \mathbf{e}_i \quad (3.20)$$

$$\hat{\mathbf{h}}_i = \hat{\mathbf{h}}_{i-1} + \mathbf{G}_i \mathbf{e}_i \quad (3.21)$$

where

$$\mathbf{e}_i = \mathcal{Y}(i) - \sqrt{\rho} \hat{\mathbf{X}}_a(i) \hat{\mathbf{h}}_{i-1} \quad (3.22)$$

$$\mathbf{\Gamma}_i = \left[\mathbf{I}_2 + \rho \hat{\mathbf{X}}_a(i) \mathbf{P}_{i-1} \hat{\mathbf{X}}_a(i)^H \right]^{-1} \quad (3.23)$$

$$\mathbf{G}_i = \sqrt{\rho} \mathbf{P}_{i-1} \hat{\mathbf{X}}_a(i)^H \mathbf{\Gamma}_i \quad (3.24)$$

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{G}_i \mathbf{\Gamma}_i^{-1} \mathbf{G}_i^H \quad (3.25)$$

The RLS recursions are initialized by

$$M_{\hat{\chi}_{(i-1)}} = 0, \quad \hat{\mathbf{h}}_{-1} = \mathbf{0} \quad \text{and} \quad \mathbf{P}_{-1} = \mathbf{R}_h.$$

Before introducing our algorithm, we first number the $|\Omega|^2$ combinations of the constellation points from two antennas by $1, 2, \dots, |\Omega|^2$ and treat them as a big constellation set Ψ , where the k^{th} ($1 \leq k \leq |\Omega|^2$) vector constellation point is denoted by $\Psi(k)$. We then perform the depth-first search of signal tree for joint ML solution as shown in Algorithm 3.1.

Algorithm 3.1 Blind MAP Algorithm

Parameters: Initial search radius r , ρ and channel covariance matrix \mathbf{R}_h .

Inputs: \mathcal{Y} , constellation set Ψ and the $1 \times N$ carrier index vector I .

Outputs: Estimated channel $\hat{\mathbf{h}}$ and data vector $\hat{\mathcal{X}}$.

1. (Initialize) Set $i = 1$, $I(i) = 1$, $\hat{\mathcal{X}}(i) = \Psi(I(i))$ and construct the *Alamouti matrix* $\hat{\mathbf{X}}_a(i)$.
 2. (Compare with bound) Compute and store the metric $M_{\hat{\mathcal{X}}(i)}$. If $M_{\hat{\mathcal{X}}(i)} > r$; go to 3; else go to 4.
 3. (Backtracking) Find the largest $1 \leq j \leq i$ such that $I(j) < |\Omega|^2$. If there exists such j , set $i = j$ and go to 5; else go to 6.
 4. (Increment subcarrier) If $i < N$, set $i = i + 1$, $I(i) = 1$, $\hat{\mathcal{X}}(i) = \Psi(I(i))$ and go to 2; else store the current $\hat{\mathcal{X}}_{(N)}$, update $r = M_{\hat{\mathcal{X}}_{(N)}}$ and go to 3.
 5. (Increment constellation) Set $I(i) = I(i) + 1$ and $\hat{\mathcal{X}}(i) = \Psi(I(i))$. Go to 2.
 6. (End/Restart) If a full-length sequence $\hat{\mathcal{X}}_{(N)}$ has been found in step 4, output it as the MAP solution and terminate; otherwise, double 'r' and go to 1.
-

The algorithm essentially reduces the search space of exhaustive ML search by performing a trimmed search over the signal tree of N layers, where each tree node at the i^{th} layer corresponds to a specific partial sequence $\mathcal{X}_{(i)}$ and each tree node at the intermediate layer has $|\Omega|^2$ off-springs to the next layer.

The parameter ρ can be easily determined by estimating the noise variance, whereas for \mathbf{R}_h , our simulation results indicate that we can replace it with an identity matrix with almost no effect on the performance via carrier reordering (see the next Section). To obtain the initial guess of search radius we can use the strategy described in [44] to determine r that would guarantee a MAP solution with very high probability. Nevertheless, the algorithm itself takes care of the

value of r , in that if it is too small such that the algorithm is not able to back-track, then it doubles the value of r and if it is too large such that the algorithm reaches the last subcarrier too quickly then it reduces r to the most recent value of objective function (see step 4 and 6). Therefore any choice of r would guarantee the MAP solution.

3.4 Low-Complexity Blind Algorithm

The complexity of the algorithm is mainly due to: (i) calculation of the bound $M_{\hat{\mathbf{x}}(i)}$ in step 2, and (ii) the backtracking in step 3. The former, as can be seen from RLS recursions, depends heavily on computation of $2L \times 2L$ matrix \mathbf{P}_i in (3.25). We show how we can completely avoid computing \mathbf{P}_i by exploiting the structure of the FFT matrix and hence simply discard (3.25) from RLS recursions. This means that RLS algorithm will reduce to least mean square (LMS) in terms of complexity. The issue of backtracking will be treated in Section 3.5.

3.4.1 Reducing Complexity by Avoiding \mathbf{P}_i

Let us assume that $\mathbf{P}_{-1} = \mathbf{I}$ and the row vectors \mathbf{a}_i of $\underline{\mathbf{F}}$ are orthogonal for $i = 0, 1, 2, \dots, N - 1$, i.e., $\mathbf{a}_i^H \mathbf{a}_j = 0$ for $i \neq j$ (In fact a weaker condition that three consecutive vectors \mathbf{a}_i , \mathbf{a}_{i+1} and \mathbf{a}_{i+2} are orthogonal, would suffice). First we merge (3.24) with (3.25) and obtain,

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \rho \mathbf{P}_{i-1} \hat{\mathbf{X}}_a^H(i) \Gamma_i^H \hat{\mathbf{X}}_a(i) \mathbf{P}_{i-1} \quad (3.26)$$

Then, by using (3.26) and our assumptions, it follows by induction that

$$\mathbf{P}_i \widehat{\mathbf{X}}_a^H(i+1) = \widehat{\mathbf{X}}_a^H(i+1), \mathbf{P}_i \widehat{\mathbf{X}}_a^H(i+2) = \widehat{\mathbf{X}}_a^H(i+2) \text{ and } \mathbf{P}_{i+1} \widehat{\mathbf{X}}_a^H(i+2) = \widehat{\mathbf{X}}_a^H(i+2)$$

Hence, if the successive regressors are orthogonal we can simply replace \mathbf{P}_i with an identity matrix and hence discard equation (3.25). Moreover, from the orthogonality assumptions it also follows that:

$$\widehat{\mathbf{X}}_a(i) \widehat{\mathbf{X}}_a^H(j) = \begin{cases} 0 & \text{if } i \neq j \\ L \left(\|\widehat{\mathbf{X}}_1(i)\|^2 + \|\widehat{\mathbf{X}}_2(i)\|^2 \right) \mathbf{I}_2 & \text{if } i = j \end{cases} \quad (3.27)$$

where, \mathbf{I}_n represents an $n \times n$ identity matrix. Incorporating these results into RLS recursions, the matrices $\mathbf{\Gamma}_i$ and \mathbf{G}_i become independent of \mathbf{P}_i and are given as,

$$\mathbf{\Gamma}_i = \frac{1}{1 + \rho L \left(\|\widehat{\mathbf{x}}_1(i)\|^2 + \|\widehat{\mathbf{x}}_2(i)\|^2 \right)} \mathbf{I}_2 \quad (3.28)$$

$$\mathbf{G}_i = \sqrt{\rho} \widehat{\mathbf{X}}_a(i)^H \mathbf{\Gamma}_i \quad (3.29)$$

The resulting low-complexity blind algorithm based on (3.20)–(3.22), (3.28) and (3.29) for metric computation, requires no matrix inversion or computation of \mathbf{P}_i .

3.4.2 Reducing Complexity by Carrier Reordering

In the above approximation, we assumed that $\mathbf{P}_{-1} = \mathbf{I}$ and \mathbf{a}_i are orthogonal which allows us to use (3.27). However, \mathbf{a}_i are rows of the partial FFT matrix

$\underline{\mathbf{E}}$, so strictly speaking they are not orthogonal. Hence, the successive regressor matrices would not be orthogonal too. However, we can make them orthogonal or semi-orthogonal by carrier reordering based on the idea presented in Section 2.2. Specifically, in Fig. 3.2 we plot the magnitude of correlation of these partial vectors given by (2.17) for $N = 16$ and $L = 4$. It can be seen that rows 1, 5, 9, 13 are orthogonal to each other and so are the rows 2, 6, 10, 14 and so on. If we visit the sub-carriers in order 1, 5, 9, 13, 2, 6, 10, 14, \dots , 4, 8, 12, 16 we find that consecutive vectors will be orthogonal or approximately orthogonal. In general, (as we saw earlier Section 2.2) with $\Delta = N/L$ the vectors $\mathbf{a}_i, \mathbf{a}_{i+\Delta}, \mathbf{a}_{i+2\Delta}, \forall i$ are orthogonal. Therefore, by simple reordering the carriers we can achieve orthogonality among different sub-carriers and use that fact to reduce the complexity of our algorithm as done previously in Section 3.4.1.

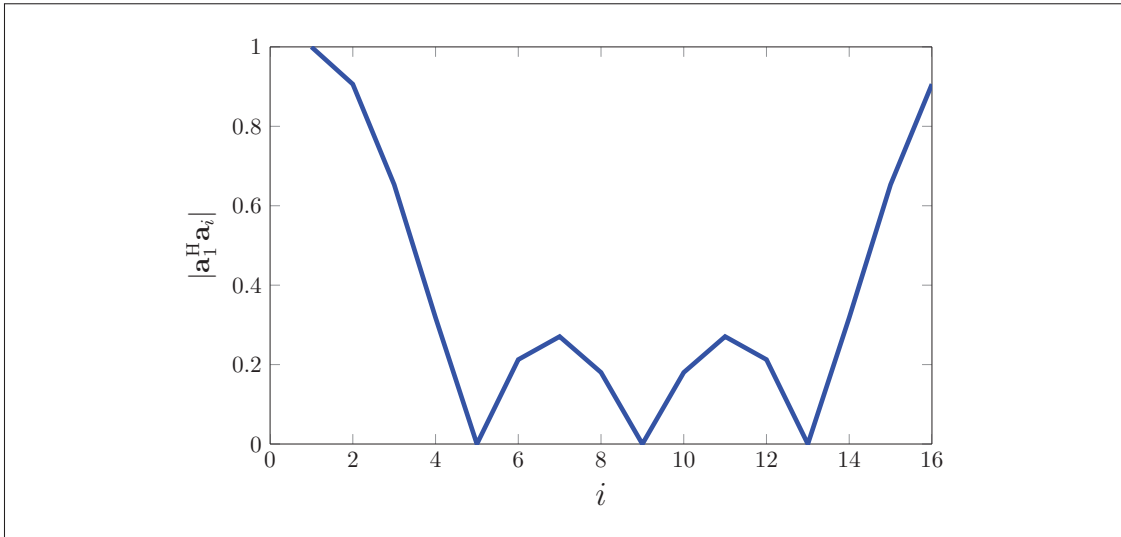


Figure 3.2: Correlation between partial vectors \mathbf{a}_1 and \mathbf{a}_i for $N=16$ and $L=4$.

3.5 Complexity Reduction by Reliable Carriers

The second major source of complexity is the backtracking, which occurs whenever the constraint $M_{\hat{\mathcal{X}}_{(i)}} < r$ is not satisfied. The algorithm then goes back either to the nearest subcarrier or to the current subcarrier whose alphabet is not exhausted and increments the alphabet (step 3 then step 5). This issue is rigorously analyzed in [44] where it is shown that the probability of backtracking is almost zero at high SNR, however, no solution is proposed in the low SNR case. Moreover, with two transmit antennas, the search space at each node grows as $|\Omega|^2$ as compared to $|\Omega|$ for SISO system. Thus the complexity of proposed algorithms due to backtracking ultimately dominates the complexity induced by computing the matrix \mathbf{P}_i (or its inverse) and becomes the real bottleneck.

Since the backtracking is inevitable in all blind search algorithms, it cannot be avoided in practice. However, its effect can be minimized and we aim to do so by using the concept of reliable carriers. The basic idea is that if we are able to arrange the data according to its reliability, starting with the most reliable data, then there would be a less chance that we need to backtrack. Since earlier data is reliable, there is no need to backtrack for this part. The later data might not be reliable but by the time we start processing this data, the algorithm would have converged. However, measuring the data reliability requires tentative channel estimates which can only be obtained by using some pilots. Therefore, we transform our blind algorithm into semi-blind algorithm that would require a short training sequence of L symbols only at the start of transmission to get a tentative estimate

of the data and its reliability and no further pilots or channel statistics would be required.

3.5.1 Measuring the Reliability

To minimize backtracking, the algorithm must devise a procedure to identify the reliable data carriers from the tentative estimates of channel and the data. Thus it is imperative to measure the reliability of data carriers. With receiver having an estimate of channel $\hat{\mathbf{H}}$ using pilots, the decoding process can be accomplished by re-writing (3.10) as follows

$$\tilde{\mathbf{y}} = \sqrt{\rho} \mathbf{H}_a \mathbf{x} + \tilde{\mathbf{w}} \quad (3.30)$$

where $\tilde{\mathbf{y}} = \left[(\mathbf{y}^{(k)})^T \quad (\mathbf{y}^{*(k+1)})^T \right]^T$, $\tilde{\mathbf{w}} = \left[(\mathbf{w}^{(k)})^T \quad (\mathbf{w}^{*(k+1)})^T \right]^T$ and \mathbf{H}_a is an *Alamouti-like* matrix defined as follows

$$\mathbf{H}_a \triangleq \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_2^* & -\Lambda_1^* \end{bmatrix} \quad (3.31)$$

Now, using the zero-forcing i.e., left multiplying both sides of (3.30) with (the estimate of) $\frac{1}{\sqrt{\rho}} \mathbf{H}_a^{-1}$ and re-arranging the terms we get,

$$\hat{\mathbf{x}} \approx \mathbf{x} + \mathbf{z} \quad (3.32)$$

where $\mathbf{z} \triangleq \frac{1}{\sqrt{\rho}} \hat{\mathbf{H}}_a^{-1} \tilde{\mathbf{W}}$ represents the distortion due to noise and channel estimation error. Given the estimate $\hat{\mathbf{H}}_a$, \mathbf{z} can be modelled as Gaussian with zero mean and covariance $\frac{1}{\rho} (\hat{\mathbf{H}}_a^{-1}) (\hat{\mathbf{H}}_a^{-1})^H$. Hence, some data-carriers $\mathbf{x}(k)$, would be severely effected by noise and channel perturbation errors i.e., $\mathbf{z}(k)$ and fall outside their correct decision regions, while for some other data-carriers the distortion is not strong enough and they are decoded correctly. All those data carriers $\mathbf{x}(k)$ which satisfy the condition $\langle \hat{\mathbf{x}}(k) \rangle = \mathbf{x}(k)$ with high probability, are termed reliable carriers.

The authors in [70] have developed a rigorous method for assessing the reliability of data carriers, based on which the expression for reliability is a vector-wise likelihood ratio defined as

$$\mathfrak{R}^{exact} = \log \frac{f_{\mathbf{z}}(\hat{\mathbf{x}} - \langle \hat{\mathbf{x}} \rangle)}{\sum_{m=1, \Omega_m \neq \langle \hat{\mathbf{x}} \rangle}^{|\Omega|} f_{\mathbf{z}}(\hat{\mathbf{x}} - \Omega_m)} \quad (3.33)$$

where $f_{\mathbf{z}}(\cdot)$ is the pdf of \mathbf{z} . Intuitively, (3.33) measures the reliability in decoding $\mathbf{x}(k)$ to the nearest constellation point $\langle \hat{\mathbf{x}} \rangle$ relative to decoding it any other constellation point. Fig. 3.3 also illustrates this concept, such that, for instance even though $\hat{\mathbf{x}}(1)$ and $\hat{\mathbf{x}}(2)$ have the same distance from \mathbf{x} (i.e., value of numerator in (3.33) is same), $\hat{\mathbf{x}}(2)$ has a higher reliability (i.e., the denominator in (3.33) is smaller) than $\hat{\mathbf{x}}(1)$ as it is farther from the nearest neighbour. Thus higher the value of $\mathfrak{R}^{exact}(k)$ the higher the probability of data-carrier to be decoded correctly and higher the reliability of the carrier.

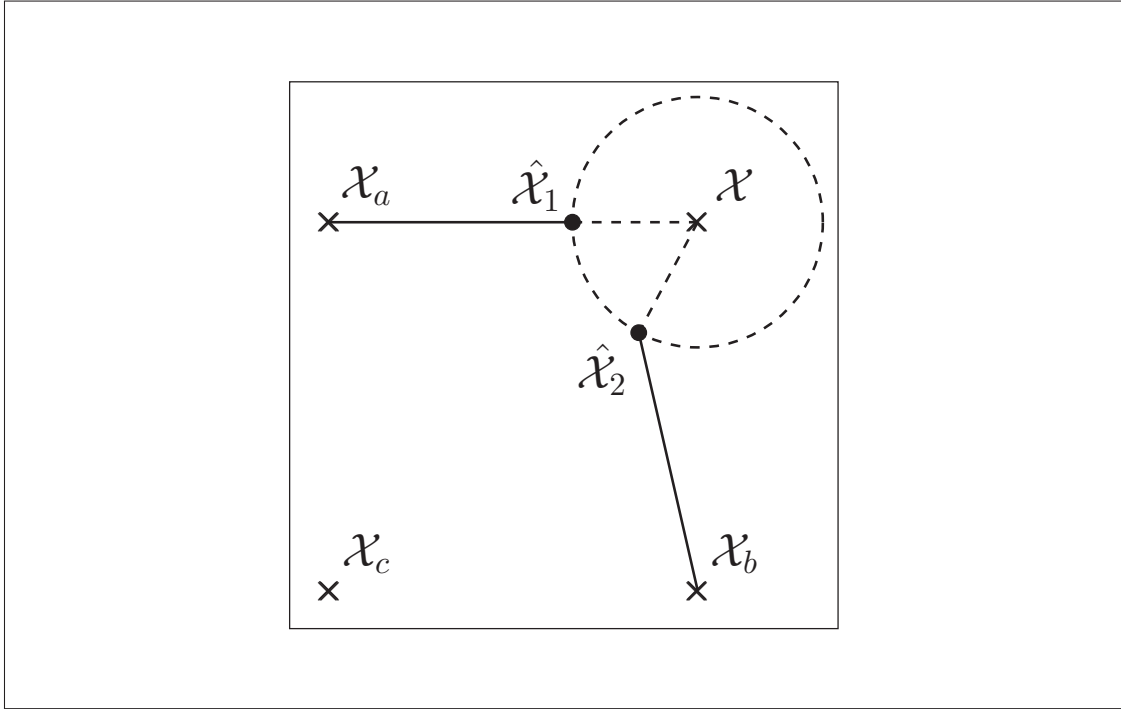


Figure 3.3: Reliability of data carriers $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ in decoding them to the nearest neighbour constellation point \mathcal{X} .

Once the vector \mathfrak{R} is computed, we can proceed to select the most reliable data tones. These reliable data tones can then be supplied to our algorithm for initial search of the ML solution. Based on the above developments, the proposed semi-blind algorithm is given below.

Remarks: The first two steps of the semi-blind algorithm serve as pre-processing steps tailored to minimizing the backtracking of blind algorithm in step 3. One can easily obtain the channel estimates from pilots to run the algorithm. The prediction step 1 is trivial and would suffer only little distortion as the channel does not change much in slow fading. To initiate the RLS recursions of blind algorithm in step 3, we initialize the channel vector with its previous

Algorithm 3.2 Semi-blind algorithm

Obtain an initial estimate of the channel vector \mathbf{h} from L training/pilot symbols at start of transmission, then repeat the following steps over two consecutive time instants.

1. Predict and decode the carriers $\hat{\mathcal{X}}$ from previous channel estimate $\hat{\mathbf{h}}$ and observation vector $\tilde{\mathcal{Y}}$ as in (3.32).
 2. Use (3.33) to compute reliability of data carriers, $\mathfrak{R}(k)_{k=1}^P$, with $1 \leq P \leq N$ and rearrange them in decreasing order of their reliability. The parameter P represents the total number of re-ordered carriers used by the algorithm.
 3. Run Blind Algorithm 3.1 starting with the most reliable data, to obtain exact ML estimates of the channel and the data.
-

estimate and set $\mathbf{P}_i = \mathbf{I}$; thus no channel statistics are required a priori. The blind algorithm is supplied with re-ordered carriers with most reliable data to start over the search (step 2), so that there will be almost no backtracking during the convergence of the algorithm. However, we observe that the carrier reordering based on reliability measures does not ensure orthogonality of successive regressors therefore the low complexity variants of RLS introduced earlier, cannot be employed.

3.6 Simulation Results

For simulation of Blind Algorithm 3.1, we assume that channels between two transmitters and a receiver are both independent Rayleigh fading, stationary over two consecutive OFDM blocks and each having an exponential power decay profile i.e. $E\{|h_i(\tau)|^2\} = e^{-0.2\tau}$. Information symbols are modulated using BPSK or 4-QAM constellations.

In Fig. 3.4, we compare the BER performance of the proposed exact blind algorithm and the proposed low-complexity variants, i.e., blind algorithm with (i) $\mathbf{P}_i = \mathbf{I}$ and (ii) $\mathbf{P}_i = \mathbf{I}$ with subcarrier reordering, against the perfectly known channel. The results shown in Fig. 3.4(a) for BPSK data symbols indicate that with $\mathbf{P}_i = \mathbf{I}$, the performance degrades and BER reaches an error floor. However, with subcarrier reordering approach we almost get the same performance as that of exact blind algorithm without requiring the channel statistics. Similar trend is observed in Fig. 3.4(b), when 4-QAM signal modulation is considered.

For Semi-blind Algorithm 3.2, we adopt the AR(1) process to model the slow rayleigh fading channels, such that the channel weight vector varies as [62]

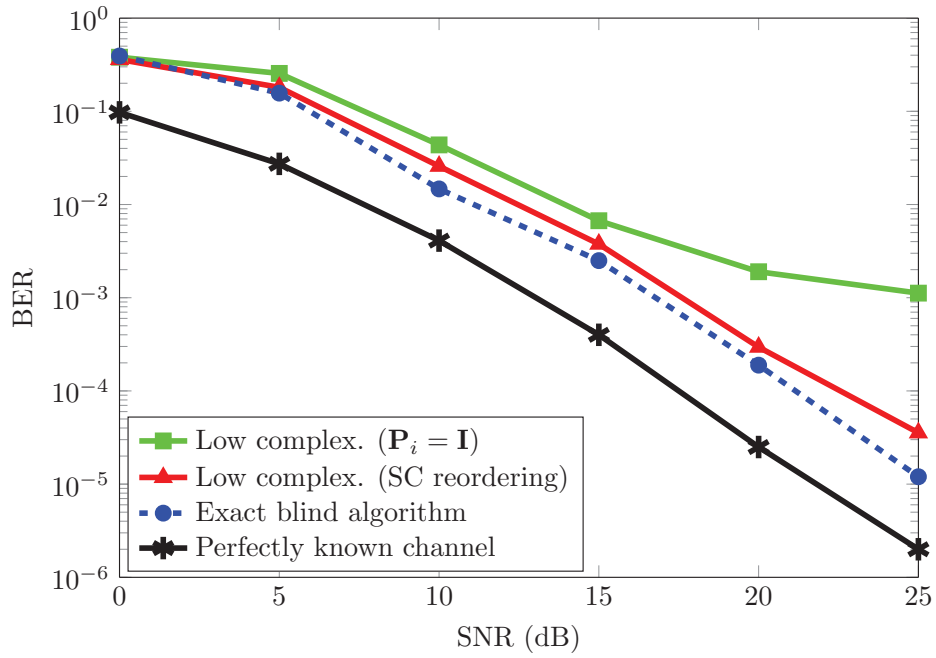
$$\mathbf{h}(n) = \alpha \mathbf{h}(n-1) + \mathbf{q}(n)$$

where, $\alpha = J_0(2\pi f_d T_s)$ and \mathbf{q} is a complex normal vector with covariance matrix $(1 - \alpha^2)\mathbf{I}$. The product of maximum Doppler frequency f_d and sampling time T_s , referred to as normalized doppler frequency F_d , controls the amount of time variations of the channel taps. Two different values of normalized doppler frequency; 0.1 and 0.001 corresponding to relatively fast and slow varying channel, are considered. Results for semi-blind algorithm are presented against perfect coherent detection in Fig. 3.5 which show favourable performance of the proposed algorithm under different modulation constellations and fading conditions.

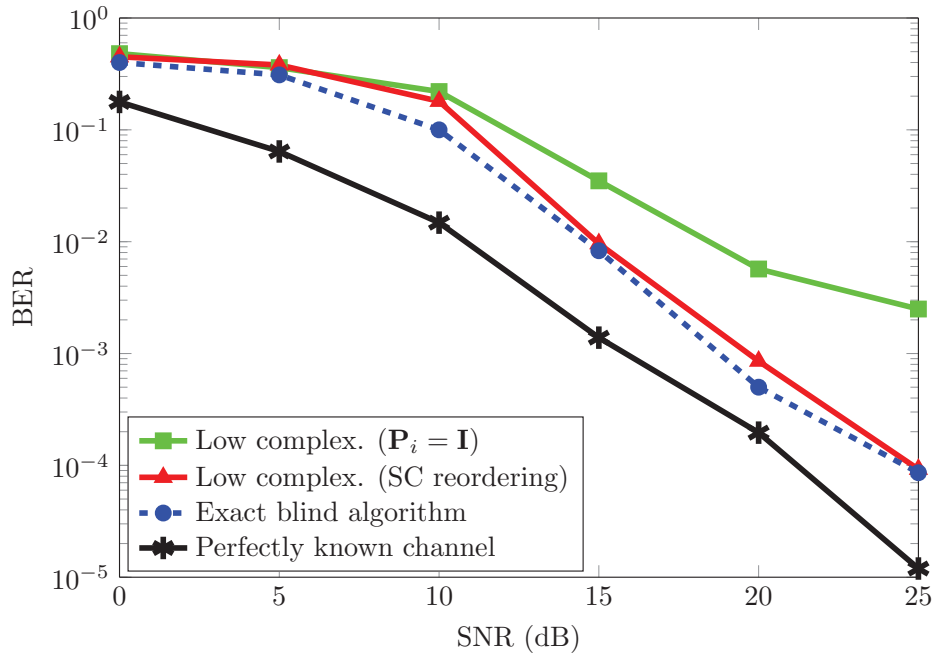
To assess the computational complexity of the proposed algorithm, we compare the average number of nodes visited by the algorithm with various reliability

measures in Fig. 3.6(a). It is clearly observed that the proposed reliability scheme offers significantly lower complexity at lower SNR values. At higher SNR the complexity is constant, confirming the fact that there is almost no backtracking. Fig. 3.6(b) shows that the performance for various degrees of reliability measures is almost identical which means that computational advantages are attained without degrading the performance. Through simulations it has also been observed that the reliability of around 50-60 percent is enough for a good performance, although more importantly, the algorithm doesn't disfavour the usage of more reliable carriers.

Finally in Fig. 3.7 the complexity comparison of semi-blind algorithm with and without using reliable carriers is presented for different modulation schemes such as BPSK, 4-QAM and 16-QAM. The results clearly indicate the computational advantages gained by the proposed method using reliable carriers for minimizing the backtracking of the algorithm. In comparison, for $N = 16$ and 4-QAM constellation, the exhaustive search would require to examine $4^{(32)} \approx 1.84 \times 10^{19}$ hypothesis for each coherence time.

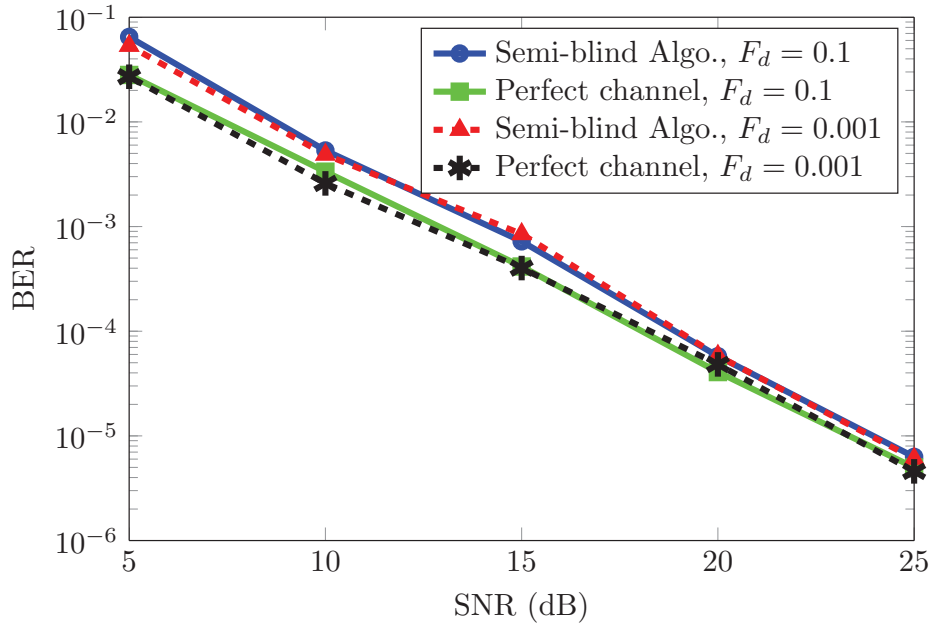


(a) BPSK symbols

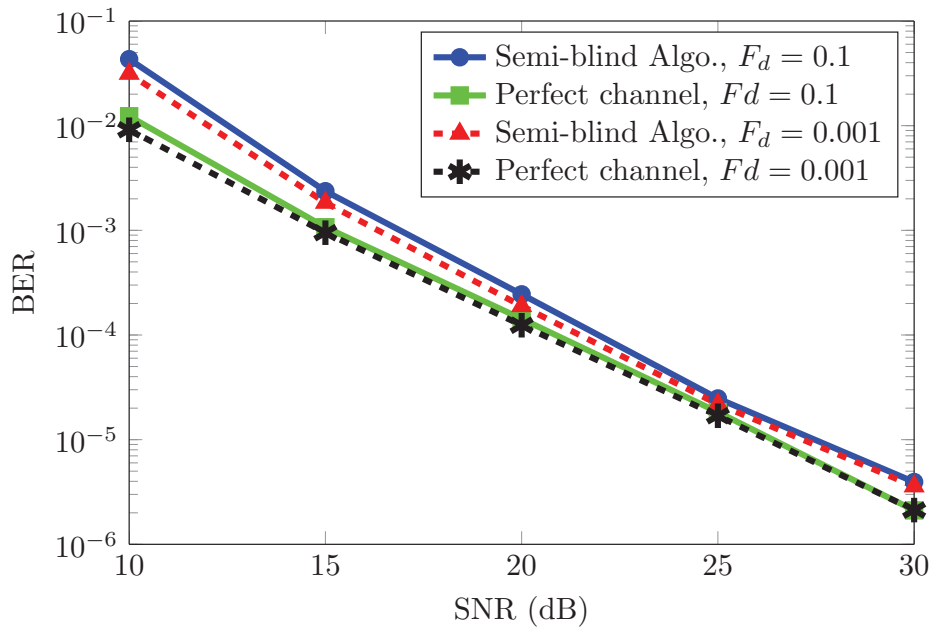


(b) 4-QAM symbols

Figure 3.4: BER Performance of blind algorithm over Rayleigh fading channel with $N = 16$ and $L = 4$.

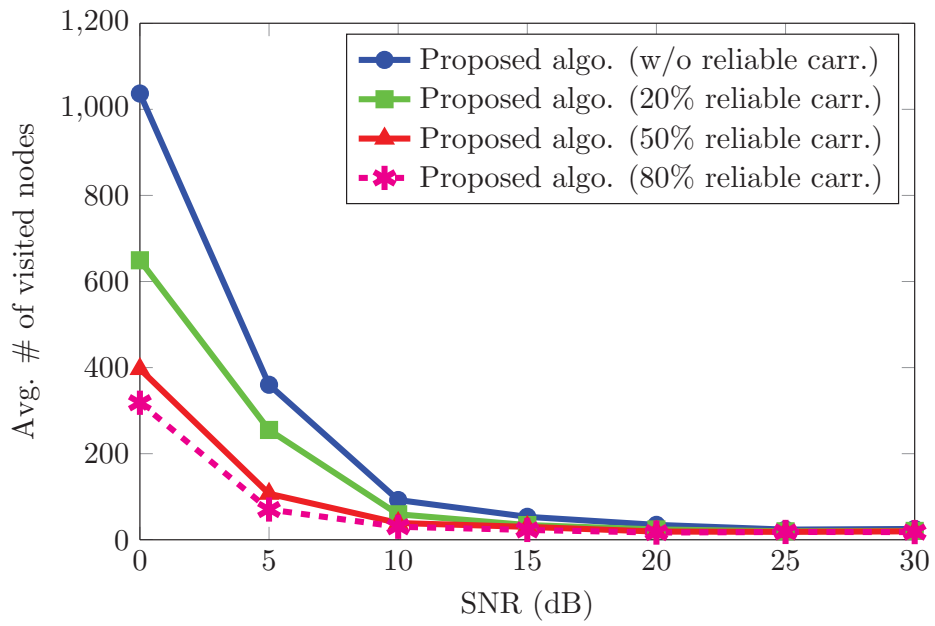


(a) BPSK symbols

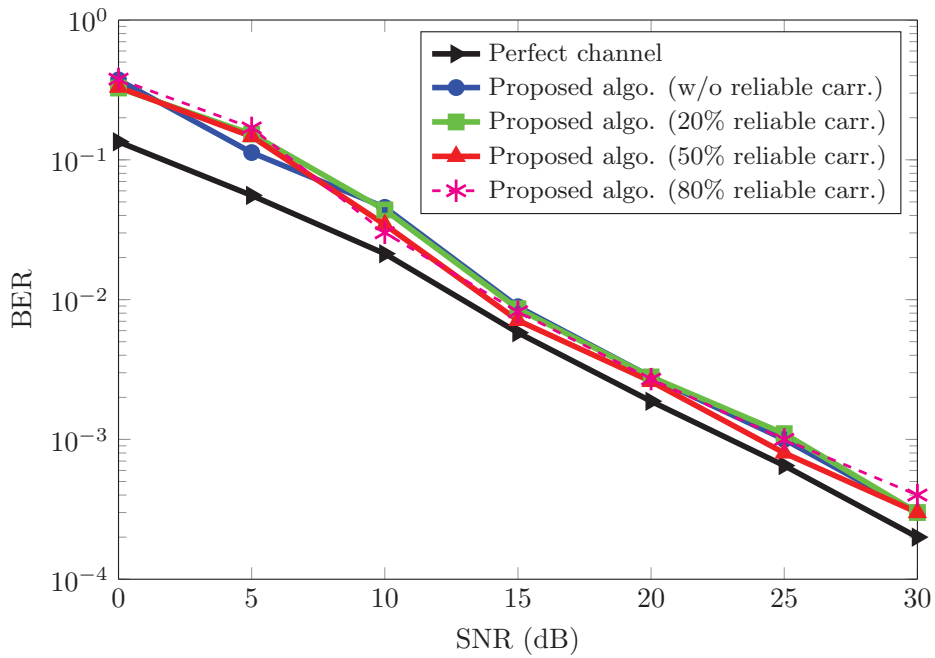


(b) 4-QAM symbols

Figure 3.5: BER performance of semi-blind algorithm over Rayleigh fading channel with $N = 32$ and $L = 4$



(a)



(b)

Figure 3.6: (a) Computational complexity and (b) BER performance of the proposed algorithm with various degrees of reliability measurements using BPSK symbols with $N=16$ and $L=4$.

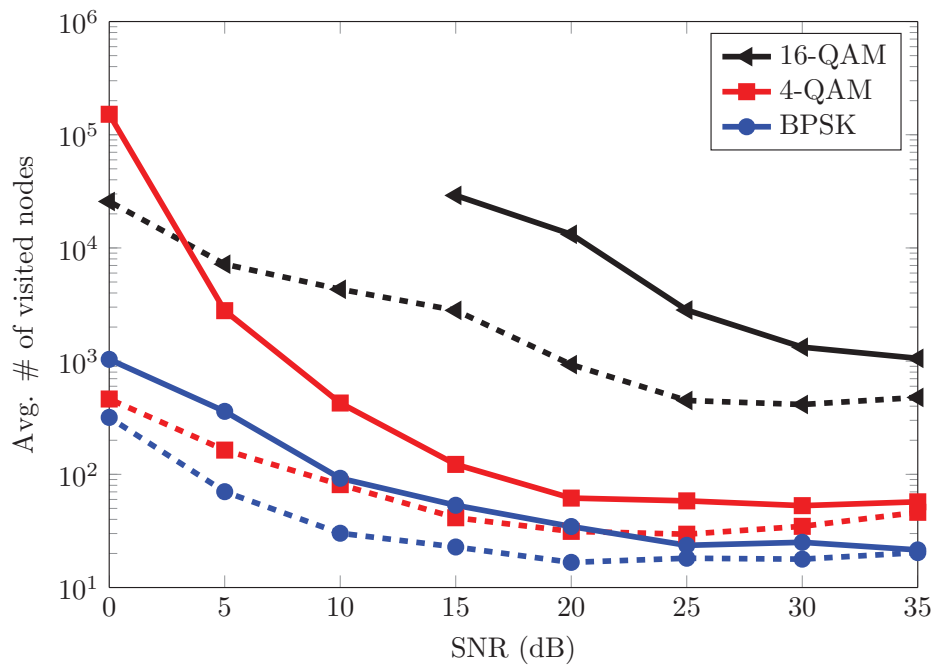


Figure 3.7: Complexity of the proposed semi-blind algorithm without (solid lines) and with reliable carriers (dashed lines) for different modulations with $N=16$ and $L=4$.

CHAPTER 4

CHANNEL ESTIMATION FOR MASSIVE MIMO OFDM SYSTEMS

The demand for wireless data traffic has increased rapidly since the past few years mainly due to unprecedented use of smart phones, tablets, laptops etc. With growing number of mobile devices and wireless internet connections, the demand for future wireless data would increase even more. Basically, the key parameter for wireless data traffic is the wireless throughput (bits/sec) defined as:

$$\text{Throughput (bits/sec)} = \text{Bandwidth (Hz)} \times \text{Spectral efficiency (bits/sec/Hz)}$$

Hence, to improve the throughput the new technologies which can either increase the bandwidth or the spectral efficiency or both should be exploited. This

chapter focuses on massive MIMO systems which have recently emerge as one solution to meet the demands for the next generation wireless communications as they can significantly improve the spectral efficiency and provide huge gains in throughput over SISO and conventional MIMO systems.

Specifically, we focus on estimation of correlated Rayleigh fading channels in the uplink of large antenna or massive MIMO OFDM systems. In massive MIMO, with increased number of BS antennas, the number of channel parameters to be estimated also grows large. This makes the conventional MMSE solution almost impractical due to very high complexity. We propose an efficient distributed MMSE algorithm that can achieve near optimal channel estimates at very low complexity by exploiting the strong spatial correlations and symmetry of large antenna array elements. The proposed method involves solving a (fixed) reduced dimensional MMSE problem at each antenna element followed by a repetitive sharing of information through collaboration among neighboring elements. To further enhance the channel estimates and/or reduce the number of reserved pilot tones, we propose a data-aided estimation technique that relies on finding a set of most reliable data carriers. Simulation results validate the near optimal performance of proposed estimation algorithm.

4.1 System Model

We consider a multi-cell massive MIMO-OFDM wireless system as shown in Fig. 4.1(a), where the BS in each cell is equipped with uniform planar array (UPA)

consisting of a large number of antennas. We assume that each BS serves a number of single antenna user terminals. The antennas on UPAs are distributed across M rows and G columns with horizontal and vertical spacing of d_x and d_y , respectively. We define the linear (column-wise) index of (m, g) th antenna as $r = m + M(g - 1)$ where $1 \leq m \leq M$, $1 \leq g \leq G$, $1 \leq r \leq R$ and $R = MG$ is the total number of antennas in a UPA. Fig. 4.1(b) shows an example of a $M \times G$ UPA structure with antenna indexing. Note that, depending on values of G and M , the antennas could have linear or a rectangular configuration. However, we confine our attention to rectangular UPA structure which is a viable configuration in deployment scenarios for massive MIMO [71].

Each user communicates with the BS using OFDM and transmits uplink pilots for channel estimation. We assume that all users in a particular cell are assigned orthogonal frequency tones so that there is no intra-cell interference. However, there are (interfering) users in the neighboring cells that transmit pilots at the same frequency tones, resulting in an inter-cell interference or pilot contamination. In this Chapter, we assume that there is no inter-cell interference and hence without loss of generality, we focus on a single-cell single-user scenario (the case of multi-cell will be treated in Chapter 5).

4.1.1 Channel Model

We assume that multi-path channel between user and receive antenna r is modeled by a Gaussian L -tap CIR vector defined by $\mathbf{h}_r \triangleq [h_r(0), h_r(1), \dots, h_r(L - 1)]^T \in$

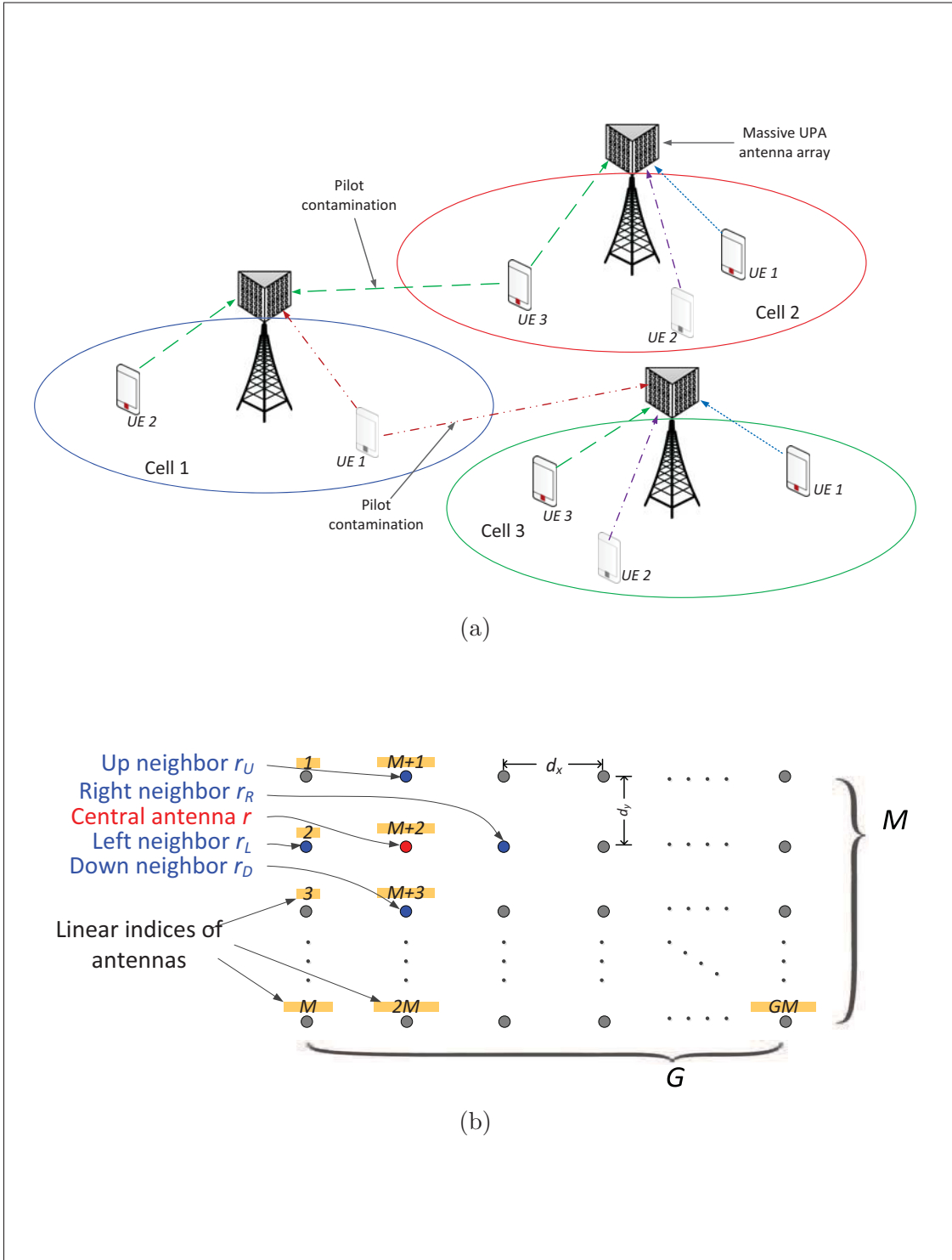


Figure 4.1: (a) Multi-cell massive MIMO system layout (b) An example of $M \times G$ UPA structure with antenna indexing.

$\mathbb{C}^{L \times 1}$. We append all the CIR vectors from a user to the BS to form a composite channel vector $\mathbf{h} \triangleq [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_R^T]^T$ of size $RL \times 1$. Further, we collect the l th tap of all CIRs to form the l th tap vector $\mathbf{h}^{(l)} \triangleq [h_1(l), h_2(l), \dots, h_R(l)]^T$ of size $R \times 1$. Then, the $RL \times RL$ dimensional composite channel correlation matrix can be written as

$$\mathbf{R}_h \triangleq \mathbb{E}\{\mathbf{h}\mathbf{h}^H\} = \mathbf{R}_a \otimes \mathbf{R}_{tap}, \quad (4.1)$$

which is the Kronecker product of two components: (i) The $R \times R$ dimensional antenna spatial correlation matrix, $\mathbf{R}_a = \mathbb{E}\{\mathbf{h}^{(l)}\mathbf{h}^{(l)H}\}$, $\forall l = 0, 1, \dots, L-1$ and (ii) The $L \times L$ dimensional channel tap correlation matrix, $\mathbf{R}_{tap} = \mathbb{E}\{\mathbf{h}_r\mathbf{h}_r^H\}$, $\forall r = 1, 2, \dots, R$, that depends on channel PDP. In the channel correlation model in (4.1), we implicitly assume \mathbf{R}_a to be identical across the l taps and \mathbf{R}_{tap} to be identical across the array. For the spatial correlation matrix \mathbf{R}_a , we adopt a ray-based 3D channel model from [72] which is appropriate for rectangular arrays. Accordingly, the spatial correlation between array elements $r=(m, g)$ and $r'=(p, q)$ is given by,

$$[\mathbf{R}_a]_{r,r'} = \frac{D_1}{\sqrt{D_5}} e^{-\frac{D_7 + (D_2(\sin\phi)\sigma)^2}{2D_5}} e^{j\frac{D_2 D_6}{D_5}}, \quad (4.2)$$

where the D_i 's are defined as,

$$\begin{aligned}
D_1 &= e^{j\frac{2\pi d_x}{\nu}(p-m)\cos(\theta)} e^{-\frac{1}{2}\left(\xi\frac{2\pi d_x}{\nu}\right)^2(p-m)^2\sin^2\theta} , \\
D_2 &= \frac{2\pi d_x}{\nu}(q-g)\sin(\theta) , \\
D_3 &= \xi\frac{2\pi d_x}{\nu}(q-g)\cos(\theta) , \\
D_4 &= \frac{1}{2}\left(\xi\frac{2\pi}{\nu}\right)^2(p-m)(q-g)\sin(2\theta) , \\
D_5 &= (D_3)^2(\sin(\phi)\sigma)^2 + 1 , \\
D_6 &= D_4(\sin(\phi)\sigma)^2 + \cos(\phi) , \\
D_7 &= (D_3)^2\cos^2\phi - (D_4)^2(\sin(\phi)\sigma)^2 - 2D_4\cos\phi .
\end{aligned}$$

Here, ν is the carrier-frequency wavelength in meters, ϕ and θ are the mean horizontal angle-of-departure (AoD) and the mean vertical AoD in radians, respectively, σ and ξ are the standard deviation of horizontal AoD and the standard deviation of vertical AoD, respectively. As shown in [72], the spatial correlation matrix can be well approximated as

$$\mathbf{R}_a \approx \mathbf{R}_{az} \otimes \mathbf{R}_{el} , \tag{4.3}$$

where \mathbf{R}_{az} and \mathbf{R}_{el} are the correlation matrices in azimuth (horizontal) and elevation (vertical) directions, respectively, defined as below:

$$[\mathbf{R}_{el}]_{m,p} = e^{j\frac{2\pi d_x}{\nu}(p-m)\cos(\theta)} e^{-\frac{1}{2}(\xi\frac{2\pi d_x}{\nu})^2(p-m)^2\sin^2\theta}, \quad (M \times M)$$

$$[\mathbf{R}_{az}]_{g,q} = \frac{1}{\sqrt{D_5}} e^{-\frac{D_3^2 \cos^2 \phi}{2D_5}} e^{j\frac{D_2 \cos \phi}{D_5}} e^{-\frac{1}{2}\frac{(D_2 \sigma)^2}{D_5}}, \quad (G \times G).$$

4.1.2 Signal Model

We assume that there are N OFDM sub-carriers and let \mathcal{X} represent the N -dimensional information symbol whose entries are drawn from a bi-dimensional constellation (e.g., QPSK or QAM). The equivalent time-domain symbol is obtained by taking inverse Fourier transform, i.e., $\mathbf{x} = \mathbf{F}^H \mathcal{X}$. The time-domain symbol is then transmitted after inserting a cyclic prefix (CP) of length at least $L-1$ to avoid inter-symbol-interference (ISI). After discarding the CP at the receiver, the frequency-domain OFDM symbol at r th antenna can be represented as

$$\mathcal{Y}_r = \sqrt{N} \text{diag}(\mathcal{X}) \underline{\mathbf{F}} \mathbf{h}_r + \mathcal{W}_r = \mathbf{A} \mathbf{h}_r + \mathcal{W}_r, \quad (4.4)$$

where $\mathbf{A} \triangleq \sqrt{N} \text{diag}(\mathcal{X}) \underline{\mathbf{F}}$, $\underline{\mathbf{F}}$ is truncated Fourier matrix formed by selecting the first L columns of \mathbf{F} and \mathcal{W}_r is frequency domain noise vector of zero mean and covariance $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_N$, assumed to be uncorrelated with the channel vector \mathbf{h}_r . For a set of K pilot indices denoted by vector \mathcal{P} , the system equation (4.4) reduces to

$$\mathcal{Y}_r(\mathcal{P}) = \mathbf{A}(\mathcal{P}) \mathbf{h}_r + \mathcal{W}_r(\mathcal{P}), \quad (4.5)$$

where $\mathbf{y}_r(\mathcal{P})$ and $\mathbf{w}_r(\mathcal{P})$ are formed by selecting the entries of \mathbf{y}_r and \mathbf{w}_r indexed by \mathcal{P} while $\mathbf{A}(\mathcal{P})$ is a $K \times L$ matrix formed by selecting the rows of \mathbf{A} indexed by \mathcal{P} . We can now collect the pilot measurements (4.5) received by all antennas into a single system of equations as follows

$$\mathbf{y}(\mathcal{P}) = [\mathbf{I}_R \otimes \mathbf{A}(\mathcal{P})]\mathbf{h} + \mathbf{w}(\mathcal{P}), \quad (4.6)$$

where $\mathbf{y}(\mathcal{P})$ and $\mathbf{w}(\mathcal{P})$ are formed by column-wise stacking of pilots and noise observations at each antenna while \mathbf{I}_R is an $R \times R$ identity matrix. For convenience, we assume the noise variance to be identical across the array so that $\mathbf{w}(\mathcal{P}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_w = \sigma_w^2 \mathbf{I}_{RK})$. Note that the number of unknown channel coefficients in (4.6) is RL whereas the total number of equations is RK . Therefore, a necessary condition to solve (4.6) for \mathbf{h} (and also (4.5) for \mathbf{h}_r) using least squares, is that the number of pilots be at least equal to L i.e., $K \geq L$. However, K could be reduced if we utilize the correlation information. With the models defined above, we are ready to estimate the CIRs between the user and each BS antenna.

4.2 LMMSE and LS based Channel Estimation

In this Section, we pursue different approaches, based on LS and LMMSE that can be adopted for channel estimation in massive MIMO setup depending on whether the information processing takes place independently at each antenna element (local processing) or jointly at a centralized processor. We also discuss

their limitations which motivate us to propose a novel distributed approach for channel estimation.

4.2.1 The Localized LMMSE (L-LMMSE) Estimation

In this approach, all CIRs are estimated independently based on the observations received at each antenna element by using the classical LMMSE solution. Using the linear system model in (4.5), the LMMSE estimate of \mathbf{h}_r is obtained by minimizing the (local) MSE, $\mathbb{E}\{\|\mathbf{h}_r - \hat{\mathbf{h}}_r\|^2\}$, over $\hat{\mathbf{h}}_r$ as follows [62]

$$\hat{\mathbf{h}}_r = (\mathbf{R}_{tap}^{-1} + \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{y}_r, \quad (4.7)$$

where we drop the index vector \mathcal{P} for convenience. Similarly, it follows that the (minimum) MSE is,

$$\text{mse}_r = \text{trace} (\mathbf{R}_{tap}^{-1} + \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1}. \quad (4.8)$$

The overall global MSE can be obtained by taking summation over all array elements i.e., $\text{MSE}^{(L)} = \sum_{r=1}^R \text{mse}_r$, which after simplifying (4.8), can be expressed as,

$$\text{MSE}^{(L)} = R \sum_{i=1}^L \left(\frac{\delta_i}{1 + \rho K \delta_i} \right), \quad (4.9)$$

where $\{\delta_i\}_{i=1}^L$ are eigenvalues of \mathbf{R}_{tap} , $\rho \triangleq E_x / \sigma_w^2$ is the SNR with E_x representing the average signal energy per symbol and the superscript (L) indicates L-LMMSE.

Observe from (4.9) that channel delay spread L , has an adverse effect on MSE per-

formance, which can be reduced by increasing the number of pilot tones. Despite the fact that the computational complexity of L-LMMSE increases linearly with the number of BS antennas (see Table 4.1), the CIR estimates are not optimal in the sense of minimizing the overall or global MSE. The estimates would have been optimal, had the antennas been placed sufficiently apart so that the channel vectors were effectively uncorrelated. But for massive MIMO with extremely large number of antennas, it is expected that antennas are located in close proximity, so the channel vectors are highly likely to be correlated with each other.

4.2.2 The Optimal LMMSE (O-LMMSE) Solution

In this strategy all the channel vectors are estimated simultaneously by minimizing the global MSE, $\mathbb{E}\{\|\mathbf{h} - \hat{\mathbf{h}}\|^2\}$ over the composite channel vector $\hat{\mathbf{h}}$. This could be realized by sending all observations to a central processor and then invoking the LMMSE estimation based on the composite system model in (4.6). The solution to this problem is given by

$$\hat{\mathbf{h}} = \left(\mathbf{R}_{\mathbf{h}}^{-1} + \hat{\mathbf{A}}^H \mathbf{R}_w^{-1} \hat{\mathbf{A}} \right)^{-1} \hat{\mathbf{A}}^H \mathbf{R}_w^{-1} \mathcal{Y}, \quad (4.10)$$

where $\hat{\mathbf{A}} = \mathbf{I}_R \otimes \mathbf{A}$, $\mathbf{R}_{\mathbf{h}}$ is as given in (4.1) and for notational convenience we dropped the index \mathcal{P} . The corresponding MSE is

$$\text{MSE}^{(O)} = \text{trace} \left(\mathbf{R}_{\mathbf{h}}^{-1} + \hat{\mathbf{A}}^H \mathbf{R}_w^{-1} \hat{\mathbf{A}} \right)^{-1}, \quad (4.11)$$

which can be simplified to yield

$$\text{MSE}^{(0)} = \sum_{j=1}^R \sum_{i=1}^L \frac{\eta_j \delta_i}{1 + \rho K \eta_j \delta_i}, \quad (4.12)$$

where η_j and δ_i are eigenvalues of \mathbf{R}_a and \mathbf{R}_{tap} , respectively. By comparing (4.12) with (4.9), we conclude that in presence of spatial correlation, the optimal solution yields better MSE performance than the localized strategy, however, it has the following two major drawbacks:

1. Realization of optimal strategy requires global sharing of information to/from the central processor that results in communication overhead (as it requires complex signalling which can be very expensive).
2. As evident from (4.10), the computation complexity of optimal LMMSE grows with cubic power of the number of BS antennas as it requires inverting a non-trivial matrix of very high dimension $RK \times RK$ (see Table 4.1).

In a massive MIMO scenario where R is of the order of few hundreds, both of the above mentioned operations are very expensive and almost impractical.

4.2.3 Estimation Using Least Squares

If the channel statistics are unknown, one can employ simple LS based estimation. In the absence of correlation, we can let the inverse of channel correlation matrix go to zero, i.e., $\mathbf{R}_{tap}^{-1} \rightarrow \mathbf{0}$, thereby ignoring the channel statistics. Therefore, the

localized LS solution from (4.7) is

$$\hat{\mathbf{h}}_r^{\text{ls}} = (\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}_w^{-1} \mathcal{Y}_r, \quad (4.13)$$

and the resulting MSE is given by

$$\text{mse}_r^{\text{ls}} = \text{trace} (\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1}. \quad (4.14)$$

In this case, the overall MSE simplifies to

$$\text{MSE}^{(\text{LS})} = \sum_{r=1}^R \text{mse}_r^{\text{ls}} = \frac{RL}{\rho K}. \quad (4.15)$$

From (4.15) and (4.9), it is obvious that LS has poor performance in comparison with the LMMSE as it does not utilize the channel statistics. It is for this reason that the centralized LS (C-LS) solution would achieve the same MSE performance as the localized one as shown below.

$$\begin{aligned} \text{MSE}^{(\text{C-LS})} &= \text{trace} \left((\mathbf{I}_R \otimes \mathbf{A})^H (\mathbf{I}_R \otimes \mathbf{R}_w)^{-1} (\mathbf{I}_R \otimes \mathbf{A}) \right)^{-1} \\ &= \text{trace} (\mathbf{I}_R \otimes \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1}, \\ &= \sum_{r=1}^R \text{trace} (\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1} \\ &= \text{MSE}^{(\text{LS})}, \end{aligned}$$

where we have used the Kronecker product identities, $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ and $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$.

In short, the L-LMMSE estimation has the advantage of low complexity (and better performance than LS) but it is unable to exploit the strong spatial correlation among array elements. On the other hand, O-LMMSE exploits the spatial correlations but at a significantly higher computational cost. This motivates us to propose a method that can overcome the shortcomings of aforementioned techniques without affecting the estimation quality. Specifically, we propose a distributed estimation of CIRs based on antenna coordination that attains near optimal performance with tractable complexity. The proposed distributed LMMSE estimation is described next and is further extended in Section 4.4 via a data-aided technique.

4.3 The Proposed Distributed LMMSE Estimation

It is well known from equivalence results in linear estimation theory [73] that the O-LMMSE solution (4.10) could be obtained by solving an RL dimensional optimization problem

$$\underset{\mathbf{h}}{\operatorname{argmin}} \|\mathcal{Y} - \mathbf{A}'\mathbf{h}\|_{\mathbf{R}_w^{-1}}^2 + \|\mathbf{h}\|_{\mathbf{R}_h^{-1}}^2, \quad (4.16)$$

where all the variables are defined as before. Instead of solving (4.16) globally, we aim to solve it in a distributed manner over R antennas in which the r th antenna has access to \mathcal{Y}_r only. Moreover, the antenna r is interested only in determining its own CIR (i.e., \mathbf{h}_r) without worrying about other \mathbf{h}_j 's. Here, we would like to mention that this problem is fundamentally different from those considered in the context of adaptive networks (see [74] and references therein). In particular, most of the existing distributed estimation techniques in adaptive networks deal with single task problems devoted to estimating a single common parameter of interest and rely on full cooperation between the nodes, i.e., exchanging both the estimates and the observations. Our proposed solution, the distributed LMMSE (D-LMMSE) algorithm, as will become clear, is much simpler in that it exploits the structure of spatial correlation matrix \mathbf{R}_a and relies only on exchanging the (partial) weighted estimates of CIRs with immediate neighbors, thus significantly reducing the communication and computational cost. The working principle of the proposed D-LMMSE algorithm is depicted in Fig. 4.2 which is composed of three main steps namely, estimation, sharing and updating, as explained below.

Estimation

In the estimation step, each antenna acting as a center antenna r_C , estimates not only its own CIR but also the CIRs of its neighborhood. The neighborhood of r_C consists of 4-direct neighbors represented by the set $\mathcal{N}=\{r_L, r_R, r_U, r_D\}^1$ on the left, right, top and bottom positions respectively as shown in Fig. 4.3(a). Now, let the corresponding channel vectors be represented by $\mathbf{h}_C, \mathbf{h}_L, \mathbf{h}_R, \mathbf{h}_U$ and \mathbf{h}_D , re-

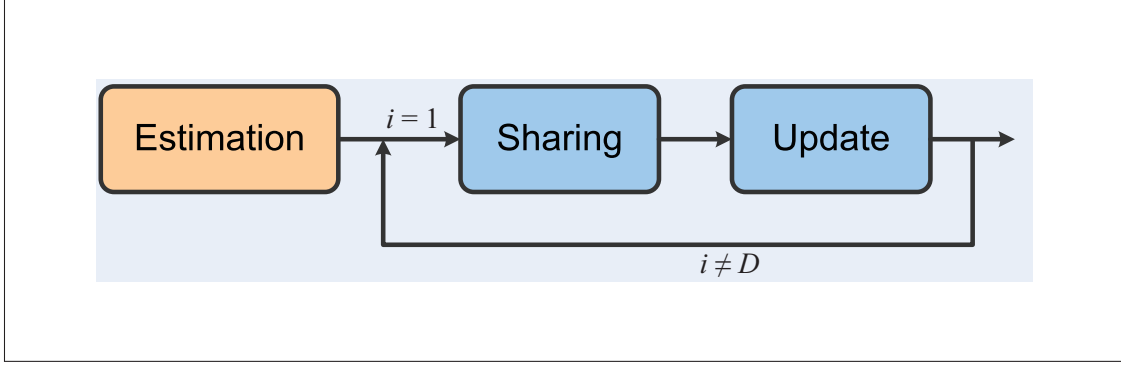


Figure 4.2: The working principle of D-LMMSE Algorithm

spectively, and let \mathbf{h}^c represent $|\mathcal{N}^+|L \times 1$ dimensional composite channel vector of the central antenna and its $|\mathcal{N}|$ direct neighbors (i.e., $\mathbf{h}^c = [\mathbf{h}_C^T, \mathbf{h}_L^T, \mathbf{h}_R^T, \mathbf{h}_U^T, \mathbf{h}_D^T]^T$). During the estimation, each antenna acting as a central element computes the estimate of \mathbf{h}^c by solving a reduced dimensional weighted least squares (WLS) optimization problem

$$\hat{\mathbf{h}}^c = \underset{\mathbf{h}^c}{\operatorname{argmin}} \|\mathcal{Y}_C(\mathcal{P}) - \mathbf{A}(\mathcal{P})\mathbf{h}_C\|_{\mathbf{R}_w}^2 + \|\mathbf{h}^c\|_{\mathbf{R}_{\mathbf{h}^c}}^2, \quad (4.17)$$

where $\mathcal{Y}_C(\mathcal{P})$ represents pilot observations at the central element, $\mathbf{R}_{\mathbf{h}^c}$ is channel correlation matrix defined as $\mathbf{R}_{\mathbf{h}^c} \triangleq \mathbb{E}\{\mathbf{h}^c(\mathbf{h}^c)^H\}$, which is the subset of full correlation matrix $\mathbf{R}_{\mathbf{h}}$ and $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_K$ is the noise covariance matrix at the central element. From (4.17) it is clear that information is processed locally at each antenna as each antenna uses only its own observations and interacts with its neighborhood only through $\mathbf{R}_{\mathbf{h}^c}$. The solution to the above WLS minimization problem can be obtained by first re-writing (4.17) explicitly in terms of \mathbf{h}^c as

follows

$$\hat{\mathbf{h}}^c = \underset{\mathbf{h}^c}{\operatorname{argmin}} \left\| \bar{\mathbf{y}} - \bar{\mathbf{A}}\mathbf{h}^c \right\|_{\mathbf{R}_w^{-1}}^2 + \|\mathbf{h}^c\|_{\mathbf{R}_{\mathbf{h}^c}^{-1}}^2, \quad (4.18)$$

where $\bar{\mathbf{y}} = \mathcal{Y}_C(\mathcal{P})$ and $\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A}(\mathcal{P}) & \mathbf{0}_{K \times L|\mathcal{N}|} \end{bmatrix}$. Then, by invoking the equivalence between LMMSE and WLS estimation problems we obtain,

$$\hat{\mathbf{h}}^c = (\mathbf{R}_{\mathbf{h}^c}^{-1} + \bar{\mathbf{A}}^H \mathbf{R}_w^{-1} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^H \mathbf{R}_w^{-1} \bar{\mathbf{y}}. \quad (4.19)$$

Now $\hat{\mathbf{h}}^c$ can be re-written as $\hat{\mathbf{h}}^c = (\mathbf{P}^c)^{-1} \hat{\mathbf{h}}_w^c$, where

$$\mathbf{P}^c = \mathbf{R}_{\mathbf{h}^c}^{-1} + \bar{\mathbf{A}}^H \mathbf{R}_w^{-1} \bar{\mathbf{A}}, \quad (4.20)$$

corresponds to the inverse error covariance matrix and $\hat{\mathbf{h}}_w^c$ represents the weighted estimate given by

$$\hat{\mathbf{h}}_w^c = \mathbf{P}^c \hat{\mathbf{h}}^c = \bar{\mathbf{A}}^H \mathbf{R}_w^{-1} \bar{\mathbf{y}}. \quad (4.21)$$

Above weighting by inverse error covariance matrix asserts that we put more confidence into the estimates which are more reliable and vice versa. After computing \mathbf{P} matrix and the weighted estimate $\hat{\mathbf{h}}_w^c$, each antenna is ready to move to the sharing step.

Sharing

The sharing step is the key to the proposed distributed algorithm where the information is shared through collaboration between antennas. Let us define the

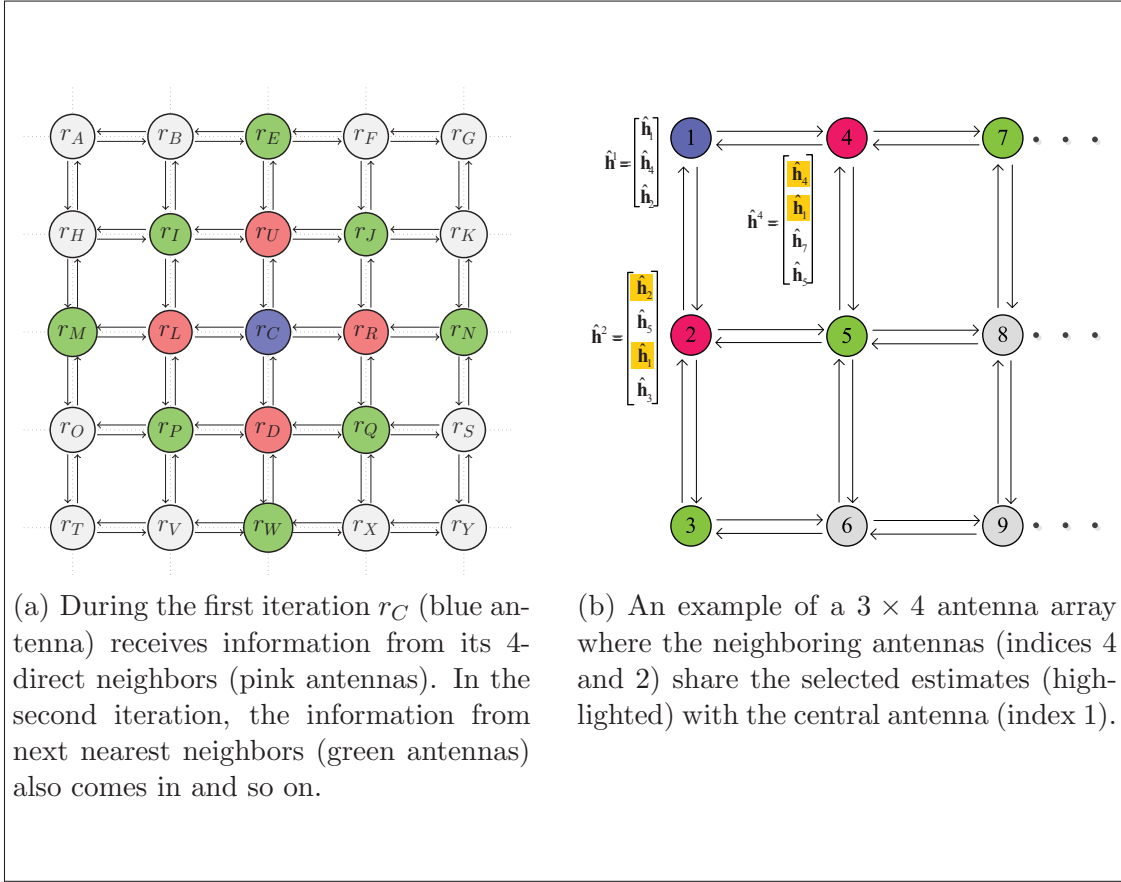


Figure 4.3: Details of (a) Information diffusion process and (b) Information sharing process.

sub-vector $\hat{\mathbf{h}}_{wj}$ of the composite vector $\hat{\mathbf{h}}_w^k$ as the (weighted) CIR estimate of antenna j computed by the antenna k . In sharing step, the antenna k would share only the selected components; its own (weighted) estimate $\hat{\mathbf{h}}_{wk}$ and the (weighted) estimate $\hat{\mathbf{h}}_{wj}$ with element $j \in \mathcal{N}$. Henceforth, the shared vectors will be termed as partial vectors and represented by an underline notation. An example of how this sharing takes place is also depicted in Fig. 4.3(b) for a 3×4 array with central element $r_C=1$ having only two neighbors; $\mathcal{N}=\{r_R=4, r_D=2\}$. As shown, each of the neighboring element shares only two sub-vectors (i.e., partial information) with the central antenna. The collaboration between the rest of the

elements takes place in similar fashion.

As a result of information sharing, each antenna acting as a central node r_C receives $|\mathcal{N}|$ partial vectors, $\hat{\mathbf{h}}_w^j, j \in \mathcal{N}$, from its neighbors, each of dimension $|\mathcal{N}^+|L \times 1$ and having only two non-zero components; $\hat{\mathbf{h}}_{wj}$ and $\hat{\mathbf{h}}_{wc}$. For the example in Fig. 4.3(b), the composite vector of the central node and the partial vectors received from its neighbors are given as follows,

$$\hat{\mathbf{h}}_w^1 = \begin{bmatrix} \hat{\mathbf{h}}_{w1} \\ \hat{\mathbf{h}}_{w4} \\ \hat{\mathbf{h}}_{w2} \end{bmatrix}, \quad \hat{\mathbf{h}}_w^4 = \begin{bmatrix} \hat{\mathbf{h}}_{w1} \\ \hat{\mathbf{h}}_{w4} \\ \mathbf{0} \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{h}}_w^2 = \begin{bmatrix} \hat{\mathbf{h}}_{w1} \\ \mathbf{0} \\ \hat{\mathbf{h}}_{w2} \end{bmatrix}. \quad (4.22)$$

Note that the estimates which are not shared are assigned as null vectors.

Update

Upon receiving the (partial) LMMSE estimates from the neighboring elements, each antenna acting as the central element updates its estimate and error covariance matrix. The update rule is summarized in the following Lemma [73]

Lemma 4.1 *Let \mathbf{y}_1 and \mathbf{y}_2 be two separate observations of a zero mean random vector \mathbf{h} , such that $\mathbf{y}_1 = \mathbf{A}_1 \mathbf{h} + \mathbf{w}_1$ and $\mathbf{y}_2 = \mathbf{A}_2 \mathbf{h} + \mathbf{w}_2$, where we assume that \mathbf{h} is uncorrelated with both \mathbf{w}_1 and \mathbf{w}_2 . Let $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ denote the LMMSE estimates of \mathbf{h} and \mathbf{C}_1 and \mathbf{C}_2 be the corresponding error covariance matrices in two experiments. Then, the optimal LMMSE estimator and the error covariance matrix of*

\mathbf{h} given both the observations are,

$$\mathbf{C}^{-1}\hat{\mathbf{h}} = \mathbf{C}_1^{-1}\hat{\mathbf{h}}_1 + \mathbf{C}_2^{-1}\hat{\mathbf{h}}_2, \quad (4.23)$$

and

$$\mathbf{C}^{-1} = \mathbf{C}_1^{-1} + \mathbf{C}_2^{-1} + \mathbf{R}_h^{-1} - \mathbf{R}_1^{-1} - \mathbf{R}_2^{-1}, \quad (4.24)$$

where $\mathbf{R}_h = \mathbb{E}\{\mathbf{h}\mathbf{h}^H\}$ and \mathbf{R}_1 and \mathbf{R}_2 are covariance matrices of \mathbf{h} in the two experiments.

Proof. See [73]. ▮

The aforementioned Lemma suggests an optimal way of combining the individual estimates and can be easily extended to more than two observations. We use this Lemma at each antenna to improve the initial channel estimate by combining it with the estimates computed and shared by $|\mathcal{N}|$ neighbors. Consequently, by treating each antenna as a central element r_C , the update equations are given by,

$$\hat{\mathbf{h}}_w^{c(i)} = \hat{\mathbf{h}}_w^{c(i-1)} + \sum_{j \in \mathcal{N}} \hat{\mathbf{h}}_w^{j(i-1)}, \quad (4.25)$$

and

$$\mathbf{P}^{c(i)} = \mathbf{P}^{c(i-1)} + \sum_{j \in \mathcal{N}} \left(\underline{\mathbf{P}}^{j(i-1)} - \underline{\mathbf{R}}_{\mathbf{h}^j}^{-1} \right), \quad (4.26)$$

respectively, where $\underline{\mathbf{P}}^j$ and $\underline{\mathbf{R}}_{\mathbf{h}^j}$ represent the partial (inverse) error covariance and correlation matrices associated with the partial estimates $\hat{\mathbf{h}}_w^j$ and i represents the iteration index.

The recursions in the update equations are initialized by (4.21) and (4.20), respectively, which are available after the estimation step. In the subsequent iterations, each antenna would also require the partial matrices, $\underline{\mathbf{P}}^j$'s and $\underline{\mathbf{R}}_{\mathbf{h}^j}$'s, for each of its $|\mathcal{N}|$ neighbors. Fortunately, they can be obtained from \mathbf{P}^c and $\mathbf{R}_{\mathbf{h}^c}$, respectively (which are available at the central antenna) by exploiting the symmetrical structure of \mathbf{R}_a . Thus, there is no need to share them across the neighboring elements, that in turn saves a significant amount of communication burden. Specifically, the matrices $\mathbf{R}_{\mathbf{h}^c}$ and \mathbf{P}^c exhibit the following two properties:

Property 1: The matrix $\mathbf{R}_{\mathbf{h}^c}$ is identical for all elements in the neighborhood of r_C i.e., $\mathbf{R}_{\mathbf{h}^c} = \mathbf{R}_{\mathbf{h}^j}, \forall j \in \mathcal{N}$

Property 2: The matrix \mathbf{P}^c is identical for all elements in the neighborhood of r_C i.e., $\mathbf{P}^c = \mathbf{P}^j, \forall j \in \mathcal{N}$

Property 1 is due to symmetric nature of the spatial correlation matrix \mathbf{R}_a which implies that the spatial correlation between any two antennas placed equidistant apart, is the same. Therefore, it is not difficult to see that property 1 holds exactly under the Kronecker model and our earlier assumption of identical tap correlation across the antenna array in Section 4.1. Property 2 is the consequence of property 1 when incorporated into (4.20). Note that these properties are generally satisfied as the spatial correlation matrix is generally symmetric, if not, then antennas can share these matrices as well.

Hence, to obtain the partial correlation matrices, $\underline{\mathbf{R}}_{\mathbf{h}^j}, j \in \mathcal{N}$, we first set $\underline{\mathbf{R}}_{\mathbf{h}^j} = \mathbf{R}_{\mathbf{h}^c}$ and then modify the off-diagonal block entries corresponding to the null vectors of partial estimates as $\underline{\mathbf{R}}_{ij} = \mathbf{0}$ if any $\hat{\mathbf{h}}_{wi}, \hat{\mathbf{h}}_{wj} = \mathbf{0}$ and the diagonal block entries as $\underline{\mathbf{R}}_{ii} = \mathbf{I}_L$ if $\hat{\mathbf{h}}_{wi} = \mathbf{0}$, where the subscript ij denotes the (i, j) th block. The matrices $\underline{\mathbf{P}}^j$'s are obtained similarly except that the diagonal block entry corresponding to null vectors is replaced by $a\mathbf{I}$ where $0 < a \ll 1$ is a small positive number, which indicates very low weight or confidence in the estimates that are not shared. In essence, the central element has the full information needed to construct $\underline{\mathbf{P}}^j$'s and $\underline{\mathbf{R}}_{\mathbf{h}^j}$'s. For the example in Fig. 4.3(b), the partial correlation and error covariance matrices associated with estimates in (4.22) are given in (4.27) and (4.28), respectively. Based on above steps and procedures, the proposed D-LMMSE algorithm is summarized in Algorithm 4.1.

$$\underline{\mathbf{R}}_{\mathbf{h}^1} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{14} & \mathbf{R}_{12} \\ \mathbf{R}_{41} & \mathbf{R}_{44} & \mathbf{R}_{42} \\ \mathbf{R}_{21} & \mathbf{R}_{24} & \mathbf{R}_{22} \end{bmatrix}, \underline{\mathbf{R}}_{\mathbf{h}^4} = \begin{bmatrix} \mathbf{R}_{44} & \mathbf{R}_{41} & \mathbf{0} \\ \mathbf{R}_{14} & \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_L \end{bmatrix} \text{ and } \underline{\mathbf{R}}_{\mathbf{h}^2} = \begin{bmatrix} \mathbf{R}_{22} & \mathbf{0} & \mathbf{R}_{21} \\ \mathbf{0} & \mathbf{I}_L & \mathbf{0} \\ \mathbf{R}_{12} & \mathbf{0} & \mathbf{R}_{11} \end{bmatrix} \quad (4.27)$$

$$\underline{\mathbf{P}}^1 = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{14} & \mathbf{P}_{12} \\ \mathbf{P}_{41} & \mathbf{P}_{44} & \mathbf{P}_{42} \\ \mathbf{P}_{21} & \mathbf{P}_{24} & \mathbf{P}_{22} \end{bmatrix}, \underline{\mathbf{P}}^4 = \begin{bmatrix} \mathbf{P}_{44} & \mathbf{P}_{41} & \mathbf{0} \\ \mathbf{P}_{14} & \mathbf{P}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & a\mathbf{I} \end{bmatrix} \text{ and } \underline{\mathbf{P}}^2 = \begin{bmatrix} \mathbf{P}_{22} & \mathbf{0} & \mathbf{P}_{21} \\ \mathbf{0} & a\mathbf{I} & \mathbf{0} \\ \mathbf{P}_{12} & \mathbf{0} & \mathbf{P}_{11} \end{bmatrix} \quad (4.28)$$

Remarks:

Algorithm 4.1 Distributive LMMSE (D-LMMSE) algorithm

1. **(Estimation)** Each antenna acting as a central element r_C computes $\hat{\mathbf{h}}_w^c$ and \mathbf{P}^c by using (4.21) and (4.20) respectively.
 2. **(Sharing)** Each antenna acting as a central element r_C shares partial estimates, $\hat{\mathbf{h}}_w^c$ with its $|\mathcal{N}|$ neighbors as described in Section 4.3.
 3. **(Pre-processing)** Using $\mathbf{R}_{\mathbf{h}^c}$, \mathbf{P}^c from step 1 and the received (partial) information $\{\hat{\mathbf{h}}_w^j\}_{j=1}^{|\mathcal{N}|}$ in step 2, each antenna, acting as a central element r_C , constructs $\{\mathbf{R}_{\mathbf{h}^j}^{-1}\}$, $\{\mathbf{P}^j\}$, $j \in \mathcal{N}$.
 4. **(Update)** Each antenna acting as a central element r_C , updates its weighted estimate and error covariance using (4.25) and (4.26) respectively.
 5. **(Iterate)** Repeat steps 2-4 D times, where D is the maximum number of iterations.
 6. **(Output)** Compute $\hat{\mathbf{h}}^c = (\mathbf{P}^c)^{-1} \hat{\mathbf{h}}_w^c$ and output the estimated CIR $\hat{\mathbf{h}}_C$.
-

1. Information sharing and update take place during each iteration of the algorithm such that after few iterations the information diffuses swiftly across the whole array resulting in fast convergence. This concept of sharing is depicted in Fig. 4.3(a).
2. The repetitive sharing enables each antenna in the array to utilize the observations from distant elements, thereby improving its estimate in each iteration till it converges to near optimal solution.
3. As opposed to the centralized processing, the proposed sharing step is more convenient and computationally more efficient as all antennas do not communicate with each other. The collaboration takes place only among the neighboring antennas. Therefore, the complexity of proposed algorithm is significantly less than the centralized approach.

4. Note that the antennas share only the partial information because only selected vectors are transmitted to the neighbors which save significant amount of communication. Also, the estimation step and repetitive sharing and update steps require simple linear block processing and have a fixed size data structure which is well suited for real implementations. In contrast, the memory and processing requirements for the centralized approach are even more challenging with large array dimensions.

4.3.1 Complexity Analysis

In Table 4.1, we compare the computational complexity of proposed D-LMMSE algorithm with LS, L-LMMSE and the centralized O-LMMSE algorithm in terms of multiply and add operations. The figures indicate that complexity of proposed algorithm is slightly higher than L-LMMSE but is significantly less than the centralized approach. It is also worth mentioning here that, the \mathbf{P} matrices in (4.20) can be computed off-line and in parallel at all antennas as they do not depend on observations. Moreover, the computation of weighted estimates in (4.21) does not involve any matrix inversion. Further, the update in (4.25) requires simple addition during each step of iteration, while (4.26) needs one time computations of inversions $\underline{\mathbf{R}}_{\mathbf{h}^j}^{-1}$ as they do not depend on iteration index. Finally, the computation of inverse, $(\mathbf{P}^c)^{-1}$ is required but only after the convergence when each antenna outputs its final estimate.

Table 4.1: Computational Complexity

Algorithm	Multiplications (\times)	Additions (+)	Complexity
LS	$RK(L + 1)$	$R(KL - 1)$	$O(RLK)$
L-LMMSE	$R[2L^3 + L^2 + K(L + 1)]$	$RL[L + K - 1]$	$O(RL^3)$
O-LMMSE	$R[(L^3 + 1)R^2 + RL(L + K) + K] + L^3$	R^2LK	$O(R^3L^3)$
D-LMMSE	$R[(5^3 + 1)L^3 + 2(5L)^2 + L(K + 1) + 5^3]$	$R[D(5L)^3 + (5L)^2 + L(K - 1) - D]$	$O(RL^3)$

Choice of Parameter D

The choice of parameter D i.e., the maximum number of required iterations, has a great influence on computational complexity and convergence of D-LMMSE algorithm. A trivial choice for D is that it can be set to the largest dimension of the array i.e., $D = \max(M, G)$, which ensures that each antenna receives information from every other antenna in the array. However, such a high value of D is very inefficient from the computational complexity point of view, particularly when the array dimensions are large. A simple loose upper bound on D can be derived by noting that total number of antennas sharing information in D iterations are $2D(D + 1) + 1$. Hence, we should have $2D(D + 1) + 1 \leq R$ which leads to

$$D \leq \sqrt{\frac{R}{2} - \frac{1}{4}} - \frac{1}{2}. \quad (4.29)$$

It must be emphasised here, that the actual value of D also depends on the spatial correlations among antennas. Specifically, if the antennas are not very strongly correlated, then we might not gain from sharing and a small number of iterations might be sufficient.

4.3.2 Linkage Between Localized and Centralized Solutions

Our distributed LMMSE approach can be seen as a hybrid of the localized (L-LMMSE) and the centralized (O-LMMSE) estimation approaches. Specifically, we derive the linkage between L-LMMSE and O-LMMSE solutions to justify our estimation and collaboration approach as a viable strategy. To this end, we take a closer look into the optimal LMMSE solution by decomposing (4.10) as follows

$$\begin{aligned}
\hat{\mathbf{h}} &= \left(\mathbf{R}_h^{-1} + \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^H \mathbf{R}_w^{-1} \mathcal{Y} \\
&= \left[(\mathbf{R}_a \otimes \mathbf{R}_{tap})^{-1} + (\mathbf{I}_R \otimes \mathbf{A})^H (\mathbf{I}_R \otimes \mathbf{R}_w)^{-1} (\mathbf{I}_R \otimes \mathbf{A}) \right]^{-1} (\mathbf{I}_R \otimes \mathbf{A})^H (\mathbf{I}_R \otimes \mathbf{R}_w)^{-1} \mathcal{Y} \\
&= \left[\mathbf{R}_a^{-1} \otimes \mathbf{R}_{tap}^{-1} + \mathbf{I}_R \otimes (\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A}) \right]^{-1} \left[\mathbf{I}_R \otimes (\mathbf{A}^H \mathbf{R}_w^{-1}) \right] \mathcal{Y} \tag{4.30}
\end{aligned}$$

For simplicity we assume that $\mathbf{R}_{tap} = \mathbf{I}_L$ and after introducing the matrix $\mathbf{I}_R \otimes (\mathbf{I}_L + \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})$ and its inverse between brackets $[\cdot]^{-1}[\cdot]$ of (4.30), we get

$$\begin{aligned}
\hat{\mathbf{h}}^{(opt)} &= \left[\mathbf{R}_a^{-1} \otimes \mathbf{I}_L + \mathbf{I}_R \otimes (\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A}) \right]^{-1} \left[\mathbf{I}_R \otimes (\mathbf{I}_L + \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A}) \right] \hat{\mathbf{h}}^{(L)} \\
&= \left[\mathbf{R}_a^{-1} \otimes \mathbf{I}_L + \mathbf{I}_R \otimes (K\rho) \mathbf{I}_L \right]^{-1} \left[\mathbf{I}_R \otimes (\mathbf{I}_L + K\rho \mathbf{I}_L) \right] \hat{\mathbf{h}}^{(L)} \\
&= \left[(\mathbf{R}_a^{-1} + K\rho \mathbf{I}_R)^{-1} \otimes \mathbf{I}_L \right] \left[(1 + K\rho) \mathbf{I}_R \otimes \mathbf{I}_L \right] \hat{\mathbf{h}}^{(L)} \\
&= \left[(\mathbf{R}_a^{-1} + K\rho \mathbf{I}_R)^{-1} (1 + K\rho) \otimes \mathbf{I}_L \right] \hat{\mathbf{h}}^{(L)} \tag{4.31}
\end{aligned}$$

To get further insight, we use the EVD $\mathbf{R}_a = \mathbf{V}\mathbf{S}\mathbf{V}^H$, where \mathbf{S} is diagonal matrix of eigenvalues $\{\eta_i\}$, to obtain

$$\begin{aligned}\hat{\mathbf{h}} &= \left[\mathbf{V} (\mathbf{S}^{-1} + K\rho\mathbf{I}_R)^{-1} (1 + K\rho)\mathbf{V}^H \otimes \mathbf{I}_L \right] \hat{\mathbf{h}}^{(L)} \\ &= [\mathbf{V}\mathbf{\Delta}\mathbf{V}^H \otimes \mathbf{I}_L] \hat{\mathbf{h}}^{(L)}\end{aligned}\quad (4.32)$$

where $\mathbf{\Delta} = (\mathbf{S}^{-1} + K\rho\mathbf{I}_R)^{-1} (1 + K\rho)$ is a diagonal matrix with entries, $\eta_i(1 + K\rho)/(1 + K\rho\eta_i), i=1, \dots, R$. Relation in (4.32) provides a linkage between L-LMMSE and O-LMMSE solutions and suggests that collaboration among antennas (through sharing the local estimates) is necessary to get the optimal solution which is the key step of our distributed strategy. Moreover, the level of collaboration depends on the spatial correlation values. Specifically, in case of no correlation, i.e., $\eta_i=1, \forall i$, there would be no advantage of collaboration and the optimal solution would converge to the localized one.

4.4 Data-Aided D-LMMSE Estimation

The basic idea of data-aided channel estimation is to exploit the decoded data-carriers in order to improve the initial channel estimates obtained using pilots. It is possible that some of the data-carriers are erroneous due to noise and channel estimation errors, while some of the other data-carriers are reliable i.e., they are likely to be decoded correctly. An important problem is how to down-select a subset of the most reliable data-carriers to be used as data-pilots. For this purpose,

we shall use the same idea presented earlier in Section 3.5. The working principle of the proposed Data-Aided D-LMMSE (DAD-LMMSE) algorithm is described in Fig. 4.4 which is essentially the same as D-LMMSE with two additional steps tailored to improving initial channel estimates by selecting the most reliable data-carriers as explained next.

4.4.1 Reliable Carriers Selection

Consider the received OFDM symbol at any antenna as shown in (4.4), and let $\hat{\mathbf{h}}$ and $\hat{\mathcal{H}} \triangleq \sqrt{N}\mathbf{F}\hat{\mathbf{h}}_r$ be the CIR and CFR estimates obtained using pilots. Then, the tentative estimates of the data symbols are obtained by equalizing the received OFDM symbol using zero-forcing (ZF) as follows

$$\begin{aligned}\hat{\boldsymbol{\chi}}(k) &= \frac{\boldsymbol{\mathcal{Y}}(k)}{\hat{\mathcal{H}}(k)}, \quad k \in \{1, 2, \dots, N\} \setminus \mathcal{P} \\ &\approx \boldsymbol{\mathcal{X}}(k) + \frac{\boldsymbol{\mathcal{W}}(k)}{\hat{\mathcal{H}}(k)} = \boldsymbol{\mathcal{X}}(k) + \boldsymbol{\mathcal{Z}}(k),\end{aligned}\tag{4.33}$$

where $\boldsymbol{\mathcal{Z}}(k)$ represents the distortion on k -th data-carrier due to noise and channel estimation error. Given the CFR estimate, $\boldsymbol{\mathcal{Z}}(k)$ can be modelled as Gaussian with zero mean and variance $\sigma_z^2 = \hat{\mathcal{H}}(k)^{-2}\sigma_w^2$. The recovery of data symbols is then performed by simple hard decisions on estimated symbols $\hat{\boldsymbol{\chi}}(k)$ denoted by $\langle \hat{\boldsymbol{\chi}}(k) \rangle$. Clearly, the errors in the decoding process occur due to both noise and inaccurate channel estimates. Hence, some data-carriers would be severely effected by these distortions and fall outside their correct decision regions. All those data

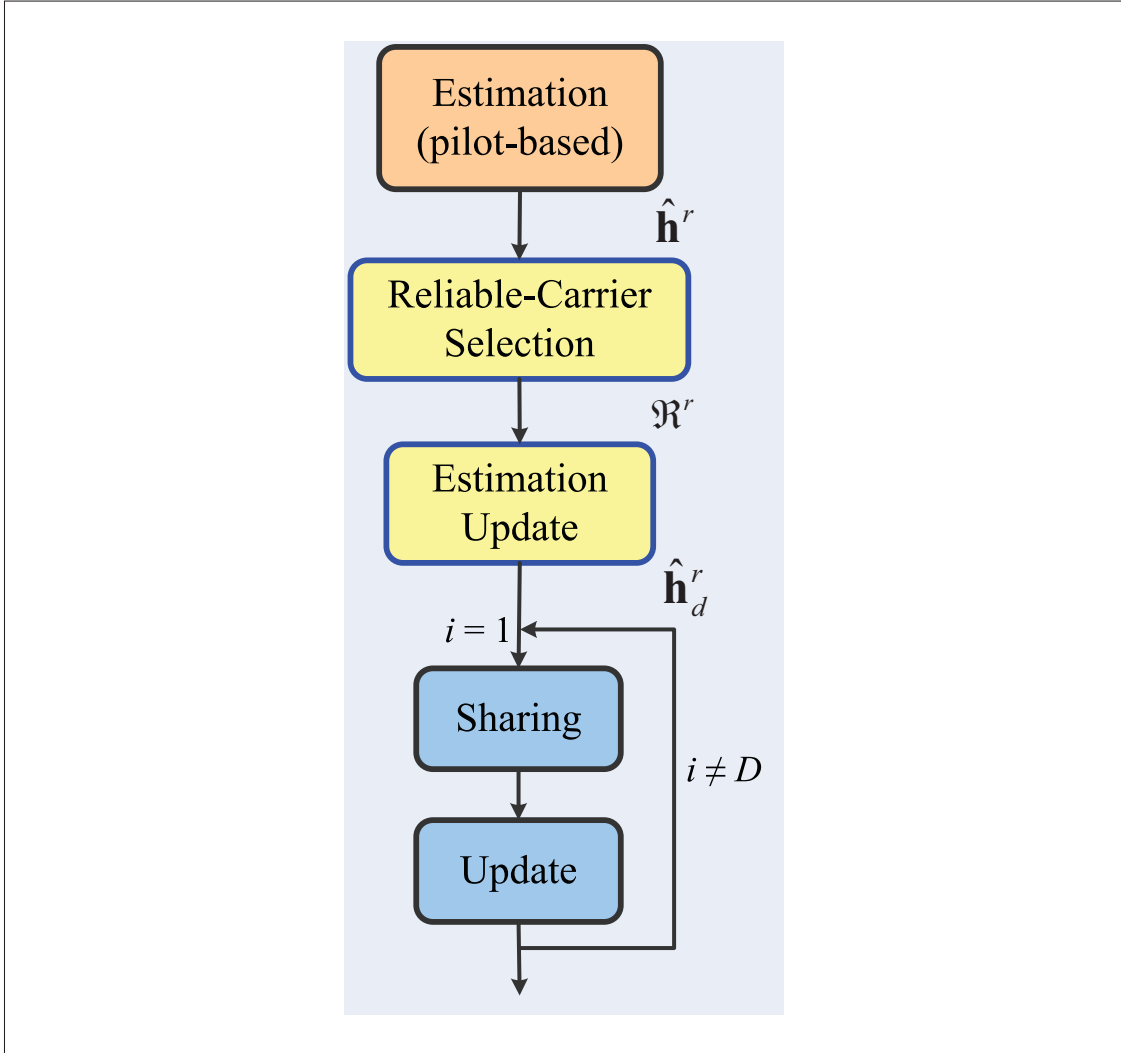


Figure 4.4: The working principle of DAD-LMMSE Algorithm

carriers $\hat{\mathcal{X}}(k)$ which satisfy the condition $\langle \hat{\mathcal{X}}(k) \rangle = \mathcal{X}(k)$ with high probability, are termed reliable carriers.

The proposed strategy for selecting the subset $\mathcal{R} \subset N \setminus \mathcal{P}$ of the most reliable data-carriers is based on the criteria [70]

$$\mathfrak{R}(k) = \frac{f_z \left(\mathcal{Z}(k) = \mathcal{X}(k) - \langle \hat{\mathcal{X}}(k) \rangle \right)}{\sum_{m=1, \Omega_m \neq \langle \hat{\mathcal{X}}(k) \rangle}^{|\Omega|} f_z \left(\mathcal{Z}(k) = \mathcal{X}(k) - \Omega_m \right)}, \quad (4.34)$$

where $f_z(\cdot)$ is the *pdf* of $\mathcal{Z}(k)$ and $\{\Omega_m\}$ represents the set of constellation al-

phabets. Note that the numerator in (4.34) is the probability that $\mathcal{X}(k)$ will be decoded correctly while the denominator sums the probabilities of all possible incorrect decisions due to distortion $\mathcal{Z}(k)$. The subset \mathcal{R} is formed by selecting only those data-carriers for which $\mathfrak{R}(k) > 1$, i.e.,

$$\mathcal{R} = \{k \mid \mathfrak{R}(k) > 1\} . \quad (4.35)$$

4.4.2 Revisiting the Estimation Step

We now revisit the estimation step of the proposed Algorithm 4.1 using both the pilots and reliable carriers in order to enhance the initial estimates. Let \mathcal{R}^r be the set of indices of reliable data carriers for antenna r . Each antenna could revisit the estimation step by solving (4.17) using an extended set of indices, $\mathcal{P} \cup \mathcal{R}^r$ corresponding to pilots and reliable data carriers. However, since the pilots have already been utilized to obtain an estimate $\hat{\mathbf{h}}^r$ of CIR, we simply need to update this estimate based on reliable data. Thus, using the block form of RLS [62], we can write

$$\hat{\mathbf{h}}_d^r = \hat{\mathbf{h}}^r + \mathbf{C}_e^r \bar{\mathbf{A}}_d^H \mathbf{G} \left(\bar{\mathbf{y}}_d - \bar{\mathbf{A}}_d \hat{\mathbf{h}}^r \right) , \quad (4.36)$$

where the gain matrix \mathbf{G} is defined as

$$\mathbf{G} = \left(\mathbf{R}_w + \bar{\mathbf{A}}_d \mathbf{C}_e^r \bar{\mathbf{A}}_d^H \right)^{-1} . \quad (4.37)$$

and the corresponding error covariance matrix is evaluated as

$$\mathbf{C}_{ed}^r = \mathbf{C}_e^r - \mathbf{C}_e^r \bar{\mathbf{A}}_d^H \mathbf{G} \bar{\mathbf{A}}_d. \quad (4.38)$$

Here $\bar{\mathcal{Y}}_d = \mathcal{Y}_r(\mathcal{P} \cup \mathcal{R}^r)$ is an extended set of observations and $\bar{\mathbf{A}}_d = \begin{bmatrix} \mathbf{A}(\mathcal{P} \cup \mathcal{R}^r) & \mathbf{0}_{|\mathcal{P} \cup \mathcal{R}^r| \times |\mathcal{N}|L} \end{bmatrix}$ is the extended data matrix. The data-aided approach is described in Algorithm 4.2.

Algorithm 4.2 Data-aided Distributed LMMSE (DAD-LMMSE) Algorithm

1. Each antenna acting as a central element r_C computes $\hat{\mathbf{h}}^c$ and \mathbf{C}_e^r by using (4.19) and (4.20) respectively.
 2. Each antenna uses its CIR estimate, $\hat{\mathbf{h}}^r$ to form the subset \mathcal{R}^r of the most reliable data-carriers.
 3. Update the estimates and error covariance in step (1) using (4.36)-(4.38).
 4. Run steps (2)-(6) of Algorithm 4.1, with $\mathbf{P}^r = (\mathbf{C}_e^r)^{-1}$ and $\hat{\mathbf{h}}_w^r = \mathbf{P}^r \hat{\mathbf{h}}^r$.
-

4.5 Simulation Results

We adopt the channel model in (4.1) with spatial correlation matrix given in (4.3) whose parameters are: $\phi = \pi/3$ (mean horizontal AoD in radians), $\theta = 3\pi/8$ (mean vertical AoD in radians), $\sigma = \pi/12$ (standard deviation of horizontal AoD) and $\xi = \pi/36$ (standard deviation of vertical AoD). The channel tap correlation matrix follows an exponentially decaying PDP, $\mathbb{E}\{|h_r(\tau)|^2\} = e^{-\tau}$, while rest of the parameters are given in Table 4.2, where ν represents the carrier frequency wavelength in meters. It is also assumed that receiver has the knowledge of channel

correlations.

To assess the performance of different algorithms we use the following MSE performance criterion:

$$\text{MSE} = \frac{1}{\Theta} \sum_{i=1}^{\Theta} \|\mathbf{h}^i - \hat{\mathbf{h}}^i\|^2 \quad (4.39)$$

where \mathbf{h}^i and $\hat{\mathbf{h}}^i$ are true and estimated CIR vectors in the i th trial respectively, each of size $RL \times 1$ and Θ represents the total number of trials. We used $\Theta=100$ in our simulations.

We conduct different experiments to study the performance of our proposed distributed approach and compare it with the three methods i.e., LS, L-LMMSE and O-LMMSE described earlier in Section 4.2.

How Many Iterations (D)?

In this experiment we are interested in finding the number of iterations, required for convergence of the proposed distributed LMMSE algorithm. We plot the MSE of proposed D-LMMSE algorithm (red curve) against the parameter D (i.e., number of iterations) in Fig. 4.5(a). The SNR was fixed at 0 dB. The MSE values of other algorithms, which do not depend on parameter D , are also shown. It can be seen that the proposed algorithm converges very closely to the optimal in 3 iterations. Note that, when the antennas do not collaborate (i.e., $D=0$), the MSE of distributed algorithm coincides with that of L-LMMSE because no information sharing takes place. As the information from neighbors comes in during the next few iterations, the MSE decays exponentially until it converges to near optimal

Table 4.2: Simulation Parameters

Parameter	Value
Array Size ($M \times G$)	10×10
Array element spacing d_x, d_y	$0.3\nu, 0.5\nu$
Number of OFDM sub-carriers (N)	256
Number of pilots (K)	32
Signal constellation modulation	4/16/64 – QAM
Channel length (L)	8

solution. Fig. 4.5(b) also suggests that there would be hardly any improvement in MSE for $D > 3$.

The impact of antenna correlation on convergence is depicted in Fig. 4.6, where we plot MSEs with different antenna spacings. It is evident that as the correlation gets larger (i.e., element spacing decreases) the number of iterations required for convergence gets closer to the bound in (4.29) depicted by vertical dashed blue line. In case of weak correlation, the collaboration may not be beneficial and the algorithm requires less number of iterations for convergence.

MSE Performance in AWGN

In this experiment, we compare the MSE performance of proposed distributed algorithms with various algorithms in the presence of AWGN using the parameters in Table 4.2. The results given in Fig. 4.7, show that O-LMMSE performs better than both LS and L-LMMSE in terms of MSE as it is able to utilize the antenna spatial correlations. As shown, the proposed D-LMMSE algorithm (Algorithm 4.1) achieves near optimal results in just 3 iterations. It is obvious that data-

aided approach has the best performance compared to all others and that the effect of using reliable carriers is more pronounced at higher SNR. The analytical MSE expressions given in Section 4.2, for LS, L-LMMSE and O-LMMSE under AWGN are also plotted with legends (Th.), which agree with simulation results.

Fig. 4.8 demonstrates the MSE behaviour of different algorithms with varying number of pilots K at SNR of 20 dB. As is shown, increasing the pilot tones yields better estimation performance but this comes at the cost of lower spectral efficiency. The data-aided algorithm however, is able to achieve the best performance even for a small number of pilot observations.

Computational Complexity

Finally we compare the average runtime of various algorithms that can be regarded as a measure of computational complexity. Fig. 4.9 shows the average runtime with increasing number of BS antennas under the default simulation parameters of Table 4.2. It is clear that computational requirements for proposed D-LMMSE algorithm, with different values of parameter D , grow at much slower pace than that of the O-LMMSE algorithm as the number of BS antenna increases. Further, in terms of memory requirements and communication overhead (not shown here), the advantages of D-LMMSE are even more tangible.

4.6 Concluding Remarks

Channel estimation is a challenging problem in massive MIMO systems as the conventional techniques applicable to MIMO systems cannot be employed owing to an exceptionally large number of unknown channel coefficients. We proposed a distributed algorithm where, each antenna estimates its own CIR but in a collaborative manner by sharing information with the neighbors. As a result, the proposed algorithm attains near optimal solution at a significantly reduced complexity than the centralized strategy. To reduce the pilots overhead, the distributed LMMSE algorithm is extended using data-aided estimation based on reliable carriers.

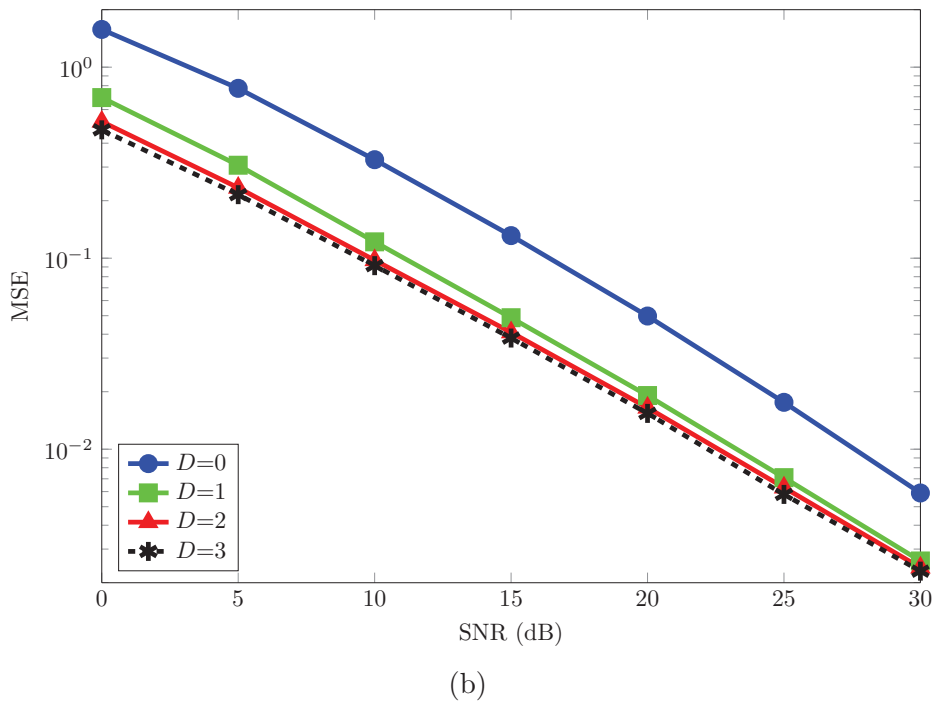
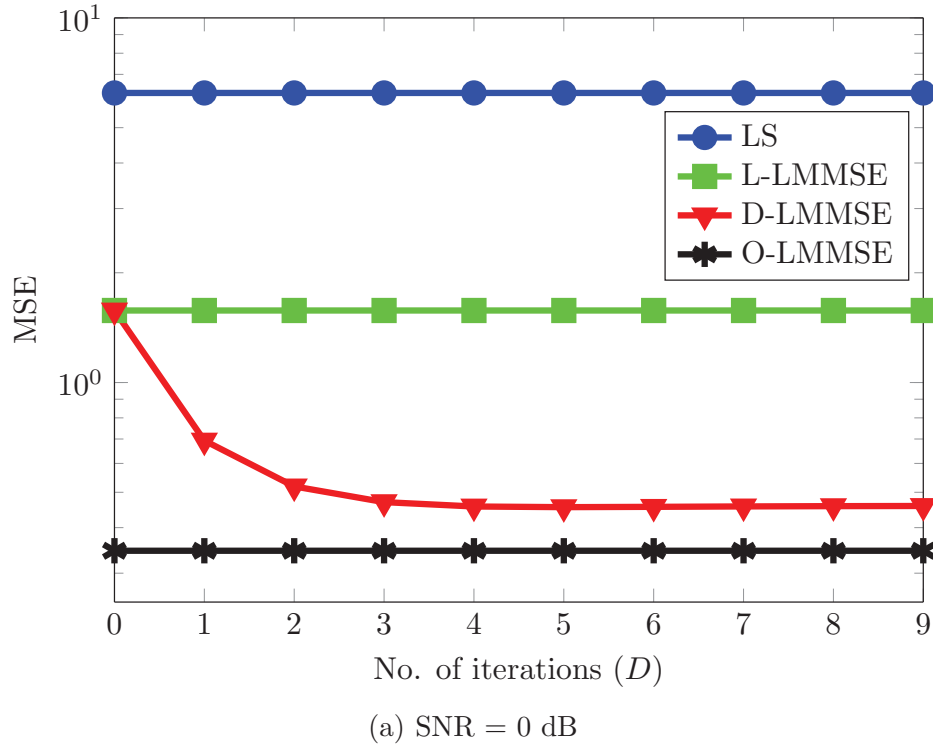


Figure 4.5: Number of iterations (D) required for convergence of the proposed distributed algorithm

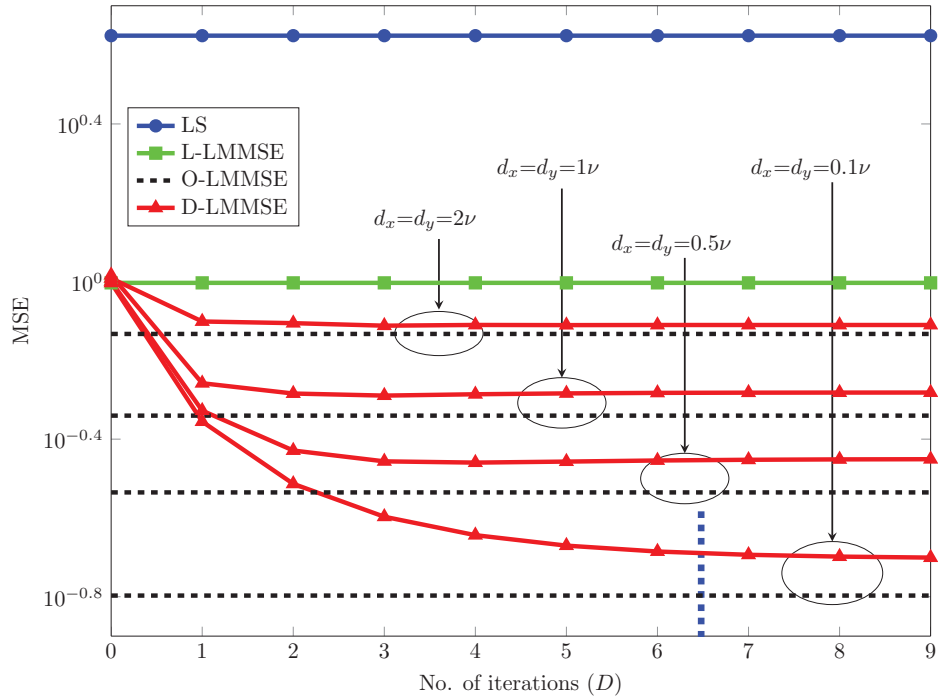


Figure 4.6: Effect of spatial correlation on convergence of the proposed distributed algorithm.

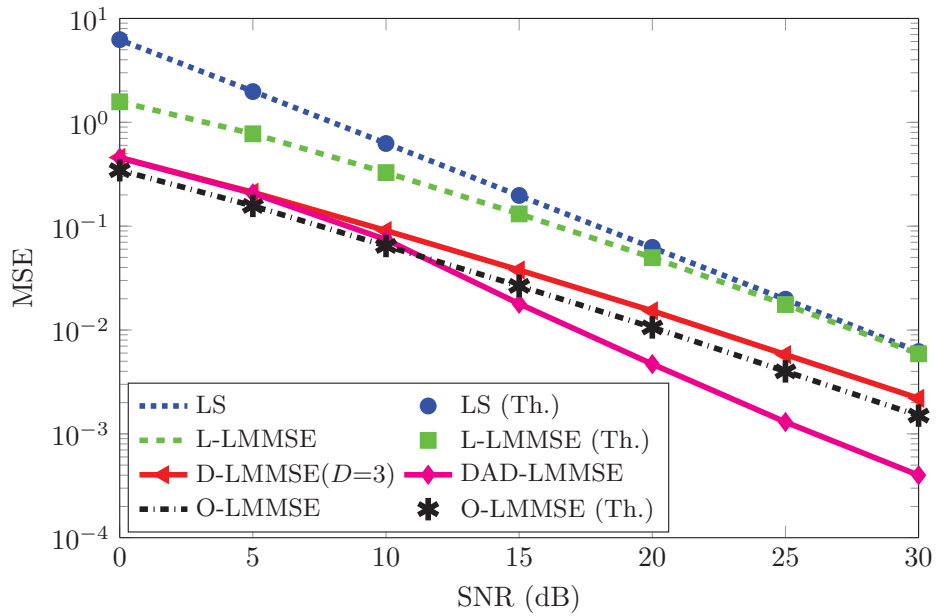


Figure 4.7: MSE performance comparison of different algorithms.

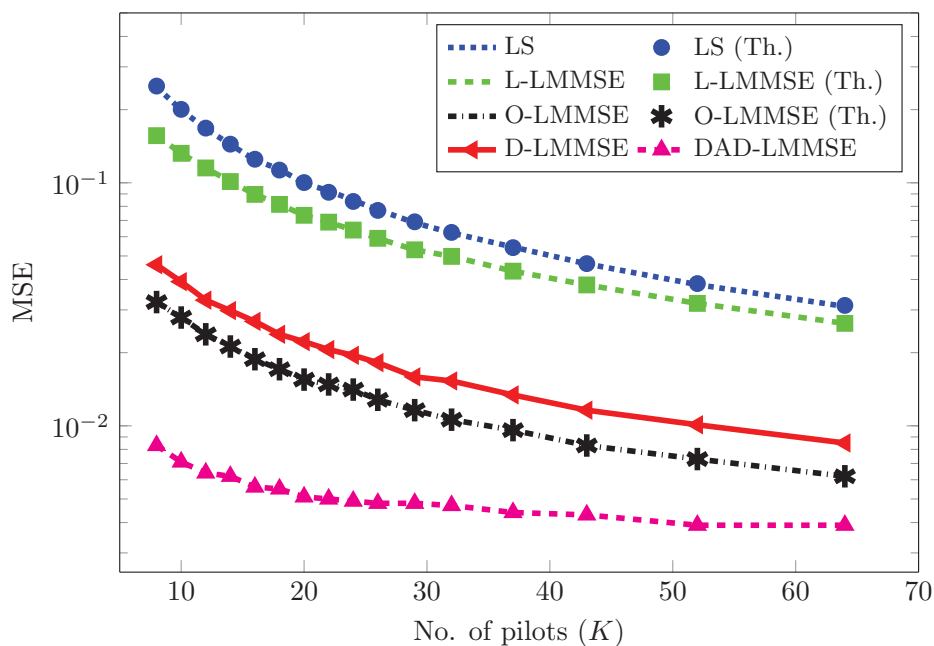


Figure 4.8: MSE performance comparison of the proposed data-aided algorithm with pilot-based techniques.

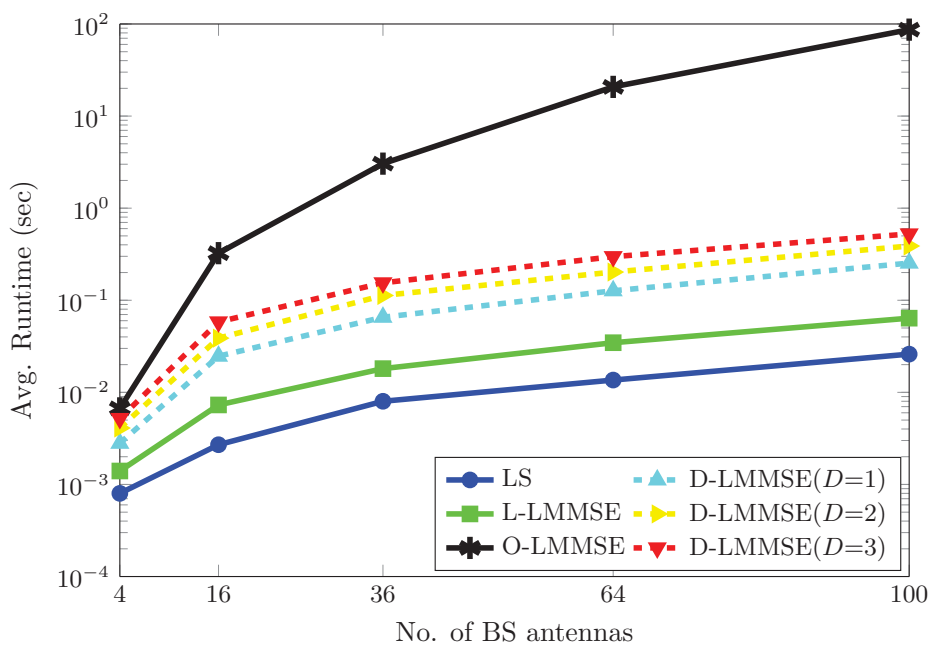


Figure 4.9: Average runtime of various algorithms.

CHAPTER 5

EFFECT OF PILOT CONTAMINATION ON CHANNEL ESTIMATION

In Chapter 4 we assumed a single-cell scenario where all the users are allocated orthogonal resources for channel estimation, thus the pilot observations were corrupted only by AWGN. In multi-cell environment, the pilot contamination due to aggressive reuse of the pilots across different cells, has severe impact on channel estimation performance. In fact, it was shown in [75] that the effect of uncorrelated interference and fast Rayleigh fading diminishes as the number of BS antennas increase while the effect of pilot contamination is not eliminated. This Chapter is exclusively devoted to investigating the effect of pilot contamination on channel estimation performance. The effect of pilot contamination on system performance has been analysed by many researches [54, 55], but only few studies have analysed

its impact on channel estimation performance [47]. Moreover, in these works, the analysis is carried out for fixed locations of (interference) users. As such, the existing studies cannot analytically answer how the randomness of user's locations would effect MSE performance under pilot contamination. In contrast, we approach the problem by using concepts from the stochastic geometry. By assuming that interfering users are distributed according to a homogeneous poisson point process (PPP), we obtain an analytical expression for the interference variance (or power) across OFDM frequency tones and use it to derive MSE expressions for LS and LMMSE based channel estimation algorithms under both noise and pilot contaminated regimes. Analytical expressions are validated by simulations.

5.1 Pilot Contamination and Implications

An important characteristic of the wireless channel, as already pointed out in Section 1.2, is its finite coherence interval, i.e., the number of time-frequency resources over which the channel can be assumed to be approximately constant (or static) is finite. Therefore, channel estimates obtained via pilots are valid only within this interval, after which the channel must be estimated again. Moreover, considering the preferred TDD mode in massive MIMO, the number of mutually orthogonal pilot sequences are limited by this coherence interval.

The increasing number of users in massive MIMO will quickly exhaust the availability of orthogonal pilot sequences, especially when the coherence interval is short. This means that only the users within a particular cell can be assigned or-

thogonal pilot sequences to avoid intra-cell interference and pilots must be reused across different cells. The consequence of pilot reuse leads to inter-cell interference or pilot contamination, resulting in impairments of channel estimates during uplink as depicted in Fig. 5.1(a). The worst case pilot contamination occurs when the same set of orthogonal pilot sequences are used in each cell.

Since the interference takes place only with the users sharing the same pilots, we assume single user in each cell without loss of generality and that each user is transmitting identical pilots which represents the worst case scenario from pilot contamination point of view. Let us assume that cell i is of interest and all other cells $\{j\}_{j=1, j \neq i}^J$ are interfering cells as depicted in Fig. 5.1(a) for $J = 2$. The received OFDM symbol at antenna r of i -th BS, after omitting the index \mathcal{P} for convenience, can be written as,

$$\mathbf{y}_r = \mathbf{A}\mathbf{h}_{r,i} + \underbrace{\sum_{j=1, j \neq i}^J \beta_j \mathbf{A}\mathbf{h}_{r,j}}_{\text{Pilot contamination}} + \mathbf{w}_r \quad (5.1)$$

which is the pilot contaminated version of (4.5) and where $\mathbf{h}_{r,j}$ is the channel vector from user (or cell) j to the antenna r of i -th BS and the scaling factor β_j represents the path-loss and shadow fading coefficient for cell j where $0 \leq \beta_j \leq 1, \forall j \neq i$. Thus $\beta_j = 0$ corresponds to no interference and $\beta_j = 1$ represents the strongest interference from user j . From (5.1) we observe that pilot observations received at the BS are adversely affected due to pilot contamination. In the conventional approaches, such as [47], the effect of pilot contamination and noise is analysed

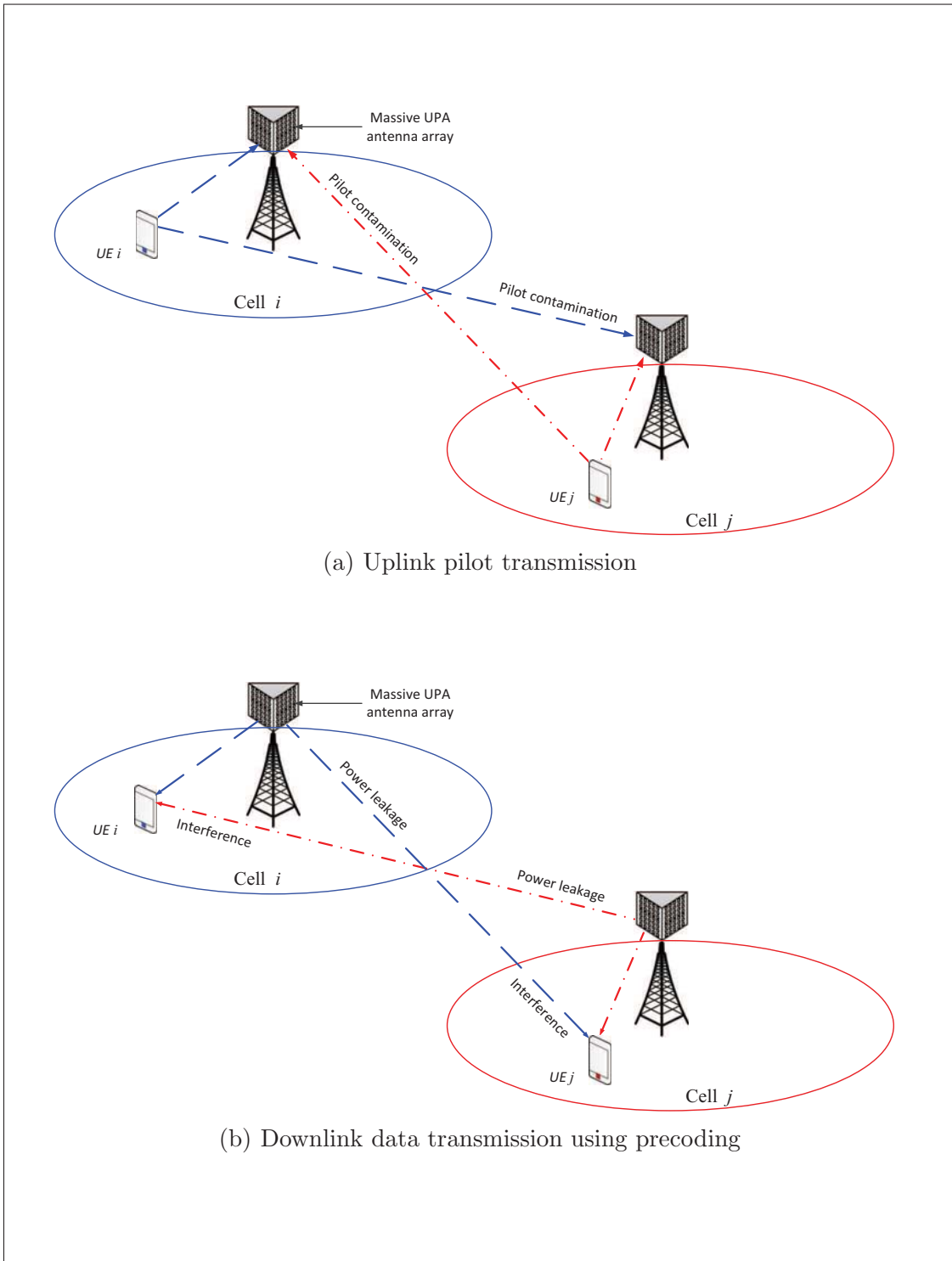


Figure 5.1: Effect of pilot contamination due to reuse of pilots, on the uplink and downlink data transmissions.

by computing the correlation of interference term, which is zero-mean complex Gaussian random vector. In contrast, we will use a different approach based on stochastic geometry for the analysis of pilot contamination in the next two Sections.

Besides the uplink, pilot contamination has also severe implications in the downlink transmission when the BS performs precoding, such as maximum ratio transmission (MRT) or conjugate beamforming, to focus the signal energy towards intended users. Due to contaminated channel estimates, some of the transmission power will leak to other cell's users, causing interference with those users, as well as power attenuation towards the desired users. Thus the effect of pilot contamination on downlink is two-fold; power loss and interference, as depicted in Fig. 5.1(b).

5.2 Mathematical Preliminaries

In this Section we provide brief mathematical preliminaries and important results which are used in the analysis presented in this Chapter.

5.2.1 Point Processes

A point refers to a simple geometric object in some Euclidean space \mathbb{R}^d . A point process is the random collection of the points in Euclidean space that can be represented by set $\Psi = \{\psi_1, \psi_2, \psi_3, \dots\} \subset \mathbb{R}^d$ with random points ψ_i . The point process in 1D (i.e., $d = 1$) is useful to model events occurring at random

time instants, for example arrivals of calls or customers. The point processes in real 2D (or 3D) space, called spatial point processes, are useful to model random patterns or locations of location of users in wireless and/or cellular networks.

5.2.2 Poisson Point Process (PPP)

A simple but an important type of point process is the poisson point process (PPP). Due to its simplicity and tractability, the PPP has been widely used in stochastic geometry for modelling of interference in cellular networks [76, 77, 78, 79]. A spatial PPP is a point process with intensity $\lambda > 0$, that satisfies the following two conditions:

1. If $\mathcal{B} \subset \mathbb{R}^d$, then the number of points in \mathcal{B} , denoted by $N(\mathcal{B})$, has poisson distribution with mean $\mu(\mathcal{B})$.
2. If $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m$ are disjoint regions in \mathbb{R}^d , then $N(\mathcal{B}_1), N(\mathcal{B}_2), \dots, N(\mathcal{B}_m)$ are independent random variables.

Thus for the poisson process we can write,

$$\mathbb{P}\{N(\mathcal{B}) = k\} = \frac{\mu(\mathcal{B})^k}{k!} e^{-\mu(\mathcal{B})} \quad (5.2)$$

where,

$$\mu(\mathcal{B}) = \int_{\mathcal{B}} \lambda(\mathbf{x}) d\mathbf{x} \quad (5.3)$$

$$= \lambda \mathcal{A}(\mathcal{B}) . \quad (5.4)$$

Eq. (5.4) is a special case of (5.3) for a homogeneous PPP where λ is uniform. Here $\mathcal{A}(\mathcal{B})$ denotes the Lebesgue measure of the subset \mathcal{B} , for example in 2D space, it is the area of region \mathcal{B} . A homogeneous PPP is stationary and isotropic i.e., it is invariant to translation and rotation.

Next, we state one of the most important Theorem with regard to PPP, that will be used later in the Chapter, to derive statistics of the interference. This Theorem is known as Campbell's Theorem [80].

Theorem 5.1 (Campbell's Theorem) *Let Ψ be a Poisson Point Process on \mathbb{R}^d with intensity λ and let $f : \mathbb{R}^d \rightarrow \mathbb{R}^+$ be a non-negative measurable function.*

Then the random sum

$$S = \sum_{\mathbf{x} \in \Psi} f(\mathbf{x}), \quad (5.5)$$

is a random variable with

$$\mathbb{E} \left\{ \sum_{\mathbf{x} \in \Psi} f(\mathbf{x}) \right\} = \int_{\mathbb{R}^d} f(\mathbf{x}) \lambda(\mathbf{x}) d\mathbf{x} \quad (5.6)$$

Proof. See [80]. █

5.3 Modified Network Model

To characterise inter-cell interference resulting from pilot contamination, we modify our previous 2D network model of Fig. 4.1(a) (and also Fig. 5.1(a)) by introducing interferers that are assumed to be distributed according to PPP. Specifically, without loss of generality, we assume a single user in a reference cell of radius γ_o ,

communicating with the BS located at the origin O in a 2-D plane. The interfering users (outside radius γ_o) are distributed over a circular region of radius γ_m according to a homogenous PP, denoted by Ψ and having intensity λ . The interfering space is thus an annular region with radii γ_o and γ_m , where the distance of i -th interferer from BS satisfies $\gamma_o < \gamma_i < \gamma_m$. Fig. 5.2 shows a realization of interferes distributed according to homogeneous PP of $\lambda=0.1$ with $\gamma_o=2\text{m}$ and $\gamma_m=5\text{m}$.

5.4 Interference Characterization

The knowledge of the interference statistics is critical in studying the effect of pilot contamination on channel estimation. It is assumed that during uplink pilot transmission for channel estimation in massive MIMO-OFDM, all users (i.e., the desired user in the reference cell and the interfering users) communicate with the BS by using identical pilots, which is the worst case scenario from pilot contamination point of view. Since each OFDM subcarrier acts as an independent narrow-band channel, it is enough to characterize the interference at a single OFDM tone. This makes the analysis quite simple and tractable. Thus, consider the complex received interference at any given sub-carrier (at the BS antenna r) due to all interfering users, which can be represented as [81]

$$\mathcal{I} = \sum_{i \in \Psi} \sqrt{E_x} x_i h_i \quad (5.7)$$

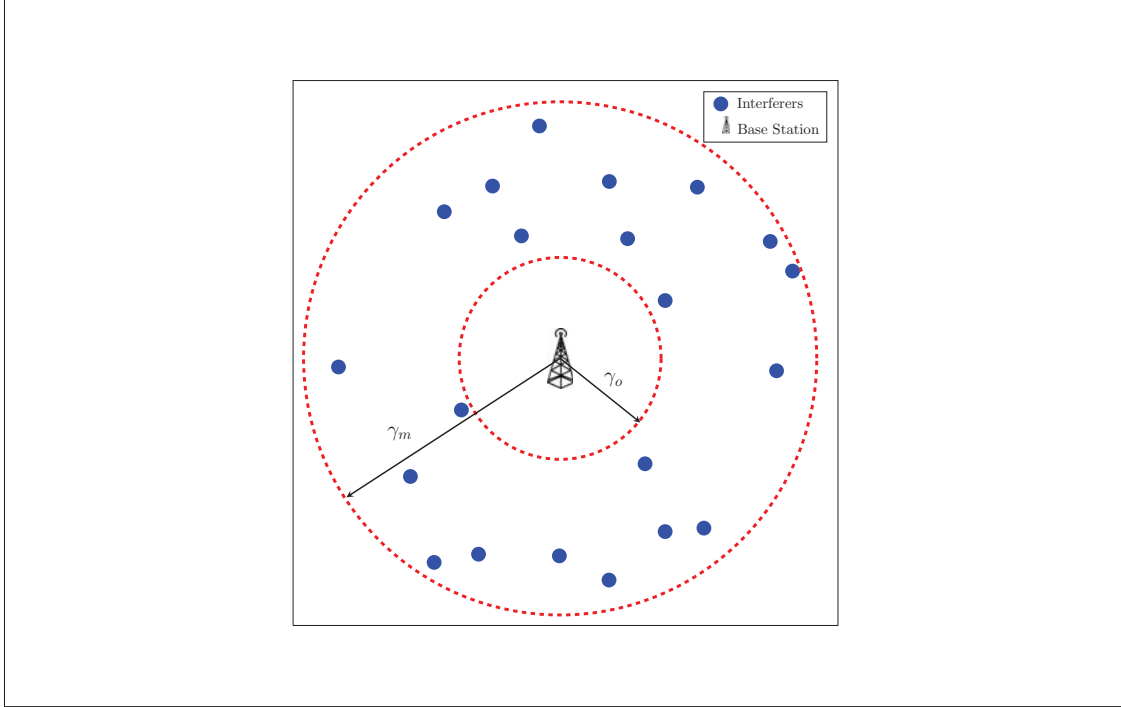


Figure 5.2: Realization of interferes distributed according to PPP of $\lambda=0.3$, $\gamma_o=2m$ and $\gamma_m=5m$ with BS at the origin.

where, $x_i=a_i\exp\{j\theta_i\}$ is the interfering symbol, $h_i=\gamma_i^{-b}\alpha_i\exp\{j\phi_i\}$ is the interfering channel, where $b > 1$ is the pathloss exponent, α_i is an independent Rayleigh distributed random variable with $\zeta=\mathbb{E}\{\alpha_i^2\}=1$ and ϕ_i is independent random variable that is uniformly distributed over $[0, 2\pi)$. The symbols x_i are generated from a general bi-dimensional constellation with $|\Omega|$ equiprobable symbols $\Omega_m=a^{(m)}\exp\{j\theta^{(m)}\}$, $m=1, 2, \dots, |\Omega|$. We assume that all interfering users transmit with the same average energy per symbol E_x and that the transmission constellation is normalized so that $\mathbb{E}\{|x_i|^2\}=1$. Therefore, (5.7) can be written as

$$\mathcal{I} = \sum_{i \in \Psi \setminus \mathcal{O}} \frac{\sqrt{E_x} a_i \alpha_i \exp\{j(\theta_i + \phi_i)\}}{\gamma_i^b} = \sum_{i \in \Psi \setminus \mathcal{O}} \frac{\sqrt{E_x} z_i}{\gamma_i^b} \quad (5.8)$$

where, $z_i = a_i \alpha_i \exp\{j(\theta_i + \phi_i)\}$ are independent spherically symmetric (SS) random variables.

For LS/LMMSE based channel estimation, we need mean and variance of interference \mathcal{I} , which are given in the following Lemma.

Lemma 5.1 *Using the network model in Section 5.3, the mean and variance of interference \mathcal{I} are given by,*

$$\mu_{\mathcal{I}} = \mathbb{E}\{\mathcal{I}\} = 0 \quad (5.9)$$

and

$$\sigma_{\mathcal{I}}^2 = \mathbb{E}\{|\mathcal{I}|^2\} = \pi \lambda (b-1)^{-1} \mathbb{E}\{|x|^2\} E_x \zeta \left(\frac{1}{\gamma_o^{2b-2}} - \frac{1}{\gamma_m^{2b-2}} \right) \quad (5.10)$$

respectively.

Proof. See Appendix A.1. ▮

Although (5.10) is derived by considering that the interference space is annular, it can be extended for an infinite interference space by taking the limit as $\gamma_m \rightarrow \infty$ yielding

$$\sigma_{\mathcal{I}}^2 = \pi \lambda \gamma_o^2 (b-1)^{-1} \mathbb{E}\{|x|^2\} \left(\frac{E_x \zeta}{\gamma_o^{b-1}} \right). \quad (5.11)$$

5.5 Effect of PC on MSE Performance

The knowledge of interference statistics at single OFDM frequency tone enables us to evaluate the aggregate interference correlation over all OFDM tones and/or

across the entire BS antenna array using known channel statistics. It is assumed that all the users have similar channel characteristics as described in Section 4.1. Thus, consider the received OFDM symbol at r th BS antenna, after omitting the index \mathcal{P}

$$\begin{aligned}\mathbf{y}_r &= \mathbf{A}\mathbf{h}_r + \mathcal{I}_r + \mathcal{W}_r \\ &= \mathbf{A}\mathbf{h}_r + \mathcal{E}_r\end{aligned}\tag{5.12}$$

where, \mathcal{I}_r is the interference due to pilot contamination and \mathcal{E}_r is the combined interference term due to both pilot contamination and noise. The correlation matrix of interference \mathcal{E}_r is given by

$$\mathbf{R}_{\mathcal{E}_r} = \mathbf{R}_{\mathcal{I}_r} + \mathbf{R}_w\tag{5.13}$$

$$= \sigma_{\mathcal{I}}^2 \mathbf{A}\mathbf{R}_{tap}\mathbf{A}^H + \sigma_w^2 \mathbf{I}_K\tag{5.14}$$

where (5.13) follows due to independence of interferences due to PC and AWGN, and (5.14) is based on the assumption that all user channels have identical correlations (as in Section 4.1) and use the same pilots, and $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_K$ is noise covariance matrix.

Similarly, in the multi-antenna case, based on system model of (4.6), the in-

terference correlation matrix for the entire BS array can be obtained by

$$\begin{aligned}\mathbf{R}_\varepsilon &= \mathbf{R}_\mathcal{I} + \mathbf{R}_w \\ &= \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_h \mathbf{A}^H + \sigma_w^2 \mathbf{I}_{RK}\end{aligned}\tag{5.15}$$

where $\mathbf{A} = \mathbf{I}_R \otimes \mathbf{A}$ as defined in Chapter 4. Using these interference correlations, we can derive the MSE expressions for LS, L-LMMSE and O-LMMSE algorithms in the presence of noise and pilot contamination. The results are presented in following Theorems.

Theorem 5.2 *For the system model described in Section 4.1 and pilot contamination as characterised in Section 5.5, the MSE expression for LS estimation algorithm of Section 4.2.3 under both AWGN and pilot contamination is given by*

$$\text{MSE}^{(\text{PC-LS})} = \frac{RL}{\rho K} + R\sigma_{\mathcal{I}}^2 \text{trace}(\mathbf{\Lambda}),\tag{5.16}$$

where $\sigma_{\mathcal{I}}^2$ is given in (5.10) and $\mathbf{\Lambda}$ is a diagonal matrix with eigenvalues of \mathbf{R}_{tap} spread along the diagonal and all users are assumed to have similar channel characteristics.

Proof. See Appendix A.2. |

Theorem 5.2, shows that MSE is composed of two terms. The first term due to AWGN can be suppressed by increasing the number of pilot tones but the second term due to pilot contamination cannot be reduced by adding more pilots and even persists at high SNR (i.e., $\rho \rightarrow \infty$).

Theorem 5.3 *For the system model described in Section 4.1 and pilot contamination as characterised in Section 5.5, the MSE expression for L-LMMSE estimation algorithm presented in Section 4.2.1 under both AWGN and pilot contamination is given by*

$$\text{MSE}^{(\text{PC-L})} = R \sum_{i=1}^L \frac{\delta_i (1 + \rho K \delta_i \sigma_{\mathcal{I}}^2)}{1 + \rho K \delta_i + \rho K \delta_i \sigma_{\mathcal{I}}^2}, \quad (5.17)$$

where $\sigma_{\mathcal{I}}^2$ is given in (5.10), δ_i are the eigenvalues of \mathbf{R}_{tap} and all users are assumed to have similar channel characteristics.

Proof. Replace \mathbf{R}_w with $\mathbf{R}_w + \mathbf{R}_{\mathcal{I}_r}$ in MSE expression (4.8), then invoke the eigenvalue decomposition (EVD) of \mathbf{R}_{tap} . We skip the detailed proof due to its similarity to Theorem 5.2 given in Appendix A.2. ▮

Note that (5.17) reduces to MSE expression (4.8) for AWGN had there been no pilot contamination. At high SNR (i.e. $\rho \gg 1$), the MSE expression (5.17) reduces to

$$\text{MSE}^{(\text{PC-L})} \xrightarrow{\text{high SNR}} R \left(\frac{\sigma_{\mathcal{I}}^2}{1 + \sigma_{\mathcal{I}}^2} \right) \text{trace}(\mathbf{\Lambda}), \quad (5.18)$$

which shows that the MSE is independent of the number of pilots and that LMMSE estimation is more robust to pilot contamination compared to LS.

Theorem 5.4 *For the system model described in Section 4.1 and pilot contamination as characterised in Section 5.5, the MSE expression for O-LMMSE estimation algorithm presented in Section 4.2.2 under both AWGN and pilot contamination is given by*

$$\text{MSE}^{(\text{PC-O})} = \sum_{j=1}^R \sum_{i=1}^L \frac{\eta_j \delta_i (1 + \rho K \eta_j \delta_i \sigma_{\mathcal{I}}^2)}{1 + \rho K \eta_j \delta_i + \rho K \eta_j \delta_i \sigma_{\mathcal{I}}^2}, \quad (5.19)$$

where $\sigma_{\mathcal{I}}^2$ is given in (5.10), η_j and δ_i are the eigenvalues of \mathbf{R}_a and \mathbf{R}_{tap} respectively, and all users are assumed to have similar channel characteristics.

Proof. See Appendix A.3. ▮

Note that (5.19) reduces to the MSE expression for AWGN given in (4.11) in absence of pilot contamination. Again observe that, under the assumption of high SNR, this simplifies to,

$$\text{MSE}^{(\text{PC-O})} \xrightarrow{\text{high SNR}} \left(\frac{\sigma_{\mathcal{I}}^2}{1 + \sigma_{\mathcal{I}}^2} \right) \text{trace}(\mathbf{R}_a) \text{trace}(\mathbf{\Lambda}). \quad (5.20)$$

This indicates that MSE depends strongly on interference power and is independent of number of pilots K . Since $\text{trace}(\mathbf{R}_a) \leq R$, the O-LMMSE seems to be more robust to pilot contamination compared to both LS and L-LMMSE. The MSE expression also gives a clue that effect of pilot contamination can be minimized by exploiting the spatial correlations and by optimizing the BS antenna array design.

Above Theorems quantify the effect of pilot contamination on MSE performance of channel estimation in terms of interference power (or variance) which in turn depends on various network parameters described in Lemma 5.1. These results are summarized in Table 5.1. The MSE performance against various parameters will be numerically analysed through simulations.

Table 5.1: Summary of MSE expressions

Algorithm	AWGN only	AWGN + Pilot Contamination
LS	$\text{MSE} = \frac{RL}{\rho K}$	$\text{MSE} = \frac{RL}{\rho K} + R\sigma_{\mathcal{I}}^2 \text{trace}(\mathbf{\Lambda})$
L-LMMSE	$\text{MSE} = R \sum_{i=1}^L \left(\frac{\delta_i}{1+\rho K \delta_i} \right)$	$\text{MSE} = R \sum_{i=1}^L \frac{\delta_i (1+\rho K \delta_i \sigma_{\mathcal{I}}^2)}{1+\rho K \delta_i + \rho K \delta_i \sigma_{\mathcal{I}}^2}$
O-LMMSE	$\text{MSE} = \sum_{j=1}^R \sum_{i=1}^L \frac{\eta_j \delta_i}{1+\rho K \eta_j \delta_i}$	$\text{MSE} = \sum_{j=1}^R \sum_{i=1}^L \frac{\eta_j \delta_i (1+\rho K \eta_j \delta_i \sigma_{\mathcal{I}}^2)}{1+\rho K \eta_j \delta_i + \rho K \eta_j \delta_i \sigma_{\mathcal{I}}^2}$

5.6 Simulation Results

For simulations, we use the same parameters and the MSE criterion to evaluate the performance as described in Chapter 4 with a notable difference of modified network model based on stochastic geometry introduced earlier at the beginning of this Chapter. Different experiments are conducted to validate the analysis of pilot contamination presented in this Chapter.

Mean and variance of interference

We first validate the mean and variance of the interference given in Lemma 5.1 by numerical simulations. In order to mimic the setup described in Section 5.3, we use single antenna BS and assume that CIRs from each user to the BS has a uniform PDP. Further, we assume that BS is located at the origin, the desired user at a distance of 1m from BS with $\gamma_o=2\text{m}$ while interfering users are distributed in a region of radius 5m according to a PPP with density λ and pathloss exponent b . All users communicate with BS using OFDM with $N=256$, $L=8$ and $K=32$ identical pilot symbols drawn from a 4-QAM constellation. Fig. 5.3 compares the

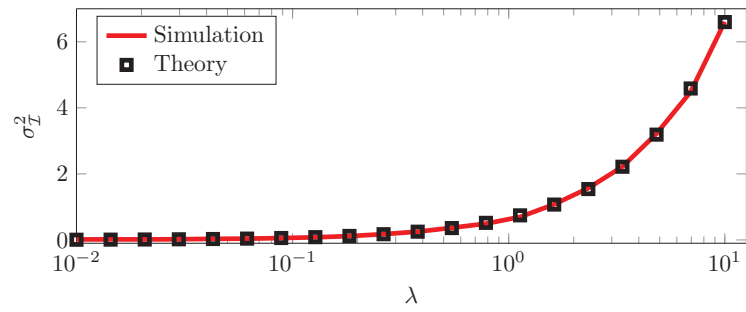
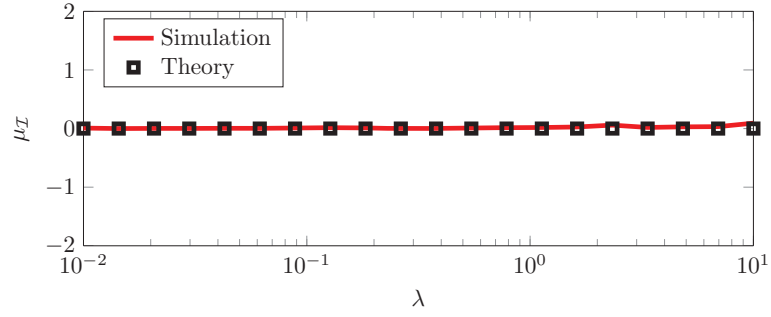
mean and variance of interference observed on single OFDM carrier (randomly picked) due to simulated sources with expressions given in Lemma 5.1, as a function of λ and b . In Fig. 5.3(a) pathloss exponent is kept fixed at $b = 2$ and λ is varied while in Fig. 5.3(b) the parameter $\lambda = 1$ and pathloss exponent b is varied. The results indicate a close match between simulation and theory for a wide range of both parameters.

MSE Performance Under AWGN and Pilot Contamination

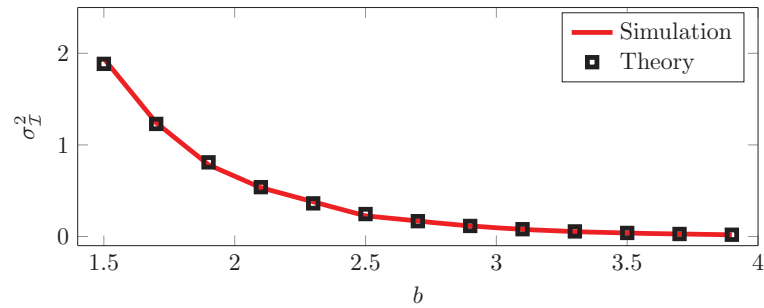
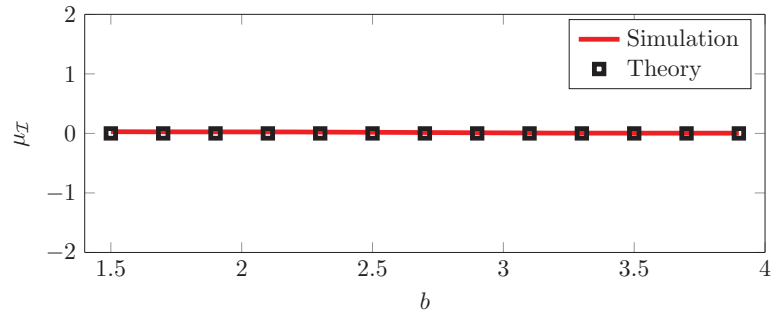
We now compare the MSE performance of different algorithms in the presence of both AWGN and pilot contamination. For simulations, we use the parameters given in Table 4.2 with the interfering users distributed according to a PPP of $\lambda=0.1$ in circular region of radius 5m, the desired user is assumed 1m away from BS located at origin with $\gamma_o=2$ m and pathloss $b=2$. In Fig. 5.4, the simulated MSE performance of different algorithms is compared over a wide range of SNR with the analytical expressions given in Theorems 5.2, 5.3 and 5.4 (see Section 5.5). Observe that all MSE curves decrease with increasing SNR in lower range but reach an error floor at higher SNR. This is in stark contrast to AWGN case (see Fig. 4.7), which indicates that pilot contamination persists even at high SNR and its effect on MSE is more severe than AWGN.

We present similar analysis in Fig. 5.5(a), where the MSE is plotted as a function of λ with SNR fixed at 10 dB. It is obvious that all algorithms perform well for small values of λ . However when λ increases, the interference due to pilot contamination dominates AWGN, thus severely degrading the performance

as indicated by a sharp increase in MSE curves. Note that LMMSE channel estimation is more robust to pilot contamination than simple LS based channel estimation. The effect of pathloss is portrayed in Fig. 5.5(b), where a close match can be observed between simulation and theoretical analysis over a wide range of λ and b .



(a)



(b)

Figure 5.3: Mean and variance of interference as a function of: (a) Intensity λ , with $b = 2$ and (b) Pathloss exponent b , with $\lambda = 1$.

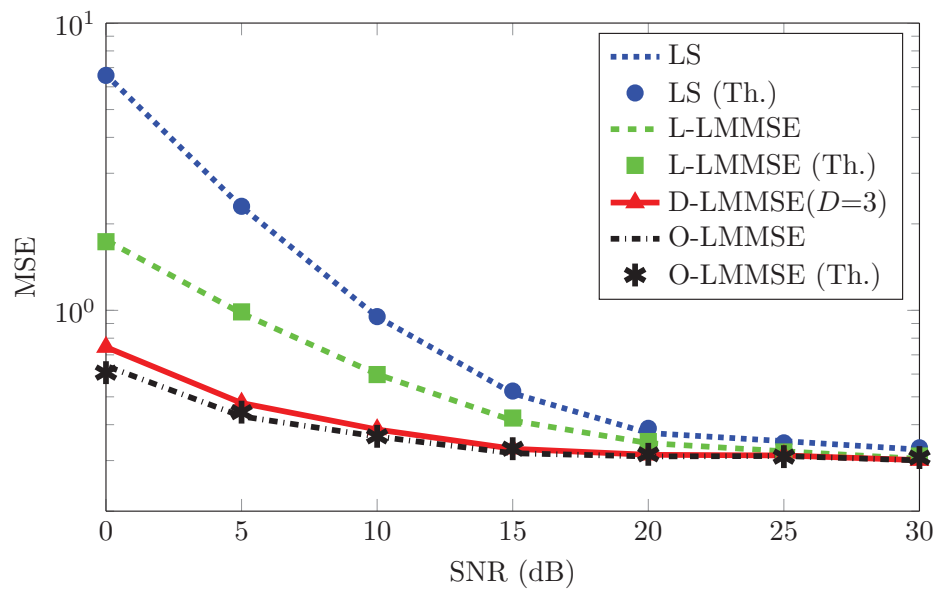


Figure 5.4: MSE performance in presence of AWGN and PC with parameters $\lambda=0.1$ and $b=2$.

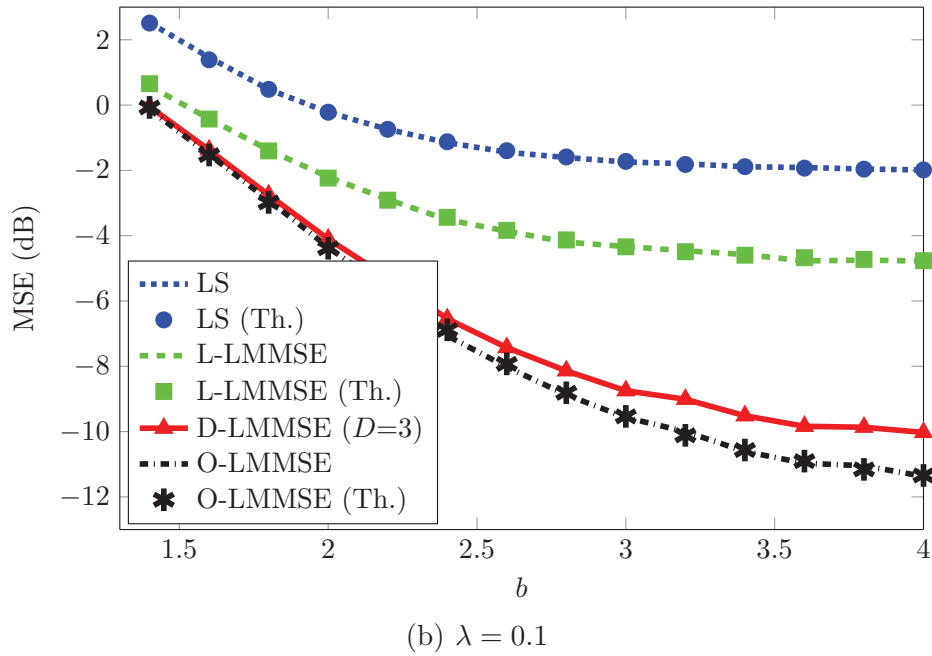
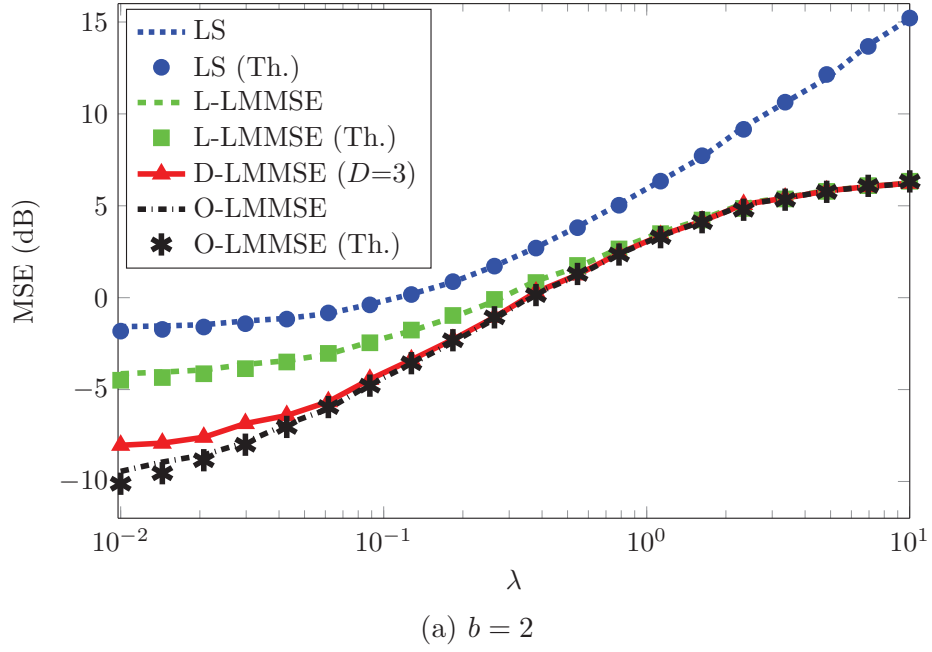


Figure 5.5: Effect of interfering user intensity λ and pathloss exponent b on MSE performance at SNR = 10 dB.

CHAPTER 6

CONCLUSIONS AND FUTURE RECOMMENDATIONS

6.1 Achievements of the Work

In this thesis we have investigated several low complexity channel estimation techniques for OFDM based wireless systems by utilizing the inherent structure and constraints of the communication problem at hand. Specifically, in Chapter 2, we exploited the structure of FFT matrices induced by OFDM and the frequency correlation of the channel in developing a low complexity MMSE based channel estimation algorithm for SISO-OFDM systems.

In Chapter 3, we presented a blind and semi-blind channel estimation techniques for MIMO-OFDM systems employing Alamouti coding. The proposed blind algorithm performed joint channel estimation and data-detection for both constant modulus and nonconstant modulus constellations. The computational

complexity of proposed algorithms was substantially reduced by utilizing the structure of FFT matrices, frequency and time correlations of the channel and using the finite alphabet property of transmitted symbols.

The thesis also investigated channel estimation in massive MIMO-OFDM systems where exceptionally large number of unknown channel coefficients needs to be estimated. A distributed LMMSE algorithm was presented in Section 4.3 which was further extended with a data-aided approach. By relying on antenna coordination, the distributed algorithm turned out to be an efficient strategy for estimating large number of channel coefficients, and that too at significantly reduced complexity. The structure of channel frequency and spatial antenna correlations were both exploited to reduce the communication overhead. Further, the finite alphabet constraint was utilized to reduce the large pilot overhead in massive MIMO systems.

Finally, the impact of pilot contamination on channel estimation performance was studied in Chapter 5. In order to quantify the interference resulting from neighboring cell users, a stochastic geometry based analysis was carried out which culminated in closed form expressions for MSE under both additive white noise and pilot contamination.

6.2 Summary of Contributions

The main contributions of this thesis are summarised as follows:

- Development of optimal low-complexity MMSE channel estimation scheme

for SISO-OFDM systems based on pilots.

- Development blind and semi-blind methods for MIMO-OFDM systems employing Alamouti coding.
- Development of distributed LMMSE algorithm for massive MIMO-OFDM systems.
- Analysis of the effect of pilot contamination on MSE performance of channel estimators by using stochastic geometry.

6.3 Future Recommendations

The research work on channel estimation carried out in this thesis can be extended in many dimensions. Some recommendations for future work are listed below.

- Pilot based algorithm of Chapter 2 can be extended by a data-aided approach using the concept of reliable carriers selection method presented in Chapter 4. Further, an exact performance analysis of data-aided algorithm utilizing ZF detector, can be carried out based on statistics of the channel and known results on ratio distributions. Moreover, the proposed channel estimation algorithm, although presented for SISO-OFDM systems, can also be adopted for MIMO-OFDM systems.
- The distributed LMMSE algorithm presented in Chapter 4 can be investigated for solving multi-task problems over a wireless sensor network, where

each sensor in the network tries to estimate a different but correlated set of unknown parameters.

- Transformation of distributed LMMSE algorithm into stochastic algorithms based on LMS/RLS algorithms could be an interesting future work as it would not only alleviate the requirement of exact channel statistics but also cope with time variations of the channel parameters. However, the LMS based stochastic gradient algorithm might have convergence issues with finite data records.
- Based on stochastic geometry, the analysis of pilot contamination can be extended to investigate both uplink and downlink throughput of massive MIMO systems by focusing on received signal-to-interference-noise-ratio (SINR). Recall that current work only deals with the MSE performance analysis but not the throughput analysis.

APPENDIX A

A.1 Proof of Lemma 5.1

The mean of \mathcal{I} can be determined as follows,

$$\begin{aligned}\mu_{\mathcal{I}} &= \mathbb{E}\{\mathcal{I}\} \\ &= \mathbb{E}\left\{\sum_{i \in \Psi} \frac{\sqrt{E_x} z_i}{\gamma_i^b}\right\} \\ &= \mathbb{E}_{\Psi}\left\{\sum_{i \in \Psi} \frac{\sqrt{E_x} \mathbb{E}_z\{z_i\}}{\gamma_i^b}\right\} \\ &= \sqrt{E_x} \mathbb{E}\{z_i\} \int_{\mathbb{R}^2} \frac{1}{r^b} r dr d\theta = 0\end{aligned}\tag{A.1}$$

where, (A.1) results from Campbell's Theorem given in Section 5.2.2 and then the fact, $\mathbb{E}\{z_i\} = 0$ yields the zero mean.

Similarly, the variance of interference can be computed as follows,

$$\begin{aligned}
\sigma_{\mathcal{I}}^2 &= \mathbb{E}\{|\mathcal{I}|^2\} \\
&= \mathbb{E}_{\Psi} \left\{ \mathbb{E}_z \sum_{i \in \Psi} \frac{\sqrt{E_x} z_i}{\gamma_i^b} \sum_{j \in \Psi} \frac{\sqrt{E_x} z_j^*}{\gamma_j^b} \right\} \\
&= \mathbb{E}_{\Psi} \left\{ \sum_{i \in \Psi} \frac{E_x \mathbb{E}_z\{|z_i|^2\}}{\gamma_i^{2b}} \right\} \tag{A.2}
\end{aligned}$$

$$= \lambda E_x \mathbb{E}\{|z_i|^2\} \int_0^{2\pi} \int_{\gamma_o}^{\gamma_m} \frac{1}{r^{2b}} r dr d\theta \tag{A.3}$$

$$= \pi \lambda (b-1)^{-1} E_x \zeta \mathbb{E}\{|x|^2\} \left(\frac{1}{\gamma_o^{2b-2}} - \frac{1}{\gamma_m^{2b-2}} \right) \tag{A.4}$$

where, (A.2) is due to the fact that z_i are independent SS random variables, in (A.3) we employed Campbell's Theorem and in (A.4) we used the result $\mathbb{E}\{|z_i|^2\} = \mathbb{E}\{a_i^2 \alpha_i^2\} = \zeta \mathbb{E}\{|x|^2\}$, where we note that a_i and α_i are independent random variables, which completes the proof.

A.2 Proof of Theorem 5.2

By replacing \mathbf{R}_w with $\mathbf{R}_w + \mathbf{R}_{\mathcal{I}_r}$ in MSE expression of (4.14), we obtain

$$\begin{aligned}
\text{mse}_r^{\text{ls}} &= \text{trace} \left(\mathbf{A}^H (\mathbf{R}_w + \mathbf{R}_{\mathcal{I}_r})^{-1} \mathbf{A} \right)^{-1} \\
&= \text{trace} \left(\mathbf{A}^H (\mathbf{R}_w + \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_{\text{tap}} \mathbf{A}^H)^{-1} \mathbf{A} \right)^{-1} \\
&\stackrel{(a)}{=} \text{trace} \left(\mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A} - \sigma_{\mathcal{I}}^2 \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A} (\mathbf{R}_{\text{tap}}^{-1} + \sigma_{\mathcal{I}}^2 \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}_w^{-1} \mathbf{A} \right)^{-1} \tag{A.5}
\end{aligned}$$

where, (A.5) follows from matrix inversion Lemma. Now, using the EVD of the channel correlation matrix $\mathbf{R}_{tap} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$ and the fact that $\mathbf{A}^H\mathbf{R}_w^{-1}\mathbf{A} = \frac{KE_x}{\sigma_w^2}\mathbf{I}_L$ we obtain,

$$\text{mse}_r^{\text{ls}} = \text{trace} \left(\frac{KE_x}{\sigma_w^2} \mathbf{I}_L - \sigma_{\mathcal{I}}^2 \left(\frac{KE_x}{\sigma_w^2} \right)^2 \left(\mathbf{\Lambda}^{-1} + \frac{\sigma_{\mathcal{I}}^2 KE_x}{\sigma_w^2} \mathbf{I}_L \right)^{-1} \right)^{-1} \quad (\text{A.6})$$

$$= \sum_{i=1}^L \left(\frac{KE_x}{\sigma_w^2} - \sigma_{\mathcal{I}}^2 \left(\frac{KE_x}{\sigma_w^2} \right)^2 \left(\delta_i^{-1} + \frac{\sigma_{\mathcal{I}}^2 KE_x}{\sigma_w^2} \right)^{-1} \right)^{-1} \quad (\text{A.7})$$

where, (A.6) follows from the property that $\text{trace}(\mathbf{Q}\mathbf{R}\mathbf{Q}^H) = \text{trace}(\mathbf{R})$ if \mathbf{Q} is unitary. After simple algebraic manipulations, the term inside the summation of (A.7) simplifies to $\frac{\sigma_w^2 L}{KE_x} + \sigma_{\mathcal{I}}^2 \sum_{i=1}^L \delta_i$, which completes the proof.

A.3 Proof of Theorem 5.4

Under both AWGN and pilot contamination, we replace \mathbf{R}_w with $\mathbf{R}_{\mathcal{E}} = \mathbf{R}_w + \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_h \mathbf{A}^H$ in the MSE expression (4.11) to get,

$$\begin{aligned} \text{MSE}^{(0)} &= \text{trace} \left(\mathbf{R}_h^{-1} + \mathbf{A}^H \left(\mathbf{R}_w + \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_h \mathbf{A}^H \right)^{-1} \mathbf{A} \right)^{-1} \\ &= \text{trace} \left(\mathbf{R}_h^{-1} + \mathbf{A} \mathbf{R}_h \mathbf{A}^H - \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_h \mathbf{A}^H \left(\mathbf{R}_h^{-1} + \sigma_{\mathcal{I}}^2 \mathbf{A} \mathbf{R}_h \mathbf{A}^H \right)^{-1} \mathbf{A} \mathbf{R}_h \mathbf{A}^H \right)^{-1} \end{aligned} \quad (\text{A.8})$$

where (A.8) follows from matrix inversion Lemma. Using the properties of kronecker product, it can be shown that $\mathbf{A} \mathbf{R}_h \mathbf{A}^H = \frac{KE_x}{\sigma_w^2} (\mathbf{I}_R \otimes \mathbf{I}_L)$. Further, the channel correlation matrix $\mathbf{R}_h = \mathbf{R}_a \otimes \mathbf{R}_{tap}$ can be decomposed as

$\mathbf{R}_h = (\mathbf{V} \otimes \mathbf{Q})(\mathbf{S} \otimes \mathbf{\Lambda})(\mathbf{V} \otimes \mathbf{Q})^H$, where we introduced the EVDs, $\mathbf{R}_a = \mathbf{V}\mathbf{S}\mathbf{V}^H$ and $\mathbf{R}_{tap} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$. Incorporating these results in (A.8) yields,

$$\text{MSE}^{(0)} = \text{trace} \left(\mathbf{S}^{-1} \otimes \mathbf{\Lambda}^{-1} + \frac{KE_x}{\sigma_w^2} (\mathbf{I}_R \otimes \mathbf{I}_L) - \sigma_{\mathcal{I}}^2 \left(\frac{KE_x}{\sigma_w^2} \right)^2 \left(\mathbf{S}^{-1} \otimes \mathbf{\Lambda}^{-1} + \frac{\sigma_{\mathcal{I}}^2 KE_x}{\sigma_w^2} (\mathbf{I}_R \otimes \mathbf{I}_L) \right)^{-1} \right) \quad (\text{A.9})$$

$$= \sum_{j=1}^R \sum_{i=1}^L \left(\frac{1}{\eta_j \delta_i} + \frac{KE_x}{\sigma_w^2} - \sigma_{\mathcal{I}}^2 \left(\frac{KE_x}{\sigma_w^2} \right)^2 \left(\frac{1}{\eta_j \delta_i} + \frac{\sigma_{\mathcal{I}}^2 KE_x}{\sigma_w^2} \right)^{-1} \right)^{-1} \quad (\text{A.10})$$

where, (A.9) follows from property, $\text{trace}(\mathbf{Q}\mathbf{R}\mathbf{Q}^H) = \text{trace}(\mathbf{R})$ when \mathbf{Q} is unitary and (A.10) is due to the diagonal nature of the matrix inside the trace operator, η_j and δ_i represent the eigenvalues of matrices \mathbf{R}_a and \mathbf{R}_{tap} respectively. After some algebraic manipulations, (A.10) simplifies to the result given in Theorem 5.4.

Bibliography

- [1] Emre Telatar. Capacity of Multi-Antenna Gaussian Channels. *European transactions on telecommunications*, 10(6):585–595, 1999.
- [2] G. J. Foschini and M. J. Gans. On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas. *Wireless Personal Communications*, 6:311–335, 1998.
- [3] A.J. Paulraj, D.A. GORE, R.U. Nabar, and H. Bolcskei. An Overview of MIMO Communications - A Key to Gigabit Wireless. *Proceedings of the IEEE*, 92(2):198–218, Feb 2004.
- [4] S. Weinstein and P. Ebert. Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform. *Communication Technology, IEEE Transactions on*, 19(5):628–634, October 1971.
- [5] B.P. Crow, I. Widjaja, Jeong Geun Kim, and P.T. Sakai. IEEE 802.11 Wireless Local Area Networks. *Communications Magazine, IEEE*, 35(9):116–126, Sep 1997.
- [6] *IEEE Standard 802.11n-2009: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE, 2009.

- [7] *IEEE Standard 802.16-2009: Air Interface for Broadband Wireless Access Systems*. IEEE, 2009.
- [8] *Overview of 3GPP Release 12 V0.0.3*. IEEE, 2012.
- [9] Jayakumari. J. MIMO-OFDM for 4G Wireless Systems. *International Journal of Engineering Science and Technology*, 2:2886–2889, Jul 2010.
- [10] Nokia Siemens Networks. 2020: Beyond 4G: Radio Evolution for the Gigabit Experience, 2011. white paper.
- [11] K. E. Skouby WWRF, L. Sorensen. 2020: Beyond 4G: Radio Evolution for the Gigabit Experience, July 2009.
- [12] T.L. Marzetta. Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas. *Wireless Communications, IEEE Transactions on*, 9(11):3590–3600, November 2010.
- [13] F. Rusek, D. Persson, Buon Kiong Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, and F. Tufvesson. Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays. *Signal Processing Magazine, IEEE*, 30(1):40–60, Jan 2013.
- [14] J. Hoydis, S. ten Brink, and M. Debbah. Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need? *Selected Areas in Communications, IEEE Journal on*, 31(2):160–171, February 2013.
- [15] Hoon Huh, G. Caire, H.C. Papadopoulos, and S.A. Ramprasad. Achieving Massive MIMO Spectral Efficiency with a Not-so-Large Number of Antennas. *Wireless Communications, IEEE Transactions on*, 11(9):3226–3239, September 2012.

- [16] F Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta, and P. Popovski. Five Disruptive Technology Directions for 5G. *Communications Magazine, IEEE*, 52(2):74–80, February 2014.
- [17] Theodore Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2nd edition, 2001.
- [18] Arogyaswami Paulraj, Rohit Nabar, and Dhananjay Gore. *Introduction to Space-Time Wireless Communications*. Cambridge University Press, New York, NY, USA, 1st edition, 2008.
- [19] J.-J. van de Beek, O. Edfors, M. Sandell, S.K. Wilson, and P. Ola Borjesson. On Channel Estimation in OFDM Systems. In *Vehicular Technology Conference, 1995 IEEE 45th*, volume 2, pages 815–819 vol.2, Jul 1995.
- [20] J. Rinne and M. Renfors. Pilot Spacing in Orthogonal Frequency Division Multiplexing Systems on Practical Channels. In *Consumer Electronics, 1996. Digest of Technical Papers., International Conference on*, pages 4–, June 1996.
- [21] C. R N Athaudage and A. D S Jayalath. Low-Complexity Channel Estimation for Wireless OFDM Systems. In *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, volume 1, pages 521–525 Vol.1, Sept 2003.
- [22] R. Weber. Low-Complexity Channel Estimation for WCDMA Random Access. In *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, volume 1, pages 344–351 vol.1, 2000.

- [23] Xiaowen Wang and K.J.R. Liu. OFDM Channel Estimation Based on Time-Frequency Polynomial Model of Fading Multipath Channel. In *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, volume 1, pages 460–464 vol.1, 2001.
- [24] Ningping Sun, T. Ayabe, and T. Nishizaki. Efficient Spline Interpolation Curve Modeling. In *Intelligent Information Hiding and Multimedia Signal Processing, 2007. IHHMSP 2007. Third International Conference on*, volume 2, pages 59–62, Nov 2007.
- [25] Ming Jiang, Siji Huang, and Wenkun Wen. Adaptive Polar-Linear Interpolation Aided Channel Estimation for Wireless Communication Systems. *Wireless Communications, IEEE Transactions on*, 11(3):920–926, March 2012.
- [26] P. Hoeher, S. Kaiser, and P. Robertson. Two-Dimensional Pilot-Symbol-Aided Channel Estimation by Wiener Filtering. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 3, pages 1845–1848 vol.3, Apr 1997.
- [27] Pei-Yun Tsai and Tzi-Dar Chiueh. Adaptive Raised-Cosine Channel Interpolation for Pilot-Aided OFDM Systems. *Wireless Communications, IEEE Transactions on*, 8(2):1028–1037, Feb 2009.
- [28] Seog Geun Kang, Yong Min Ha, and Eon Kyeong Joo. A Comparative Investigation on Channel Estimation Algorithms for OFDM in Mobile Communications. *Broadcasting, IEEE Transactions on*, 49(2):142–149, June 2003.

- [29] Jia Lei, Cao Dazhong, and Hou Chunping. Two Novel Transform Domain Estimation Methods for OFDM System and Their Application Environment. In *Electrical and Computer Engineering, 2004. Canadian Conference on*, volume 1, pages 377–380 Vol.1, May 2004.
- [30] O. Edfors, M. Sandell, J.-J. van de Beek, S.K. Wilson, and P.O. Borjesson. OFDM Channel Estimation by Singular Value Decomposition. In *Vehicular Technology Conference, 1996. Mobile Technology for the Human Race., IEEE 46th*, volume 2, pages 923–927 vol.2, Apr 1996.
- [31] M.K. Ozdemir, H. Arslan, and E. Arvas. Toward Real-Time Adaptive Low-Rank LMMSE Channel Estimation of MIMO-OFDM Systems. *Wireless Communications, IEEE Transactions on*, 5(10):2675–2678, Oct 2006.
- [32] C. Mehlhruer, S. Caban, and M. Rupp. An Accurate and Low Complex Channel Estimator for OFDM WiMAX. In *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on*, pages 922–926, March 2008.
- [33] Michal Simko, Di Wu, Christian Mehlhruer, Johan Eilert, and Dake Liu. Implementation Aspects of Channel Estimation for 3GPP LTE Terminals. In *Wireless Conference 2011 - Sustainable Wireless Technologies (European Wireless), 11th European*, pages 1–5, April 2011.
- [34] B. Muquet, M. De Courville, and P. Duhamel. Subspace-Based Blind and Semi-Blind Channel Estimation for OFDM Systems. *Signal Processing, IEEE Transactions on*, 50(7):1699–1712, 2002.

- [35] Chao-Cheng Tu and B. Champagne. Subspace-Based Blind Channel Estimation for MIMO-OFDM Systems with Reduced Time Averaging. *Vehicular Technology, IEEE Transactions on*, 59(3):1539–1544, 2010.
- [36] H. Bolcskei, R.W. Heath, and A.J. Paulraj. Blind Equalization in OFDM-Based Multi-Antenna Systems. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 58–63, 2000.
- [37] Jinho Choi and Cheng-Chew Lim. A Cholesky Factorization Based Approach for Blind FIR Channel Identification. *Signal Processing, IEEE Transactions on*, 56(4):1730–1735, 2008.
- [38] S.A. Banani and R.G. Vaughan. OFDM with Iterative Blind Channel Estimation. *Vehicular Technology, IEEE Transactions on*, 59(9):4298–4308, 2010.
- [39] Tao Cui and C. Tellambura. Joint Data Detection and Channel Estimation for OFDM Systems. *Communications, IEEE Transactions on*, 54(4):670–679, 2006.
- [40] E.G. Larsson, Petre Stoica, and Jian Li. On Maximum-Likelihood Detection and Decoding for Space-Time Coding Systems. *Signal Processing, IEEE Transactions on*, 50(4):937–944, 2002.
- [41] Tsung-Hui Chang, Wing-Kin Ma, and Chong-Yung Chi. Maximum-Likelihood Detection of Orthogonal Space-Time Block Coded OFDM in Unknown Block Fading Channels. *Signal Processing, IEEE Transactions on*, 56(4):1637–1649, 2008.
- [42] T.Y. Al-Naffouri and A.A. Quadeer. Cyclic Prefix Based Enhanced Data Recovery in OFDM. *Signal Processing, IEEE Transactions on*, 58(6):3406–3410, 2010.

- [43] Weiyu Xu, M. Stojnic, and B. Hassibi. Low-Complexity Blind Maximum-Likelihood Detection for SIMO Systems with General Constellations. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 2817–2820, 2008.
- [44] T.Y. Al-Naffouri, A.A. Dahman, M.S. Sohail, Weiyu Xu, and B. Hassibi. Low-Complexity Blind Equalization for OFDM Systems with General Constellations. *Signal Processing, IEEE Transactions on*, 60(12):6395–6407, 2012.
- [45] E. Bjornson and B. Ottersten. A Framework for Training-Based Estimation in Arbitrarily Correlated Rician MIMO Channels with Rician Disturbance. *Signal Processing, IEEE Transactions on*, 58(3):1807–1820, March 2010.
- [46] N. Shariati, E. Bjornson, M. Bengtsson, and M. Debbah. Low-Complexity Channel Estimation in Large-Scale MIMO using Polynomial Expansion. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pages 1157–1162, Sept 2013.
- [47] P. Xu, J. Wang, J. Wang, and F. Qi. Analysis and Design of Channel Estimation in Multicell Multiuser MIMO OFDM Systems. *IEEE Transactions on Vehicular Technology*, PP(99):1–11, 2014.
- [48] Haifan Yin, D. Gesbert, M. Filippou, and Yingzhuang Liu. A Coordinated Approach to Channel Estimation in Large-Scale Multiple-Antenna Systems. *Selected Areas in Communications, IEEE Journal on*, 31(2):264–273, February 2013.
- [49] Hien Quoc Ngo and E.G. Larsson. EVD-Based Channel Estimation in Multicell Multiuser MIMO Systems with Very Large Antenna Arrays. In *Acoustics, Speech*

- and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3249–3252, March 2012.
- [50] Sinh Le Hong Nguyen and Ali Ghayeb. Compressive Sensing-Based Channel Estimation for Massive Multiuser MIMO Systems. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 2890–2895, April 2013.
- [51] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli. Estimating Sparse MIMO Channels having Common Support. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 2920–2923, May 2011.
- [52] M. Masood, L. H. Afify, and T. Y. Al-Naffouri. Efficient Coordinated Recovery of Sparse Channels in Massive MIMO. *IEEE Transactions on Signal Processing*, 63(1):104–118, Jan 2015.
- [53] T.L. Marzetta. How Much Training is Required for Multiuser MIMO? In *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pages 359–363, Oct 2006.
- [54] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath. Pilot Contamination and Precoding in Multi-Cell TDD Systems. *IEEE Transactions on Wireless Communications*, 10(8):2640–2651, August 2011.
- [55] B. Gopalakrishnan and N. Jindal. An Analysis of Pilot Contamination on Multi-User MIMO Cellular Systems with Many Antennas. In *IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 381–385, June 2011.

- [56] M. Filippou, D. Gesbert, and H. Yin. Decontaminating Pilots in Cognitive Massive MIMO Networks. In *International Symposium on Wireless Communication Systems (ISWCS)*, pages 816–820, Aug 2012.
- [57] F. Fernandes, A. Ashikhmin, and T. L. Marzetta. Inter-Cell Interference in Non-cooperative TDD Large Scale Antenna Systems. *IEEE Journal on Selected Areas in Communications*, 31(2):192–201, February 2013.
- [58] R. R. Mueller, Mikko Vehkaperae, and L. Cottatellucci. Blind Pilot Decontamination. In *17th International ITG Workshop on Smart Antennas (WSA)*, pages 1–6, March 2013.
- [59] Anzhong Hu, Tiejun Lv, and Yueming Lu. Subspace-Based Semi-Blind Channel Estimation for Large-Scale Multi-Cell Multiuser MIMO Systems. In *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pages 1–5, June 2013.
- [60] R. Negi and J. Cioffi. Pilot Tone Selection for Channel Estimation in a Mobile OFDM System. *Consumer Electronics, IEEE Transactions on*, 44(3):1122–1128, Aug 1998.
- [61] Xiaodong Cai and G.B. Giannakis. Error Probability Minimizing Pilots for OFDM with M-PSK Modulation over Rayleigh-Fading Channels. *Vehicular Technology, IEEE Transactions on*, 53(1):146–155, Jan 2004.
- [62] A.H. Sayed. *Fundamentals of Adaptive Filtering*. Wiley, 2003.
- [63] Magnus Sandell and Ove Edfors. A Comparative Study of Pilot-Based Channel Estimators for Wireless OFDM. 1996.

- [64] P. Hammarberg and O. Edfors. A Comparison of DFT and SVD Based Channel Estimation in MIMO OFDM Systems. In *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, pages 1–5, Sept 2006.
- [65] M.K. Ozdemir and H. Arslan. Channel Estimation for Wireless OFDM Systems. *Communications Surveys Tutorials, IEEE*, 9(2):18–48, Second 2007.
- [66] Petre Stoica and G. Ganesan. Space-Time Block Codes: Trained, Blind and Semi-Blind Detection. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages II–1609–II–1612, 2002.
- [67] A.L. Swindlehurst and G. Leus. Blind and Semi-Blind Equalization for Generalized Space-Time Block Codes. *Signal Processing, IEEE Transactions on*, 50(10):2489–2498, 2002.
- [68] S. Alamouti. A Simple Transmit Diversity Technique for Wireless Communications. *Selected Areas in Communications, IEEE Journal on*, 16(8):1451–1458, 1998.
- [69] N. Al-Dhahir. Single-Carrier Frequency-Domain Equalization for Space-Time Block-Coded Transmissions over Frequency-Selective Fading Channels. *Communications Letters, IEEE*, 5(7):304–306, 2001.
- [70] E.B. Al-Safadi and T.Y. Al-Naffouri. Pilotless Recovery of Clipped OFDM Signals by Compressive Sensing over Reliable Data Carriers. In *Signal Processing Advances in Wireless Communications (SPAWC), 2012 IEEE 13th International Workshop on*, pages 580–584, 2012.

- [71] E. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive MIMO for Next Generation Wireless Systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.
- [72] Y. Dawei, F. W. Vook, T. A. Thomas, D. J. Love, and A. Ghosh. Kronecker Product Correlation Model and Limited Feedback Codebook Design in a 3D Channel Model. In *IEEE International Conference on Communications (ICC)*, pages 5865–5870, June 2014.
- [73] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice-Hall information and system sciences series. Prentice Hall, 2000.
- [74] A. H. Sayed. Adaptation, Learning, and Optimization over Networks. *Foundations and Trends in Machine Learning*, 7(4-5):311–801, 2014.
- [75] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson. Analysis of the Pilot Contamination Effect in Very Large Multicell Multiuser MIMO Systems for Physical Channel Models. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3464–3467, May 2011.
- [76] H. ElSawy, E. Hossain, and M. Haenggi. Stochastic Geometry for Modeling, Analysis, and Design of Multi-Tier and Cognitive Cellular Wireless Networks: A Survey. *IEEE Communications Surveys Tutorials*, 15(3):996–1019, Third 2013.
- [77] M. Di Renzo and W. Lu. The Equivalent-in-Distribution (EiD)-Based Approach: On the Analysis of Cellular Networks Using Stochastic Geometry. *IEEE Communications Letters*, 18(5):761–764, May 2014.

- [78] R. K. Ganti and M. Haenggi. Spatial and Temporal Correlation of the Interference in ALOHA Ad Hoc Networks. *IEEE Communications Letters*, 13(9):631–633, Sept 2009.
- [79] S.N. Chiu, D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. Wiley Series in Probability and Statistics. Wiley, 2013.
- [80] Spatial Point Processes and their Applications. In *Stochastic Geometry*, pages 1–75. Springer Science + Business Media, 2007.
- [81] A. Ali, H. Elsayy, T. Y. Al-Naffouri, and M. Alouini. Narrowband Interference Parameterization for Sparse Bayesian Recovery. In *IEEE International Conference on Communications (ICC)*, June 2015 (Accepted).

Vitae

- Name: Alam Zaib
- Nationality: Pakistani
- Date of Birth: December 20, 1978
- Present Address: EE Department, KFUPM, Dhahran 31261, Saudi Arabia.
- Permanent Address: Madina Colony, Dub#2, Mansehra, 21300, Pakistan.
- Telephone: (+966) 50 155 3720, (+92) 321 511 8058
- Email: *alamzaib78@gmail.com*
- M.Sc. Electrical Engineering & Information Technology – Karlsruhe Institute of Technology (KIT) Germany, October 2009
- B.Sc. Electrical Engineering – NWFP University of Engineering & Technology (UET), Peshawar Pakistan, October 2002

- **Research Interests**

My research interests are broad area of statistical signal processing with applications to wireless communications over fading channels. Specifically, I am interested in low-complexity, adaptive and distributed algorithms for channel estimation and data detection for massive MIMO-OFDM systems, and in exploring Random Matrix Theory & Stochastic Geometry for analysis of these systems. I have been interested in solving Blind Source Separation (BSS) problems using Independent Component Analysis (ICA), Deconvolution and ill-posed problems and would also like to explore Sparse Signal Processing for seismic applications.

Publications

- [J1] **A. Zaib** and T. Y. Al-Naffouri, Blind and semi-blind ML detection for space-time block-coded OFDM wireless systems, *EURASIP J. Adv. Sig. Proc.*, vol. 2014, p. 131, 2014.
- [J2] **A. Zaib**, M. Masood, A. Ali, W. Xu and T. Y. AlNaffouri, Distributed channel estimation and pilot contamination analysis for massive MIMO-OFDM systems, *submitted to IEEE. Trans. Communications*, 2015.
- [J3] H. A. J. Alshamary, T. Y. Al-Naffouri, **A. Zaib**, and W. Xu, Optimal non-coherent data detection for massive SIMO wireless systems: A polynomial complexity solution, *submitted to IEEE Trans. Signal Processing*, 2015.
- [J4] **A. Zaib** and T. Y. Al-Naffouri, An optimal low-complexity interpolation technique for channel estimation in OFDM systems, *to be submitted*.
- [C1] **A. Zaib**, M. Masood, M. Ghogho and T. Y. Al-Naffouri, Distributive estimation of frequency selective channels for massive MIMO systems, *23rd European Signal Processing Conference (EUSIPO2015), Nice, France, Aug-Sep, 2015*.
- [C2] H. A. J. Alshamary, T. Y. Al-Naffouri, **A. Zaib**, and W. Xu, Optimal non-coherent data detection for massive SIMO wireless systems: A polynomial complexity solution, *IEEE Signal Processing Workshop (SPW2015), Utah, Aug. 2015*.