2006

Diversity receiver design and channel statistic estimation in fading channels

Jinghua Jin
Iowa State University

Diversity receiver design and channel statistic estimation in fading channels

by

Jinghua Jin

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
Aleksandar Dogandžić, Co-major Professor
Yao Ma, Co-major Professor
Zhengdao Wang
Yong Guan
Huaiqing Wu

Iowa State University
Ames, Iowa
2006

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Jinghua Jin

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES ix

LIST OF ACRONYMS xi

ACKNOWLEDGEMENTS xiii

ABSTRACT xiv

CHAPTER 1. INTRODUCTION 1

1.1 Overview of wireless communications 1

1.2 Diversity combining 2

1.3 Fading channel side information estimation 4

1.4 Problem formulations and main results 6

CHAPTER 2. EQUAL GAIN COMBINING 12

2.1 Introduction 12

2.2 System and channel model 13

2.2.1 Signal model 13

2.2.2 PSAM-based channel estimator 14

2.2.3 Channel estimation error model 16

2.2.4 Symbol detection 19

2.3 Error probability analysis 20

2.3.1 I-PAM 20

2.3.2 M-QAM 22

2.4 Conditional error rate 23

2.4.1 MRC effective SNRs with ICE 23

2.4.2 EGC effective SNRs with ICE 25

2.4.3 EGC amplitude CHF with ICE 26
2.4.4 Discussion 28
2.5 Numerical examples 30
2.6 Summary 34
CHAPTER 3. HYBRID SELECTION/EQUAL GAIN COMBINING . . . 37
3.1 Introduction 37
3.2 Signal model 38
3.3 MGF of HS/EGC output amplitude 40
3.4 Efficient evaluation for TMGF of fading amplitudes 42
3.4.1 Closed-form TMGF for Rayleigh fading channels 43
3.5 Error and outage probabilities of HS/EGC 46
3.5.1 Error probability 46
3.5.2 Outage probability and SNR statistics 46
3.6 Numerical examples 47
3.7 Summary 52
CHAPTER 4. ESTIMATION OF COMPOSITE GAMMA-LOGNORMAL FADING CHANNEL STATISTICS 56
4.1 Introduction 56
4.2 Measurement model and ML estimation 57
4.2.1 Newton-Raphson method 58
4.2.2 EM algorithm 60
4.2.3 Choosing the initial values 62
4.2.4 Cramer-Rao bounds 62
4.3 Numerical examples 63
4.4 Summary 66
CHAPTER 5. ESTIMATION OF MIMO RICIAN FADING CHANNEL STATISTICS 67
5.1 Introduction 67
5.2 Measurement model 68
5.3 ML and REML estimation 71
5.3.1 Correlated fading with arbitrary $\Phi_k \Phi_k^H$ 73

5.3.2 Correlated fading with constant $\Phi_k \Phi_k{}^H$ 75
5.3.3 Independent fading 78
5.4 Numerical examples 80
5.4.1 Block-fading scenario 80
5.4.2 Continuous-fading scenario 83
5.5 Summary 86
CHAPTER 6. CONCLUSIONS AND FUTURE WORK 89
APPENDIX A. DERIVATION OF $\hat{m}$ 92
APPENDIX B. CDF AND TMGF FOR FADING AMPLITUDES 94
B.1 CDF of Nakagami-q random variable 94
B.2 Closed-form TMGF for Nakagami-m Fading Channels 94
APPENDIX C. EM ALGORITHM 98
APPENDIX D. ECME ALGORITHM 100
D.1 The Log-likelihood function 100
D.2 Restricted maximum likelihood 100
D.3 ECME algorithms for correlated fading 101
D.3.1 ML estimation 101
D.3.2 REML estimation 103
D.4 ML estimation for correlated fading with constant $\Phi_k \Phi_k{}^H$ 103
D.5 Cramer-Rao bound 105
D.5.1 CRB for correlated fading with constant $\Phi_k \Phi_k{}^H$ 106
D.5.2 CRB for independent fading with $\Phi_k \Phi_k{}^H = I_{n_T}$ 107
BIBLIOGRAPHY 109

LIST OF FIGURES

Figure 2.1 Illustration of PSAM-based channel estimator. 15

Figure 2.2 Illustration of the error model. 18

Figure 2.3 Receiver Structure for M-QAM with EGC and PSAM in fading channels. 20

Figure 2.4 Decision boundaries (B_1, B_2) and bit-symbol mapping ($b_1 b_2$) for 4-PAM. 21

Figure 2.5 BER (averaged over all i_{off}'s) v.s. bit ASNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE- and Sinc-CEs (with Hamming and rectangular windows), in a Rayleigh fading channel. $L = 3$, $F_1 = F_2 = 4$, $P = 10$, and $B_f T = 0.02$. 31

Figure 2.6 BER (averaged over all i_{off}'s) v.s. bit ASNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE-CE and Sinc-CEs(with Hamming and rectangular windows), in a Rayleigh fading channel. $L = 3$, $F_1 = F_2 = 4$, $P = 15$, and $B_f T = 0.03$. 32

Figure 2.7 $1 - |\rho|^2$ ($i_{\text{off}} = 8$) v.s. bit SNR per branch for diversity 16-QAM with MMSE-CE and Sinc-CEs (with Hamming and Rectangular windows), in a Rayleigh fading channel. $P = 15$, $F_1 = F_2 = 6$, $B_f T = [0.02, 0.03]$. 33

Figure 2.8 $\hat{m}$ vs. bit SNR per branch for 16-QAM with MMSE- and Sinc-CEs (with Hamming and rectangular windows), in a slow Nakagami fading channel. $L = 2$, $P = 10$, $F_1 = F_2 = 3$, $B_f T = 0$. 34

Figure 2.9 Average BER vs. bit SNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE- and Sinc-CEs (with Hamming and rectangular windows) channel estimators, respectively, in a slow Nakagami fading channel. $m = [0.7, 2.7]$, $L = 2$, $P = 15$, $F_1 = F_2 = 3$, $B_f T = 0$. 35

Figure 2.10 Average BER vs. bit SNR per branch for 16-QAM MRC and EGC receivers with perfect CSI and MMSE-CE respectively, in a slow Nakagami fading channel. $m = 2.7$, $L = [1, 2, 3]$, $P = 15$, $F_1 = F_2 = 3$, $B_f T = 0$. 36

Figure 2.11 Average BER vs. bit SNR per branch for M-QAM (with $M = [16, 32, 64]$) MRC and EGC receivers with perfect CSI and MMSE-CE respectively, in a slow Nakagami fading channel. $m = 2.7$, $L = 2$, $P = 10$, $F_1 = F_2 = 5$, $B_f T = 0$. 36

Figure 3.1 Receiver Structure valid for both constant modulus format and non-constant modulus format signals with pre-detection HS/EGC in fading channels. 39

Figure 3.2 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.n.d. Rayleigh fading channel, with $L_c = 1 \sim 4$ and $L = 4$. . . 48

Figure 3.3 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.i.d. Nakagami fading channel, with $m = 2.1$, $L_c = 1 \sim 4$ and $L = 4$. 49

Figure 3.4 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.n.d. Nakagami fading channel, with $m = [3, 2.4, 1.8, 1.2]$, $L_c = 1 \sim 4$ and $L = 4$. 50

Figure 3.5 Average SER vs. bit ASNR for 64-QAM with HS/EGC and HS/MRC over an i.i.d. Rayleigh fading channel, with $L_c = 2$ and $L = 2 \sim 6$. . . 51

Figure 3.6 Average SER vs. bit ASNR for 16-QAM with HS/EGC and HS/MRC over an i.i.d. Weibull fading channel, with $L_c = 2$, $L = 2 \sim 6$, and Weibull parameter $c = 3.1$. 52

Figure 3.7 Average SER vs. bit ASNR for QPSK with HS/EGC over i.i.d. and i.n.d. Nakagami-q fading channels, with $L = 4$ and $L_c = 1 \sim 4$. 53

Figure 3.8 Average SER vs. bit ASNR for 16-QAM with HS/EGC over i.i.d. and i.n.d. Rician fading channels with $L = 4$. For i.i.d. Rician channels, $K = 3$ dB; and for i.n.d. Rician channels $[K_1, K_2, K_3, K_4] = [8, 6, 4, 2]$ dB. 54

Figure 3.9 Outage probability vs. bit ASNR of the first branch for HS/EGC and HS/MRC over an i.n.d. Weibull fading channel with $L = 5$ and $L_c = [1, 2, 3, 5]$. The Weibull c-parameters are given by [5,4,3,2,2], respectively. 54

Figure 3.10 Outage probability vs. bit ASNR of the first branch for HS/EGC over i.i.d. and i.n.d. Nakagami-m fading channels with $L = 5$ and $L_c = [1, 2, 5]$. For i.i.d. Nakagami channels $m = 2.1$, and for i.n.d. channels $m = [3, 2.4, 1.8, 1.2, 0.6]$. 55

Figure 4.1 Mean-square error and Cramér-Rao bound for the proposed estimator of m as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$. 64

Figure 4.2 Mean-square error and Cramér-Rao bound for the proposed estimator of μ as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$. 64

Figure 4.3 Mean-square error and Cramér-Rao bound for the proposed estimator of σ^2 as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$. 65

Figure 4.4 Average number of iterations (left) and CPU time (right) of the EM, Newton-Raphson, and BFGS algorithms as a function of K. 65

Figure 5.1 MSEs and CRBs of ML and REML estimates of some variance components under the unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario. 81

Figure 5.2 MSEs and CRBs of ML and REML estimates of LOS coefficients under the unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario. 82

Figure 5.3 Biases for the ML and REML estimates of the sum of all elements of ψ under the correlated block-fading scenario and unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario. 84

Figure 5.4 MSEs and block-fading CRBs of ML and REML estimates of some variance components under correlated Rician continuous-fading scenario and unstructured (left) and structured (right) LOS array response models. 85

Figure 5.5 MSEs and block-fading CRBs of ML, REML, AML, and LS estimates of the LOS coefficients under correlated Rician continuous-fading scenario and unstructured (left) and structured (right) LOS array response models. 86

Figure 5.6 MSEs and block-fading CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, and sum of all elements of ψ under the correlated Rician continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and unstructured (left) and structured (right) LOS array response models, as functions of K. 87

Figure 5.7 MSEs and block-fading CRBs for the ML, REML, AML, and LS estimates of the LOS coefficients under the correlated continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and unstructured (left) and structured (right) LOS array response models, as functions of K. 88

LIST OF TABLES

Table 2.1 Coefficients for BER calculation for 4-PAM and 8-PAM. 22

Table 3.1 PDF and CDF expressions of fading channel amplitude in Rayleigh, Rician, Nakagami-q, Nakagami-m and Weibull fading channels. $Q_1(\sqrt{2a}, \sqrt{2b}) = \int_b^\infty e^{-t-a} I_0(2\sqrt{at})dt$ is the Marcum-q function of the first order. 42

Table 5.1 Percentages of trials in which the alternating-projection and closed-form estimators for the structured and unstructured LOS array response models in (5.28) and (5.29b) were not positive semidefinite, as functions of K. 82

LIST OF ACRONYMS

AF	amount of fading
AWGN	additive white Gaussian noise
BER	bit error rate
CDMA	code-division multiple-access
CHF	characteristic function
CRB	Cramer-Rao bound
CSI	channel state information
DV	decision variable
ECME	expectation-conditional maximization either
EGC	equal gain combining
EM	expectation-maximization
GCQ	Gauss-Chebyshev quadrature
GSC	generalized selection combining
HS/EGC	hybrid selection/EGC
HS/MRC	hybrid selection/MRC
ICE	imperfect channel estimation
LOS	line-of-sight
MGF	moment generating function
MIMO	multiple-input multiple-output
ML	maximum likelihood
MMSE	minimum mean square error
MRC	maximal ratio combining
MSE	mean square error
NCM	nonconstant modulus

OFDM	orthogonal freq.-division multiplex
PAM	pulse amplitude modulation
PDF	probability density function
PSAM	pilot symbol assisted modulation
PSK	phase shift keying
QAM	quadrature amplitude modulation
REML	restricted maximum likelihood
SER	symbol error rate
SNR	signal-to-noise ratio
UWB	ultra-wideband communication
WLAN	wireless local area network
i.i.d.	independent identically distributed
i.n.d.	independent non-identically distributed

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of my study and life at Iowa State University.

First and foremost, I would like to thank Dr. Aleksandar Dogandžić and Dr. Yao Ma for their guidance, patience and support throughout this research and the writing of this thesis. Their insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. Without their constructive suggestions and knowledgeable guidance, this work could not have been successfully complete.

I gratefully appreciate other committee members for their efforts and contributions to this work: Dr. Zhengdao Wang, Dr. Yong Guan, and Dr. Huaiqing Wu.

I am grateful to my friends in CSP group, Dongbo Zhang, Benhong Zhang, Wei Mo, Lei Zhao, Wei Zhang, and Xiaofan Yang, for making my experience at Iowa State both memorable and fun. Special thanks go to my friends, Ling Li, Jia Jia, and Licheng Jin, for their friendship and fun time we had together.

I am thankful to my friends in control group, Hullas Seghal, Pranav Agarwal, Deepak Sahoo and Vikas Yadav, for interesting discussion of Indian tradition and culture.

Last but not least, I am grateful to my family, specially my husband, my parents, and my brother who were always behind me with endless emotional support in this endurance study.

ABSTRACT

Wireless communication has experienced a tremendous evolution over the past few decades. In wireless communication systems, the channels exhibit phenomena such as fading, multipath, and time-variations. These phenomena often lead to performance degradations or data rate loss and hence present many challenges to the analysis and design of wireless communication systems compared with wired ones. To ensure reliable high-speed communications over insufficient channel bandwidth is one of the major challenges in harsh wireless environments.

The main goal of this thesis is to provide an in-depth study of two important techniques that are effective in improving the performance, data rate, or bandwidth-efficiency in wireless communication systems. The two techniques are, first, diversity combining equipped with quadrature amplitude modulation (QAM), and second, the estimation of fading channel statistical properties.

To effectively combat the adverse effect of fading and to improve the error rate performance, one of the major approaches is to employ diversity combining techniques. In the first part of this thesis, we focus on the equal gain combining (EGC) and hybrid-selection equal gain combining (HS/EGC) for bandwidth-efficient wireless systems (i.e. QAM systems). For EGC QAM systems, we propose the receiver structure and the corresponding decision variables, and then study the effects of imperfect channel estimation (ICE) and quantify the loss of the signal-to-noise ratio (SNR) gain caused by ICE. For HS/EGC QAM system, we develop a general approach to derive unified error rate and outage probability formulas over various types of fading channels based on the proposed HS/EGC receiver. The main contribution of this work lies in that it provides effective hybrid diversity schemes and new analytical approaches to enable thorough analysis and effective design of bandwidth efficient wireless communication systems which suffer from ICE and operate in realistic multipath channels.

Channel side information is of fundamental significance in wireless communication systems, proven to be critical in determining the systems design, achievable data rate, and achievable

performance. A relatively less costly form of side information is the fading channel statistical properties. It is widely assumed that these statistics are available for analysis/design, but little is known about how to obtain them effectively. In the second part of this thesis, we study the estimation of the fading channel statistics. We propose several iterative algorithms to estimate the first- and second-order statistics of general fading or composite fading-shadowing channels and derive the Cramer-Rao bounds (CRBs) for all the cases. We demonstrate that these iterative methods are efficient in the sense that they achieve their corresponding CRBs. The main contribution of this work is that it bridges the gap between the broad utilization of fading channel statistical properties and the lack of systematic study that makes such statistical properties available.

CHAPTER 1. INTRODUCTION

1.1 Overview of wireless communications

Over decades, wireless communication has experienced tremendous popularity and attracted intensive research efforts. Indeed, many existing applications, such as wireless cellular systems, wireless local area networks (WLANs), and satellite communication networks, are gaining maturity and generating increasing impact on virtually every aspect of industry, business, and daily life. In the meantime, many emerging applications, such as wireless sensor networks, are developing from research ideas to realistic systems and are expected to create appealing opportunities and profound changes in future.

There are still many technical issues, however, in the design of wireless communication systems. First of all, wireless systems operate over complicated time-varying channels with possibly multipath fading, shadow fading, interference, and so on. This may lead to severe data rate loss or performance degradations, or both. For example, in a multipath fading scenario, there might exist multiple paths that the radio waves may propagate between the transmitters and the receivers, and the paths may undergo rapid and drastic changes due to the user mobility and environmental variations. Therefore, multipath fading can cause time delay, unpredictable amplitude/phase fluctuations, cancellation of the received waves, and hence sometimes considerable signal strength reduction. Second, only limited radio resources are available for wireless applications, which makes bandwidth efficiency a primary concern in wireless systems designs. As the total number of wireless devices is expected to reach into the trillions soon, the bandwidth limitation becomes more and more critical and demanding. Finally, the communication devices should be small in volume and light in weight with low power consumption, which inevitably restricts their capabilities of transmitting and processing information. These characteristics and constraints essentially differentiate wireless communications from wired communications, and lead to many new challenges to the design of wireless com-

munication systems; among these challenges, a major one is how to ensure reliable high-speed information transmission over unreliable channels with fading and with limited bandwidth.

Due to the above challenges in wireless communication systems, we are required to apply spectral-efficient and power-efficient techniques that are able to mitigate fading impairments and to improve communication reliability. These techniques include, but not limited to, diversity techniques (such as spatial, temporal, frequency, and multipath diversity combining), channel side information monitoring and utilization (such as power and rate adaptation and beam-forming, based on channel statistics/state estimation), space-time techniques (such as space-time coding), near optimal coding (such as low-density parity-check (LDPC) coding and turbo coding), spectral-efficient modulation (such as multilevel quadrature amplitude modulation, QAM for short), equalization, feedback from receivers to transmitters, relaying, clustering, etc. Each of these techniques may improve one or more aspect in wireless communications, and it is usually seen in a wireless communication system that several techniques are jointly employed to achieve better tradeoff among sometimes conflicting requirements. In what follows, we briefly introduce diversity combining and channel side information estimation.

1.2 Diversity combining

Diversity combining is considered as a good option to improve the performance of wireless systems towards the goal of supporting emerging high speed data applications. In its general form, diversity combining includes combinations of spatial diversity, frequency diversity, temporal diversity, or multipath diversity [19, 90].

The main intuition behind diversity combining is to take advantage of the low probability of concurrent deep fades in all the independent paths to lower the probability of error and outage. In other words, multiple replicas of the same information-bearing signal that are obtained from independent paths are combined by appropriate methods, which provides higher possibility of increased overall SNR at the receiver side and hence better detection performance. A remarkable feature of the improvement given by diversity combining is that no increase of either the transmission bandwidth or power is required, but the expense is higher system complexity [94].

We can classify diversity combining according to the methods of combining employed at the receiver, i.e. maximum ratio combining (MRC), equal gain combining (EGC), selection

combining (SC), and hybrid selection (HS) combining.

Maximum ratio combining (MRC): MRC receiver performs detection based on a linear combination of the co-phased path signals, where the weight of each path signal is the fading amplitude of that path. MRC is the *optimal* scheme since it provides maximum receiver-side SNR, and it can yield 20-40 dB performance gains compared to non-diversity system [41]. The disadvantage of MRC is mainly its complexity. Perfect knowledge of the fading amplitudes and phases is needed. In addition, MRC is not considered to be "scalable" as its implementation complexity increases along with the number of available signal paths; note that the number of available signal paths could be rather high in some application, e.g. wideband code-division multiple access (CDMA). Finally, MRC might be sensitive to channel estimation errors, especially in the low SNR regime [94].

Equal gain combining (EGC) [1]: In EGC receiver, all paths are co-phased and equally weighted. It does not require the amplitude information about the fades and has relatively less complexity. Consequently, EGC is only sub-optimal. However, EGC performance is comparable with MRC performance; in the Rayleigh fading case, EGC is only about 1 dB inferior to MRC [41]. Therefore, EGC provides a rather appealing practical solution. On the other hand, similar to the MRC scenario, EGC suffers from 1) the scalability issue, and 2) sensitivity to channel estimation errors. In addition, the analysis of EGC receiver is more difficult than that of MRC receiver [7, 11, 12].

Selection combining (SC): SC receiver performs detection based on only one signal with the highest SNR, and thus eliminates a large portion of the complexity. SC is the least complicated scheme among those described in this section, and it overcomes the scalability difficulty. The drawback of SC is, however, that it does not fully exploit the amount of diversity offered by the channel since only one path signal is used.

Hybrid selection (HS) combining: HS combining receiver combines a predetermined number of the strongest (in the sense of the highest SNR) paths among all the available paths according to either MRC or EGC rule (referred to as HS/MRC or HS/EGC). The major advantage of it is that it provides wireless systems designers the flexibility of striking a *desired* balance between the performance and complexity not offered by MRC, EGC, or SC. More specifically, HS combining complexity does not increase as the number of available paths

[1] In this thesis, we mainly consider pre-detection EGC (or coherent EGC).

increases and is normally significantly lower than MRC and EGC, but it exploits sufficient amount of diversity offered by the channel which is not exploited sufficiently in SC and hence has better performance than SC. On the other hand, performance analysis of HS combining may be technically involved, similar as MRC and EGC.

More detailed descriptions of diversity combining may be found in [2, 35, 41, 72, 89, 94] and therein references.

1.3 Fading channel side information estimation

Channel side information is of fundamental importance in wireless communication systems design since almost all the efficient techniques require knowledge of side information, mainly in two forms, namely, *channel state information (CSI)* and *channel distribution information (CDI)*. This is especially needed when wireless communication designers are really pushed to the extreme of utilizing the limited bandwidth to achieve the highest possible rate, which in turn demands channel side information estimation, given that the channel is fluctuating in a non-deterministic way. Below, we outline the utilization of side information in wireless communications, followed by the description of the estimation of side information, namely the estimation of CSI and CDI.

Side information about fading channel gains is usually represented in two different forms: information about the channel states and information about the channel statistic properties. In either form, such information may be assumed as perfect or noisy. Nonetheless, the availability of side information has significant effects on the systems design, achievable data rate, and communication performance. As an example of the effects, let us recall that in order to achieve the highest receiver-side SNR to improve the data rate and performance, MRC needs to be used at the receiver side, which requires perfect knowledge of CSI. CSI can also be used at the transmitter side, allowing the transmitter to adapt its transmission strategy relative to the channel state variation. More specifically, if the transmitter knows that the channel will experience a larger channel gain (i.e. good channel realization), it would use larger power to transmit faster, whereas if the transmitter knows that the channel will experience a smaller channel gain (i.e. poor channel realization), it would use smaller power to transmit slower. This adaptive strategy improves the efficiency of power utilization. In fact, the solution to the *optimal* (in the sense of achieving the Shannon capacity or ergodic capacity) transmitter

design is waterfilling-in-time based on perfect CSI at the transmitter side [45].

Often times, CSI may be impossible or impractical to obtain, especially in fast fading scenarios. In such cases, knowledge about the channel state distribution, i.e. CDI, may be used. Once CDI is obtained, one might keep using that information as long as the channel statistic properties are not changed. This simplifies the systems design, but it leads to lower data rate and worse performance, compared with cases employing CSI. An example of using CDI is the optimal eigen-beamforming in transmitter design. Based on the second-moments of the channel fades, one can design the transmitter such that the transmitter concentrates its power along the direction with the highest gain (highest not in the instantaneous sense but in probabilistic average sense) [108].

Note that, in a large amount of references, it was assumed that the perfect CSI or CDI is available at the receiver, and the side information comes for free. In reality, such quantities have to be estimated via estimation methods and hence are never obtained without errors and for free. Unfortunately, there are only a small portion of references discussing channel side information estimation issues and imperfect side information issues. The importance of channel estimation is clear: this is necessary if one wishes to use the side information. Moreover, what is of particular interest to us is the estimation of CDI, to which we would like to devote the second half of the present thesis. To the best of our knowledge, the study of estimating CDI is especially rare in the literature, despite its fundamental importance, ironically; see the next section for more detailed description of motivation and problem statement.

Finally, we point out that it is also practically significant to realize that side information is non-perfect: no "separation principle" between channel estimation and systems design based on perfect side information is available, so systems design based on perfect side information may well lead to unexpected results if used in imperfect side information situations. This suggests us to study, first, channel side information estimation methods, and second, the impact of channel estimation error on existing systems design. Note that rather than directly incorporating imperfect side information into systems design, which may be a very difficult task, we characterize the loss due to imperfect side information in a system designed for perfect side information. We remark that both the issues of channel side information estimation and effects of imperfect estimation will be studied in the present thesis.

To summarize, side information about the fading channel has crucial effects on the wireless

communication system, that is, it can increase the data rate, improve the performance, and to some extend determine how one may design the transmitter and receiver. Channel side information estimation is the procedure that fulfils the role of providing side information. In general, channel side information estimation also provides us how accurate the estimations are and hence allows analysis of a communication system with ICE.

1.4 Problem formulations and main results

The description in Section 1.2 indicates that diversity combining techniques and channel distribution information estimation play a significant role in wireless communication systems, in the sense of effectively improving the performance, data rate, or bandwidth-efficiency. The main goal of this thesis is to provide an in-depth study of these two techniques. In this section, we motivate the specific problems we wish to address, briefly describe the approaches, informally present our main findings, and summarize the contributions.

In the first half of this thesis, we focus on the EGC and HS/EGC for bandwidth-efficient wireless systems (i.e. QAM systems). Note that, due to the increasingly restricting bandwidth requirement in the wireless channels, spectrally efficient multilevel modulations, such as QAM, have been widely employed in wireless communication systems, for example in WLAN [49] and orthogonal frequency-division multiplex (OFDM) [100]. The main advantage of QAM is that it achieves high data rate transmission without increasing the bandwidth. For this reason, it has been incorporated into diversity schemes and generated intensive interest for the design of future wireless communication systems. We remark that the main drawback of QAM is that it is not power efficient, especially when the constellation size is large.

The advantages of MRC, EGC, or HS/EGC described in Section 1.2 and those of QAM described above suggest that combining QAM with diversity schemes, such as MRC, EGC, or HS/EGC, may provide appealing solutions to meet the requirement of high data rate transmission over wireless fading channels. In particular, QAM offers high data rate under bandwidth constraint, MRC offers optimal receiver design which maximizes receiver-side SNR and hence guarantees high data rate and low error rate, and EGC or HS/EGC offers low error rate with low complexity.

In a diversity QAM receiver, the measured CSI needs to be used, which is usually imperfect and represented stochastically. In other words, the imprecise CSI is available instead of the

precise CSI. It is of practical importance to study the performance in the presence of channel estimation errors, and the effects of imperfect channel estimation (ICE) on the performance must be taken into account in this design. To the best of our knowledge, an accurate bit error rate (BER) result for QAM with EGC and ICE over general fading channels is not presented in the literature yet even for the Rayleigh fading channel. This calls for the further investigation of performance analysis of EGC QAM with ICE. Therefore, the first problem we wish to solve is

Problem 1. *Design EGC receiver and evaluate the BER performance for square and rectangular M-QAM over Rayleigh or Nakagami fading channels, and characterize the performance loss due to ICE.*

Despite the conceptual and practical significance, the study on MRC QAM and EGC QAM presents theoretical challenges due to the facts that MRC/EGC output SNR depends on the statistics of the sum of the faded signal envelopes and that finding the closed form expression of the probability density function (PDF) of the sum of more than two Rayleigh, Rician, Nakagami-m distributed random variables is extremely difficult [5, 15, 16]. To circumvent the difficulty of deriving the PDFs of the fading amplitude, some researchers proposed to make use of the moment generating function (MGF) or characteristic function (CHF) of the fading amplitude, see [6, 11]. The MGF- or CHF-based approach simplifies the analytical evaluation of diversity receiver. Hence, our approach to attack the above problem is to adopt a MGF or CHF based analysis. In order for us to estimate the channel state information, we apply pilot-symbol assisted modulation (PSAM) channel estimator.

In addition, EGC QAM is particularly challenging compared to MRC QAM. Among other reasons, one specific difficulty is that traditional EGC does not involve any modulus information about the channel fades, which makes it directly applicable to constant modulus modulation (such as PSK), and not directly applicable to constant modulus modulation (such as QAM). A non-trivial extension to account for some amount of modulus information is needed to study EGC QAM.

The above approaches and extension help us to establish the following results to Problem 1. We obtain a novel receiver design for EGC QAM and a formulation of BER of M-QAM with ICE in terms of the signal constellation-dependent effective SNRs or amplitudes, which allows us to derive the general, accurate, and easy-to-evaluate BER formulas for square and

rectangular diversity M-QAM with channel estimation errors. Our result shows that the performance loss caused by ICE may be manifested by the signal decision space distortion and a scaling of the effective SNR. Using our analytical result, we evaluate the performance of M-QAM with PSAM and present some insightful findings. We also verify our results using numerically simulations.

Our contributions of this work are: 1) proposing the first EGC receiver for M-QAM system and presenting the related decision variables; 2) providing general and easy-to-evaluate BER expressions for diversity M-QAM with channel estimation errors; and 3) quantifying SNR loss caused by channel estimation errors and showing the relation between performance loss and system parameters. To summarize its importance, this work enables wireless communication systems designers to completely and easily analyze the performance and effectively design bandwidth efficient systems that can achieve both high diversity gain and low complexity, and it sheds light on how the designed systems may behave practically under ICE. As an example, it allows us to better study adaptive modulation scheme that has the ability to achieve better data rate and performance tradeoff: if the channel is good, then a larger constellation may be used for higher data rate, and if the channel is bad, then a smaller constellation may be used for smaller BER. Such a scheme would require a thorough study of the underline non-adaptive modulation schemes.

In order to provide a scheme sufficiently exploring diversity and yet with a lower implementation complexity that is desired not to increase as the number of diversity branches increases, we are interested in studying HS/EGC with QAM. This may be regarded as a promising practical candidate diversity format for multipath channels or multiple antenna systems. Note that to enable HS/EGC diversity for QAM, which is a non-constant modulus (NCM) modulation, this would in turn require designing HS/EGC receiver structure and proposing a decision variable.

Problem 2. *Propose an HS/EGC receiver structure and decision variable for QAM, and evaluate the error and outage probabilities over general fading channels.*

HS/EGC is a new diversity format in which a subset of branches with the largest SNRs are selected and combined according to EGC rule. It provides a performance close to HS/MRC, and yet with a lower implementation complexity.

The main approach we apply is based on MGF of fading amplitudes and Parseval's theorem. And our main results state as follows. We propose the relevant HS/EGC receiver structure and decision variable. We develop a general analytical framework to evaluate the error and outage probabilities of pre-detection HS/EGC for a large class of modulation formats and versatile system and channel parameters, such as different fading types and non-identically distributed diversity branches. Performance comparison with HS/MRC suggests that HS/EGC achieves the same diversity order as HS/MRC.

The contributions of this work could be summarized as: 1) proposing HS/EGC receiver for QAM system and presenting the related decision variables; 2) providing computation efficient error rate expressions for a large class of signal constellations with HS/EGC diversity in general fading channels; and 3) deriving closed-form expressions of truncated MGF for Rayleigh and Nakagami-m fading channels. The significance of this work is that it offers wireless communication systems designers another alternative method to completely analyze the performance and effectively design bandwidth efficient systems that can achieve both sufficiently high diversity gain and fairly low complexity.

In the second half of this thesis, we focus more on the CDI estimation issues. If the estimation of CSI is too costly or impractical, the above described diversity combining schemes (as well as most CSI based schemes) cannot be used. In such cases, we are motivated to estimate the channel statistic properties or CDI instead. Depending on different radio propagation environments, different models describing the channel statistic properties are used. There are only a few references scattered in the literature that studied the CDI estimation problem, see [25, 60, 82, 107]. In this thesis, we aim at providing a generic framework to estimate CDI for rather general fading channels. We first investigate the problem of estimating composite gamma-lognormal fading channel. A composite fading-shadowed environment consists of multipath fading superimposed on lognormal shadowing. This is often the scenario in congested downtown areas with slow moving pedestrians and vehicles [48, 95], landmobile satellite systems subject to urban shadowing [79, 101], and distributed antenna systems [87, 88]. In this work, we study the composite gamma-lognormal fading, whose PDF is obtained by averaging the gamma distributed signal power over the conditional density of the lognormally distributed mean signal power.

Problem 3. *Find efficient estimators to estimate the fading and shadowing parameters of the*

composite gamma-lognormal fading channels.

We know that maximum likelihood (ML) estimates are asymptotically efficient. Therefore, we focus on ML based algorithms. However, analytic solutions to ML estimation problems usually elude the study, due to the nonlinearity involved in the likelihood functions. Instead of looking for analytic solutions, recursive numerical approaches are widely used.

Our main results are as follows. Newton-Raphson and expectation-maximization (EM) [2] algorithms are developed to compute the ML estimates of the mean and variance of the shadowing component, and the Nakagami-m parameter of the fading component. Among others, the EM algorithm has been proven as computationally efficient and numerically stable. We also derive Cramer-Rao bounds (CRBs) for the unknown parameters. Numerical simulations demonstrate the proposed methods asymptotically.

Finally we study the CDI estimation for Rician and Rayleigh fading channels. Rician fading is frequently used if there is one line-of-sight (LOS) component with other scattered components, and if the strength of the LOS component is zero, it reduces to Rayleigh fading. Furthermore, we allow the channel to have multiple transmitting and receiving antennas, i.e. multi-input multi-output (MIMO) channel.

Problem 4. *Find efficient estimators to estimate the mean and covariance parameters of MIMO Rician and Rayleigh fading channels.*

Again we are interested in proposing ML based methods to obtain the efficient estimates. It is well known that *expectation-conditional maximization either* (ECME) algorithm is an extension of EM algorithm, which shares with EM its monotone increase in likelihood and stable convergence to an ML estimate, but converges more quickly than EM [66].

In particular, we propose ML and restricted maximum likelihood (REML) methods based on ECME algorithms for estimating the mean and covariance parameters of MIMO Rician and Rayleigh block-fading channels using measurements from multiple coherent intervals containing both amplitudes and phases of the received signal. Correlated and independent fading scenarios with structured and unstructured LOS array response models are considered. Computationally efficient ML and approximate ML (AML) estimators are proposed for unitary

[2]The EM algorithms converge monotonically to a local or the global maximum of the likelihood function, see e.g. [77, ch. 3] and [81, ch. 12.4].

space-time modulation schemes and orthogonal designs in correlated fading. We also derive CRBs for the unknown parameters, discuss initialization of the proposed algorithms, and show that they asymptotically attain CRB via numerical simulations under both the block- and continuous-fading scenarios.

Our main contribution is that, the solutions to Problems 3 and 4 bridge the gap between the broad utilization of CDI and lack of systematic study that makes this information available. The results also enable us to characterize the estimation error and hence to evaluate the effects of imperfect estimates on wireless systems behavior, and furthermore, make it possible to improve existing transmission and receiving schemes in the imperfect estimation situation.

In the rest of this thesis, each chapter consists of the study of one problem shown above, followed by conclusions and future work. Throughout this thesis, we use superscripts $*$, T, H, -1 to represent the scalar conjugate, matrix transpose, conjugate transpose, and matrix inversion, respectively. We use $\mathrm{Re}(x)$, $\mathrm{Im}(x)$, $E[x]$, and $\mathrm{var}\{x\}$ to denote the real part, imaginary part, expected value, and variance of x, respectively. $\det(A)$ is the determinant of matrix A. $\mathbf{0}_{M\times N}$ and $\mathbf{1}_{M\times N}$ are the $M \times N$ all-zero and all-one matrices, respectively; and I_L represents the $L \times L$ identity matrix.

CHAPTER 2. EQUAL GAIN COMBINING

2.1 Introduction

To ensure reliable transmission in wireless communication, one of the major approaches is to employ diversity techniques. Diversity can effectively combat the adverse effects of fading and improve the error rate performance. Among popular diversity receivers, MRC is optimal in the sense that it maximizes the receiver output SNR. However, the implementation of MRC is more complicated than other schemes, such as EGC, because it needs the knowledge of both moduli and phases of the channel gains. Compared to MRC, EGC provides comparable performance with a lower complexity. Therefore, there has been an increasing interest in the study of EGC, for example, in the design of low-complexity receiver for direct sequence code-division multiple access (DS/CDMA) systems [46] and ultra-wideband (UWB) communication systems [37, 91].

On the other hand, due to the limited bandwidth resource in the wireless channels, spectrally efficient multilevel constellations, such as M-QAM, has been widely employed in wireless communication systems. Therefore, the accurate error probability of the M-QAM receivers in various fading conditions is important for high spectral-efficiency communication and adaptive modulation design [9, 42, 68].

For coherent detection, CSI is needed at the receiver. In practice, CSI cannot be perfect in fading channels, and thus the effect of ICE must be considered for accurate performance evaluation and system design [21, 22, 23, 35, 71, 72, 73, 93, 105]. In [13], performance of binary phase-shift keying (BPSK) with EGC and Gaussian channel estimation errors in a Rayleigh fading channel was studied. And the error rate of EGC diversity M-ary PSK (M-PSK) with ICE in several types of fading channels was derived in [71].

Performance of diversity M-QAM with ICE has also been studied in the past. In [103], BER for pulse amplitude modulation (PAM) and QAM signals in an MRC diversity Rayleigh fading

channel and a non-diversity Rician fading channel with ICE was obtained. BER expressions for MRC diversity square M-QAM with PSAM [23, 59] for independent and identically distributed (i.i.d.) Rayleigh and Rician fading channels were derived in [21, 105] and [22], respectively. An exact BER result for square and rectangular M-QAM with MRC in arbitrarily correlated Rician fading channels was developed in [73].

An accurate BER result accounting for the effect of ICE on the performance of EGC M-QAM, however, has not been obtained yet even for the Rayleigh fading channel. Furthermore, the above mentioned performance analysis results for MRC M-QAM with ICE have been confined to Rayleigh and Rician fading channels. In this chapter, we present a unified performance analysis for M-QAM with both MRC and EGC diversity formats in Nakagami and Rayleigh fading channels.

Although the decision variable for the EGC with constant modulus modulation formats (such as M-PSK) is well known [94, 95], the decision variable for EGC with non-constant modulus modulation formats is virtually not known. Consequently, it is difficult to model and simulate the effect of ICE on EGC M-QAM in fading channels. To facilitate the analysis of EGC M-QAM with ICE, we propose the EGC receiver structure and related decision variable, which are valid for NCM formats. In addition, we propose a channel estimation error model which is valid for arbitrary linear channel estimators in Rayleigh fading channels, and for near-minimum mean square error (MMSE) channel estimators in Nakagami fading channels.

We provide general, accurate and easy-to-evaluate BER expressions for square and rectangular diversity M-QAM with channel estimation errors. Our results analytically show the performance loss caused by ICE compared to the case of perfect CSI. Our analysis method can be readily extended to the design of adaptive modulation schemes using diversity M-QAM in the presence of ICE over Nakagami fading channels.

2.2 System and channel model

2.2.1 Signal model

We use (I, J)-QAM to denote the modulation with I and J signal levels in the horizontal and vertical directions, respectively, for $I, J \in \{2^k\}$, $k = \{1, 2, \ldots\}$. The total number of signal levels M is given by $M = I \times J$. When $I = J$, the $(\sqrt{M}, \sqrt{M})$-QAM is the square M-QAM;

and when $I \neq J$, the (I, J)-QAM is called the rectangular QAM [27, 94]. We denote the received signals in the ith symbol interval over all L diversity branches by

$$\mathbf{y}(i) = \mathbf{c}(i)d(i) + \mathbf{n}(i) \tag{2.1}$$

where $d(i) = \sqrt{E_a}(A_I + jA_J)$ is an M-QAM symbol, with $j = \sqrt{-1}$, $A_I \in \{-I+1, \ldots, -1, 1, \ldots, I-1\}$, and $A_J \in \{-J+1, \ldots, -1, 1, \ldots, J-1\}$. E_a is used to normalize the data symbol energy to unity, i.e. $E_d = E[|d(i)|^2] = 1$. For square QAM $E_a = \frac{3}{2(M-1)}$; and for rectangular (I, J)-QAM $E_a = \frac{3}{I^2+J^2-2}$ [27]. For I-ary PAM $E_a = \frac{3}{I^2-1}$.

The channel vector $\mathbf{c}(i) = [c_1(i), \ldots, c_L(i)]^T$ contains i.i.d. channel gains with $\sigma_c^2 = E[|c_l(i)|^2]$ for all l branches. We assume a fast Rayleigh fading channel and a slow Nakagami fading channel. For time-selective Rayleigh fading channels, we define B_f as the Doppler fading bandwidth, and T as a symbol duration. The normalized temporal channel correlation coefficient (assumed to be identical for all the branches) is defined as

$$\tilde{R}_c(n) = \frac{E[c_l(i)c_l^*(i-n)]}{\sigma_c^2}. \tag{2.2}$$

For Jakes' fading spectrum, $\tilde{R}_c(n) = J_0(2\pi n B_f T)e^{j2\pi f_o nT}$; and for the Gaussian fading spectrum, $\tilde{R}_c(n) = \exp(-(\pi n B_f T)^2)e^{j2\pi f_o nT}$, where $J_0(x)$ is the zeroth order Bessel function of the first kind and f_o is the frequency offset.

For the Nakagami fading channel, we assume a block static fading model. That is, the channel remains constant for a block of symbols' duration (related to the PSAM parameters) and changes independently from one block to another. The additive background noise vector, $\mathbf{n}(i) = [n_1(i), \ldots, n_L(i)]^T$, is a zero-mean circularly symmetric complex Gaussian process with average power $E[|n_l(i)|^2] = N_0$ for $l = 1, \ldots, L$. The average bit SNR at each signal branch is given by $\bar{\gamma}_b = \sigma_c^2/[(\log_2 M)N_0]$.

2.2.2 PSAM-based channel estimator

In order to utilize PSAM to estimate the channel coefficient $\mathbf{c}(i)$ of the desired symbol $d(i)$, a pilot symbol is inserted into the data stream every P symbol intervals. And we employ F pilot symbols to estimate $\mathbf{c}(i)$. The pilot symbols are known to both the transmitter and the receiver.

The pilot symbols may be collected in an $F \times 1$ vector

$$\mathbf{d}_{\text{PS}} = [d(i - PF_1 + i_{\text{off}}), \ldots, d(i - P + i_{\text{off}}), d(i + i_{\text{off}}), \ldots, d(i + P(F_2 - 1) + i_{\text{off}})]^T,$$

where F_1 and F_2 (with $F_1 + F_2 = F$) are the numbers of pilot symbols on the left and right sides of $d(i)$, respectively, and i_{off} ($i_{\text{off}} = 1, 2, \ldots, P-1$) is the offset of the desired symbol $d(i)$ to the closest pilot symbol on its right side. We call P the insertion interval, and F_1 and F_2 the pre- and post-interpolation orders, respectively. See Figure 2.1. For slow Nakagami fading channel, we assume the channel keeps constant within FP symbol durations.

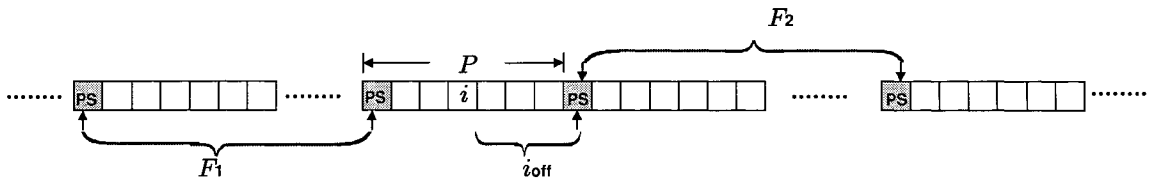

Figure 2.1 Illustration of PSAM-based channel estimator.

The received signals at the pilot symbols' positions for estimating channel $c_l(i)$ may be written as an $F \times 1$

$$\mathbf{y}_{l,\text{PS}} = \text{diag}(\mathbf{d}_{\text{PS}})\mathbf{c}_{l,\text{PS}} + \mathbf{n}_{l,\text{PS}}, \tag{2.3}$$

where

$$\mathbf{c}_{l,\text{PS}} = [c_l(i - PF_1 + i_{\text{off}}), \ldots, c_l(i + i_{\text{off}}), \ldots, c_l(i + P(F_2 - 1) + i_{\text{off}})]^T,$$

$$\mathbf{n}_{l,\text{PS}} = [n_l(i - PF_1 + i_{\text{off}}), \ldots, n_l(i + i_{\text{off}}), \ldots, n_l(i + P(F_2 - 1) + i_{\text{off}})]^T.$$

are the channel gain and noise components of the lth branch at the pilot symbols' positions for estimating $c_l(i)$, respectively.

Without loss of generality, we assume $\mathbf{d}_{\text{PS}} = \sqrt{P_{\text{PS}}}\mathbf{1}_{F\times 1}$, where P_{PS} is the transmitted power of the pilot symbols. Thus, equation (2.3) simplifies to

$$\mathbf{y}_{l,\text{PS}} = \sqrt{P_{\text{PS}}}\mathbf{c}_{l,\text{PS}} + \mathbf{n}_{l,\text{PS}}. \tag{2.4}$$

The channel estimate for $c_l(i)$, denoted as $\hat{c}_l(i)$, is given by

$$\hat{c}_l(i) = \mathbf{w}_l^H \mathbf{y}_{l,\text{PS}}, \tag{2.5}$$

where $\mathbf{w}_l$ is an $F \times 1$-size channel estimator filter vector.

For the sinc-interpolator-based channel estimator (Sinc-CE) with rectangular window, we have

$$\mathbf{w}_{l,\text{sinc}} = [h_{\text{sinc}}(-F_1 + i_{\text{off}}/P), \ldots, h_{\text{sinc}}(i_{\text{off}}/P), \ldots, h_{\text{sinc}}((F_2 - 1) + i_{\text{off}}/P)]^T \tag{2.6}$$

where $h_{\text{sinc}}(x) = \sin(\pi x)/(\pi x)$. Other than rectangular window, a Hamming window may also be employed.

For the MMSE channel estimator (MMSE-CE) we have

$$\mathbf{w}_{l,\text{mmse}} = \mathbb{R}_{y_{l_{\text{PS}}}}^{-1} \mathbf{r}_{y_{l_{\text{PS}}},c} \tag{2.7}$$

where $\mathbb{R}_{y_{l_{\text{PS}}}} = E[\mathbf{y}_{l,\text{PS}}\mathbf{y}_{l,\text{PS}}^H]$ and $\mathbf{r}_{y_{l_{\text{PS}}},c} = E[\mathbf{y}_{l,\text{PS}}c_l^*(i)]$ are the auto-correlation matrix of $\mathbf{y}_{l,\text{PS}}$ and cross-correlation vector between $\mathbf{y}_{l,\text{PS}}$ and $c_l(i)$, respectively. After some manipulations, we can show that

$$\mathbf{r}_{y_{l_{\text{PS}}},c} = \sqrt{P_{\text{PS}}}\sigma_c^2[\tilde{R}_c(-PF_1 + i_{\text{off}}), \ldots, \tilde{R}_c(i_{\text{off}}), \ldots, \tilde{R}_c(P(F_2 - 1) + i_{\text{off}})] \tag{2.8a}$$

$$\mathbb{R}_{y_{l_{\text{PS}}}} = \begin{bmatrix} P_{\text{PS}}\sigma_c^2\tilde{R}_c(0) + N_0 & P_{\text{PS}}\sigma_c^2\tilde{R}_c^*(P) & \ldots & P_{\text{PS}}\sigma_c^2\tilde{R}_c^*(P(F-1)) \\ P_{\text{PS}}\sigma_c^2\tilde{R}_c(P) & P_{\text{PS}}\sigma_c^2\tilde{R}_c(0) + N_0 & \ldots & \\ \vdots & & \ddots & \\ P_{\text{PS}}\sigma_c^2\tilde{R}_c(P(F-1)) & & \ldots & P_{\text{PS}}\sigma_c^2\tilde{R}_c(0) + N_0 \end{bmatrix} \tag{2.8b}$$

where $\tilde{R}_c(n)$ is the temporal channel correlation coefficient defined in Section 2.2.1.

In the case of slow fading, it holds that $\tilde{R}_c(n) = 1$ for all $n = 0, \cdots, PF - 1$. Hence, it can be shown that

$$\mathbf{r}_{y_{l_{\text{PS}}},c} = \sqrt{P_{\text{PS}}}\sigma_c^2\mathbf{1}_F, \quad \mathbb{R}_{y_{l_{\text{PS}}}} = P_{\text{PS}}\sigma_c^2\mathbf{1}_F\mathbf{1}_F^T + N_0 I \tag{2.9}$$

where $\mathbf{1}_N$ is the $N \times 1$ all-one vector.

2.2.3 Channel estimation error model

The estimate for channel vector $\mathbf{c}(i)$ is denoted by $\hat{\mathbf{c}}(i) = [\hat{c}_1(i), \ldots, \hat{c}_L(i)]^T$. For i.i.d. signal branches, without loss of generality, we assume $E[|\hat{c}_l(i)|^2] = \sigma_{\hat{c}}^2$, for all l. Let the normalized correlation coefficient between $\hat{c}_l(i)$ and $c_l(i)$ be defined as

$$\rho = \frac{E[c_l(i)\hat{c}_l^*(i)]}{\sigma_c\sigma_{\hat{c}}}, \tag{2.10}$$

where σ_c and $\sigma_{\hat{c}}$ are the standard deviations of c_l and $\hat{c}_l$, respectively. We can show that

$$\sigma_{\hat{c}} = \sqrt{\mathbf{w}_l^H\mathbb{R}_{y_{l_{\text{PS}}}}\mathbf{w}_l}, \quad \text{and} \quad \rho = \frac{(\mathbf{w}_l^H\mathbf{r}_{y_{l_{\text{PS}}},c})^*}{\sigma_c\sqrt{\mathbf{w}_l^H\mathbb{R}_{y_{l_{\text{PS}}}}\mathbf{w}_l}}.$$

In general, ρ is a function of the average SNR, the fading channel parameters (e.g. $B_f T$, and the Nakagami m parameter), and the PSAM parameters. For MMSE-CE, we can show

that

$$\rho_{\text{mmse}} = \frac{\mathbf{r}^H_{y_{l_{\text{PS}}},c}\mathbb{R}^{-1}_{y_{l_{\text{PS}}}}\mathbf{r}_{y_{l_{\text{PS}}},c}}{\sigma_c\sqrt{\mathbf{r}^H_{y_{l_{\text{PS}}},c}\mathbb{R}^{-1}_{y_{l_{\text{PS}}}}\mathbf{r}_{y_{l_{\text{PS}}},c}}} = \frac{\sqrt{\mathbf{r}^H_{y_{l_{\text{PS}}},c}\mathbb{R}^{-1}_{y_{l_{\text{PS}}}}\mathbf{r}_{y_{l_{\text{PS}}},c}}}{\sigma_c}.$$

The channel estimation error at the lth branch is given by $e_l(i) = \hat{c}_l(i) - c_l(i)$. Normally, only for MMSE-CE $E[e_l(i)\hat{c}_l^*(i)] = 0$ holds, i.e., $e_l(i)$ and $\hat{c}_l(i)$ are orthogonal [20, 71, 93]. In other words, in general the channel estimation error is *not* orthogonal to the estimate, which may complicates our performance analysis to provide an accurate BER result.

In order to facilitate our analysis and to account for the non-orthogonal channel estimation errors for arbitrary linear channel estimators, the first step we need to follow is to construct an *equivalent* channel estimation error term $z_l(i)$ (based on a given arbitrary linear channel estimator) which is uncorrelated to the linear estimate $\hat{c}_l(i)$. In the following, we propose a versatile channel estimation error model as

$$c_l(i) = \rho\frac{\sigma_c}{\sigma_{\hat{c}}}\hat{c}_l(i) + z_l(i) \tag{2.11}$$

where ρ is the correlation coefficient defined in (2.10), $\{z_l(i)\}_{l=1}^L$ are the i.i.d. equivalent estimation error terms with zero mean and variance $\sigma_z^2 = (1 - |\rho|^2)\sigma_c^2$.

To see that for this error model, it holds that $E[z_l(i)\hat{c}_l^*(i)] = 0$, let us compute $E[c_l(i)\hat{c}_l^*(i)]$ using both sides of (2.11).

$$\begin{aligned} E[c_l(i)\hat{c}_l^*(i)] &= \rho\frac{\sigma_c}{\sigma_{\hat{c}}}E[|\hat{c}_l(i)|^2] + E[z_l(i)\hat{c}_l^*(i)] \\ &= \rho\sigma_c\sigma_{\hat{c}} + E[z_l(i)\hat{c}_l^*(i)] \\ &= E[c_l(i)\hat{c}_l^*(i)] + E[z_l(i)\hat{c}_l^*(i)] \end{aligned}$$

which leads to $E[z_l(i)\hat{c}_l^*(i)] = 0$.

We remark that our proposed error model (2.11) is in essence the Gram-Schmidt orthogonalization of the random variables $c_l(i)$ and $\hat{c}_l(i)$. That is, to obtain a random variable $z_l(i)$ from $c_l(i)$ and its estimate $\hat{c}_l(i)$ and $z_l(i)$ is orthogonal to $\hat{c}_l(i)$, we may subtract the projection of $c_l(i)$ from $\hat{c}_l(i)$. See Figure 2.2. Note that the cosine of the angle between $c_l(i)$ and $\hat{c}_l(i)$ in the figure is indeed the correlation coefficient ρ.

If the channel gain amplitude is distributed according to Rayleigh fading model, i.e. $c_l(i)$ is complex Gaussian, then $\hat{c}_l(i)$ is complex Gaussian, and therefore $z_l(i)$ is Gaussian distributed. Thus, equation (2.11) can be used to model channel estimation errors caused by arbitrary

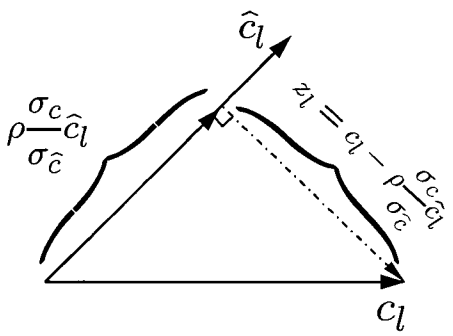

Figure 2.2 Illustration of the error model.

linear channel estimators in the Rayleigh fading channels, where $\{z_l(i)\}_{l=1}^{L}$ follow a complex Gaussian distribution and are independent of the channel estimates $\{\hat{c}_l(i)\}_{l=1}^{L}$.

As a special case for MMSE-CE, using the orthogonality principle we obtain that $\sigma_{\hat{c}}^2 = |\rho|^2\sigma_c^2$, and (2.11) reduces to $c_l(i) = \hat{c}_l(i) + z_l(i)$.

We note that a popular channel estimation error model, shown below in (2.12), can be included in (2.11) as a special case.

$$\hat{c}_l(i) = c_l(i) + e_l(i) \tag{2.12}$$

where $e_l(i)$ is the zero-mean channel estimation error due to the noise and is independent of the channel fade $c_l(i)$ (note it does not claim the error is independent of the estimation $\hat{c}_l(i)$). Equation (2.12) is valid for least-square (LS) channel estimators [14, 86] over slow fading channels, but generally does not hold for fast fading channels. In fact, in the slow fading case, equation (2.4) changes to

$$\mathbf{y}_{l,\mathrm{PS}} = \sqrt{P_{\mathrm{PS}}}\mathbf{1}_F c_l(i) + \mathbf{n}_{l,\mathrm{PS}}, \tag{2.13}$$

it can be easily computed that the LS estimator is

$$\begin{aligned} \hat{c}_l(i) &= (\sqrt{P_{\mathrm{PS}}}\mathbf{1}_F^T\sqrt{P_{\mathrm{PS}}}\mathbf{1}_F)^{-1}\sqrt{P_{\mathrm{PS}}}\mathbf{1}_F^T\mathbf{y}_{l,\mathrm{PS}} \\ &= \frac{1}{\sqrt{P_{\mathrm{PS}}}L}\mathbf{1}_F^T\mathbf{y}_{l,\mathrm{PS}} = c_l + \frac{1}{\sqrt{P_{\mathrm{PS}}}L}\mathbf{1}_F^T\mathbf{n}_{l,\mathrm{PS}}. \end{aligned} \tag{2.14}$$

That is, the estimation error depends only on the channel noise. Since the channel noise and the channel fades are independent, we have verified that $e_l(i)$ and $c_l(i)$ are independent. For the model (2.12) we find that $\rho = \sigma_c/\sqrt{\sigma_c^2 + \sigma_e^2}$, where σ_e^2 is the variance of $e_l(i)$. Thus, for

the case of (2.12) our model in (2.11) reduces to

$$c_l(i) = \frac{\sigma_c^2}{\sigma_{\hat{c}}\sqrt{\sigma_c^2 + \sigma_e^2}}\hat{c}_l(i) + z_l(i) \tag{2.15}$$

where $z_l(i)$ has zero mean and variance $\sigma_e^2\sigma_c^2/[\sigma_c^2 + \sigma_e^2]$.

Furthermore, (2.11) is valid for modelling ICE caused by near-MMSE channel estimators in Nakagami fading channels, where the estimation error $z_l(i)$ can be approximated as a Gaussian random variable. Here, near MMSE-CE refers to a channel estimator $\mathbf{w}_l$ for which the normalized vector norm $|\mathbf{w}_l - \mathbf{w}_{l,\text{mmse}}|/|\mathbf{w}_{l,\text{mmse}}| \ll 1$ holds. For a channel estimator which differs significantly from the MMSE-CE, $z_l(i)$ in (2.11) may not be accurately approximated as a complex Gaussian variable, and the BER analysis result involves an approximation.

2.2.4 Symbol detection

2.2.4.1 MRC

The decision variable (DV) at the MRC QAM receiver output is given by [21] $\tilde{D} = \sum_{l=1}^{L} \hat{c}_l^*(i)y_l(i)/\sum_{l=1}^{L} |\hat{c}_l(i)|^2$. The transmitted symbol $d(i)$ can be recovered by comparing $\tilde{D}$ with the horizontal and vertical QAM decision boundaries [21, 27, 94].

2.2.4.2 EGC

Though the BER performance of EGC diversity QAM with perfect CSI have been studied [7, 12, 94], to our knowledge, the receiver structure to realize EGC for M-QAM (even with perfect CSI) has not been derived in the literature. Here, for theoretical interest and practical importance, we propose the EGC receiver structure shown in Figure 2.3 valid for NCM formats taking the ICE into account. In our proposed EGC receiver, a NCM power normalization section (a factor of $1/[\sum_{l=1}^{L} |\hat{c}_l(i)|]$) is included. This makes it different from conventional EGC, which is valid only for constant modulus formats.

The corresponding decision variable for EGC M-QAM in the presence of ICE is

$$\tilde{D} = \frac{\sum_{l=1}^{L} \hat{c}_l^*(i)y_l(i)/|\hat{c}_l(i)|}{\sum_{l=1}^{L} |\hat{c}_l(i)|} \tag{2.16}$$

We verify the validity of (2.16) by analysis and simulation given later in this paper. We note that for EGC only the phase estimate $\hat{c}_l(i)/|\hat{c}_l(i)|$ for all the branches and $\sum_{l=1}^{L} |\hat{c}_l(i)|$, the modulus of sum of the channel gain estimates, are required; whereas for MRC both the phases

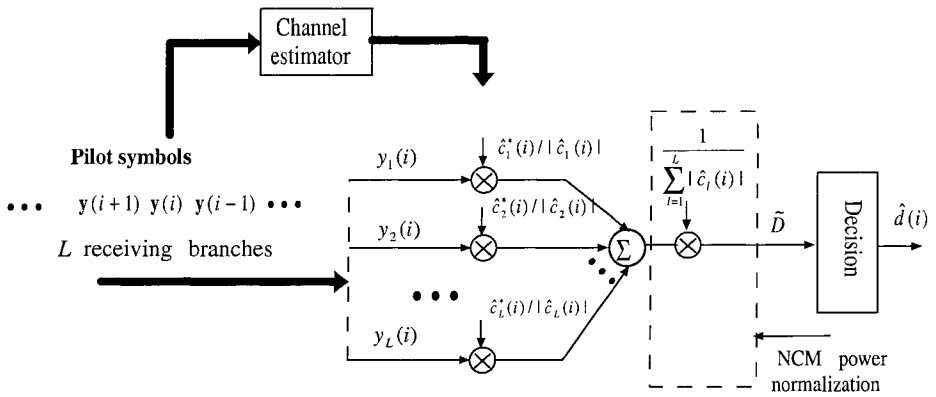

Figure 2.3 Receiver Structure for M-QAM with EGC and PSAM in fading channels.

and moduli of the channel gain estimates $\{\hat{c}_l(i)\}_{l=1}^{L}$ at all the branches are required. Thus, EGC is still simpler than MRC for NCM formats. However, compared to constant modulation format, EGC for QAM is more complex as knowledge of $\sum_{l=1}^{L}|\hat{c}_l(i)|$ is required. This is not surprising because for NCM signalling a channel gain normalization is required. For EGC M-QAM only the phases of $\{c_l(i)\}_{l=1}^{L}$ and $\sum_{l=1}^{L}|c_l(i)|$ have to be estimated (i.e., $L+1$ real values) instead of $\{c_l(i)\}_{l=1}^{L}$ (i.e., $2L$ real values). And in order to enable a fair comparison between MRC and EGC with ICE, we adopt the estimation error model in (2.11) in both cases.

2.3 Error probability analysis

To evaluate the performance of (I, J)-QAM, we derive the BER for horizontal PAM first, and then extend the result to both square and rectangular M-QAM.

2.3.1 I-PAM

Consider horizontal I-PAM with $d(i) = aA_I$, where a is given by $a = \sqrt{E_a}$. For $I = 4$, the two bits b_1b_2 are gray-coded and mapped to four possible transmitted symbols, as shown in Fig. 2.4. Let $\tilde{D}_R = \mathrm{Re}(\tilde{D})$ be the real part of $\tilde{D}$ defined in (2.16).

For horizontal PAM, the BER conditioned on data $d(i) = aA_I$ and the decision boundary $a \cdot B_n$ may be defined as the probability that $\tilde{D}_R$ and $d(i)$ fall on the different sides of the

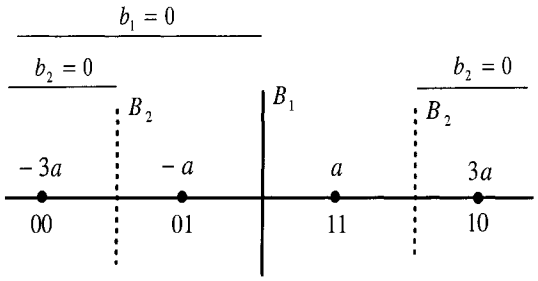

Figure 2.4 Decision boundaries (B_1, B_2) and bit-symbol mapping ($b_1 b_2$) for 4-PAM.

decision boundary. Thus, we define the conditional BERs as

$$\tilde{P}_H(B_n|A_n) = \begin{cases} P\{\tilde{D}_R < B_n \cdot a | d(i) = aA_n\} & A_n \geq B_n \\ P\{\tilde{D}_R > B_n \cdot a | d(i) = aA_n\} & A_n < B_n \end{cases} \tag{2.17}$$

where $B_n \cdot a$ is a decision boundary, and $P\{A|B\}$ denotes the conditional probability of event A given event B. The error probabilities for b_1 and b_2 are, respectively, given by

$$\begin{aligned} P_{4,H}(1) &= \frac{1}{2}[\tilde{P}_H(0|3) + \tilde{P}_H(0|1)] \\ P_{4,H}(2) &= \frac{1}{2}[\tilde{P}_H(2|3) - \tilde{P}_H(-2|3) + \tilde{P}_H(-2|1) + \tilde{P}_H(2|1)] \end{aligned}$$

To summarize, we can express the average BER for horizontal 4-PAM as

$$\bar{P}_{4,H} = \frac{1}{4} \sum_{n=1}^{N_4} w_n \tilde{P}_H(B_n|A_n) \tag{2.18}$$

where $N_4 = 6$, and the coefficients w_n, B_n, and A_n are given in Table 2.1 (a). Note that Table 2.1 given in this section is different from that defined in [73], due to the differences in the definition of the conditional BERs in (2.17). Similarly, we derive the average BER for 8-PAM as

$$\bar{P}_{8,H} = \frac{1}{12} \sum_{n=1}^{N_8} w_n \tilde{P}_H(B_n|A_n) \tag{2.19}$$

where $N_8 = 28$, and w_n, B_n, and A_n are given in Table 2.1 (b). The results for 16, 32-PAM can be obtained using a similar procedure, but are omitted here due to space limitations.

Table 2.1 Coefficients for BER calculation for 4-PAM and 8-PAM.

(a) 4-PAM

n	1	2	3	4	5	6
B_n	2	-2	-2	2	0	0
A_n	3	3	1	1	3	1
w_n	1	-1	1	1	1	1

(b) 8-PAM

n	B_n	A_n	w_n	n	B_n	A_n	w_n
1	0	1	1	15	-2	7	-1
2	0	3	1	16	-6	7	-1
3	0	5	1	17	6	5	1
4	0	7	1	18	2	5	1
5	4	5	1	19	-2	5	-1
6	-4	5	-1	20	-6	5	1
7	4	7	1	21	6	3	1
8	-4	7	-1	22	2	3	1
9	-4	3	1	23	-2	3	-1
10	-4	1	1	24	-6	3	1
11	4	3	1	25	2	1	1
12	4	1	1	26	6	1	-1
13	6	7	1	27	-2	1	1
14	2	7	-1	28	-6	1	-1

2.3.2 M-QAM

For rectangular (I,J)-QAM (with $M = I \times J$), conditioned on the transmitted symbol $d(i) = a(A_I + jA_J)$, we define the conditional BER for the horizontal components of QAM as

$$\tilde{P}_H(B_n|A_I, A_J) = \begin{cases} P\{\tilde{D}_R < B_n \cdot a|d\} & A_I \geq B_n \\ P\{\tilde{D}_R > B_n \cdot a|d\} & A_I < B_n \end{cases} \tag{2.20}$$

Note that for gray-coded (I,J)-QAM, the decision boundaries for horizontal signals are independent of the vertical signal levels, and vice versa [27]. Thus, we can extend the BER result for PAM to the case of M-QAM, and obtain

$$\bar{P}_{I,H} = \frac{2}{IJ\log_2 I} \sum_{n=1}^{J} \sum_{n=1}^{N_I} w_n \tilde{P}_H(B_n|A_n, J - 2n + 1),$$

$$\bar{P}_{J,V} = \frac{2}{IJ\log_2 J}\sum_{n=1}^{I}\sum_{n=1}^{N_J} w_n \tilde{P}_H(B_n|A_n, I-2n+1),$$

where $N_4 = 6$ and $N_8 = 28$, and the coefficients w_n, B_n, and A_n are given in Table 2.1 (a), (b) for $I, J = 4$ and 8, respectively. Here, $\bar{P}_{J,V}$ (besides $\bar{P}_{I,H}$) is also expressed in terms of function $\tilde{P}_H$ since we assumed the channel gains and the estimates are circularly-symmetric processes. The average BER (valid for both square and rectangular QAM) is given by

$$\bar{P}_M = \frac{1}{\log_2(I \cdot J)}[(\log_2 I)\bar{P}_{I,H} + (\log_2 J)\bar{P}_{J,V}]. \tag{2.21}$$

We underscore that for non-gray-coded bit-mapping, our result is applicable with slight modifications. Next, we derive the conditional BER $\tilde{P}_H(B_n|A_I, A_J)$ for M-QAM by evaluating the signal-constellation-dependent effective SNRs with ICE. $\tilde{P}_H(B_n|A_n)$ for PAM can be obtained as a special case.

2.4 Conditional error rate

In this section, we evaluate the distribution of MRC/EGC M-QAM effective output SNRs with ICE, which allows us to obtain the conditional BER $\tilde{P}_H(B_n|A_I, A_J)$.

2.4.1 MRC effective SNRs with ICE

Given that[1] $d = (A_I + jA_J)a$, $y_l = c_l d + n_l$, and $c_l = \rho\frac{\sigma_c}{\sigma_{\hat{c}}}\hat{c}_l + z_l$, for $l = 1, \ldots, N$, we obtain the real part of the decision variable for MRC M-QAM as

$$\begin{aligned}
\tilde{D}_R &= \frac{\sum_{l=1}^{L}\mathrm{Re}\,(\hat{c}_l^* y_l)}{\sum_{l=1}^{L}|\hat{c}_l|^2} = \frac{\sum_{l=1}^{L}\mathrm{Re}\,(\hat{c}_l^*(c_l d + n_l))}{\sum_{l=1}^{L}|\hat{c}_l|^2} \\
&= \frac{\sum_{l=1}^{L}\mathrm{Re}\left(\hat{c}_l^*((\rho\frac{\sigma_c}{\sigma_{\hat{c}}}\hat{c}_l + z_l)d + n_l)\right)}{\sum_{l=1}^{L}|\hat{c}_l|^2} = \frac{\sum_{l=1}^{L}\mathrm{Re}\left(\rho d\frac{\sigma_c}{\sigma_{\hat{c}}}|\hat{c}_l|^2 + \hat{c}_l^*(z_l d + n_l)\right)}{\sum_{l=1}^{L}|\hat{c}_l|^2} \\
&= \frac{\sigma_c}{\sigma_{\hat{c}}}\mathrm{Re}(\rho d) + \frac{\sum_{l=1}^{L}\mathrm{Re}\,(\hat{c}_l^*(z_l d + n_l))}{\sum_{l=1}^{L}|\hat{c}_l|^2} = \frac{\sigma_c}{\sigma_{\hat{c}}}\mathrm{Re}(\rho d) + \tilde{n}_R
\end{aligned} \tag{2.22}$$

where $\tilde{n}_R$ is the MRC output noise given by

$$\tilde{n}_R = \frac{\mathrm{Re}\left(\sum_{l=1}^{L}\hat{c}_l^*(z_l d + n_l)\right)}{\sum_{l=1}^{L}|\hat{c}_l|^2}.$$

[1]Below, we drop the symbol index i of $y_l(i)$, $d(i)$, and $\hat{c}_l(i)$ when no confusion arises.

The conditional BER in (2.20) may be rewritten as

$$\tilde{P}_H(B_n|A_I, A_J) = P\left\{\tilde{n}_R < -\left|B_n a - \frac{\sigma_c}{\sigma_{\hat{c}}}\mathrm{Re}(\rho d)\right|\right\} \tag{2.23}$$

Using (2.22) and (2.23), we can define the effective MRC output SNR with ICE as

$$\gamma^{\mathrm{MRC}}_{\mathrm{ICE}|d,B_n} = \frac{|B_n a - \frac{\sigma_c}{\sigma_{\hat{c}}}\mathrm{Re}(\rho d)|^2}{2\mathrm{var}\{\tilde{n}_R\}} \tag{2.24}$$

Using the property that $\hat{c}_l$, z_l, and n_l are mutually independent, and that $(z_l d + n_l)$ is a zero-mean circularly symmetric Gaussian noise with variance $(1-|\rho|^2)\sigma_c^2|d|^2 + N_0$, we obtain that conditioned on $\hat{c}_l$,

$$\mathrm{var}\{\mathrm{Re}(\sum_{l=1}^{L}\hat{c}_l^*(z_l d + n_l))\} = \frac{1}{2}[(1-|\rho|^2)\sigma_c^2|d|^2 + N_0](\sum_{l=1}^{L}|\hat{c}_l|^2),$$

Thus,

$$\mathrm{var}\{\tilde{n}_R\} = \frac{1}{2}[(1-|\rho|^2)\sigma_c^2|d|^2 + N_0]/[\sum_{l=1}^{L}|\hat{c}_l|^2],$$

and (2.24) reduces to

$$\gamma^{\mathrm{MRC}}_{\mathrm{ICE}|d,B_n} = \frac{(B_n a - \frac{\sigma_c}{\sigma_{\hat{c}}}\mathrm{Re}(\rho d))^2 \sum_{l=1}^{L}|\hat{c}_l|^2}{(1-|\rho|^2)\sigma_c^2|d|^2 + N_0}. \tag{2.25}$$

To gain more insight into the effects of different system parameters, we assume that ρ is real below. It follows from (2.25) that

$$\begin{aligned}\gamma^{\mathrm{MRC}}_{\mathrm{ICE}|d,B_n} &= \frac{(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2 E_a \sum_{l=1}^{L}|\hat{c}_l|^2}{(1-|\rho|^2)\sigma_c^2|d|^2 + N_0} \\ &= \frac{(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2 E_a \sum_{l=1}^{L}\hat{\gamma}_l}{(1-|\rho|^2)\bar{\gamma}|d|^2 + 1} = \tilde{\beta}\sum_{l=1}^{L}\hat{\gamma}_l\end{aligned} \tag{2.26}$$

where $\bar{\gamma} = \sigma_c^2/N_0$ is the average SNR (ASNR) per branch, and $\hat{\gamma}_l = |\hat{c}_l|^2/N_0$ is the SNR of the channel gain estimate at the lth branch, and $\tilde{\beta}$ is given by

$$\tilde{\beta} = \frac{(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2 E_a}{(1-|\rho|^2)\bar{\gamma}|d|^2 + 1}. \tag{2.27}$$

It may appear that as $\hat{\gamma}_l = \sigma_{\hat{c}}^2/N_0$ increases the effective SNR given by (2.26) may improve, which is, however, not necessarily true. In (2.26) the factor $(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2$ specifies that the value of $\sigma_{\hat{c}}$ should be chosen such that $\frac{\sigma_c}{\sigma_{\hat{c}}}\rho = 1$, i.e., the decision region is not distorted. When $\sigma_{\hat{c}}$ (or $\hat{\gamma}_l$) increases or decreases such that $\frac{\sigma_c}{\sigma_{\hat{c}}}\rho \neq 1$, $(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2$ will decrease for some values of A_I and B_n and the average BER will become worse.

Using (2.25) or (2.26), the MGF of $\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}$ over Nakagami-m channels can be obtained as

$$\Phi_{\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}}(s) = E[e^{s\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}}] = \prod_{l=1}^{L} \Phi_{\hat{\gamma}_l}(\tilde{\beta}s) \simeq (1 - s\tilde{\beta}\bar{\hat{\gamma}}/\hat{m})^{-\hat{m}L} \tag{2.28}$$

where $\hat{m}$ is the m-parameter for $\hat{c}_l$ given by (A.7), and $\bar{\hat{\gamma}} = E[\hat{\gamma}] = E[|\hat{c}_l|^2]/N_0$. Note that (2.28) is exact for Rayleigh channels (with $\hat{m} = 1$).

Using the effective SNR obtained in (2.28) we have

$$\begin{aligned} \tilde{P}_H(B_n|A_I, A_J) &= \int_0^{\infty} Q(\sqrt{2\gamma}) f_{\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}}(\gamma) d\gamma \\ &= \frac{1}{\pi} \int_0^{\pi/2} \Phi_{\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}}\left(-\frac{1}{\sin^2\theta}\right) d\theta. \end{aligned} \tag{2.29}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(-y^2/2) dy$ is the Gaussian-Q function and $f_{\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}}(x)$ is the PDF of $\gamma^{\text{MRC}}_{\text{ICE}|d,B_n}$. With the MGF of the effective MRC output SNR being derived in (2.28), the conditional BERs given in (2.29) can be accurately evaluated, e.g., by employing a Gauss-Chebyshev quadrature (GCQ) formula [1].

2.4.2 EGC effective SNRs with ICE

For EGC M-QAM with ICE based on the decision variable given in (2.16) we obtain

$$\begin{aligned} \tilde{D}_R &= \frac{\sum_{l=1}^{L} \text{Re}\,(y_l \hat{c}_l^*/|\hat{c}_l|)}{\sum_{l=1}^{L} |\hat{c}_l|} = \frac{\sum_{l=1}^{L} \text{Re}\left(\hat{c}_l^*((\rho\frac{\sigma_c}{\sigma_{\hat{c}}}\hat{c}_l + z_l)d + n_l)/|\hat{c}_l|\right)}{\sum_{l=1}^{L} |\hat{c}_l|} \\ &= \frac{\sum_{l=1}^{L} \text{Re}\left(|\hat{c}_l|\rho d\frac{\sigma_c}{\sigma_{\hat{c}}} + \hat{c}_l^*(z_l d + n_l)/|\hat{c}_l|\right)}{\sum_{l=1}^{L} |\hat{c}_l|} = \frac{\sigma_c}{\sigma_{\hat{c}}}\text{Re}(\rho d) + \tilde{n}_R \end{aligned} \tag{2.30}$$

where $\tilde{n}_R$ is the EGC output noise given by

$$\tilde{n}_R = \frac{\text{Re}\left(\sum_{l=1}^{L} \hat{c}_l^*(z_l d + n_l)/|\hat{c}_l|\right)}{\sum_{l=1}^{L} |\hat{c}_l|}$$

Using a procedure similar to that for the case of MRC, we define the EGC QAM effective output SNR with ICE as

$$\gamma^{\text{EGC}}_{\text{ICE}|d,B_n} = \frac{|B_n a - \frac{\sigma_c}{\sigma_{\hat{c}}}\text{Re}(\rho d)|^2}{2\text{var}\{\tilde{n}_R\}}. \tag{2.31}$$

Conditioned on $\hat{c}_l$, we obtain that

$$\text{var}\{\tilde{n}_R\} = \frac{L}{2}[(1 - |\rho|^2)\sigma_c^2|d|^2 + N_0]/[\sum_{l=1}^{L} |\hat{c}_l|]^2$$

The effective EGC SNR in (2.31) is then given by

$$\gamma^{\text{EGC}}_{\text{ICE}|d,B_n} = \frac{|B_n a - \frac{\sigma_c}{\sigma_{\hat{c}}}\text{Re}(\rho d)|^2 \frac{1}{L}[\sum_{l=1}^{L}|\hat{c}_l|]^2}{(1-|\rho|^2)\sigma_c^2|d|^2 + N_0} \tag{2.32}$$

Assume that ρ is real below. We obtain

$$\gamma^{\text{EGC}}_{\text{ICE}|d,B_n} = \frac{(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2 E_a(\sum_{l=1}^{L}|\hat{c}_l|)^2/L}{(1-|\rho|^2)\sigma_c^2|d|^2 + N_0} = \tilde{\beta}\left[\sum_{l=1}^{L}\sqrt{\hat{\gamma}_l}\right]^2/L \tag{2.33}$$

where $\tilde{\beta}$ is given by (2.27). Equation (2.33) verifies that our receiver structure in Fig. 2.3 indeed attains the EGC diversity performance for both the perfect CSI (by setting $\rho = 1$ and $\frac{\sigma_c}{\sigma_{\hat{c}}}\rho = 1$) and the ICE cases. A comparison between (2.26) and (2.33) also reveals that the SNR losses (shown by the factor $\tilde{\beta}$) for both MRC and EGC formats caused by ICE are the same. Finally, from (2.26) and (2.33) we observe that the effects of ICE are manifested in two aspects: a decision space distortion (shown by the factor $\frac{\sigma_c}{\sigma_{\hat{c}}}\rho$ in $(B_n - \frac{\sigma_c}{\sigma_{\hat{c}}}\rho A_I)^2$) and a scaling of the effective SNR (shown by the denominator $[(1-|\rho|^2)\bar{\gamma}|d|^2+1]$). For MMSE-CE, $\frac{\sigma_c}{\sigma_{\hat{c}}}\rho = 1$ and so only the SNR scaling occurs.

Using the effective SNR given by (2.31) we have

$$\tilde{P}_H(B_n|A_I, A_J) = \frac{1}{\pi}\int_0^{\pi/2} \Phi_{\gamma^{\text{EGC}}_{\text{ICE}|d,B_n}}\left(-\frac{1}{\sin^2\theta}\right) d\theta, \tag{2.34}$$

where is the MGF of $\gamma^{\text{EGC}}_{\text{ICE}|d,B_n}$.

By approximating several i.i.d. Nakagami random variables with one Nakagami random variable [7, eqs. (41)-(45)], we obtain an accurate approximation for $\Phi_{\gamma^{\text{EGC}}_{\text{ICE}|d,B_n}}(s)$ in Nakagami-m fading channels as

$$\Phi_{\gamma^{\text{EGC}}_{\text{ICE}|d,B_n}}(s) \simeq \left(1 - \frac{s\bar{\gamma}_{\text{eq}|d,B_n}}{\hat{m}}\right)^{-\hat{m}L} \tag{2.35}$$

where $\bar{\gamma}_{\text{eq}|d,B_n}$ is given by $\bar{\gamma}_{\text{eq}|d,B_n} = \tilde{\beta}\bar{\hat{\gamma}}[1 + (L-1)\Gamma^2(\hat{m}+0.5)/\Gamma^2(\hat{m})]/L$, and $\Gamma(x)$ is the Gamma function [1]. To avoid the approximation involved in (2.35) for the evaluation of the MGF of the EGC effective SNR, we may obtain accurate BER results by deriving CHF for the EGC output amplitude with ICE, as given in next section.

2.4.3 EGC amplitude CHF with ICE

Let us define the EGC output amplitude with ICE (conditioned on d and B_n) as

$$\hat{x}^{\text{EGC}}_{\text{ICE}|d,B_n} = \sqrt{\gamma^{\text{EGC}}_{\text{ICE}|d,B_n}} = \sum_{l=1}^{L}\sqrt{\hat{\gamma}_l\tilde{\beta}/L}$$

The CHF of $\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}$ is defined as $\Phi_{\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}}(\omega) = \int_0^\infty f_{\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}}(x)e^{j\omega x}dx$. It follows that

$$\Phi_{\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}}(\omega) = \prod_{l=1}^{L} \Phi_{\sqrt{\tilde{\gamma}_l}}\left(\omega\sqrt{\tilde{\beta}/L}\right) = \left[\Phi_{\sqrt{\tilde{\gamma}}}\left(\omega\sqrt{\tilde{\beta}/L}\right)\right]^L,$$

where $\Phi_{\sqrt{\tilde{\gamma}}}(\omega) = \int_0^\infty f_{\sqrt{\tilde{\gamma}_l}}(x)e^{j\omega x}dx$ is the CHF of the amplitude $\sqrt{\tilde{\gamma}_l}$, for $l = 1, \ldots, L$.
$\gamma^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n} = \frac{|B_n a - \frac{\sigma_e}{\sigma_{\hat{e}}}\mathrm{Re}(\rho d)|^2}{2\mathrm{var}\{\tilde{n}_R\}} = \frac{(B_n - \frac{\sigma_e}{\sigma_{\hat{e}}}\rho A_I)^2 E_a \frac{1}{L}[\sum_{l=1}^{L}|\hat{c}_l|]^2}{(1-|\rho|^2)\sigma_e^2|d|^2 + N_0}$

The explicit expression of $\Phi_{\sqrt{\tilde{\gamma}_l}}(\omega)$ for Rayleigh and Nakagami-m channels may be obtained following a CHF result in [12, 16]. In detail, for Rayleigh fading

$$\Phi_{\sqrt{\tilde{\gamma}}}(\omega) = \phi(1; \frac{1}{2}; -\frac{\bar{\tilde{\gamma}}\omega^2}{4}) + j\omega\sqrt{\frac{\pi\bar{\tilde{\gamma}}}{4}}\exp(-\frac{\bar{\tilde{\gamma}}\omega^2}{4}),$$

where $\bar{\tilde{\gamma}}_l = E[\tilde{\gamma}_l]$ and $\phi(a, b; x) = \sum_{n=0}^{\infty}\frac{(a)_n}{(b)_n}\frac{x^n}{n!}$ is the confluent hypergeometric function of the first kind, and $(a)_k = a(a+1)\cdots(a+k-1)$ is the Pochhammer symbol [1]. For Nakagami-m fading channels,

$$\Phi_{\sqrt{\tilde{\gamma}}}(\omega) = \phi(\hat{m}; \frac{1}{2}; -\frac{\bar{\tilde{\gamma}}\omega^2}{4\hat{m}}) + j\omega\frac{\Gamma(\hat{m}+0.5)}{\Gamma(\hat{m})}\sqrt{\frac{\bar{\tilde{\gamma}}}{\hat{m}}}\phi(\hat{m}+\frac{1}{2}; \frac{3}{2}; -\frac{\bar{\tilde{\gamma}}\omega^2}{4\hat{m}}), \tag{2.36}$$

Applying Parseval's theorem, we obtain the conditional BER in Nakagami fading channels as

$$\begin{aligned}\tilde{P}_H(B_n|A_I, A_J) &= \int_0^\infty Q(\sqrt{2}x)f_{\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}}(x)dx = \frac{1}{\pi}\int_0^\infty \frac{\Phi_Q(\omega/\sqrt{2})}{\sqrt{2}}\Phi^*_{\hat{x}^{\mathrm{EGC}}_{\mathrm{ICE}|d,B_n}}(\omega)d\omega \\ &= \frac{1}{\pi}\int_0^\infty \frac{\Phi_Q(\omega/\sqrt{2})}{\sqrt{2}}\left[\Phi^*_{\sqrt{\tilde{\gamma}}}\left(\omega\sqrt{\tilde{\beta}/L}\right)\right]^L d\omega \end{aligned} \tag{2.37}$$

where $\Phi_Q(\omega) = \mathrm{FT}[Q(x)]$ is the Fourier transform of $Q(x)$, and $\Phi_{\sqrt{\tilde{\gamma}}}(\omega)$ is given by (2.36). Using [11, eq. (42)] and the equality that $Q(x) = \mathrm{erfc}(\sqrt{2}x)/2$, where $\mathrm{erfc}(x)$ is the complementary error function, we obtain

$$\Phi_Q(\omega) = \frac{1}{2\omega}\left[\frac{2}{\sqrt{\pi}}F\left(\frac{\omega}{\sqrt{2}}\right) + j\left(1 - \exp(-\frac{\omega^2}{2})\right)\right], \tag{2.38}$$

where $F(\omega) = \omega\phi(1, \frac{3}{2}; -\omega^2)$ is Dawson's integral.

We note that the effects of signal points, boundaries, and system parameters, are included in $\tilde{\beta}$, and thus the term $\Phi^*_{\sqrt{\tilde{\gamma}}}\left(\omega\sqrt{\tilde{\beta}/L}\right)$ in (2.37) has to be evaluated for many different values of $\tilde{\beta}$ to yield the average BER. Since it is much more time-consuming to calculate the CHF $\Phi_{\sqrt{\tilde{\gamma}}}(\omega)$ than $\Phi_Q(\omega)$, to improve the numerical efficiency, we derive the following expression equivalent to (2.37) as

$$\tilde{P}_H(B_n|A_I, A_J) = \frac{1}{\pi}\int_0^\infty \Phi_Q\left(\omega/\sqrt{2\tilde{\beta}}\right)/\sqrt{2\tilde{\beta}}\left[\Phi^*_{\sqrt{\tilde{\gamma}}}\left(\omega/\sqrt{L}\right)\right]^L d\omega \tag{2.39}$$

In (2.39), for a given ASNR, $\Phi_{\sqrt{\tilde{\gamma}}}(\omega)$ can be pre-computed for several sampled values of ω, and only $\Phi_Q\left(\omega/\sqrt{2\tilde{\beta}}\right)$ should be calculated frequently for the BER evaluation. Consequently, (2.39) is much faster to evaluate than (2.37). A GCQ formula for (2.39) can be obtained as

$$\tilde{P}_H(B_n|A_I,A_J) = \frac{1}{2N_s}\sum_{n=1}^{N_s}\left[\Phi^*_{\sqrt{\tilde{\gamma}}}\left(\tan(\theta_n)/\sqrt{L}\right)\right]^L \frac{\Phi_Q\left(\tan(\theta_n)/\sqrt{2\tilde{\beta}}\right)}{\sqrt{2\tilde{\beta}}\cos^2(\theta_n)} \tag{2.40}$$

where $\theta_n = \frac{\pi(n-0.5)}{2N_s}$, for $n = 1,\ldots,N_s$ and N_s is GCQ order.

2.4.4 Discussion

The MMSE channel estimator is a particular important channel estimator [20, 71, 93]. Below, we focus on MMSE channel estimator and provide some results which put new insight into the SNR loss caused by ICE.

2.4.4.1 Comparison of the ICE and Perfect-CSI cases

For MMSE channel estimator, $\rho\sigma_c/\sigma_{\hat{c}} = 1$ and $E[|\hat{c}_l|^2] = |\rho|^2E[|c_l|^2]$ are valid. Below, we explicitly show the relation between the input and output SNRs for diversity M-QAM (valid for both MRC and EGC) with ICE in Nakagami channels for the MMSE channel estimator case.

Equations (2.26) and (2.33) reduce to

$$\gamma^{\text{MRC}}_{\text{ICE}|d,B_n} = \tilde{\beta}\sum_{l=1}^{L}\hat{\gamma}_l \simeq \beta\sum_{l=1}^{L}\gamma_l \tag{2.41}$$

$$\gamma^{\text{EGC}}_{\text{ICE}|d,B_n} = \tilde{\beta}(\sum_{l=1}^{L}\sqrt{\hat{\gamma}_l})^2/L \simeq \beta(\sum_{l=1}^{L}\sqrt{\gamma_l})^2/L \tag{2.42}$$

where $\beta = \frac{|\rho|^2(B_n-A_I)^2E_a}{(1-|\rho|^2)|d|^2\tilde{\gamma}+1}$ is the SNR scaling factor accounting for the M-QAM signal-boundary distance, ICE, and the input SNR. The approximation involved in (2.41) and (2.42) is due to the assumption that $\hat{c}_l$ and c_l have the same distributions ($\hat{m} \simeq m$), which is a tight approximation for medium-to-high ASNRs over Nakagami channels. For Rayleigh fading channels, (2.41) and (2.42) are exact. We underscore that the purpose of deriving (2.41) and (2.42) is to explicitly show the ASNR loss caused by the channel estimation errors. For a more accurate BER evaluation, the results given in Sections 2.4.1 – 2.4.3 should be used instead.

For the perfect CSI case, the constellation-dependent SNR for the conditional BER at the lth branch is given by

$$\gamma_{\mathrm{CSI},l|d,B_n} = (B_n - A_I)^2 E_a \gamma_l \tag{2.43}$$

Employing the results for the MRC (or EGC) output SNR in the perfect CSI case, we obtain the MGFs of (2.41) and (2.42) as (we use "Div" here to denote "MRC/EGC")

$$\gamma^{\mathrm{Div}}_{\mathrm{ICE}|d,B_n} = \frac{|\rho|^2}{(1-|\rho|^2)|d|^2\bar{\gamma}+1}\gamma^{\mathrm{Div}}_{\mathrm{CSI}|d,B_n} = \frac{\gamma^{\mathrm{Div}}_{\mathrm{CSI}|d,B_n}}{\kappa(\gamma)}$$

where

$$\kappa(\bar{\gamma}) = \frac{(1-|\rho|^2)|d|^2\bar{\gamma}+1}{|\rho|^2}. \tag{2.44}$$

$\kappa(\bar{\gamma})$ may be regarded as the ASNR loss per branch caused by ICE at the MRC/EGC output compared to the perfect CSI case. Some comments are in order. For the signal constellation point with a larger modulus (i.e., larger $|d|^2$), the loss is larger. Only when $\rho = 1$ and $\sigma_c = \sigma_{\hat{c}}$, there is no performance loss. Second, when the ASNR $\bar{\gamma}$ increases, ρ may also increase (depending on the channel estimator employed), and the factor $(1-|\rho|^2)\bar{\gamma}$ will determine whether or not there is an error floor. If $(1-|\rho|^2)\bar{\gamma}$ is not upper-bounded for $\bar{\gamma} \to \infty$, a detection error floor occurs. Finally, a smaller $1-|\rho|^2$ means higher channel estimation accuracy and less SNR loss.

2.4.4.2 Some closed-form expressions

Following a derivation procedure in [12], we present some closed form expressions for the BER formulas for EGC M-QAM with ICE in Nakagami-m fading channels. Using the Gil-Pelaez inversion theorem, the conditional error probability $\tilde{P}_H(B_n|A_I, A_J)$ can be reformulated as

$$\begin{aligned}\tilde{P}_H(B_i|A_I,A_J) &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty t^{-1}e^{-t^2/2}\mathrm{Im}\{\Phi_{x_{\mathrm{EGC}}}(\sqrt{2\tilde{\beta}}t)\}dt \\ &= \frac{1}{2} - \frac{1}{2\pi}\int_0^\infty t^{-1}e^{-t}\mathrm{Im}\{\Phi_{x_{\mathrm{EGC}}}(2\sqrt{\tilde{\beta}}t)\}dt.\end{aligned} \tag{2.45}$$

1) $L = 2$: By utilizing the identity in [47, eq.(9.220.2) and (7.622.3)], the error probability is given by

$$\begin{aligned}\tilde{P}_H(B_i|A_I,A_J) &= \frac{1}{2} - \frac{\Gamma(\hat{m}+0.5)}{\Gamma(\hat{m})}\sqrt{\frac{2\tilde{\beta}\bar{\gamma}}{\pi\hat{m}(1+2a)}} \\ &\quad \times F_2\left(\frac{1}{2}, \frac{1}{2}-\hat{m}, 1-\hat{m}, \frac{1}{2}, \frac{3}{2}; \frac{a}{1+2a}, \frac{a}{1+2a}\right),\end{aligned} \tag{2.46}$$

where $\hat{m}$ is the Nakagami-parameter for $\hat{c}_l$, $a = \frac{\tilde{\beta}\bar{\tilde{\gamma}}}{2\hat{m}}$ and $F_2(\cdot,\cdot,\cdot,\cdot,\cdot;x,y)$ is the Appel hypergeometric series of the second kind [47, eq.(9.180.2)].

When $\hat{m} = 1$ (i.e., Rayleigh fading), equation (2.46) can be further simplified to

$$\tilde{P}_H(B_i|A_I, A_J) = \frac{1}{2} - \frac{\sqrt{\bar{\tilde{\gamma}}\tilde{\beta}(\bar{\tilde{\gamma}}\tilde{\beta}+2)}}{2(\bar{\tilde{\gamma}}\tilde{\beta}+1)}. \tag{2.47}$$

We have verified that for the perfect CSI case (2.46) and (2.47) reduce to [12, eq. (29)] and [12, eq. (32)] (when assuming i.i.d. branches), respectively.

2) $L = 3$: Similarly, the error rate can be expressed in terms of the Appel hypergeometric function [47, eq.(9.19)]

$$\tilde{P}_H(B_i|A_I, A_J) = \frac{1}{2} - \frac{c_1}{2\pi} I_1 + \frac{c_2}{2\pi} I_2, \tag{2.48}$$

where $c_1 = \frac{\Gamma(\hat{m}+0.5)}{\Gamma(\hat{m})} 6\sqrt{\hat{a}}$ and $c_2 = \left(\frac{\Gamma(\hat{m}+0.5)}{\Gamma(\hat{m})} 2\sqrt{\hat{a}}\right)^3$, with $\hat{a} = \frac{\tilde{\beta}\bar{\tilde{\gamma}}}{3\hat{m}}$. I_1 and I_2 are given in terms of Appel hypergeometric function of three variables as,

$$\begin{aligned} I_1 &= \sqrt{\frac{\pi}{(1+3\hat{a})}} F_A\left(\frac{1}{2}; \frac{1}{2}-\hat{m}, \frac{1}{2}-\hat{m}, 1-\hat{m}; \frac{1}{2}, \frac{1}{2}, \frac{3}{2}; \frac{\hat{a}}{1+3\hat{a}}, \frac{\hat{a}}{1+3\hat{a}}, \frac{\hat{a}}{1+3\hat{a}}\right), \\ I_2 &= \frac{\sqrt{\pi}}{2}(1+3\hat{a})^{-\frac{3}{2}} F_A\left(\frac{3}{2}; 1-\hat{m}, 1-\hat{m}, 1-\hat{m}; \frac{3}{2}, \frac{3}{2}, \frac{3}{2}; \frac{\hat{a}}{1+3\hat{a}}, \frac{\hat{a}}{1+3\hat{a}}, \frac{\hat{a}}{1+3\hat{a}}\right). \end{aligned} \tag{2.49}$$

2.5 Numerical examples

In Fig. 2.5 and Fig. 2.6, we present the simulated and calculated BERs for 16-QAM with EGC ($L = 3$) and PSAM, for perfect CSI, MMSE- and Sinc-interpolator-based (with Hamming and rectangular windows, respectively) channel estimators in Rayleigh fading channels. We note that a quasi-analytical (QA) approach, which calculates the BER by averaging the conditional BERs over the amplitudes of the generated fading channel samples, was used in [12] to calculate the BER of EGC M-QAM with perfect CSI. However, the true BER simulation was not given.

By using the EGC structure that we proposed in Fig. 2.3, we can truly simulate the BER performance of the QAM EGC receiver based on the decision variable in (2.16), which is particularly important for verifying our analysis, especially for the ICE case. Simulation results in Figs. 2.5 and 2.6 confirm the validity and accuracy of our BER analysis for Rayleigh channels. For Fig. 2.5 we assumed that $F_1 = F_2 = 4$, $B_f T = 0.02$ and $P = 10$. The MMSE-CE has a performance close to that of the perfect CSI case, and the SNR loss is within 2.5 dB for

all SNRs shown in the figure. The BER of Sinc-CE with Hamming window (Sinc-CE-Hamm) is worse than that of MMSE-CE by about 1 dB, and Sinc-CE with rectangular window (Sinc-CE-Rect) has the worst performance. For Fig. 2.6 we assumed a larger Doppler bandwidth $B_fT = 0.03$ and a larger interpolation interval $P = 15$. For this case, the error floors for EGC with MMSE-CE and Sinc-CEs occur. Contrary to the observation of Fig. 2.5, the BER for Sinc-CE-Rect is now very close to that of MMSE-CE, and is significantly lower than that of the Sinc-CE-Hamm.

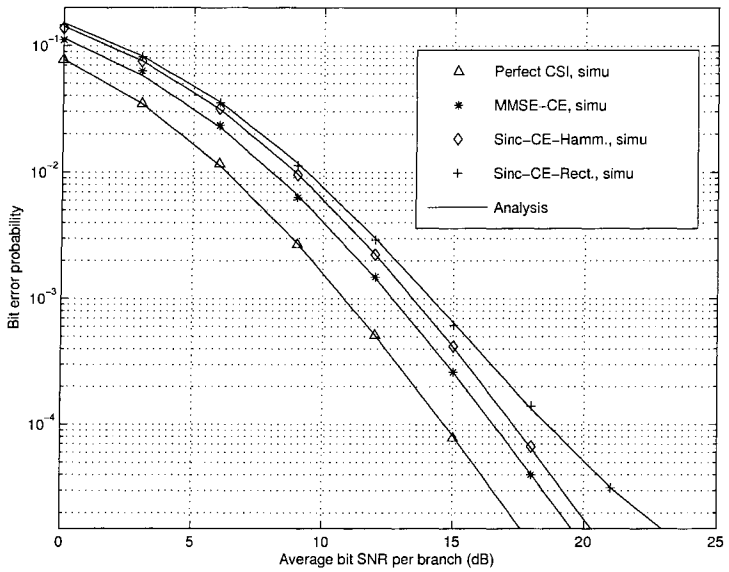

Figure 2.5 BER (averaged over all i_{off}'s) v.s. bit ASNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE- and Sinc-CEs (with Hamming and rectangular windows), in a Rayleigh fading channel. $L = 3$, $F_1 = F_2 = 4$, $P = 10$, and $B_fT = 0.02$.

In Fig. 2.7, we show $1 - |\rho|^2$ vs. the SNR per branch in Rayleigh fading channels (the result is valid for both MRC and EGC and any L), with $P = 15$ and $F = 12$. Note that $1 - |\rho|^2$ is closely related to the SNR loss per branch $\kappa(\bar{\gamma})$, as we discussed in Section 2.4.4. A smaller $1 - |\rho|^2$ relates to a better channel estimation accuracy and a smaller SNR loss. When $B_fT = 0.02$, Sinc-CE-Hamm causes a smaller loss than Sinc-CE-Rect; but for $B_fT = 0.03$, Sinc-CE-Hamm has a much larger loss than Sinc-CE-Rect, and consequently a poorer BER performance. Observation from Fig. 2.7 excellently agrees with those from Figs. 2.5 and 2.6, and suggests that the values of $1 - |\rho|^2$ can reliably reflect the performance of different channel

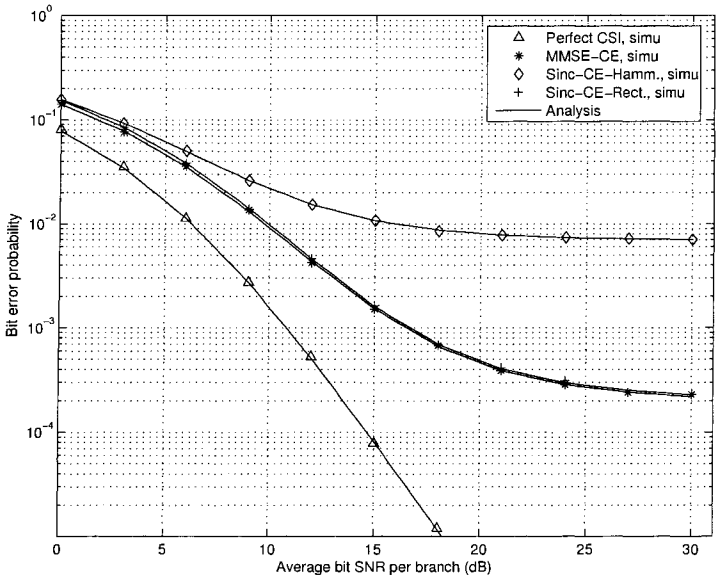

Figure 2.6 BER (averaged over all i_{off}'s) v.s. bit ASNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE-CE and Sinc-CEs(with Hamming and rectangular windows), in a Rayleigh fading channel. $L = 3$, $F_1 = F_2 = 4$, $P = 15$, and $B_f T = 0.03$.

estimators. The result also show that the relative performance of Hamming and rectangular windows crucially depends on the system parameters. We also note that for a large range of PSAM and channel parameters that we tested, Sinc-CE-Hamm achieves a better BER performance than Sinc-CE-Rect, though a case-by-case study is necessary.

In Fig. 2.8, we show the calculated and simulated $\hat{m}$, the m-parameter of the channel estimate $\hat{c}_l$, for MMSE-CE and Sinc-CE-Rect, with $m = [0.7, 2.7]$. The simulation result is obtained by calculating the values of $\hat{m}$ using (A.2) given a large number of samples of $\hat{c}_l$, where the expectations in (A.2) are replaced by the sample means. The result verifies the accuracy of our analysis, and shows that as the SNR increases, $\hat{m}$ approaches m. Also, when $m > 1$, $\hat{m} \leq m$; and when $m < 1$, $\hat{m} \geq m$, which is an interesting observation as expected from (A.7) and (A.8).

In Fig. 2.9 we present the BER for EGC 16-QAM with perfect CSI, MMSE-CE, and Sinc-CEs in slow Nakagami fading channels, with $m = [0.7, 2.7]$. It is observed that when $m = 0.7$ the analysis for all the studied channel estimators is accurate. For $m = 2.7$, the analytical

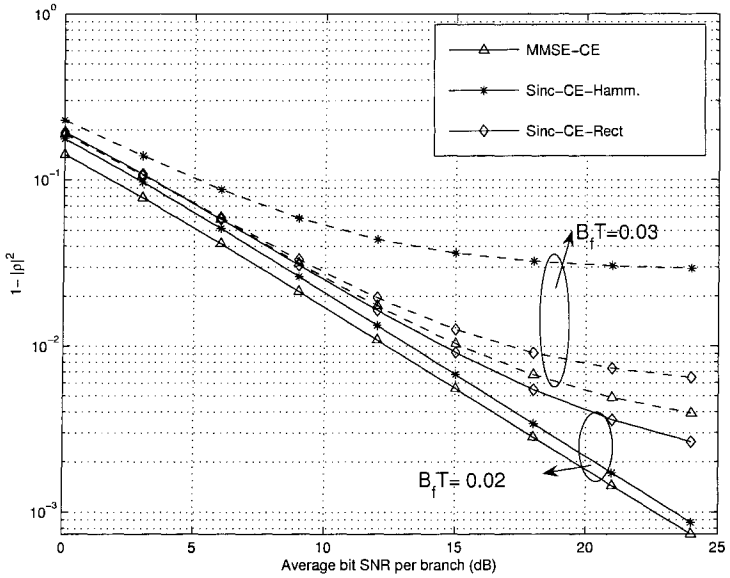

Figure 2.7 $1-|\rho|^2$ ($i_{\text{off}} = 8$) v.s. bit SNR per branch for diversity 16-QAM with MMSE-CE and Sinc-CEs (with Hamming and Rectangular windows), in a Rayleigh fading channel. $P = 15$, $F_1 = F_2 = 6$, $B_fT = [0.02, 0.03]$.

and simulated results for MMSE-CE and Sinc-CE-Hamm match well, but the analytical BER values for Sinc-CE-Rect slightly deviates from the simulated ones for high ASNRs, but may still be regarded as a tight approximation. This result suggests that our analysis is accurate for any linear channel estimators in Nakagami fading for m being close to 1, and for near-MMSE channel estimator(e.g., Sinc-CE-Hamm) when m is substantially larger than one.

In Fig. 2.10 we present the average BERs for the 16-QAM MRC and EGC receivers with perfect CSI and MMSE-CE in slow Nakagami fading channels, with $m = 2.7$ and $L = [1, 2, 3]$. The results verify our analysis for both MRC and EGC M-QAM in Nakagami channels. We observe that increasing the diversity order L improves the BER significantly.

Finally, in Fig. 2.11 we present the BERs of the 16-, 32-, and 64-QAM MRC and EGC receivers in Nakagami channels. The result shows that as M increases, the BER performance becomes worse for both perfect CSI and ICE cases. Results in Fig. 2.10 and Fig. 2.11 reveal that the SNR gap between MRC and EGC in the ICE case (with MMSE-CE) is approximately the same as in the perfect CSI case, which verifies our analysis that given the same system parameters, the SNR losses for MRC and EGC caused by ICE are approximately identical, see

Section 2.4.4.

2.6 Summary

By deriving the MGF/CHF expression for the effective SNRs/amplitudes with ICE for diversity M-QAM, we have developed a novel approach for the accurate BER evaluation of M-QAM MRC/EGC diversity receivers with PSAM in Rayleigh and Nakagami fading channels. We proposed the EGC structure valid for NCM format, and proved its validity by analytical and simulation results. We have derived the m-parameter of the estimated channel gain, and revealed its relation to that of the true channel gain. Our analytical result is valid for linear channel estimators in Rayleigh fading with arbitrary fading spectra, and for near-MMSE channel estimators in slow Nakagami fading channels. We have evaluated the performance of MMSE- and Sinc-interpolator-based channel estimators with PSAM, and presented some insightful findings.

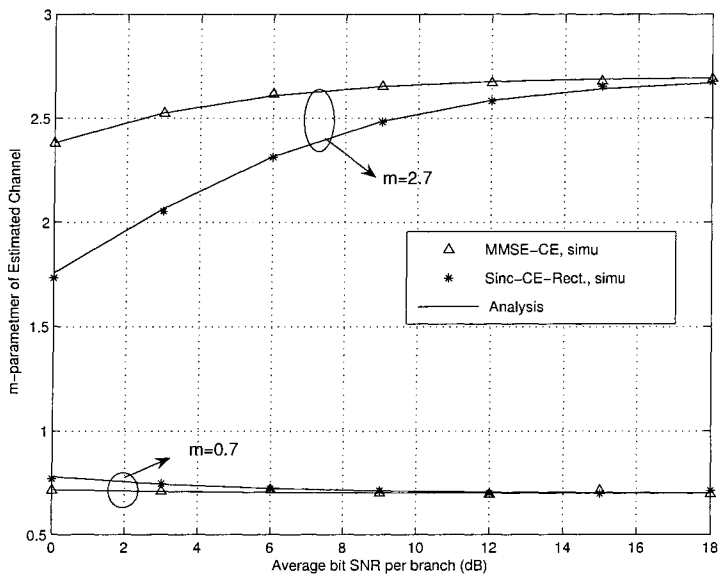

Figure 2.8 $\hat{m}$ vs. bit SNR per branch for 16-QAM with MMSE- and Sinc-CEs (with Hamming and rectangular windows), in a slow Nakagami fading channel. $L = 2$, $P = 10$, $F_1 = F_2 = 3$, $B_f T = 0$.

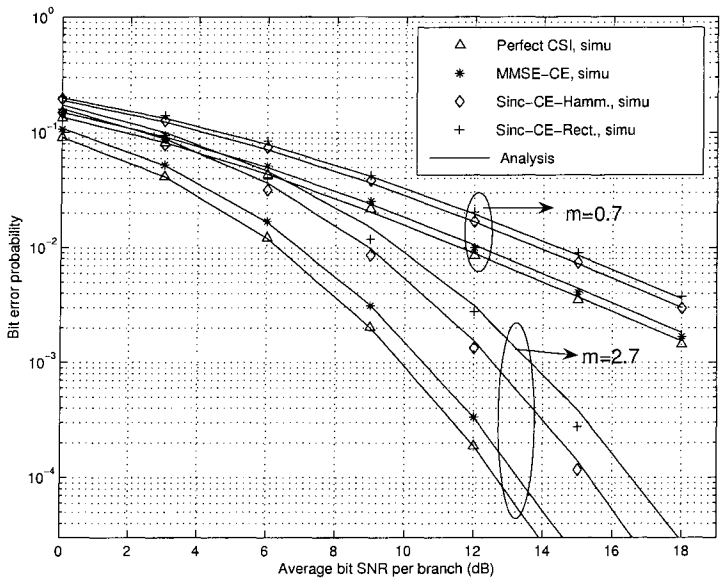

Figure 2.9 Average BER vs. bit SNR per branch for 16-QAM EGC receivers with perfect CSI, MMSE- and Sinc-CEs (with Hamming and rectangular windows) channel estimators, respectively, in a slow Nakagami fading channel. $m = [0.7, 2.7]$, $L = 2$, $P = 15$, $F_1 = F_2 = 3$, $B_f T = 0$.

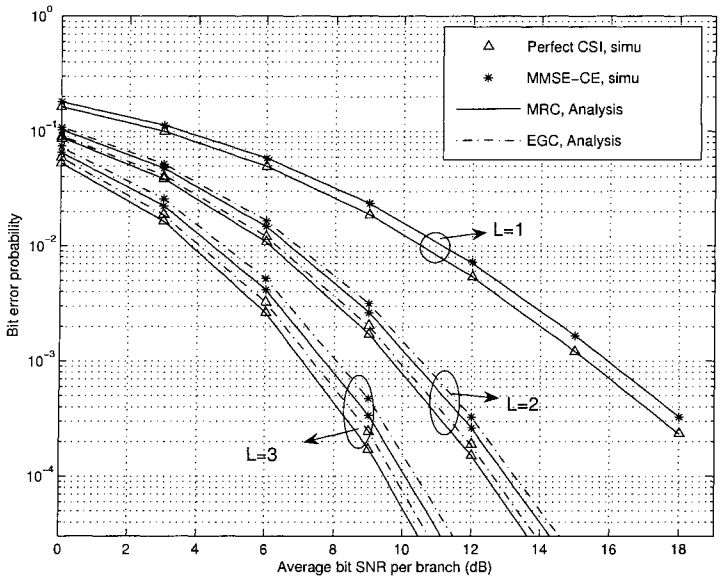

Figure 2.10 Average BER vs. bit SNR per branch for 16-QAM MRC and EGC receivers with perfect CSI and MMSE-CE respectively, in a slow Nakagami fading channel. $m = 2.7$, $L = [1, 2, 3]$, $P = 15$, $F_1 = F_2 = 3$, $B_f T = 0$.

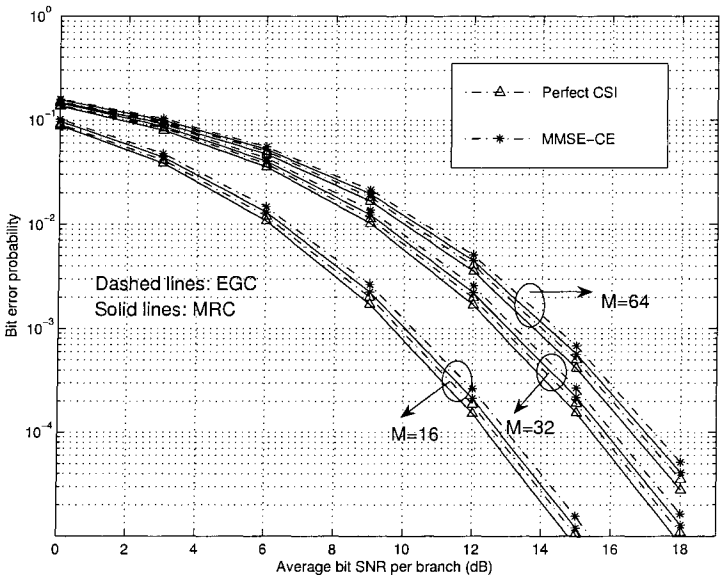

Figure 2.11 Average BER vs. bit SNR per branch for M-QAM (with $M = [16, 32, 64]$) MRC and EGC receivers with perfect CSI and MMSE-CE respectively, in a slow Nakagami fading channel. $m = 2.7$, $L = 2$, $P = 10$, $F_1 = F_2 = 5$, $B_f T = 0$.

CHAPTER 3. HYBRID SELECTION/EQUAL GAIN COMBINING

3.1 Introduction

In wireless communication scenarios with dense multipath or multiple antennas, the hybrid diversity schemes which employ subset combining have gained more attention in past. HS/MRC [6, 10, 70], also known as generalized selection combining (GSC) [104], in which L_c branches with the largest instantaneous SNRs are selected from a total of L branches and combined according to MRC rule, is a practical and useful diversity combining scheme for many wireless applications. In this scheme, both the phases and amplitudes of the selected signal branches are required to be known for MRC subset combining.

As an attractive alternative to HS/MRC, HS/EGC is a new hybrid diversity scheme which combines the signals of the selected branches according to EGC rule. This scheme may avoid the necessity of amplitude estimation for all the branches Considering the trade-off between implementation complexity and performance, HS/MRC bridges the gaps between SC and MRC. Likewise, HS/EGC includes SC and EGC as two special cases.

The post-detection HS/EGC, also known as non-coherent GSC [38, 69], is suitable for differential coherent and non-coherent modulation formats, and such schemes have been studied in [10, 38, 69, 70]. The pre-detection HS/EGC is suitable for coherent modulation formats and has been studied in [54, 83]. However, [54] is valid for M-ary modulations with HS/EGC for $L_c = 2$ and for the Nakagami-m fading channel with integer m parameters, while the result in [83] was confined to M-PSK modulation over i.i.d. Rayleigh fading channels. Analytical error rate and outage probability formulas of HS/EGC over different types of generalized fading channels, versatile fading parameters (for example, non-integer m-parameters and arbitrary L_c and L), is not available yet.

Besides the performance analysis, the HS/EGC receiver structure valid for non-constant modulus modulation format, to our knowledge, is not known in the literature.

In this chapter, we first propose the HS/EGC receiver structure valid for NCM modulation format and the relevant decision variable, and verify its validity by both analysis and simulations. Then, by deriving the MGF of the HS/EGC output amplitude and with the help of the Parseval's theorem, we develop a general approach for unified error rate and outage probability formulas in different types of fading channels (such as Rician, Nakagami-m, Nakagami-q, and Weibull models).

To efficiently evaluate the MGF of the HS/EGC amplitude for different types of fading channels , we provide a numerical quadrature technique to jointly evaluate the HS/EGC MGF and the relevant truncated MGF (TMGF) expressions, which only involves the computation of several single integrals and avoids the complexity of a two-dimensional summation as in convention. Furthermore, for theoretical interests and practical use, we derive some closed-form TMGF expressions for Rayleigh and Nakagami-m fading channels. Our new results are readily applicable to performance analysis for a wide class of modulations with HS/EGC diversity over general fading channels.

3.2 Signal model

We adopt the same channel model as given in (2.1). That is,

$$\mathbf{y}(i) = \mathbf{c}(i)d(i) + \mathbf{n}(i) \tag{3.1}$$

where $d(i)$ is the transmitted data symbol with energy $E_d = E[|d(i)|^2]$, Without loss of generality, we let $E_d = 1$. We assume that the Gaussian noise element $n_l(i)$ has zero mean and variance N_0 for all branches. The instantaneous SNR for the lth branch is $\gamma_l = |c_l(i)|^2/N_0$.

We arrange the elements in $|\mathbf{c}(i)|$ in descending order such that $|c_{(1)}| \geq |c_{(2)}| \geq \cdots \geq |c_{(L)}|$, where $|c_{(l)}|$ is the lth largest channel amplitude modulus. Define $\gamma_{(l)} = |c_{(l)}|^2/N_0$. We can show that the corresponding SNR set also satisfies $\gamma_{(1)} \geq \gamma_{(2)} \geq \cdots \geq \gamma_{(L)}$.

Let $\theta_l(i)$ be the phase of $c_l(i)$, such that $c_l(i) = |c_l(i)|e^{j\theta_l(i)}$. In HS/EGC (L_c, L), the L_c branches with largest signal amplitudes are selected and combined according to EGC rule. For constant modulus modulation format, such as M-PSK, the output decision variable is given by

$$D = \sum_{l=1}^{L_c} e^{-j\theta_{(l)}(i)} y_{(l)}(i) = \sum_{l=1}^{L_c} \left[|c_{(l)}(i)|d(i) + e^{-j\theta_{(l)}(i)} n_{(l)}(i) \right] \tag{3.2}$$

where $y_{(l)}(i) = c_{(l)}(i)d(i) + n_{(l)}(i)$. We can readily show that the SNR of D is given by

$$\gamma_{\text{HSEGC}} = \frac{\left(\sum_{l=1}^{L_c} |c_{(l)}(i)|\right)^2}{L_c N_0} = \frac{\left(\sum_{l=1}^{L_c} \sqrt{\gamma_{(l)}}\right)^2}{L_c}. \tag{3.3}$$

The HS/EGC receiver structure for non-constant modulus modulation format, to our knowledge, is not known in the literature. To fill this gap, we propose the HS/EGC receiver structure valid for NCM modulation format, as shown in Fig. 3.1.

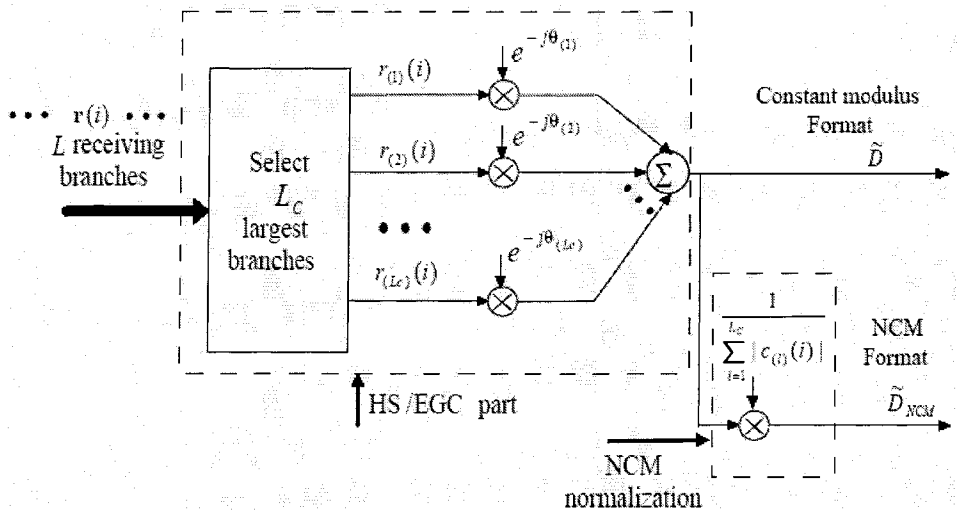

Figure 3.1 Receiver Structure valid for both constant modulus format and non-constant modulus format signals with pre-detection HS/EGC in fading channels.

The output decision variable is given by

$$D = \frac{\sum_{l=1}^{L_c} e^{-j\theta_{(l)}(i)} y_{(l)}(i)}{\sum_{l=1}^{L_c} |c_{(l)}(i)|} = \frac{\sum_{l=1}^{L_c} c^*_{(l)}(i) y_{(l)}(i)/|c_{(l)}(i)|}{\sum_{l=1}^{L_c} |c_{(l)}(i)|} \tag{3.4}$$

We can show that the SNR for D in (3.4) is given by $\gamma_D = (\sum_{l=1}^{L_c} \sqrt{\gamma_{(l)}})^2/L_c$, and the proof follows below.

Proof: Equation (3.4) leads to

$$\begin{aligned} D &= \frac{\sum_{l=1}^{L_c} c^*_{(l)}(c_{(l)}d + n_{(l)})/|c_{(l)}|}{\sum_{l=1}^{L_c} |c_{(l)}|} = \frac{\sum_{l=1}^{L_c} \left[|c_{(l)}|^2 d/|c_{(l)}| + c^*_{(l)} n_{(l)}/|c_{(l)}|\right]}{\sum_{l=1}^{L_c} |c_{(l)}|} \\ &= \frac{\sum_{l=1}^{L_c} \left[|c_{(l)}|d + e^{-j\theta_{(l)}} n_{(l)}\right]}{\sum_{l=1}^{L_c} |c_{(l)}|} = d + \frac{\sum_{l=1}^{L_c} e^{-j\theta_{(l)}} n_{(l)}}{\sum_{l=1}^{L_c} |c_{(l)}|} = d + \tilde{n}^1, \end{aligned} \tag{3.5}$$

where

$$\tilde{n} = [\sum_{l=1}^{L_c} e^{-j\theta_{(l)}} n_{(l)}]/[\sum_{l=1}^{L_c} |c_{(l)}|].$$

The SNR of D can be calculated as

$$\gamma_D = \frac{E_d}{\text{var}\{\tilde{n}\}} = \frac{1}{\text{var}\{\tilde{n}\}} \tag{3.6}$$

Conditioned on $\{c_{(l)}\}_{l=1}^{L_c}$, we obtain

$$\text{var}\{\tilde{n}\} = \frac{\sum_{l=1}^{L_c} \text{var}\{n_{(l)}\}|e^{-j\theta_{(l)}}|^2}{\left[\sum_{l=1}^{L_c} |c_{(l)}|\right]^2} = \frac{L_c N_0}{\left[\sum_{l=1}^{L_c} |c_{(l)}|\right]^2}. \tag{3.7}$$

Combining (3.6) with (3.7), we have $\gamma_D = [\sum_{l=1}^{L_c} |c_{(l)}|]^2/(L_c N_0) = (\sum_{l=1}^{L_c} \sqrt{\gamma_{(l)}})^2/L_c$, which is the same as γ_{HSEGC} given in (3.3). ■

To evaluate the performance of HS/EGC based on (3.4), we need to derive the MGF of HS/EGC output amplitudes over fading channels, as given next.

3.3 MGF of HS/EGC output amplitude

For the HS/EGC (L_c, L) receiver, we define the SNR vector from the L receiving channels as $\boldsymbol{\gamma} = [\gamma_{n_1}, \gamma_{n_2}, \ldots, \gamma_{n_L}]^T$, where γ_{n_k} represents the instantaneous SNR in the n_kth diversity branch; $n_k \in (1, 2, \ldots, L)$ for $1 \leq k \leq L$, and $n_1, n_2, \ldots, n_L$ is a permutation of $(1, 2, \ldots, L)$. We arrange the elements in $\boldsymbol{\gamma}$ in descending order as $\tilde{\boldsymbol{\gamma}} = [\gamma_{(1)}, \gamma_{(2)}, \ldots, \gamma_{(L)}]^T$, so $\tilde{\boldsymbol{\gamma}}$ is an ordered SNR set. Define $\alpha_{(l)} = \sqrt{\gamma_{(l)}}$ for $l = 1, \ldots, L$. Here, α_l is the squared-root of the SNR γ_l, and we call it the channel amplitude for later derivation. It is related to the true channel amplitude modulus $|c_l|$ by $\alpha_l = |c_l|/\sqrt{N_0}$.

If the MGF of γ_{HSEGC} can be derived, the error and outage probabilities of a large class of modulation formats over different types of fading channels can be readily evaluated [6, 70, 94]. However, γ_{HSEGC} contains cross-correlation terms between different branches and thus its MGF is difficult to be evaluated analytically. In this work, we evaluate the distribution of the output amplitude instead, which is given by

$$\alpha_{\text{HSEGC}} = \sqrt{\gamma_{\text{HSEGC}}} = \frac{\sum_{l=1}^{L_c} \alpha_{(l)}}{\sqrt{L_c}}.$$

The MGF for α_{HSEGC} is defined as

$$\Phi_{\alpha_{\text{HSEGC}}}(s) = E\left[\exp(-s\alpha_{\text{HSEGC}})\right] = E\left[\exp\left(\frac{-s\sum_{l=1}^{L_c} \alpha_{(l)}}{\sqrt{L_c}}\right)\right]. \tag{3.8}$$

Once $\Phi_{\alpha_{\text{HSEGC}}}(s)$ is obtained, the error probabilities for different modulation formats may be evaluated with the help of the Parseval's theorem, as will be shown later in Section 3.5.

The procedure to analytically evaluate (3.8) is developed below. Using a result on the joint PDF of the order statistics of HS/MRC given in [63, 70], we may express the joint PDF of $[\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(L_c)}]$ for HS/EGC as

$$\begin{aligned} &f_{\alpha_{(1)},\ldots,\alpha_{(L_c)}}(y_1, y_2, \ldots, y_{L_c}) \\ &= \sum_{\substack{n_1,\ldots,n_{L_c} \\ n_1 \neq n_2 \neq \ldots \neq n_{L_c}}} f_{\alpha_{n_1}}(y_1) f_{\alpha_{n_2}}(y_2) \cdots f_{\alpha_{n_{L_c}}}(y_{L_c}) \prod_{l'=L_c+1}^{L} F_{\alpha_{n_{l'}}}(y_{L_c}), \end{aligned} \tag{3.9}$$

where $y_1, \ldots, y_{L_c}$ are the instantaneous amplitudes satisfying $y_1 \geq y_2 \geq \ldots y_{L_c} \geq 0$, and $n_1, \ldots, n_{L_c}$ $(1 \leq n_1, \ldots, n_{L_c} \leq L)$ are the L_c branches selected with the largest instantaneous amplitudes; $f_{\alpha_{n_l}}(y)$ is the PDF for the output amplitude in the n_lth branch, for $l = 1, \ldots, L$. $F_{\alpha_{n_{l'}}}(y)$ is the cumulative distribution function (CDF) of the amplitude $\alpha_{n_{l'}}$ in the $n_{l'}$th branch. In (3.9), we used $\sum_{\substack{n_1,\ldots,n_{L_c} \\ n_1 \neq n_2 \neq \ldots \neq n_{L_c}}}$ as a short-hand for $\sum_{n_1=1}^{L} \sum_{\substack{n_2 \\ n_2 \neq n_1}}^{L} \sum_{\substack{n_{L_c} \\ n_{L_c} \neq n_1, \ldots n_{L_c-1}}}^{L}$.

To obtain a general and efficient expression for the evaluation of $\Phi_{\alpha_{\text{HSEGC}}}(s)$, it is useful to define TMGF as

$$\Phi_{\alpha_l}(s, x) = \int_x^{\infty} f_{\alpha_l}(y) e^{-sy} dy \tag{3.10}$$

where $f_{\alpha_l}(x)$ is the PDF expression for the instantaneous amplitude α_l. Note that $\Phi_{\alpha_l}(s, 0) = \Phi_{\alpha_l}(s)$, and $\Phi_{\alpha_l}(0, x) = 1 - F_{\alpha_l}(x)$, where $F_{\alpha_l}(x) = \int_0^x f_{\alpha_l}(y) d\, y$ is the CDF of the amplitude α_l.

Using a procedure similar to that for deriving the output SNR for HS/MRC [70], we obtain the MGF of the HS/EGC output amplitude valid for independent non-identically distributed (i.n.d.) channels as

$$\begin{aligned} &\Phi_{\alpha_{\text{HSEGC}}}(s) \\ &= \sum_{n_1,\ldots,n_L \in \mathcal{I}} \int_0^{\infty} e^{-sx} f_{\alpha_{n_{L_c}}}(x) \left[\prod_{l=1}^{L_c-1} \Phi_{\alpha_{n_l}}(s, x)\right] \left[\prod_{l'=L_c+1}^{L} F_{\alpha_{n_{l'}}}(x)\right] dx \end{aligned} \tag{3.11}$$

where $\mathcal{I}$ is the set for all combinations of $\{n_1, \ldots, n_L\}$ in which $\{n_1, \ldots, n_{L_c}\}$ are the indices of selected branches for HS/EGC (L_c, L). Therefore, the summation in (3.11) contains $L_c \begin{pmatrix} L \\ L_c \end{pmatrix}$ terms in total. For i.i.d. branches, (3.11) simplifies to

$$\Phi_{\alpha_{\text{HSEGC}}}(s) = \begin{pmatrix} L \\ L_c \end{pmatrix} L_c \int_0^{\infty} e^{-sx} f_{\alpha}(x) [\Phi_{\alpha}(s, x)]^{L_c-1} F_{\alpha}(x)^{L-L_c} dx. \tag{3.12}$$

To evaluate (3.11) for HS/EGC, the PDF, CDF and TMGF expressions for all the diversity branches are required. For convenience, some PDF and CDF results of channel amplitude α_l over several types of fading channels, such as Rayleigh, Rician, and Nakagami-m, Nakagami-q and Weibull fading models, are listed in Table 3.1. For the Nakagami-q fading channel, we are not aware of a closed-form CDF expression of fading amplitude available in the open literature yet. Thus, we provide a closed-form CDF expression for Nakagami-q fading in Appendix B.1.

Channel model	Formulas for PDF $f_{\alpha_l}(x)$ and CDF $F_{\alpha_l}(x)$
Rayleigh	$f_{\alpha_l}(x) = \frac{2x}{\bar{\gamma}_l}\exp(-x^2/\bar{\gamma}_l)$, $F_{\alpha_l}(x) = 1-\exp(-x^2/\bar{\gamma}_l)$.
Rician	$f_{\alpha_l}(x) = \frac{2x(1+K_l)}{\bar{\gamma}_l}\exp(-K_l-\frac{(1+K_l)x^2}{\bar{\gamma}_l})\cdot I_0\left(2x\sqrt{\frac{(1+K_l)K_l}{\bar{\gamma}_l}}\right)$ $F_{\alpha_l}(x) = 1-Q_1(\sqrt{2K_l},\sqrt{2(K_l+1)x^2/\bar{\gamma}_l})$
Nakagami-q	$f_{\alpha_l}(x) = \frac{2x}{\bar{\gamma}_l\sqrt{1-b_l^2}}\exp(-\frac{x^2}{\bar{\gamma}_l(1-b_l^2)})I_0(\frac{b_l x^2}{\bar{\gamma}_l(1-b_l^2)})$, $F_{\alpha_l}(x) = 1-\sqrt{1-b_l^2}\sum_{k=0}^{\infty}\Gamma\left(2k+1,\frac{x^2}{\bar{\gamma}_l(1-b_l^2)}\right)b_l^{2k}/[4^k(k!)^2]$. where $b_l=(1-q_l^2)/(1+q_l^2)$, and $0\le q_l\le 1$.
Nakagami-m	$f_{\alpha_l}(x) = \frac{2}{\Gamma(m_l)}(m_l/\bar{\gamma}_l)^{m_l}x^{2m_l-1}\exp(-m_l x^2/\bar{\gamma}_l)$ $F_{\alpha_l}(x) = 1-\frac{1}{\Gamma(m_l)}\Gamma(m_l,m_l x^2/\bar{\gamma}_l)$, where $m_l\in(0.5,\infty)$.
Weibull	$f_{\alpha_l}(x) = c_l\left[\frac{\Gamma(1+2/c_l)}{\bar{\gamma}_l}\right]^{c_l/2}x^{c_l-1}\cdot\exp\left(-[\frac{\Gamma(1+2/c_l)}{\bar{\gamma}_l}x^2]^{c_l/2}\right)$. $F_{\alpha_l}(x) = 1-\exp\left(-\left[\frac{\Gamma(1+2/c_l)}{\bar{\gamma}_l}x^2\right]^{c_l/2}\right)$, where $c_l\in(0,\infty)$.

Table 3.1 PDF and CDF expressions of fading channel amplitude in Rayleigh, Rician, Nakagami-q, Nakagami-m and Weibull fading channels. $Q_1(\sqrt{2a},\sqrt{2b}) = \int_b^{\infty} e^{-t-a}I_0(2\sqrt{at})dt$ is the Marcum-q function of the first order.

3.4 Efficient evaluation for TMGF of fading amplitudes

In general, the TMGF expressions for the signal amplitude α_l over different fading channels, may be evaluated using a a trapezoidal summation as shown by

$$\Phi_{\alpha_l}(s,x) = \int_x^{\infty} f_{\alpha_l}(y)e^{-sy}dy \simeq \Delta_y\left(0.5f_{\alpha_l}(y_0)e^{-sy_0} + \sum_{n=1}^{N_y-1} f_{\alpha_l}(y_n)e^{-sy_n}\right) \tag{3.13}$$

where $y_0 = x$, $y_n = y_0 + n\Delta_y$, and Δ_y and N_y are the step-size and the number of summation points, respectively. We may decrease Δ_y and/or increase N_y such that the summation in (3.13) converges. As a rule of thumb, we may choose $y_0 + N_y\Delta_y \geq 10\sqrt{\bar{\gamma}_l}$ for an accurate evaluation. We also recommend (3.11) to be evaluated using the trapezoidal summation quadrature, as given by

$$\begin{aligned}&\Phi_{\alpha_{\text{HSEGC}}}(s)\\ &\simeq \sum_{n_1,\ldots,n_L\in\mathcal{I}} \Delta_x \left\{\sum_{k=1}^{N_x} e^{-sx_k} f_{\alpha_{n_{L_c}}}(x_k) \left[\prod_{l=1}^{L_c-1} \Phi_{\alpha_{n_l}}(s,x_k)\right] \left[\prod_{l'=L_c+1}^{L} F_{\alpha_{n_{l'}}}(x_k)\right]\right\}\end{aligned} \tag{3.14}$$

where $x_n = n\Delta_x$, and Δ_x and N_x are the step-size and the number of summation points, respectively.

It may appear the calculation of all the TMGF values required in (3.13) and (3.14) involves two-dimensional summation, with the complexity of N_xN_y and LN_xN_y for i.i.d. and i.n.d. channels, respectively. For efficient evaluation of (3.14), we propose to set $\Delta_x = \Delta_y$ and $N_x = N_y$ for (3.13) and (3.14), and use the Matlab "cumsum" command to calculate the TMGF points[2] specified in (3.13). The TMGF values for $\{x_k\}_{k=1}^{N_x}$ required in (3.14) can be pre-computed and stored, which involves only L separate summations (for all the L branches) for the i.n.d. fading case, and only one summation for the i.i.d. case. Thus, with this technique the calculation of all the TMGF values for (3.14) involves a complexity of N_y and LN_y, respectively, which is a significant reduction compared to the 2-dimensional summation as in convention. The MGF of HS/EGC amplitude can now be accurately evaluated with a very low computational complexity.

For theoretical interests and comparison with some known results in the literature, we also derive some closed-form TMGF expressions for the Rayleigh and Nakagami-m fading model. For brevity, we only list the derivation for Rayleigh fading scenario, while TMGF results in Nakagami fading channel is given in Appendix B.2.

3.4.1 Closed-form TMGF for Rayleigh fading channels

The PDF of Rayleigh fading amplitude at the lth branch is given by

$$f_{\alpha_l}(x) = \frac{2x}{\bar{\gamma}_l} e^{-\frac{x^2}{\bar{\gamma}_l}}, \tag{3.15}$$

[2]Here, the operation using command "cumsum" has a complexity of the N_y-point summation, and returns N_y (with $N_y = N_x$) number of TMGF values $\Phi_{\alpha_l}(s, x_k)$, for $k = 1, \ldots, N_y$.

The CDF of α_l is given by $F_{\alpha_l}(x) = 1 - e^{-\frac{x^2}{\bar{\gamma}_l}}$.

Assume a real-valued s first. The truncated MGF of α_l is computed as follows:

$$\Phi_{\alpha_l}(s,x) = \int_x^\infty f(y)e^{-sy}dy = \frac{2}{\bar{\gamma}_l}e^{\frac{s^2\bar{\gamma}_l}{4}}\int_x^\infty ye^{-\frac{1}{\bar{\gamma}_l}(y+\frac{s\bar{\gamma}_l}{2})^2}dy. \tag{3.16}$$

Substituting $t = y + \frac{s\bar{\gamma}_l}{2}$ into (3.16) we have

$$\Phi_{\alpha_l}(s,x) = \frac{2}{\bar{\gamma}_l}e^{\frac{s^2\bar{\gamma}_l}{4}}\int_{x+\frac{s\bar{\gamma}_l}{2}}^\infty (t-\frac{s\bar{\gamma}_l}{2})e^{-\frac{1}{\bar{\gamma}_l}t^2}dt = \frac{2}{\bar{\gamma}_l}e^{\frac{s^2\bar{\gamma}_l}{4}}[\Phi_1(s,x) - \Phi_2(s,x)] \tag{3.17}$$

where

$$\Phi_1(s,x) = \int_{x+\frac{s\bar{\gamma}_l}{2}}^\infty te^{-\frac{1}{\bar{\gamma}_l}t^2}dt = \frac{\bar{\gamma}_l}{2}e^{-\frac{1}{\bar{\gamma}_l}(x+\frac{s\bar{\gamma}_l}{2})^2}, \tag{3.18}$$

and

$$\Phi_2(s,x) = \frac{s\bar{\gamma}_l}{2}\int_{x+\frac{s\bar{\gamma}_l}{2}}^\infty e^{-\frac{1}{\bar{\gamma}_l}t^2}dt = \frac{s\bar{\gamma}_l\sqrt{\pi\bar{\gamma}_l}}{2}Q\left(\sqrt{\frac{2}{\bar{\gamma}_l}}\left(x+\frac{s\bar{\gamma}_l}{2}\right)\right). \tag{3.19}$$

Because in (3.19) the standard Gaussian-Q function $Q(y)$ does not allow a complex argument y. Below, we derive another closed-form expression for $\Phi_2(s,x)$ so that complex s is included. Recall that the Gaussian-Q function has an alternative expression, which valid for complex y, given by [12, eq. (18)] and [47]

$$Q(y) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}e^{-\frac{t^2}{2}}\sin(ty)dt, \tag{3.20}$$

which leads to

$$\begin{aligned} Q(jy) &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}e^{-\frac{t^2}{2}}\sin(tjy)dt = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}e^{-\frac{t^2}{2}} \times \frac{e^{-yt}-e^{yt}}{2j}dt \\ &= \frac{1}{2} + \frac{1}{j\pi}\int_0^\infty \frac{1}{t}e^{-\frac{t^2}{2}} \times \mathrm{sh}(yt)dt = \frac{1}{2} + \frac{1}{j\pi}\int_0^\infty \frac{1}{t}e^{-\frac{t^2}{2}} \times \sum_{k=0}^\infty \frac{(yt)^{2k+1}}{(2k+1)!}dt \\ &= \frac{1}{2} + \frac{1}{j\pi}\sum_{k=0}^\infty \frac{y^{2k+1}}{(2k+1)!}\int_0^\infty e^{-\frac{t^2}{2}}t^{2k}dt. \end{aligned}$$

where $\mathrm{sh}(x) = \frac{e^x-e^{-x}}{2}$ is the hyperbolic sine function. Using a change of variable $t^2 = z$ and the Gamma integral that $\int_0^\infty z^{m-1}e^{-z/\beta}dz = \beta^m\Gamma(m)$, we obtain

$$\int_0^\infty e^{-\frac{t^2}{2}}t^{2k}dt = 2^{k-0.5}\Gamma(k+0.5) = \frac{(2k)!\sqrt{\pi/2}}{2^k k!}.$$

Thus, we derive a closed-form expression for $Q(jy)$ valid for complex y as

$$Q(jy) = \frac{1}{2} + \frac{\sqrt{2\pi}}{j2\pi}\sum_{k=0}^\infty \frac{(y)^{2k+1}}{(2k+1)!}\frac{(2k)!}{k!2^k} = \frac{1}{2} - \frac{j}{\sqrt{2\pi}}\sum_{k=0}^\infty \frac{(y)^{2k+1}}{(2k+1)k!2^k} \tag{3.21}$$

which is equivalent to (valid for complex z)

$$Q(z) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)k!2^k} = \frac{1}{2} - \frac{z}{\sqrt{2\pi}} \phi(\frac{1}{2}, \frac{3}{2}; -\frac{z^2}{2}), \tag{3.22}$$

where $\phi(a, b; z)$ is confluent hypergeometric function defined in Section 2.4.3. Based on the equality that

$$\phi(1, \frac{1}{2}, -z) = 1 - 2z \exp(-z) \cdot \phi(\frac{1}{2}, \frac{3}{2}, z),$$

we obtain another closed-form formula for $Q(z)$ as

$$Q(z) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \left[\phi\left(1, \frac{1}{2}; \frac{z^2}{2}\right) - 1 \right] z^{-1} \exp(-\frac{z^2}{2}). \tag{3.23}$$

For better numerical stability, it is suggested to use the transform

$$\phi\left(1, \frac{1}{2}; x\right) = e^x \phi\left(-\frac{1}{2}, \frac{1}{2}; -x\right)$$

for small-to-medium $|x|$ [11, 15]; and to use a divergent series [15]

$$\phi\left(1, \frac{1}{2}; x\right) = \frac{1}{2x} \sum_{k=0}^{N} \frac{(2k+1)!!}{(-2x)^k}$$

for the large $|x|$ (e.g. $|x| \geq 200$), where N is a finite positive integer, and

$$k!! = \begin{cases} k(k-2)\cdots 1, & k \text{ is odd} \\ k(k-2)\cdots 2, & k \text{ is even} \end{cases}$$

is the double factorial of the positive integer k.

Consequently, by combining (3.23) with (3.18) a closed-form expression for the TMGF $\Phi_{\alpha_l}(s, x)$ valid for complex s is derived as

$$\begin{aligned} \Phi_{\alpha_l}(s, x) = {} & e^{-sx - \frac{x^2}{\bar{\gamma}_l}} - \sqrt{\pi \bar{\gamma}_l} s \left\{ \frac{\exp(s^2 \bar{\gamma}_l / 4)}{2} - \sqrt{\frac{\bar{\gamma}_l}{4\pi}} \right. \\ & \left. \times \left[\phi\left(1, \frac{1}{2}; \frac{1}{\bar{\gamma}_l}(x + s\bar{\gamma}_l/2)^2\right) - 1 \right] \frac{\exp(-(x^2 + xs\bar{\gamma}_l)/\bar{\gamma}_l)}{x + s\bar{\gamma}_l/2} \right\}. \end{aligned} \tag{3.24}$$

As a sanity check, we plug in $x = 0$ and $s = -j\omega$, (3.24) reduces to[3]

$$\Phi_{\alpha_l}(j\omega, 0) = \phi\left(1, \frac{1}{2}; -\frac{\omega^2 \bar{\gamma}_l}{4}\right) + j\omega \sqrt{\frac{\bar{\gamma}_l}{4\pi}} \exp(-\frac{\omega^2 \bar{\gamma}_l}{4}) \tag{3.25}$$

which is identical to the CHF result of the Rayleigh fading amplitude given in [12, Table II] and [15], as expected.

[3]Here, since we define the MGF of x as $E[e^{-sx}]$ in this paper, we need to set $s = -j\omega$ instead of $s = j\omega$ to compare with the CHF defined as $E[e^{j\omega x}]$ in [12, 15].

3.5 Error and outage probabilities of HS/EGC

3.5.1 Error probability

When the MGF (or CHF) $\Phi_{\alpha_{\text{HSEGC}}}(\omega)$ has been obtained, we may apply Parseval's theorem [12] to evaluate the average BER and SER of a large class of modulation formats.

For example, we obtain the average BER of BPSK in fading channels as

$$\bar{P}_{\text{BPSK}} = \int_0^\infty P_{\text{BPSK}}(\alpha) f_{\alpha_{\text{HSEGC}}}(x) dx = \frac{1}{\pi} \int_0^\infty \Phi_{P_{\text{BPSK}}}(\omega) \Phi^*_{\alpha_{\text{HSEGC}}}(\omega) d\omega \tag{3.26}$$

where $\Phi_{P_{\text{BPSK}}}(\omega) = \text{FT}[P_{\text{BPSK}}(\alpha)] = E[e^{j\omega P_{\text{BPSK}}(\alpha)}]$ is the Fourier transform of $P_{\text{BPSK}}(\alpha)$, and $P_{\text{BPSK}}(\alpha) = Q(\sqrt{2}\alpha)$ is the conditional BER of BPSK. We have

$$\Phi_{P_{\text{BPSK}}}(\omega) = \text{FT}[Q(\sqrt{2}\alpha)] = \Phi_Q(\frac{\omega}{\sqrt{2}})/\sqrt{2}. \tag{3.27}$$

where the Fourier transform of Gassuain-Q function is given in (2.38).

To evaluate the average SER for M-QAM, we simply need to replace $\Phi_{P_{\text{BPSK}}}(\omega)$ in (2.37) with the CHF of conditional SER of M-QAM, denoted as $\Phi_{P_{\text{QAM}}}(\omega)$. An analytical expression of $\Phi_{P_{\text{QAM}}}(\omega)$ was given by [12, eq. (13)], shown here for completeness,

$$\begin{aligned} \Phi_{P_{\text{QAM}}}(\omega) = & \frac{2a}{\omega\sqrt{\pi}} F\left(\frac{\omega}{2\sqrt{b}}\right) - \frac{4c}{\omega\sqrt{\pi}} \left[F\left(\frac{\omega}{2\sqrt{b}}\right) - F\left(\frac{\omega}{2\sqrt{2b}}\right) \exp\left(\frac{-\omega^2}{8b}\right)\right] \\ & + j\left(\frac{a-c}{\omega}\left[1 - \exp(-\frac{\omega^2}{4b})\right] + \frac{4a}{\omega\pi} F^2\left(\frac{\omega^2}{2\sqrt{2b}}\right)\right) \end{aligned} \tag{3.28}$$

where $a = 2(1 - \frac{1}{\sqrt{M}})$, $b = 1.5\frac{\log_2 M}{M-1}$, and $c = (1 - \frac{1}{\sqrt{M}})^2$.

In summary, we can evaluate the HS/EGC performance of a large class of modulation formats, including M-PSK, M-QAM, M-ary differential PSK (M-DPSK), M-ary frequency-shift-keying (M-FSK), and 2-dimension signals [94]. Some relevant CHF formulas of conditional error rates were listed in [12].

3.5.2 Outage probability and SNR statistics

The outage probability for the HS/EGC diversity receiver output SNR γ_{HSEGC} is defined as

$$P_{\text{OT},\gamma_{\text{HSEGC}}}(\gamma_{\text{th}}) = \int_0^{\gamma_{\text{th}}} f_{\gamma_{\text{HSEGC}}}(\gamma) d\gamma,$$

where $f_{\gamma_{\text{HSEGC}}}(\gamma)$ is the PDF of the γ_{HSEGC}, and γ_{th} is the prescribed threshold SNR. Since $\gamma_{\text{HSEGC}} = \alpha^2_{\text{HSEGC}}$, we have $\Pr(\gamma_{\text{HSEGC}} \le \gamma_{\text{th}}) = \Pr(\alpha_{\text{HSEGC}} \le \sqrt{\gamma_{\text{th}}})$, and thus

$$P_{\text{OT},\gamma_{\text{HSEGC}}}(\gamma_{\text{th}}) = P_{\text{OT},\alpha_{\text{HSEGC}}}(\sqrt{\gamma_{\text{th}}}) = \int_0^{\sqrt{\gamma_{\text{th}}}} f_{\alpha_{\text{HSEGC}}}(x) dx.$$

Using numerical inversion of the Laplace transform of the CDF of α_{HSEGC} [62, 94], and the trapezoidal summation, an efficient formula to evaluate $P_{\text{OT},\gamma_{\text{HSEGC}}}(\gamma_{\text{th}})$ is obtained as

$$P_{\text{OT},\gamma_{\text{HSEGC}}}(\gamma_{\text{th}}) \simeq \frac{e^{A/2}}{2^Q} \sum_{\beta=\pm(\pi/2-\pi/M)} \sum_{q=0}^{Q} \binom{Q}{q} \sum_{n=0}^{N+q} \frac{(-1)^n}{c_n} \times \text{Re}\left(\Phi_{\alpha_{\text{HSEGC}}}\left(-\frac{A+jn2\pi}{2\sqrt{\gamma_{\text{th}}}}\right)/(A+jn2\pi)\right) + E_{A,N,Q} \quad (3.29)$$

where $E_{A,N,Q}$ is a remainder term that vanishes when Q becomes large, A is a constant to ensure the fast convergence of (3.29), and $c_n = 1$ for $n = 0$; $c_n = 0.5$ for $n = 1, \ldots, N+Q$.

The nth-moment of γ_{HSEGC} may be obtained using $\Phi_{\alpha_{\text{HSEGC}}}(s)$ and the moment theorem, given by

$$E[\gamma_{\text{HSEGC}}^n] = E[\alpha_{\text{HSEGC}}^{2n}] = \Phi^{(2n)}_{\alpha_{\text{HSEGC}}}(s)|_{s=0} \quad (3.30)$$

where $\Phi^{(n)}_{\alpha_l}(s)$ denotes the nth-order derivative of function $\Phi_{\alpha_l}(s)$ with respect to s. Equation (3.30) may be evaluated using the numerical differentiation.

The variance of γ_{HSEGC} is simply given by $\text{var}\{\gamma_{\text{HSEGC}}\} = \Phi^{(4)}_{\alpha_{\text{HSEGC}}}(s)|_{s=0} - [\Phi^{(2)}_{\alpha_{\text{HSEGC}}}(s)|_{s=0}]^2$. The amount of fading (AF) is calculated as

$$\text{AF}_{\text{HSEGC}} = \frac{\text{var}\{\gamma_{\text{HSEGC}}\}}{E[\gamma_{\text{HSEGC}}]^2} = \frac{\Phi^{(4)}_{\alpha_{\text{HSEGC}}}(s)|_{s=0}}{[\Phi^{(2)}_{\alpha_{\text{HSEGC}}}(s)|_{s=0}]^2} - 1. \quad (3.31)$$

3.6 Numerical examples

We provide some numerical examples to compare the performance of HS/EGC with HS/MRC, and show the effect of some system and channel parameters (such as i.n.d. fading statistics, L_c and L) on the performance of HS/EGC over different types of fading channels.

We first verify the analytical error rate formulas derived and the receiver structure for HS/EGC proposed in this chapter. In Figs. 3.2 – 3.4 we present the analytical and simulated SERs for 16-QAM with HS/EGC over an i.n.d. Rayleigh channel, an i.i.d. Nakagami-m channel (with $m = 2.1$), and an i.n.d. Nakagami-m channel, respectively. For i.n.d. Rayleigh and Nakagami-m channels, we assume the ASNR decreases by 2dB successively from the first branch to the fourth one. Furthermore, following the same branch order, the m-parameters for the i.n.d. Nakagami-m channel are given by $[m_1, m_2, m_3, m_4] = [3, 2.4, 1.8, 1.2]$, respectively. The simulated SER for 16-QAM was obtained by assuming the structure given in Fig. 3.1 and

using decision variable given in (3.2). The results in Figs. 3.2 – 3.4 confirmed our error rate analysis and also verified our proposed HS/EGC structure for NCM formats.

Fig. 3.2 shows that HS/EGC (3,4) has a slight better SER performance than the HS/EGC (4,4) for the i.n.d. Rayleigh channel. We note that this observation is also true for HS/EGC over i.i.d. Rayleigh channels based on our simulations.

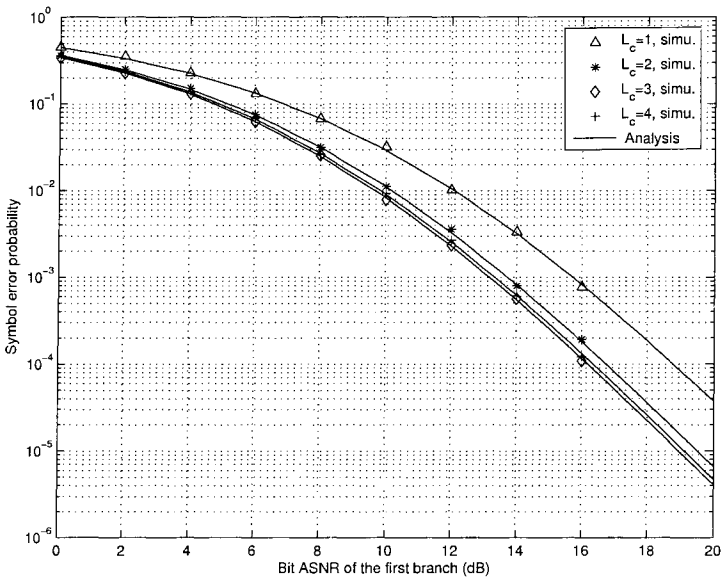

Figure 3.2 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.n.d. Rayleigh fading channel, with $L_c = 1 \sim 4$ and $L = 4$.

This phenomenon may be explained by the fact that unlike HS/MRC, in HS/EGC the EGC subset combining is only suboptimal and involves an ASNR loss, which may cause the performance of EGC (i.e., HS/EGC (L, L)) to be slightly worse than that of HS/EGC (L_c, L), with L_c being smaller than but close to L.

From Fig. 3.3 we observe that the performance of HS/EGC (4,4) is slightly better than HS/EGC (3,4) over the i.i.d. Nakagami-m channel. In Fig. 3.4, the SER result of the considered i.n.d. Nakagami channel shows that the performance of HS/EGC (3,4) is almost identical to that of HS/EGC (4,4), though slightly worse but the gap is negligible. These results suggest that when the m-parameter increases the ASNR loss caused by EGC subset combining becomes smaller, and that HS/EGC (4,4) is slightly better than HS/EGC (3,4) (e.g., when $m > 1.4$ for the i.i.d. case). Furthermore, over the i.n.d. channels the SNR gaps between

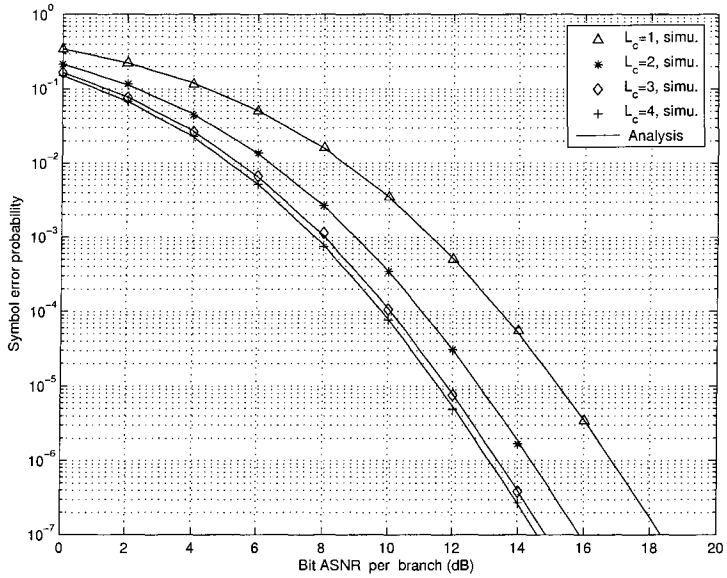

Figure 3.3 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.i.d. Nakagami fading channel, with $m = 2.1$, $L_c = 1 \sim 4$ and $L = 4$.

HS/EGC (L_c, L) and full EGC tend to be decreased compared to the i.i.d. channels.

In Fig. 3.5, we provide the average SER of 64-QAM with HS/MRC and HS/EGC vs. bit ASNR per branch over an i.i.d. Rayleigh fading channel for $L_c = 2$ with L varying form 2 to 6. The result shows that when L increases the gap between HS/EGC and HS/MRC shrinks, and the performance of both diversity formats improves substantially. It has been proved in [74] that HS/MRC (L_c, L) achieves the diversity gain of L over Rayleigh channels. Fig. 3.5 shows that HS/EGC (L_c, L) attains the same diversity gain as HS/MRC (L_c, L), though with a slightly worse SNR gain.

Besides the study for Rayleigh and Nakagami-m fading channels, performance analysis of diversity reception over Weibull fading channels has recently attracted a lot of attention [8, 26, 89]. In Fig. 3.6 we present the SER for HS/EGC over an i.i.d. Weibull fading channel with $c = 3.1$. When L increases, the diversity order increases, as expected. Also, HS/EGC achieves the same diversity gain as HS/MRC for all cases.

In Fig. 3.7 we compare the performance of quadrature PSK (QPSK) with HS/EGC over i.i.d. and i.n.d. Nakagami-q fading channels, respectively, with $L = 4$ and $L_c = 1 \sim 4$. For i.i.d. channels, we assume $q = 0.4$ for all the branches; and for i.n.d. channels, the ASNR decreases

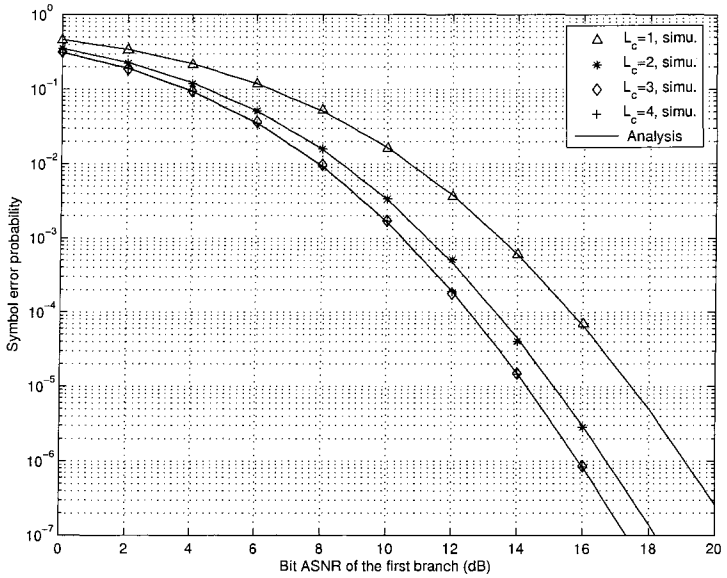

Figure 3.4 Analytical and simulated SERs vs. bit ASNR for 16-QAM with HS/EGC over an i.n.d. Nakagami fading channel, with $m = [3, 2.4, 1.8, 1.2]$, $L_c = 1 \sim 4$ and $L = 4$.

by 2dB successively from the first branch, and following the same order the q-parameters are given by $[1, 0.8, 0.6, 0.4]$. In this example, the performance of QPSK in i.i.d. channels is better than the i.n.d. channels. Furthermore, the HS/EGC (3,4) is better than HS/EGC (4,4) for both i.i.d. and i.n.d. Nakagami-q channels.

The SER of 16-QAM with HS/EGC $(L_c, 4)$ over i.i.d. and i.n.d. Rician fading channels is shown in Fig. 3.8. For i.n.d. channels, we assume the ASNR decreases by 2dB successively from the first branch, and the Rice-K factors are given by $[K_1, K_2, K_3, K_4] = [8, 6, 4, 2]$ dB. For i.i.d. channels, we assume $K = 3$ dB for all the branches. The result shows that for low-to-medium ASNRs the case of i.n.d. branches has a worse SER performance than the i.i.d. case. For high ASNR, HS/EGC in the i.n.d. case has an asymptotically better performance due to the higher Rice-K factors. The relative performances between the cases of i.i.d. and i.n.d. branches depends on the ASNR and system and channel parameters, and their high ASNR behaviors involve the asymptotical analysis and is worthy of a further investigation.

From Figs. 3.7–3.8, we also found that for both i.i.d. and i.n.d. Nakagami-q and Rician fading channels, HS/EGC (3,4) shows a slight better SER performance than the HS/EGC (4,4). Note that the same trend is observed for Rayleigh fading channels. This result shows

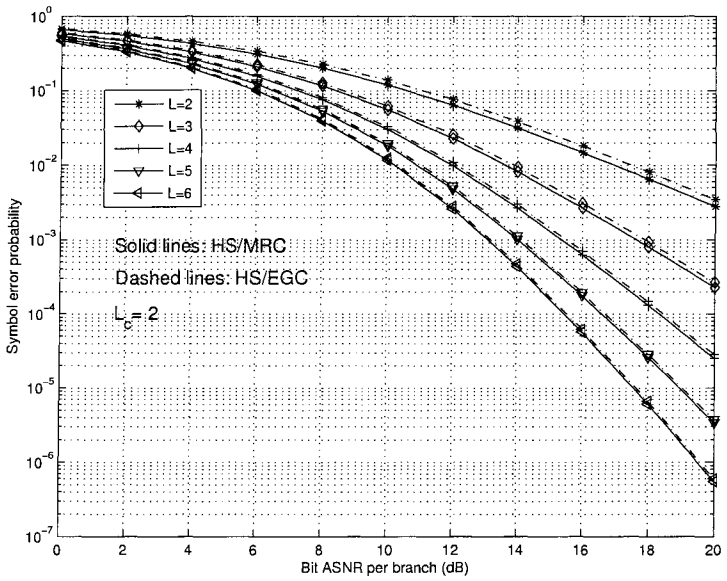

Figure 3.5 Average SER vs. bit ASNR for 64-QAM with HS/EGC and HS/MRC over an i.i.d. Rayleigh fading channel, with $L_c = 2$ and $L = 2 \sim 6$.

that when the amount of fading is larger (e.g., when the q- or m-parameter in Nakagami channels decreases), the EGC subset combining loss tends to increase, and L-fold full EGC may not be optimal compared to HS/EGC (L_c, L).

In Fig. 3.9, we provide the outage probability vs. normalized bit ASNR of the first branch for HS/EGC and HS/MRC over an i.n.d. Weibull fading channel with $L = 5$ and $L_c = [1, 2, 3, 5]$. We assume the ASNR decreases by 1.5 dB successively from the first branch to the fifth one, and the Weibull c-parameters are given by $[5, 4, 3, 2, 2]$, respectively. The result shows that when $L_c = 1$, the outage probabilities of HS/MRC and HS/EGC are identical, as expected. When $L_c > 1$, HS/EGC has a worse performance than HS/MRC and the gap slowly increases with L_c. When $L_c = L$, HS/EGC (L, L) and HS/MRC (L, L) are equivalent to EGC and MRC, respectively, and their SER gap becomes the largest for all L_c. Furthermore, the gap between HS/EGC $(3, 5)$ and HS/EGC $(5, 5)$ is almost negligible. This result suggests that HS/EGC (L_c, L) with a small L_c may be very attractive for a tradeoff between the complexity and the performance.

In Fig. 3.10 we show the outage probability of HS/EGC over i.i.d. and i.n.d. Nakagami-m fading channels with $L = 5$ and $L_c = [1, 2, 5]$. For i.i.d. Nakagami channels, we assume

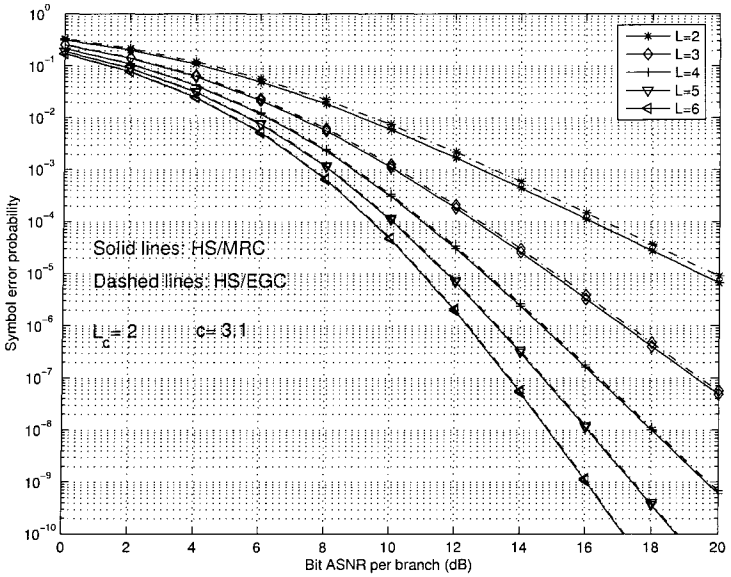

Figure 3.6 Average SER vs. bit ASNR for 16-QAM with HS/EGC and HS/MRC over an i.i.d. Weibull fading channel, with $L_c = 2$, $L = 2 \sim 6$, and Weibull parameter $c = 3.1$.

$m = 2.1$; and for i.n.d. channels, we assume the ASNR decreases by 1.5 dB successively, and the m-parameters are $[m_1, m_2, m_3, m_4, m_5] = [3, 2.4, 1.8, 1.2, 0.6]$. The gap between $L_c = 2$ and $L_c = 5$ for i.n.d. case is smaller than that for i.i.d. case. Results in both Fig. 3.9 and Fig. 3.10 suggest that for i.n.d. channels even combining a small number of branches may approach the performance of EGC.

3.7 Summary

In this chapter, we have proposed an HS/EGC receiver structure applicable to non-constant modulus modulation formats and verified its validity. By developing a general and accurate MGF expression for the HS/EGC output amplitude, we provided an unified method to evaluate the error and outage probabilities of HS/EGC with a large class of modulation formats over generalized fading channels applicable to different fading families and not necessarily identically distributed branches. Efficient numerical techniques and some closed-form expressions to evaluate the MGF and the truncated MGF expressions have been proposed. Performance comparison with HS/MRC shows that the gap between HS/EGC and HS/MRC increases as

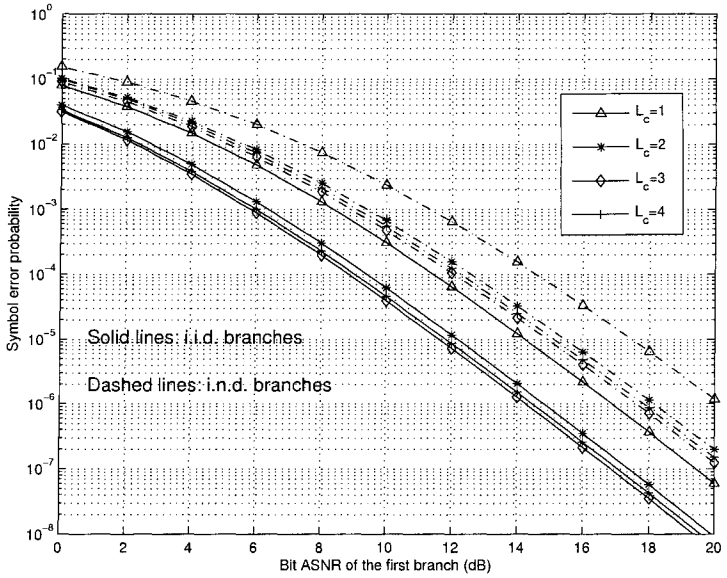

Figure 3.7 Average SER vs. bit ASNR for QPSK with HS/EGC over i.i.d. and i.n.d. Nakagami-q fading channels, with $L = 4$ and $L_c = 1 \sim 4$.

L_c increases, but decreases as L increases. The result also suggests that HS/EGC achieves the same diversity order as HS/MRC. Unlike HS/MRC, in HS/EGC the EGC subset combining is suboptimal and thus the performance may not monotonically improve as L_c increases (i.e. by combining more branches). These results provide new insight into the trade-off between performance and complexity hybrid diversity reception over generalized fading channels and aid in the design of such receivers.

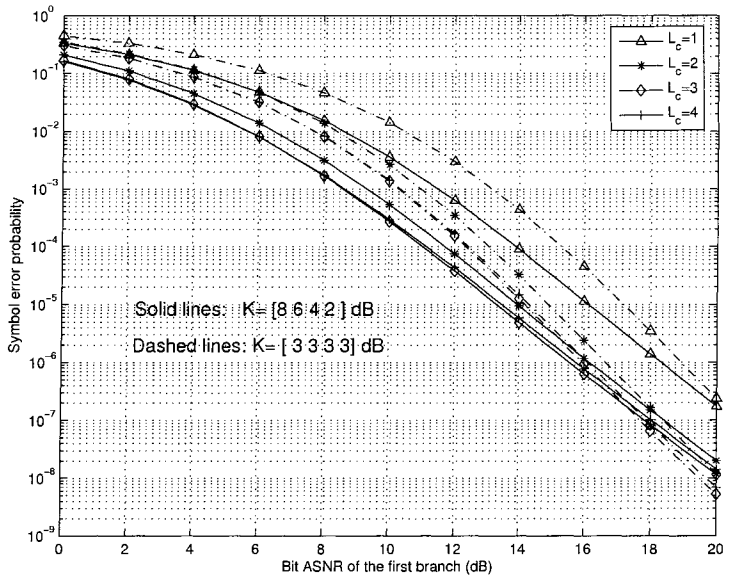

Figure 3.8 Average SER vs. bit ASNR for 16-QAM with HS/EGC over i.i.d. and i.n.d. Rician fading channels with $L = 4$. For i.i.d. Rician channels, $K = 3$ dB; and for i.n.d. Rician channels $[K_1, K_2, K_3, K_4] = [8, 6, 4, 2]$ dB.

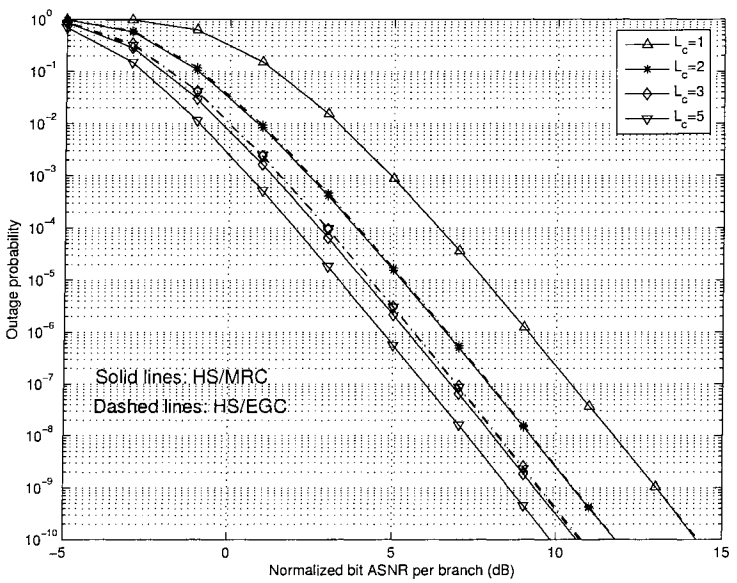

Figure 3.9 Outage probability vs. bit ASNR of the first branch for HS/EGC and HS/MRC over an i.n.d. Weibull fading channel with $L = 5$ and $L_c = [1, 2, 3, 5]$. The Weibull c-parameters are given by [5,4,3,2,2], respectively.

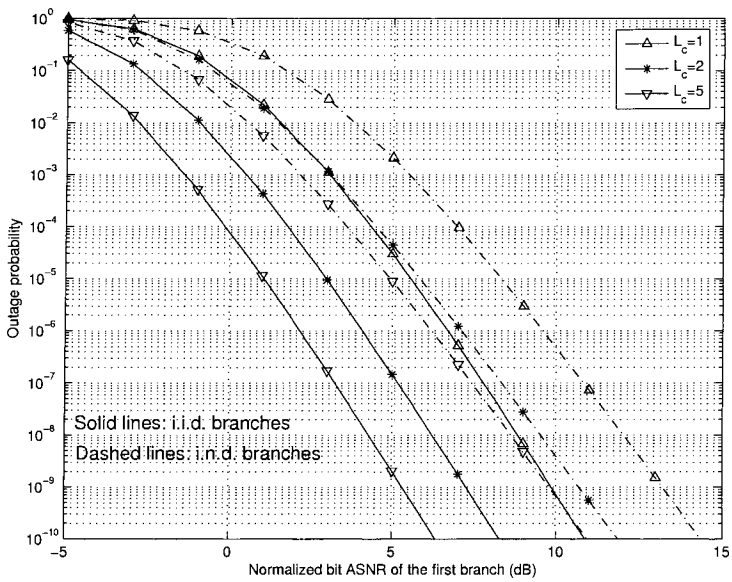

Figure 3.10 Outage probability vs. bit ASNR of the first branch for HS/EGC over i.i.d. and i.n.d. Nakagami-m fading channels with $L = 5$ and $L_c = [1, 2, 5]$. For i.i.d. Nakagami channels $m = 2.1$, and for i.n.d. channels $m = [3, 2.4, 1.8, 1.2, 0.6]$.

CHAPTER 4. ESTIMATION OF COMPOSITE GAMMA-LOGNORMAL FADING CHANNEL STATISTICS

4.1 Introduction

In the previous chapters, we have assumed that the statistic property of the fading channels are known, which is a "standard" assumption in performance evaluation and channel estimation. In addition, knowing these statistics of fading channels is beneficial for the following issues.

- implementation of space-time precoding schemes with mean and covariance feedback;
- antenna selection and adaptive modulation in spatial-multiplexing systems;
- implementation of noncoherent ML space-time receivers.

However, for fading channels with shadowing, to the best of our knowledge, no study on estimating statistic property has been done. Therefore, we conclude that there is a need for us to study the channel statistic estimation problem.

Composite fading-shadowing models are used to describe the statistic properties of wireless communication channels in congested downtown areas [48, 95], satellite communication systems [79, 101], and distributed antenna systems [87, 88]. Particularly, we consider a composite gamma-lognormal fading model. This model is fairly general and includes as special cases the Rayleigh-lognormal [48, 67] and classical Nakagami-m fading and lognormal shadowing scenarios [94, 95]. We also present ML methods based on Newton-Raphson and EM algorithms for estimating the unknown fading and shadowing parameters. Once obtained, the parameter estimates can be used to design and analyze the performance of wireless communication systems [95, 99], and to compute MMSE estimates of mean-signal powers [1], see [61].

[1] Accurate estimation of the mean-signal powers is required for implementing adaptive modulation techniques and algorithms for handoff, channel access, and power control [36, 56, 95].

4.2 Measurement model and ML estimation

Assume that N *instantaneous signal powers* $y_k(t)^2$, $t = 1, 2, \ldots, N$ have been collected in the kth observation interval, where $k = 1, 2, \ldots, K$, and define

$$\boldsymbol{y}_k = [y_k(1), y_k(2), \ldots, y_k(N)]^T, \tag{4.1}$$

We model $y_k(t)$, $t = 1, 2, \ldots, N$ as conditionally independent Gamma random variables with the following PDFs.

$$p_{y|u}(y_k(t) \,|\, u_k; m) = \frac{m^m y_k(t)^{m-1}}{u_k^m \, \Gamma(m)} \exp\Big(- \frac{m y_k(t)}{u_k} \Big), \tag{4.2}$$

where u_k is the *mean-signal (shadow) power* in the kth interval, $\Gamma(\cdot)$ denotes the gamma function, and m is the Nakagami-m fading parameter. Therefore, $y_{k_1}(t_1)|u_{k_1}$ and $y_{k_2}(t_2)|u_{k_2}$ are independent for $k_1 \neq k_2$ or $t_1 \neq t_2$ or both, where $k_1, k_2 \in \{1, 2, \ldots, K\}$ and $t_1, t_2 \in \{1, 2, \ldots, N\}$. We then model the mean-signal powers as i.i.d. random variables with lognormal PDF

$$p_u(u_k; \mu, \sigma^2) = \frac{\xi}{u_k \sqrt{2\pi\sigma^2}} \exp\Big(- \frac{(10 \log_{10} u_k - \mu)^2}{2\sigma^2} \Big). \tag{4.3}$$

In other words, u_k are assumed to be constant within an observation interval, but vary randomly from one interval to another. The assumption that the mean-signal powers u_k are independent is valid if the observation intervals are sufficiently separated in time. Here, μ (dB) and σ (dB) are the mean and standard deviation of $10 \log_{10} u_k$, also known as the *area mean* and *shadow standard deviation*, respectively (see [95]), and

$$\xi = \frac{10}{\ln 10}. \tag{4.4}$$

Our goal is to find the ML estimates of m, μ, and σ^2 using the instantaneous power observations $y_k(t)$, $t = 1, 2 \ldots, N$, $k = 1, 2, \ldots, K$. Define the unknown parameter vector

$$\boldsymbol{\theta} = [m, \mu, \sigma^2]^T. \tag{4.5}$$

The marginal distribution of $\boldsymbol{y}_k$ follows from (4.2) and (4.3):

$$p_y(\boldsymbol{y}_k; \boldsymbol{\theta}) = \int_0^\infty \Big[\prod_{t=1}^{N} p_{y|u}(y_k(t)|u; m) \Big] \cdot p_u(u; \mu, \sigma^2) \, du \tag{4.6}$$

[2]If the samples $y_k(t)$ are scaled by the noise power, they can be viewed as instantaneous SNRs.

for $k = 1, 2 \ldots, K$. The ML estimate of $\boldsymbol{\theta}$ is obtained by maximizing the log-likelihood function of all the measurements $\boldsymbol{y} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_K^T]^T$:

$$L(\boldsymbol{y}; \boldsymbol{\theta}) = \sum_{k=1}^{K} \ln p_y(\boldsymbol{y}_k; \boldsymbol{\theta}). \tag{4.7}$$

As observed in [96], the difficulty in estimating the parameters of the composite fading-shadowing models arises due to the integral form of the density function (4.6). In the following, we present Newton-Raphson and EM algorithms for finding the ML estimates of $\boldsymbol{\theta}$.

4.2.1 Newton-Raphson method

We derive the Newton-Raphson algorithm for maximizing (4.7). A quasi-Newton modification of the Newton-Raphson iteration is discussed in Section 4.2.1.1. First, we apply the change-of-variable transformation

$$x = \frac{10 \log_{10} u - \mu}{\sqrt{2\sigma^2}} \tag{4.8}$$

to (4.7):

$$\begin{aligned} L(\boldsymbol{y}; \boldsymbol{\theta}) &= -\frac{K}{2} \ln \pi + (m-1) \cdot \sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t) + KNm \ln m - KN \ln \Gamma(m) \\ &\quad + \sum_{k=1}^{K} \ln \Big\{ \int_{-\infty}^{\infty} q(x, \overline{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2)\, dx \Big\}, \end{aligned} \tag{4.9}$$

where

$$q(x, \overline{y}_k, \boldsymbol{\theta}) = \exp(-mN \cdot \overline{y}_k \cdot 10^{-(\sqrt{2\sigma^2} \cdot x + \mu)/10}) \cdot 10^{-mN \cdot (\sqrt{2\sigma^2} \cdot x + \mu)/10} \tag{4.10}$$

and

$$\overline{y}_k = \frac{1}{N} \sum_{t=1}^{N} y_k(t) \tag{4.11}$$

is the sample mean of the mean-signal power in the kth observation interval.

The gradient vector $\partial L(\boldsymbol{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta}$ and Hessian matrix $\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ can be computed using

$$\begin{aligned} \frac{\partial L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial m} &= KN \ln m + KN - KN \frac{\Gamma'(m)}{\Gamma(m)} + \sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t) \\ &\quad + \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x, \overline{y}_k, \boldsymbol{\theta}) / \partial m \exp(-x^2)\, dx}{\int_{-\infty}^{\infty} q(x, \overline{y}_k, \boldsymbol{\theta}) \exp(-x^2)\, dx}, \end{aligned} \tag{4.12a}$$

$$\frac{\partial L(\boldsymbol{y};\boldsymbol{\theta})}{\partial \mu} = \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\mu \exp(-x^2)\,dx}{\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx}, \tag{4.12b}$$

$$\frac{\partial L(\boldsymbol{y};\boldsymbol{\theta})}{\partial \sigma^2} = \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\sigma^2 \exp(-x^2)\,dx}{\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx}, \tag{4.12c}$$

and

$$\frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta})}{\partial m^2} = \frac{KN}{m} + \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial^2 q(x,\overline{y}_k,\boldsymbol{\theta})/\partial m^2 \exp(-x^2)\,dx}{\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx}$$
$$-\Big[\frac{\int_{-\infty}^{\infty} \partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial m \exp(-x^2)\,dx}{\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx}\Big]^2 - KN\frac{\Gamma(m)\Gamma''(m) - [\Gamma'(m)]^2}{[\Gamma(m)]^2}, \tag{4.13a}$$

$$\frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\Big|_{(i,j)\neq(1,1)} = \sum_{k=1}^{K}\Big\{\frac{\int_{-\infty}^{\infty} \partial^2 q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\theta_i\partial\theta_j \exp(-x^2)\,dx}{\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx} -$$
$$\frac{\int_{-\infty}^{\infty} \partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\theta_i \exp(-x^2)\,dx \int_{-\infty}^{\infty} \partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\theta_j \exp(-x^2)\,dx}{[\int_{-\infty}^{\infty} q(x,\overline{y}_k,\boldsymbol{\theta})\exp(-x^2)\,dx]^2}\Big\}. \tag{4.13b}$$

The integral expressions in (4.9), (4.12), and (4.13) are efficiently and accurately evaluated using the Gauss-Hermite quadrature formula:

$$\int_{-\infty}^{\infty} f(x)\exp(-x^2)\,dx \approx \sum_{l=1}^{L} h_{x_l} f(x_l), \tag{4.14}$$

where $f(x)$ is an arbitrary real function, L is the quadrature order, x_l are the zeroes of the Lth-order Hermite polynomial, and h_{x_l} are the Gauss-Hermite quadrature weight factors tabulated in [1]. We have omitted the expressions for the derivatives $\partial q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\theta_i$ and $\partial^2 q(x,\overline{y}_k,\boldsymbol{\theta})/\partial\theta_i\partial\theta_j$, $i,j \in \{1,2,3\}$, which are cumbersome but easy to compute. The (damped) Newton-Raphson algorithm updates the estimates of $\boldsymbol{\theta}$ as follows (see [85] and [92]):

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \lambda^{(i)} \cdot \Big[\frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta}^{(i)})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\Big]^{-1} \frac{\partial L(\boldsymbol{y};\boldsymbol{\theta}^{(i)})}{\partial\boldsymbol{\theta}}, \tag{4.15}$$

where the *damping factor* $0 < \lambda^{(i)} \leq 1$ is chosen at every step to ensure that the log-likelihood function (4.7) increases and the parameter estimates remain in the allowable parameter space (i.e. $m, \sigma^2 > 0$). The negative inverse of the Hessian matrix evaluated at the ML estimate $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(\infty)}$

$$-\Big[\frac{\partial^2 L(\boldsymbol{y};\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\Big]^{-1} \tag{4.16}$$

can be used to estimate the covariance matrix of $\widehat{\boldsymbol{\theta}}$ and to construct confidence regions for the unknown parameters [98]. The Hessian matrix formulas (4.13) will be also utilized to compute the CRB matrix for the unknown parameters, see Section 4.2.4.

4.2.1.1 BFGS quasi-Newton algorithm

To accelerate the Newton-Raphson algorithm, we propose the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method that approximates the Hessian matrices in (4.15) with the following estimates (see [98, eq. (4.3.7)] and [31, eq. (9.2.10)]):

$$H^{(i+1)} = H^{(i)} - \frac{H^{(i)}\boldsymbol{d}^{(i)}(\boldsymbol{d}^{(i)})^T H^{(i)}}{(\boldsymbol{d}^{(i)})^T H^{(i)}\boldsymbol{d}^{(i)}} + \frac{\boldsymbol{g}^{(i)}(\boldsymbol{g}^{(i)})^T}{(\boldsymbol{d}^{(i)})^T\boldsymbol{g}^{(i)}}, \tag{4.17}$$

where

$$\boldsymbol{d}^{(i)} = \boldsymbol{\theta}^{(i+1)} - \boldsymbol{\theta}^{(i)}, \tag{4.18a}$$

$$\boldsymbol{g}^{(i)} = \frac{\partial L(\boldsymbol{y};\boldsymbol{\theta}^{(i+1)})}{\partial\boldsymbol{\theta}} - \frac{\partial L(\boldsymbol{y};\boldsymbol{\theta}^{(i)})}{\partial\boldsymbol{\theta}}, \tag{4.18b}$$

and the initial value $H^{(0)}$ can be obtained by computing the exact Hessian at the initial estimate of the unknown parameter vector $\boldsymbol{\theta}^{(0)}$:

$$H^{(0)} = \frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta}^{(0)})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}. \tag{4.19}$$

Compared with the Newton-Raphson method, the BFGS quasi-Newton algorithm requires more iterations to converge, but each iteration has lower computational complexity, see the discussion in Section 4.3.

4.2.2 EM algorithm

We present an EM algorithm for computing the ML estimates of $\boldsymbol{\theta}$ [18, 30, 77]. Define the vector of the mean-signal powers $\boldsymbol{u} = [u_1, u_2, \ldots, u_K]^T$. By treating $\boldsymbol{u}$ as the *unobserved* data, we derive the following iteration between the expectation and maximization steps (see the Appendix C).

- E Step: Compute

$$\mathcal{T}_1(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) = \frac{1}{K}\sum_{k=1}^{K} \mathrm{E}_{u|y}[\ln u_k \mid \boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}], \tag{4.20a}$$

$$\mathcal{T}_2(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) = \frac{1}{K}\sum_{k=1}^{K} \mathrm{E}_{u|y}[(\ln u_k)^2 \mid \boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}], \tag{4.20b}$$

$$\mathcal{T}_3(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) = \frac{1}{K}\sum_{k=1}^{K} \{\mathrm{E}_{u|y}[u_k^{-1} \mid \boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}]\cdot\overline{y}_k\} - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t), \tag{4.20c}$$

where

$$\boldsymbol{\theta}^{(i)} = [m^{(i)}, \mu^{(i)}, (\sigma^2)^{(i)}]^T$$

is an estimate of $\boldsymbol{\theta}$ in the ith iteration and (4.20a)–(4.20c) are computed using

$$\mathrm{E}_{u|y}[t(u_k)|\boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}] = \frac{\int_0^\infty t(u)[\prod_{t=1}^N p_{y|u}(y_k(t)|u;m^{(i)})]p_u(u;\mu^{(i)},(\sigma^2)^{(i)})\,du}{\int_0^\infty [\prod_{t=1}^N p_{y|u}(y_k(t)|u;m^{(i)})]p_u(u;\mu^{(i)},(\sigma^2)^{(i)})\,du} \quad (4.21a)$$

$$\approx \frac{\sum_{l=1}^L h_{x_l} t(10^{(\sqrt{2\sigma^2}x_l+\mu)/10})\exp(-mN\overline{y}_k 10^{-(\sqrt{2\sigma^2}\cdot x_l+\mu)/10})10^{-mN\sqrt{2\sigma^2}x_l/10}}{\sum_{l=1}^L h_{x_l}\exp(-mN\overline{y}_k 10^{-(\sqrt{2\sigma^2}x_l+\mu)/10})10^{-mN\sqrt{2\sigma^2}x_l/10}} \quad (4.21b)$$

with $t(u_k) = \ln u_k$, $(\ln u_k)^2$, and u_k^{-1}, for $k = 1, 2, \ldots, K$.

- M Step: Compute

$$\mu^{(i+1)} = \xi \cdot \mathcal{T}_1(\boldsymbol{y};\boldsymbol{\theta}^{(i)}), \quad (4.22a)$$

$$(\sigma^2)^{(i+1)} = \xi^2 \cdot \{\mathcal{T}_2(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) - [\mathcal{T}_1(\boldsymbol{y};\boldsymbol{\theta}^{(i)})]^2\}\xi^2 \cdot \mathcal{T}_2(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) - (\mu^{(i+1)})^2, \quad (4.22b)$$

and find $m^{(i+1)}$ that maximizes

$$m^{(i+1)} = \arg\max_m \Big\{m\ln m - \ln[\Gamma(m)] - m\mathcal{T}_1(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) - m\mathcal{T}_3(\boldsymbol{y};\boldsymbol{\theta}^{(i)})\Big\}. \quad (4.22c)$$

Upon convergence (i.e. as $i \to \infty$), the above algorithm also provides estimated MMSE estimates of the shadow powers in decibels:

$$\mathrm{E}_{u|y}[10\log_{10} u_k \mid \boldsymbol{y}_k;\widehat{\boldsymbol{\theta}}] = \xi \cdot \mathrm{E}_{u|y}[\ln u_k \mid \boldsymbol{y}_k;\widehat{\boldsymbol{\theta}}], \quad (4.23)$$

where the unknown parameter vector $\boldsymbol{\theta}$ is replaced with its ML estimate $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(\infty)}$. Note that estimates of the shadow powers in decibels are being utilized by most handoff algorithms, as well as for channel access and power control, see [56].

We now discuss computing the conditional expectation in (4.21) and maximizing (4.22c). The approximation (4.21b) was derived by applying the change-of-variable transformation (4.8) to the numerator and denominator in (4.21a) and using the Gauss-Hermite quadrature (4.14) to numerically evaluate the obtained integrals. Due to the cancellations of the common terms in the numerator and denominator of (4.21a), the formula (4.21b) is remarkably simplified.

The computation of $m^{(i+1)}$ requires maximizing (4.22c), which was performed using the Newton-Raphson method, which is embedded within the "outer" EM iteration. We choose the initial values as [76, Ch. 8.3.6]:

$$m_{\text{init}}^{(i+1)} = \frac{3 + 2\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})}{2\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})\cdot[3+\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})]}, \quad (4.24)$$

where

$$\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) = \mathcal{T}_1(\boldsymbol{y};\boldsymbol{\theta}^{(i)}) + \mathcal{T}_3(\boldsymbol{y};\boldsymbol{\theta}^{(i)})$$

The Newton-Raphson iteration for maximizing (4.22c) converges rapidly when initialized with the approximate ML estimate in (4.24). The derivatives needed to implement this iteration are shown in (C.5) in the Appendix C, where $\varphi(\boldsymbol{y},\boldsymbol{u})$ should be replaced with $\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})$.

4.2.3 Choosing the initial values

The proposed algorithms can be initialized by fitting the simple lognormal shadowing model, which leads to the following initial estimates of the shadowing parameters:

$$\mu^{(0)} = \frac{1}{K}\sum_{k=1}^{K} 10\log_{10}\overline{y}_k, \tag{4.25a}$$

$$(\sigma^2)^{(0)} = \frac{1}{K}\sum_{k=1}^{K}[(10\log_{10}\overline{y}_k)^2] - (\mu^{(0)})^2. \tag{4.25b}$$

For $N > 1$, an initial estimate of m for starting the EM iteration can be obtained using an approximate estimator similar to (4.24):

$$m^{(0)} = \frac{3 + 2\varphi_0(\boldsymbol{y})}{2\varphi_0(\boldsymbol{y})\cdot[3 + \varphi_0(\boldsymbol{y})]},$$

where

$$\varphi_0(\boldsymbol{y}) = \frac{1}{K}\cdot\sum_{k=1}^{K}(\ln\overline{y}_k) - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)$$

is obtained by replacing the unobserved shadow powers u_k with their sample mean $\overline{y}_k$, $k = 1, 2, \ldots, K$ in the expression for $\varphi(\boldsymbol{y},\boldsymbol{u})$ in (C.6).

4.2.4 Cramer-Rao bounds

The CRB matrix for the unknown parameter vector $\boldsymbol{\theta}$ can be computed by inverting the expected negative Hessian matrix, where the expectation is performed with respect to the distribution of $\boldsymbol{y}$ (see [18] and [58]):

$$\mathrm{CRB}(\boldsymbol{\theta}) = -\left\{\mathrm{E}_y\left[\frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\right]\right\}^{-1}. \tag{4.26}$$

The above expectation requires multidimensional integration, which can be performed using Monte Carlo integration.

4.3 Numerical examples

The numerical examples presented here assess the estimation accuracy of the ML estimates of $\boldsymbol{\theta}$. Our performance metric is the Mean-square error (MSE) of an estimator, calculated using 60000 independent trials. Note that the MSEs of the Newton-Raphson and EM algorithms coincide, since the convergence points of both algorithms coincide and are equal to the ML estimate of $\boldsymbol{\theta}$.

The measurements $y_k(t)$, $t = 1, 2, \ldots, N$, $k = 1, 2, \ldots, K$ were simulated from the composite gamma-lognormal distribution with $N = 10$ samples per observation interval, $10 \leq K \leq 100$, $m = 1$ (i.e. Rayleigh fading), $\mu = 5$ dB and $\sigma = 3$ dB.

The quadrature order of the Gauss-Hermite approximation in (4.14) and (4.21b) was $L = 20$. In Figs. 4.1–4.3 we show the MSEs and corresponding CRBs [3] for the ML estimates of m, μ, and σ^2, respectively, as functions of the number of observation intervals K. The ML estimators are "almost efficient" in this scenario, i.e. their MSEs are very close to the corresponding CRBs.

In Figs. 4.1–4.3, we also show the performance of the proposed algorithms when the shadow powers u_k are correlated. We adopt the first-order autoregressive AR(1) correlation model for the shadow process in decibels (see [43, 56]):

$$10 \log_{10} u_k = \alpha \cdot 10 \log_{10} u_{k-1} + \omega_k, \tag{4.27}$$

where ω_k are i.i.d Gaussian random variables with mean $(1 - \alpha) \cdot \mu = (1 - \alpha) \cdot 5$ dB and standard deviation $\sqrt{1 - \alpha^2} \cdot \sigma = \sqrt{1 - \alpha^2} \cdot 3$ dB. The MSEs of the proposed estimators are shown for $\alpha = 0.5$ and $\alpha = 0.9$. Interestingly, the MSE performance of the estimator of m is insensitive to the value of the correlation coefficient α, see Fig. 4.1. However, the estimation of the shadowing parameters μ and σ^2 is affected by α, see Figs. 4.2 and 4.3.

We now evaluate the computational efficiency of the proposed methods. In Fig. 4.4, we show the numbers of iterations and CPU times of the EM, Newton-Raphson, and BFGS algorithms, as functions of K. The EM algorithm converged within 12 iteration steps [4], whereas the Newton-Raphson algorithm converged in four iterations.

[3] The CRB matrix was computed using (4.26), where the expectation with respect to the distribution of $\boldsymbol{y}$ was performed using Monte Carlo integration with 60000 trials.

[4] The scalar Newton-Raphson iteration embedded within the "outer" EM iteration converged within three steps and has low computational complexity compared with the expectation step in (4.20).

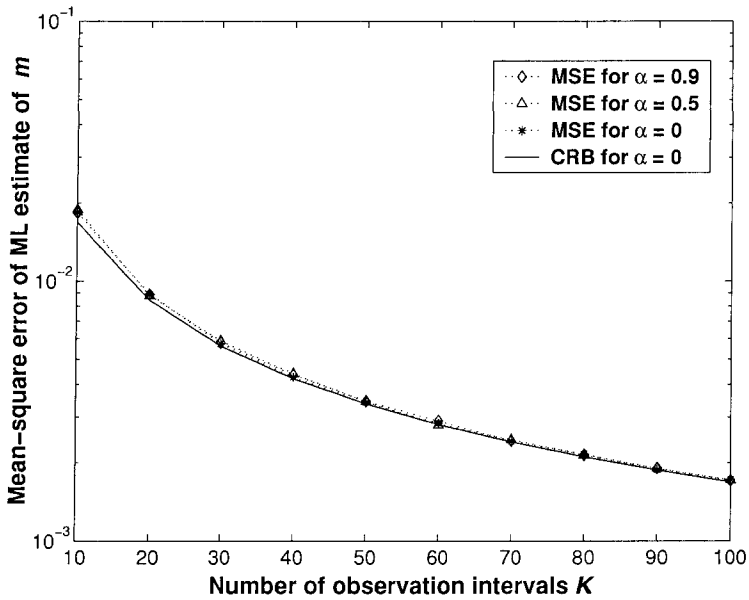

Figure 4.1 Mean-square error and Cramér-Rao bound for the proposed estimator of m as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$.

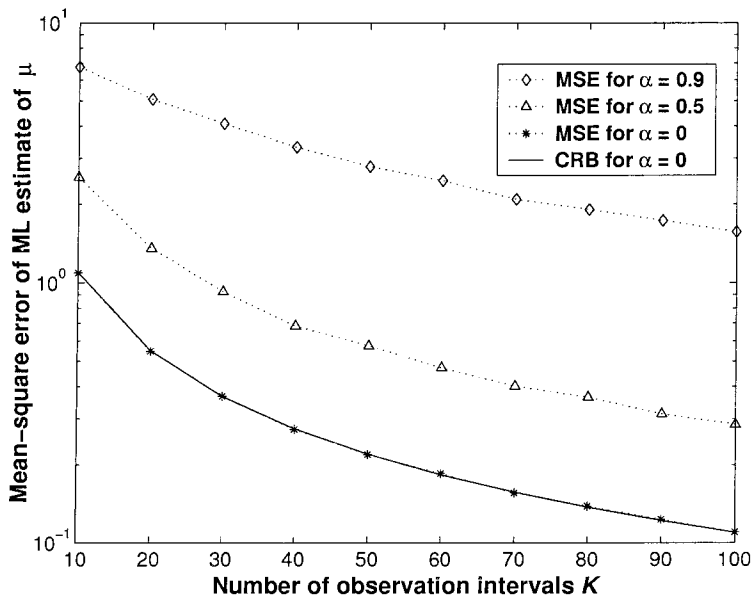

Figure 4.2 Mean-square error and Cramér-Rao bound for the proposed estimator of μ as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$.

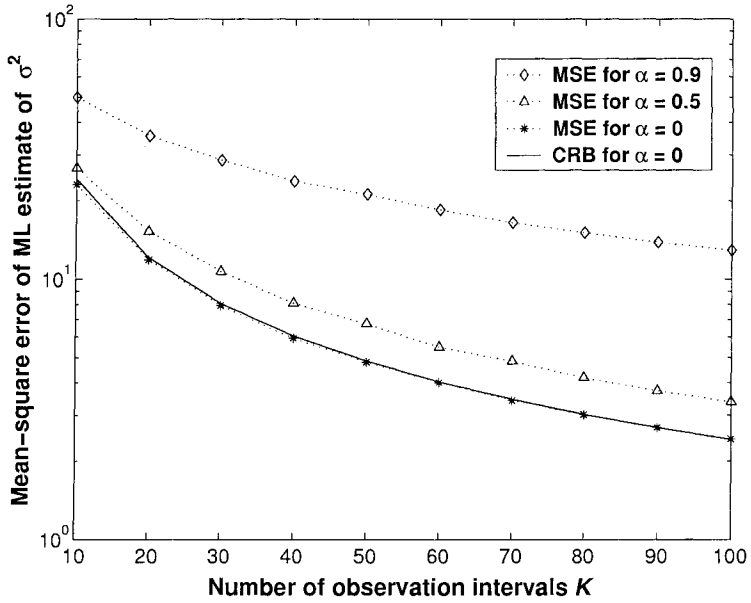

Figure 4.3 Mean-square error and Cramér-Rao bound for the proposed estimator of σ^2 as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$ and $\alpha = 0.9$.

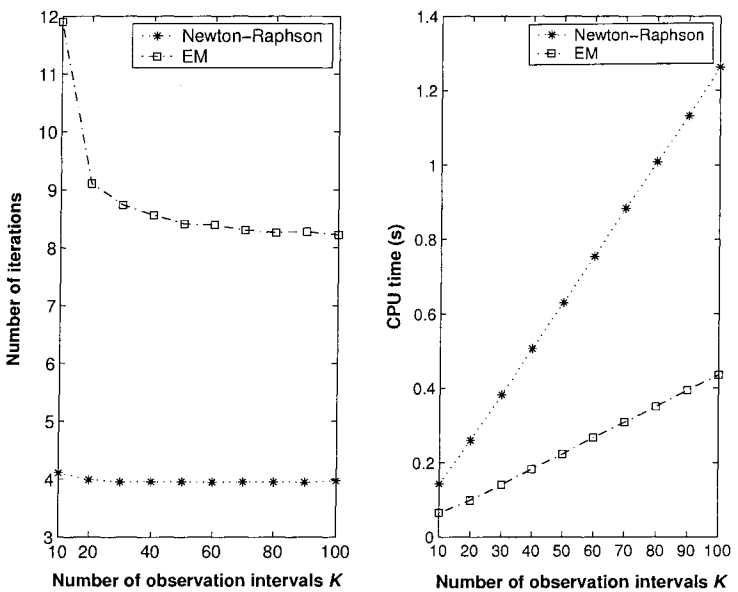

Figure 4.4 Average number of iterations (left) and CPU time (right) of the EM, Newton-Raphson, and BFGS algorithms as a function of K.

In terms of CPU time, however, the EM algorithm was faster than the Newton-Raphson method, which can be explained by the fact that a single EM iteration is significantly faster than a Newton-Raphson iteration. In particular, the Newton-Raphson algorithm requires computing and inverting the Hessian matrix, which counterbalances its advantage in speed of convergence. This is a well-known drawback of the Newton-Raphson method (see [18, Ch. 2.4.3] and [98, Ch. 4.3.2]) which can be surmounted if the derivatives in (4.12)–(4.13) are computed in parallel. The BFGS quasi-Newton algorithm converged in five iterations and outperformed the Newton-Raphson method in terms of CPU time; however it was slower than the EM algorithm.

4.4 Summary

We derived maximum likelihood methods for estimating the parameters of composite gamma-lognormal fading channels. The ML estimates of the unknown fading and shadowing parameters were computed using Newton-Raphson and EM algorithms. We also applied the BFGS quasi-Newton algorithm, discussed initializing the proposed algorithms, and computed Cramer-Rao bounds for the unknown parameters. The proposed algorithms can be extended to other composite fading-shadowing scenarios, such as Rician-lognormal [28, 101] and MIMO fading scenarios.

CHAPTER 5. ESTIMATION OF MIMO RICIAN FADING CHANNEL STATISTICS

5.1 Introduction

As we have seen in Chapter 4, the statistic properties of fading channels play an important role in wireless communication design and analysis. This problem has not received sufficient attention. It is widely assumed that the channel statistic properties are *exactly known* in the analysis; researches are more focused on estimating the channel fading gains rather than estimating the channel statistic properties, see for example [44], [55]. Compared with the channel statistic estimation, the channel estimation may be more costly and difficult to obtain at the transmitter side. Moreover, the channel estimation may not be reliable in certain cases when long training sequence cannot be used, such as wideband CDMA [24]. Therefore, some researchers have proposed communication schemes based on channel statistic properties only, without using CSI [108]. However, an efficient estimation of these statistics are not provided in the literature.

There are several references available for the channel statistic estimation problem. In [24], for fading channels without shadowing, Chaufry *et al.* proposed a method that leads to consistent estimation of the second-order statistics for a single-input single-output (SISO) frequency-selective Rayleigh fading channel. They also claimed that the same method can be extended to Rician fading channels. Marzetta [75] studied a multivariate complex Rician channel for the polarimetric synthetic aperture radar (SAR) system. However, the estimation of the mean and covariance parameters is solely based on noiseless measurements and signal energy without taking into account for the phase information in the received signals. We can see that an efficient method to estimate the statistic properties of fading channels based on *complex noisy measurements*, which contain both the phase and amplitude information of the received signals, is still lacking.

In this chapter, we study the estimation problem of MIMO Rician and Rayleigh block-fading channels. We present ML and REML methods for estimating statistic properties using *complex noisy measurements* (containing *both* the phases and amplitudes of the received signals) from multiple coherent intervals. The estimation methods developed herein are applicable to sensor array processing for moving arrays, which shares a similar measurement model [106].

5.2 Measurement model

For MIMO Rician block-fading channels, the fading coefficients are constant within a coherent interval, but vary randomly from one coherent interval to another. Assume that spatiotemporal measurements from K coherent intervals are available. At time t in the kth coherent interval, the measurement model is:

$$\boldsymbol{y}_k(t) = H_k \boldsymbol{\phi}_k(t) + \boldsymbol{e}_k(t), \qquad t = 1, \ldots, N,\ k = 1, \ldots, K \tag{5.1}$$

where

- $\boldsymbol{y}_k(t)$ an $n_{\mathrm{R}} \times 1$ data vector received by an array of n_{R} antennas;
- H_k is the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ channel response matrix;
- $\boldsymbol{\phi}_k(t)$ is an $n_{\mathrm{T}} \times 1$ vector of signals transmitted by n_{T} transmitter antennas;
- $\boldsymbol{e}_k(t)$ is additive white complex Gaussian noise with

$$\mathrm{E}[\boldsymbol{e}_{k_1}(t_1)\, \boldsymbol{e}_{k_2}(t_2)^H] = \delta_{k_1,k_2}\delta_{t_1,t_2} \cdot \sigma^2 I_{n_{\mathrm{R}}}$$

Here, δ_{ij} denotes the Kronecker delta symbol, I_n is the identity matrix of size n, and "H" is the Hermitian transpose.

We assume that the transmitted symbols are *known*, i.e., the coherent intervals contain pilot symbols. Stacking all N time samples from the kth coherent interval into a single vector and using [51, eq.(2.11)], we have

$$\boldsymbol{y}_k = Z_k \boldsymbol{h}_k + \boldsymbol{e}_k \tag{5.2}$$

where

- $\boldsymbol{h}_k = \mathrm{vec}\,\{H_k\}$ is the $n_{\mathrm{R}} n_{\mathrm{T}} \times 1$ channel response vector ;

- $\boldsymbol{y}_k = [\boldsymbol{y}_k(1)^T, \boldsymbol{y}_k(2)^T, \ldots, \boldsymbol{y}_k(N)^T]^T$ is the $n_{\mathrm{R}}N \times 1$ spatiotemporal data vector;
- $\boldsymbol{e}_k = [\boldsymbol{e}_k(1)^T, \boldsymbol{e}_k(2)^T, \ldots, \boldsymbol{e}_k(N)^T]^T$;
-

$$Z_k = \Phi_k^T \otimes I_{n_{\mathrm{R}}}, \quad \Phi_k = [\boldsymbol{\phi}_k(1) \cdots \boldsymbol{\phi}_k(N)]. \tag{5.3}$$

Here, "^{T}" denotes a transpose, $\otimes$ is the Kronecker product, and the vec operator stacks the columns of a matrix one below another into a single column vector (see [51, ch.16]). The $n_{\mathrm{T}} \times N$ matrix Φ_k is the signal matrix in the kth coherent interval and the $n_{\mathrm{R}}N \times n_{\mathrm{R}}n_{\mathrm{T}}$ matrix Z_k is the "augmented" signal matrix.

We now decompose the channel response vector $\boldsymbol{h}_k$ into a sum of the deterministic LOS component $\boldsymbol{h}_{\mathrm{LOS},k}$ and random scattering component $\boldsymbol{h}_{\mathrm{SC},k}$:

$$\boldsymbol{h}_k = \boldsymbol{h}_{\mathrm{LOS},k} + \boldsymbol{h}_{\mathrm{SC},k}. \tag{5.4}$$

For the LOS component, we use the following model:

$$\boldsymbol{h}_{\mathrm{LOS},k} = A_{\mathrm{LOS},k}\, \boldsymbol{x} \tag{5.5}$$

where $A_{\mathrm{LOS},k}$, $k = 1, 2, \ldots, K$ are $n_{\mathrm{R}}n_{\mathrm{T}} \times r$ matrices ; $\boldsymbol{x}$ is an $r \times 1$ vector of unknown complex coefficients. The model (5.5) is fairly general and can be used to describe the LOS component when dual-polarized antenna elements are employed.

(i) **unstructured LOS model**

When the transmitter and receiver LOS array responses are not known and the variation of the LOS component from one coherent interval to another can be described with a simple Doppler-shift model, we have $A_{\mathrm{LOS},k}$ as

$$A_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk) \cdot I_{n_{\mathrm{R}}n_{\mathrm{T}}} \tag{5.6}$$

where $\omega_{\mathrm{D,LOS}}$ is the LOS Doppler shift (in radians) due to the relative movement between receiver and transmitter. [1]. We assume that $\omega_{\mathrm{D,LOS}}$ is *known,* unless specified otherwise (see Section 5.4). Then, (5.5) simplifies to $\boldsymbol{h}_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk) \cdot \boldsymbol{x}$, where $\boldsymbol{x}$ is the $n_{\mathrm{R}}n_{\mathrm{T}} \times 1$ unstructured LOS array response vector.

[1]Note that $\omega_{\mathrm{D,LOS}}$ corresponds to the continuous-time LOS Doppler shift $\Omega_{\mathrm{D,LOS}}\omega_{\mathrm{D,LOS}}/\Delta t$, where Δt is the symbol duration.

(ii) **structured LOS model**

If the transmitter and receiver LOS array responses are known, we utilize the structured LOS array response model (see [32], [33], [39], and [40]):

$$A_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk) \cdot \boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}} \tag{5.7}$$

where $\boldsymbol{a}_{\mathrm{T,LOS}}$ and $\boldsymbol{a}_{\mathrm{R,LOS}}$ are the transmitter and receiver LOS array response vectors of dimensions $n_{\mathrm{T}} \times 1$ and $n_{\mathrm{R}} \times 1$, respectively. Now, (5.5) becomes

$$\boldsymbol{h}_{\mathrm{LOS},k} = \boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}} \exp(j\omega_{\mathrm{D,LOS}}Nk) \cdot x \tag{5.8}$$

where $\boldsymbol{x} = x$ is the scalar LOS complex amplitude.

To describe the channel variation from one coherent interval to another, we assume that the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}$ are zero-mean independent and identically distributed (i.i.d.) complex Gaussian with covariance matrix:

$$\Psi = \mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\boldsymbol{h}_{\mathrm{SC},k}^H], \quad k = 1, 2, \ldots, K. \tag{5.9}$$

In addition, $\boldsymbol{h}_{\mathrm{SC},k}$ and noise vectors $\boldsymbol{e}_k$ are assumed to be independent, i.e. $\mathrm{E}[\boldsymbol{e}_{k_1}\boldsymbol{h}_{\mathrm{SC},k_2}^H] = 0$, where $k_1, k_2 \in \{1, 2, \ldots, K\}$.

Our **goal** is to estimate the unknown parameters in the above model:

- LOS coefficient vector $\boldsymbol{x}$ (or scalar x);
- spatial fading covariance matrix Ψ;
- noise variance σ^2.

which are assumed to be constant over the K coherent intervals. This assumption is justified by the fact that the channel mean and covariance parameters depend on large-scale variations in the scattering environment, which are typically slow (see [64] and [80]).

Define the vector of unknown parameters:

$$\boldsymbol{\rho} = [\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T, \boldsymbol{\gamma}^T]^T \tag{5.10}$$

where $\boldsymbol{\gamma} = [\sigma^2, \boldsymbol{\psi}^T]^T$ is the vector of variance components and $\boldsymbol{\psi}$ describes a parametrization of the fading covariance matrix Ψ. We consider two models for Ψ:

(i) **unstructured model (correlated fading):**

$$\boldsymbol{\psi} = [\mathrm{Re}\{\mathrm{vech}(\Psi)\}^T, \mathrm{Im}\{\underline{\mathrm{vech}}(\Psi)\}^T]^T$$

where the correlation structure of the fading channel is completely unknown [2];

(ii) **diagonal model (independent fading):**

$$\boldsymbol{\psi} = [\psi_1, \psi_2, \ldots, \psi_{n_{\mathrm{R}} n_{\mathrm{T}}}]^T$$

where the fading-channel coefficients are independent with non-identical variances.

Note that $\boldsymbol{\psi}$ is a valid parametrization *only if* the fading covariance matrix Ψ is a positive semidefinite Hermitian matrix.

In the following sections, we derive ML and REML algorithms for estimating the unknown parameter vector $\boldsymbol{\rho}$ under correlated and independent fading scenarios described above. We also derive efficient algorithms for estimating $\boldsymbol{\rho}$ when $\Phi_k \Phi_k^H$ is constant.

5.3 ML and REML estimation

We first outline the ML and REML approaches to estimate $\boldsymbol{\rho}$ and then present the proposed algorithms. Under the measurement model in Section 5.2, the spatiotemporal data vectors $\boldsymbol{y}_k$ are independent, complex Gaussian with means and covariances

$$\mathrm{E}[\boldsymbol{y}_k] = \Upsilon_k \boldsymbol{x} = Z_k A_{\mathrm{LOS},k} \boldsymbol{x} \tag{5.11a}$$

$$\Sigma_k(\boldsymbol{\gamma}) = Z_k \Psi Z_k^H + \sigma^2 I_{n_{\mathrm{R}} N}. \tag{5.11b}$$

Thus, the log-likelihood function to be maximized is the logarithm of the joint PDF of $\boldsymbol{y}_k$, $k = 1, 2, \ldots, K$:

$$L(\boldsymbol{\rho}) = -(\boldsymbol{y} - \Upsilon \boldsymbol{x})^H \Sigma(\boldsymbol{\gamma})^{-1} (\boldsymbol{y} - \Upsilon \boldsymbol{x}) - \ln |\pi \Sigma(\boldsymbol{\gamma})| \tag{5.12}$$

where $|\cdot|$ denotes the determinant, $\boldsymbol{y} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T \cdots \boldsymbol{y}_K^T]^T$, $\Upsilon = [\Upsilon_1^T, \Upsilon_2^T \cdots \Upsilon_K^T]^T$ is an $K n_{\mathrm{R}} N \times r$ matrix of rank r, and $\Sigma(\boldsymbol{\gamma})$ is an $K n_{\mathrm{R}} N \times K n_{\mathrm{R}} N$ block-diagonal matrix:

$$\Sigma(\boldsymbol{\gamma}) = \mathrm{bdiag}\{\Sigma_1(\boldsymbol{\gamma}), \Sigma_2(\boldsymbol{\gamma}) \cdots \Sigma_K(\boldsymbol{\gamma})\}.$$

[2] Here the vech and $\underline{\mathrm{vech}}$ operators create a single column vector by stacking elements below the main diagonal columnwise; vech includes the main diagonal, whereas $\underline{\mathrm{vech}}$ omits it.

The estimate of the LOS coefficient vector $\boldsymbol{x}$ that maximizes (5.12) for any fixed $\boldsymbol{\gamma}$ is given by

$$\begin{aligned} \widehat{\boldsymbol{x}}(\boldsymbol{\gamma}) &= [\Upsilon^H \Sigma(\boldsymbol{\gamma})^{-1} \Upsilon]^{-1} \Upsilon^H \Sigma(\boldsymbol{\gamma})^{-1} \\ &= \left[\sum_{k=1}^K \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) \mathcal{A}_{\mathrm{LOS},k}\right]^{-1} \sum_{k=1}^K \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) \boldsymbol{z}_k \end{aligned} \tag{5.13}$$

where

$$W_k(\boldsymbol{\gamma}) = (\sigma^2 {Z_k}^H Z_k + {Z_k}^H Z_k \Psi {Z_k}^H Z_k)^{-1} \tag{5.14a}$$

$$\mathcal{A}_{\mathrm{LOS},k} = {Z_k}^H \Upsilon_k {Z_k}^H Z_k A_{\mathrm{LOS},k} \tag{5.14b}$$

$$\boldsymbol{z}_k = {Z_k}^H \boldsymbol{y}_k = \mathrm{vec}\,(Y_k {\Phi_k}^H) \tag{5.14c}$$

and

$$Y_k = [\boldsymbol{y}_k(1), \boldsymbol{y}_k(2) \cdots, \boldsymbol{y}_k(N)]$$

is the *spatiotemporal data matrix* in the kth coherent interval. The second equalities in (5.13) and (5.14c) follow by using (D.2) in Appendix D and [51, eq.(16.2.11)], respectively. Replacing $\boldsymbol{x}$ in (5.12) with its ML estimate in (5.13) results in the concentrated log-likelihood function:

$$L(\boldsymbol{\gamma}|\widehat{\boldsymbol{x}}(\boldsymbol{\gamma})) = -[\boldsymbol{y} - \Upsilon \widehat{\boldsymbol{x}}(\boldsymbol{\gamma})]^H \Sigma(\boldsymbol{\gamma})^{-1} [\boldsymbol{y} - \Upsilon \widehat{\boldsymbol{x}}(\boldsymbol{\gamma})] - \ln|\pi\Sigma(\boldsymbol{\gamma})| - \boldsymbol{y}^H \Pi(\boldsymbol{\gamma})\, \boldsymbol{y} - \ln|\pi\Sigma(\boldsymbol{\gamma})| \tag{5.15}$$

where

$$\Pi(\boldsymbol{\gamma}) = \Sigma(\boldsymbol{\gamma})^{-1} - \Sigma(\boldsymbol{\gamma})^{-1} \Upsilon [\Upsilon^H \Sigma(\boldsymbol{\gamma})^{-1} \Upsilon]^{-1} \Upsilon^H \Sigma(\boldsymbol{\gamma})^{-1}.$$

Note that (5.15) is a nonlinear function of the variance-component parameters $\boldsymbol{\gamma}$ that generally needs to be maximized using iterative algorithms.

Once the ML estimate $\widehat{\boldsymbol{\gamma}}$ is computed by maximizing (5.15), the ML estimate of $\boldsymbol{x}$ is obtained by substituting $\widehat{\boldsymbol{\gamma}}$ into (5.13). Interestingly, closed-form solutions for the ML estimates of $\boldsymbol{\gamma}$ exist under the correlated fading scenario with constant $\Phi_k {\Phi_k}^H$, see the discussion in Section 5.3.2.

We now introduce the REML method for estimating the unknown variance components. The REML estimate of $\boldsymbol{\gamma}$ is obtained by filtering out the deterministic Rician component from the received data, and applying ML estimation to the *error contrasts* (i.e. filtered data), which corresponds to maximizing the REML log-likelihood function (see Appendix D.2)

$$L_{\mathrm{REML}}(\boldsymbol{\gamma}) = L(\boldsymbol{\gamma}|\widehat{\boldsymbol{x}}(\boldsymbol{\gamma})) - \ln|\Upsilon^H \Sigma(\boldsymbol{\gamma})^{-1} \Upsilon| L(\boldsymbol{\gamma}|\widehat{\boldsymbol{x}}(\boldsymbol{\gamma})) - \ln\left|\sum_{k=1}^K \Upsilon_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} \Upsilon_k\right| \tag{5.16}$$

with respect to $\boldsymbol{\gamma}$, where $L(\boldsymbol{\gamma}|\widehat{\boldsymbol{x}}(\boldsymbol{\gamma}))$ is the concentrated log-likelihood in (5.15). We can also derive (5.16) by using an *integrated-likelihood approach* for eliminating nuisance parameters [17] [3]. Note that the REML method provides only estimates of the variance components $\boldsymbol{\gamma}$; however, a good estimate of $\boldsymbol{x}$ is obtained by substituting the REML estimate of $\boldsymbol{\gamma}$ into (5.13), which we call the "REML" estimate of $\boldsymbol{x}$ with a slight abuse of terminology.

Since the Rician component has been filtered out, the REML estimate of $\boldsymbol{\gamma}$ is invariant to the value of $\boldsymbol{x}$, i.e. changing $\boldsymbol{x}$ does not alter the REML estimate of $\boldsymbol{\gamma}$. Finally, we comment that the REML estimates of the variance components have *smaller bias* than the corresponding ML estimates [29, 50].

5.3.1 Correlated fading with arbitrary $\Phi_k \Phi_k^H$

In this section, we present the ML and REML estimates of the unknown parameters based on ECME algorithms for the correlated fading scenario.

5.3.1.1 ECME algorithm for ML estimation

An ECME algorithm maximizes *either* the expected complete-data log-likelihood function [4] *or* the actual observed-data log-likelihood, see [66], [77, ch. 5.7], and [78]. Here, we treat the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}$, $k = 1, 2, \ldots, K$ as the *unobserved* data and obtain the following ECME algorithm, see appendix D.3.1.

$$W_k^{(i)} = [(\sigma^2)^{(i)} Z_k{}^H Z_k + Z_k{}^H Z_k \Psi^{(i)} Z_k{}^H Z_k]^{-1} \tag{5.17a}$$

$$\boldsymbol{x}^{(i)} = \Big(\sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)} \mathcal{A}_{\mathrm{LOS},k} \Big)^{-1} \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)} \boldsymbol{z}_k \tag{5.17b}$$

$$\boldsymbol{h}_{\mathrm{SC},k}^{(i)} = \Psi^{(i)} Z_k{}^H Z_k W_k^{(i)} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}) \tag{5.17c}$$

for $k = 1, 2, \ldots, K$, and

$$(\sigma^2)^{(i+1)} = \widetilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{K n_{\mathrm{R}} N} \sum_{k=1}^{K} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)})^H W_k^{(i)} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}) \tag{5.18a}$$

$$\Psi^{(i+1)} = \Psi^{(i)} + \frac{1}{K} \sum_{k=1}^{K} \Big\{ \boldsymbol{h}_{\mathrm{SC},k}^{(i)} (\boldsymbol{h}_{\mathrm{SC},k}^{(i)})^H - \Psi^{(i)} Z_k{}^H Z_k W_k^{(i)} Z_k{}^H Z_k \Psi^{(i)} \Big\} \tag{5.18b}$$

[3] Here we treat $\boldsymbol{x}$ as a nuisance parameter vector and integrate it out using a noninformative prior.

[4] The expectation is computed with respect to the conditional distribution of the unobserved data given the observed measurements.

where

$$\begin{aligned}\widetilde{\sigma}^2 &= \frac{1}{Kn_{\mathrm{R}}N}\sum_{k=1}^{K}\left[\boldsymbol{y}_k^H\boldsymbol{y}_k - \boldsymbol{z}_k^H(Z_k{}^HZ_k)^{-1}\boldsymbol{z}_k\right] \\ &= \frac{1}{Kn_{\mathrm{R}}N}\sum_{k=1}^{K}\mathrm{tr}\left\{Y_k\left[I_N - \Phi_k{}^H(\Phi_k\Phi_k{}^H)^{-1}\Phi_k\right]Y_k{}^H\right\}.\end{aligned} \tag{5.19}$$

Note $\widetilde{\sigma}^2$ do not depend on the unknown parameters and can be computed beforehand. Also, $\boldsymbol{h}_{\mathrm{SC},k}^{(i)}$ are the (estimated) Bayesian linear-model MMSE estimators of the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}$.

Initialization: The above iteration can be initialized with $\Psi^{(-1)} = 0$, implying that the initial estimate $\boldsymbol{x}^{(-1)}$ of the LOS coefficient vector is simply its linear *least-squares (LS)* estimate.

$$\begin{aligned}\boldsymbol{x}^{(-1)} &= \Big(\sum_{k=1}^{K}\Upsilon_k^H\Upsilon_k\Big)^{-1}\sum_{k=1}^{K}\Upsilon_k^H\boldsymbol{y}_k \\ &= \Big(\sum_{k=1}^{K}A_{\mathrm{LOS},k}^HZ_k{}^HZ_kA_{\mathrm{LOS},k}\Big)^{-1}\sum_{k=1}^{K}A_{\mathrm{LOS},k}^H\boldsymbol{z}_k.\end{aligned} \tag{5.20}$$

After computing $\boldsymbol{x}^{(-1)}$, a good initial estimate of Ψ is its *modified method-of-moments estimate* (similar to [102, p. 244]):

$$\begin{aligned}\Psi^{(0)} &= \Big[\frac{1}{K}\sum_{k=1}^{K}(Z_k{}^HZ_k)^{-1}\left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\,\boldsymbol{x}^{(-1)}\right)\left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\,\boldsymbol{x}^{(-1)}\right)^H(Z_k{}^HZ_k)^{-1}\Big] \\ &\quad - \min\{(\sigma^2)^{(0)},\widehat{\lambda}\}\cdot\frac{1}{K}\sum_{k=1}^{K}(Z_k{}^HZ_k)^{-1}\end{aligned} \tag{5.21}$$

where $(\sigma^2)^{(0)}$ is a method-of-moments estimate of σ^2:

$$\begin{aligned}(\sigma^2)^{(0)} &= \frac{1}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})}\sum_{k=1}^{K}\left[\boldsymbol{y}_k^H\boldsymbol{y}_k - \boldsymbol{z}_k^H(Z_k{}^HZ_k)^{-1}\boldsymbol{z}_k\right] \\ &= \frac{Kn_{\mathrm{R}}N}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})}\cdot\widetilde{\sigma}^2\end{aligned} \tag{5.22}$$

which is a good initial estimate of σ^2, and $\widehat{\lambda}$ is the smallest generalized eigenvalue of the matrices $(1/K)\cdot\sum_{k=1}^{K}(Z_k{}^HZ_k)^{-1}\left(\boldsymbol{z}_k-\mathcal{A}_{\mathrm{LOS},k}\,\boldsymbol{x}^{(-1)}\right)\left(\boldsymbol{z}_k-\mathcal{A}_{\mathrm{LOS},k}\,\boldsymbol{x}^{(-1)}\right)^H(Z_k{}^HZ_k)^{-1}$ and $(1/K)\cdot\sum_{k=1}^{K}(Z_k{}^HZ_k)^{-1}$ [5].

The moment estimator $(\sigma^2)^{(0)}$ is obtained by pre- and post-multiplying $Z_k{}^H(\boldsymbol{y}_k-\mathrm{E}[\boldsymbol{y}_k])(\boldsymbol{y}_k-\mathrm{E}[\boldsymbol{y}_k])^HZ_k$ by $(Z_k{}^HZ_k)^{-1}$, summing over $k=1,2\ldots,K$, taking the expectation of the resulting expression, and solving for Ψ. In addition, we also applied a modification similar to [102, p. 244] to ensure that $\Psi^{(0)}$ is always a valid covariance matrix.

[5] Note that $(Z_k{}^HZ_k)^{-1}$ can be efficiently computed as $(\Phi_k{}^*\Phi_k{}^T)^{-1}\otimes I_{n_{\mathrm{R}}}$.

5.3.1.2 ECME algorithm for REML estimation

An ECME algorithm for REML estimation of $\boldsymbol{\rho}$ follows by replacing (5.18a) and (5.18b) with

$$(\sigma^2)^{(i+1)} = \frac{Kn_{\mathrm{R}}N}{Kn_{\mathrm{R}}N-r}\cdot\widetilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{Kn_{\mathrm{R}}N-r}\cdot \sum_{k=1}^{K}(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)})^H W_k^{(i)}\,(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)}) \tag{5.23a}$$

$$\begin{aligned}\Psi^{(i+1)} = \Psi^{(i)} + \frac{1}{K}\sum_{k=1}^{K}\Big\{\boldsymbol{h}_{\mathrm{SC},k}^{(i)}(\boldsymbol{h}_{\mathrm{SC},k}^{(i)})^H - \Psi^{(i)}Z_k{}^H Z_k\cdot\Big[W_k^{(i)} \\ -W_k^{(i)}\mathcal{A}_{\mathrm{LOS},k}\Big(\sum_{l=1}^{K}\mathcal{A}_{\mathrm{LOS},l}^H W_l^{(i)}\mathcal{A}_{\mathrm{LOS},l}\Big)^{-1}\mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)}\Big]\,Z_k{}^H Z_k\,\Psi^{(i)}\Big\}\end{aligned} \tag{5.23b}$$

in the iteration (5.17)–(5.18) and keeping the other steps intact. See Appendix D.3.2,

The above ML and REML ECME algorithms *always* converge to estimates that are in the parameter space: $(\sigma^2)^{(i)} \geq 0$ and $\Psi^{(i)} \geq 0$ at each iteration step i, provided that the initial values are valid. This is an important general property of the EM and related algorithms, such as ECME [81, ch. 12.4]. Similar argument applies to the REML case since the ECME REML algorithm is simply the ECME ML algorithm applied to the error contrasts.

The above algorithms for Rayleigh-fading are obtained simply by removing the step (5.17b) and setting $\mathcal{A}_{\mathrm{LOS},k} = 0$.

5.3.2 Correlated fading with constant $\Phi_k\Phi_k{}^H$

In this section, we propose a computationally efficient alternating-projection ML algorithm for the case where $\Phi_k\Phi_k{}^H$ are independent of k. This condition holds for many practically important signaling schemes, e.g. unitary space-time codes [53] and space-time block codes based on orthogonal designs [4, 97]. Assume

$$\Phi_k\Phi_k{}^H = \Gamma_\Phi \tag{5.24}$$

It follows that $Z_k{}^H Z_k$ also does not depend on k,i.e.,

$$C = Z_k{}^H Z_k = \Gamma_\Phi{}^T \otimes I_{n_{\mathrm{R}}}. \tag{5.25}$$

In this case, there exists a closed-form expression for the ML estimate of noise variance σ^2 (see Appendix D.4):

$$\begin{aligned} \widehat{\sigma}^2_{\mathrm{ML}} &= \frac{1}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})}\sum_{k=1}^{K}(\boldsymbol{y}_k^H\boldsymbol{y}_k - \boldsymbol{z}_k^H C^{-1}\boldsymbol{z}_k) \\ &= \frac{1}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})}\cdot\sum_{k=1}^{K}\mathrm{tr}\left[Y_k\left(I_N - \Phi_k{}^H\Gamma_\Phi{}^{-1}\Phi_k\right)Y_k^H\right] \end{aligned} \tag{5.26}$$

which coincides with the moment estimator in (5.22). The exact ML estimates of $\boldsymbol{x}$ and Ψ can be computed by iterating between the following two steps:

$$W^{(i)} = (\widehat{\sigma}^2_{\mathrm{ML}}C + C\Psi^{(i)}C)^{-1} \tag{5.27a}$$

$$\boldsymbol{x}^{(i)} = \Big[\sum_{k=1}^{K}\mathcal{A}^H_{\mathrm{LOS},k}W^{(i)}\mathcal{A}_{\mathrm{LOS},k}\Big]^{-1}\sum_{k=1}^{K}\mathcal{A}^H_{\mathrm{LOS},k}\,W^{(i)}\boldsymbol{z}_k \tag{5.27b}$$

$$\boldsymbol{\xi}_k^{(i)} = C^{-1}\left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\,\boldsymbol{x}^{(i)}\right) = \mathrm{vec}\left(Y_k\Phi_k{}^H\Gamma_\Phi{}^{-1}\right) - A_{\mathrm{LOS},k}\,\boldsymbol{x}^{(i)} \tag{5.27c}$$

for $k = 1, 2, \ldots, K$, and

$$\Psi^{(i+1)} = \Big[\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{\xi}_k^{(i)}\,(\boldsymbol{\xi}_k^{(i)})^H\Big] - \widehat{\sigma}^2_{\mathrm{ML}}\cdot(\Gamma_\Phi{}^T)^{-1}\otimes I_{n_{\mathrm{R}}}. \tag{5.28}$$

The above iteration increases the log-likelihood function at each cycle but may converge to solutions that are not in the parameter space. It can be shown that, if (5.24) holds, the estimators (5.26) and (5.28) are fixed points of the ECME iterations (5.18a) and (5.18b). In Appendix D.5.1, we also derive the CRB expressions for this scenario.

5.3.2.1 ML estimation for unstructured LOS array response model

Under the unstructured LOS array response model,

$$\begin{aligned} \widehat{\boldsymbol{x}}_{\mathrm{UML}} &= C^{-1}\cdot\frac{1}{K}\sum_{k=1}^{K}\exp(-j\omega_{\mathrm{D,LOS}}Nk)\,\boldsymbol{z}_k \\ &= \frac{1}{K}\sum_{k=1}^{K}\mathrm{vec}\left(Y_k\Phi_k{}^H\Gamma_\Phi{}^{-1}\right)\exp(-j\omega_{\mathrm{D,LOS}}Nk) \end{aligned} \tag{5.29a}$$

$$\widehat{\Psi}_{\mathrm{UML}} = \Big[\frac{1}{K}\sum_{k=1}^{K}\widehat{\boldsymbol{\xi}}_k\widehat{\boldsymbol{\xi}}_k^H\Big] - \widehat{\sigma}^2_{\mathrm{ML}}\cdot(\Gamma_\Phi{}^T)^{-1}\otimes I_{n_{\mathrm{R}}} \tag{5.29b}$$

are the the ML estimates of $\boldsymbol{x}$ and Ψ, where $\widehat{\sigma}^2_{\mathrm{ML}}$ is the ML estimate of σ^2 in (5.26) and

$$\widehat{\boldsymbol{\xi}}_k = \mathrm{vec}\left(Y_k\Phi_k{}^H\Gamma_\Phi{}^{-1}\right) - \exp(j\omega_{\mathrm{D,LOS}}Nk)\cdot\widehat{\boldsymbol{x}}_{\mathrm{UML}} \tag{5.30}$$

for $k = 1, 2, \ldots, K$. Here, (5.29a) follows by substituting (5.6) into (5.13) [see also (5.14b) and (5.25)] and (5.29b) is obtained by substituting $\widehat{\boldsymbol{x}}_{\mathrm{UML}}$ into (D.23b), see Appendix D.4.

If the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is *unknown*, its ML estimate can be computed by maximizing the following concentrated log-likelihood function:

$$\widehat{\omega}_{\mathrm{D,LOS}} = \arg \max_{\omega_{\mathrm{D,LOS}}} \boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}}N)^H \Big(\sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^H \Big)^{-1} \boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}}N) \tag{5.31}$$

where

$$\boldsymbol{z}_{\mathrm{DTFT}}(\omega) = \frac{1}{K} \cdot \sum_{k=1}^{K} \exp(-j\omega k)\, \boldsymbol{z}_k = \frac{1}{K} \cdot \sum_{k=1}^{K} \exp(-j\omega k)\, \mathrm{vec}\,(Y_k \Phi_k{}^H) \tag{5.32}$$

is proportional to the discrete-time Fourier transform (DTFT) of $\boldsymbol{z}_k$, $k = 1, 2, \ldots, K$. The above concentrated log-likelihood is obtained by replacing $\boldsymbol{x}$, Ψ, σ^2 and $Z_k{}^H Z_k$ in (D.4) with $\widehat{\boldsymbol{x}}_{\mathrm{UML}}$, $\widehat{\Psi}_{\mathrm{UML}}, \sigma^2_{\mathrm{ML}}$ and C, using [51, Theorem 18.1.1]), and applying a monotonic transformation. The classical algorithm for DTFT-based frequency estimation in [84, ch.6.4.4] can be easily extended and applied to maximizing (5.31).

5.3.2.2 AML estimation for structured LOS array response model

Due to the CRB decoupling between the mean and variance-component parameters, the ML estimate of Ψ for the unstructured LOS array response model in (5.29b) is asymptotically efficient under the structured LOS array response model. Hence, $\widehat{\Psi}_{\mathrm{UML}}$ is also an asymptotical ML (AML) estimate of Ψ, provided that it is positive semidefinite. Now, we obtain a closed-form AML estimate of the structured-array complex LOS amplitude by substituting (5.7) and $\widehat{\sigma}^2_{\mathrm{ML}}$ and $\widehat{\Psi}_{\mathrm{UML}}$ into (5.13):

$$\widehat{x}_{\mathrm{AML}} = \frac{(\boldsymbol{a}^H_{\mathrm{T,LOS}} \otimes \boldsymbol{a}^H_{\mathrm{R,LOS}})\, C\, \Xi^{-1}\, \boldsymbol{z}_{\mathrm{DTFT}}}{(\boldsymbol{a}^H_{\mathrm{T,LOS}} \otimes \boldsymbol{a}^H_{\mathrm{R,LOS}})\, C\, \Xi^{-1} C\, (\boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}})} \tag{5.33}$$

where

$$\Xi = \Big[\frac{1}{K} \sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k{}^H \Big] - \boldsymbol{z}_{\mathrm{DTFT}} \boldsymbol{z}_{\mathrm{DTFT}}{}^H. \tag{5.34}$$

To simplify the notation, we have omitted the dependence of $\boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}}N)$ on $\omega_{\mathrm{D,LOS}}N$ in (5.33). The above estimator is asymptotically efficient. However, it is based on the assumption that $\widehat{\Psi}_{\mathrm{UML}} \geq 0$ and performs poorly when this assumption does not hold (e.g., when K is small). See Figs. 5.2 and 5.5 in Section 5.4.

5.3.2.3 ML estimation for Rayleigh fading

Under the Rayleigh-fading scenario, the closed-form expressions for the ML estimates of σ^2 and Ψ are given by (5.26) and

$$\widehat{\Psi}_{\mathrm{ML}} = \Big[\frac{1}{K}\sum_{k=1}^{K}\mathrm{vec}\,(Y_k\,\Phi_k{}^H\Gamma_\Phi{}^{-1})\cdot\mathrm{vec}\,(Y_k\,\Phi_k{}^H\Gamma_\Phi{}^{-1})^H\Big] - \widehat{\sigma}^2_{\mathrm{ML}}\cdot(\Gamma_\Phi{}^T)^{-1}\otimes I_{n_{\mathrm{R}}}. \quad (5.35)$$

The closed-form ML estimates in (5.26), (5.29), and (5.35) can be used to implement noncoherent ML space-time receivers, which require fast estimation of the fading parameters.

The estimates of Ψ obtained using (5.27)–(5.28) and closed-form expressions (5.29b) and (5.35) are maximum likelihood only if they are positive semidefinite; otherwise, we can apply the ECME algorithm in Section 5.3.1.1 which always converges to solutions within the parameter space.

Clearly, a necessary condition for (5.28), (5.29b), and (5.35) to be positive semidefinite is: $K \geq n_{\mathrm{R}}n_{\mathrm{T}}$. The probability that (5.27)–(5.28), (5.29b), and (5.35) yield non-positive semidefinite estimates of Ψ is asymptotically zero as either $K \to \infty$ or $Nn_{\mathrm{R}} \to \infty$. For the unstructured LOS array response model, we can remove (5.17b) from the ML and REML ECME iterations and use the closed-form expression for the ML estimate of $\boldsymbol{x}$ in (5.29a).

5.3.3 Independent fading

In this section, we develop ECME ML and REML algorithms for estimating $\boldsymbol{\rho}$ under the independent fading scenario and further simplify them in the case where $\Phi_k\Phi_k{}^H$ is an identity matrix. Approximately independent fading occurs, for example, in virtual channel representations, see [64] and references therein. In [33], we also derived Henderson's methods [52] for this scenario which performed similarly to the algorithms proposed here.

5.3.3.1 ECME algorithm for ML estimation

The ECME ML algorithm for independent fading follows using arguments similar to those in Appendix D.3.1, where ECME algorithms were derived for the correlated fading scenario. It iterates between (5.17a)–(5.17c) for $k = 1, 2, \ldots, K$ and

$$(\sigma^2)^{(i+1)} = \widetilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{Kn_{\mathrm{R}}N}\cdot\sum_{k=1}^{K}(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)})^H\,W_k^{(i)}\,(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)}) \quad (5.36a)$$

$$(\psi_n)^{(i+1)} = (\psi_n)^{(i)} + \frac{1}{K}\sum_{k=1}^{K}\left\{|[\boldsymbol{h}_{\mathrm{SC},k}^{(i)}]_n|^2 - [(\psi_n)^{(i)}]^2 \cdot [Z_k{}^H Z_k W_k^{(i)} Z_k{}^H Z_k]_{n,n}\right\} \quad (5.36b)$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$, where $\widetilde{\sigma}^2$ has been defined in (5.19) and

$$\Psi^{(i)} = \mathrm{diag}\{(\psi_1)^{(i)}, (\psi_2)^{(i)}, \ldots, (\psi_{n_{\mathrm{R}} n_{\mathrm{T}}})^{(i)}\}.$$

Here $|\cdot|$ denotes absolute value, $[A]_{n,n}$ is the (n,n) element of matrix A, and $[\mathbf{a}]_n$ is the nth element of vector $\mathbf{a}$. When $\Phi_k \Phi_k^H = I_{n_{\mathrm{T}}}$, (5.17a) and (5.17c) further simplify to

$$W^{(i)} = \mathrm{diag}\left\{[(\sigma^2)^{(i)} + (\psi_1)^{(i)}]^{-1}, \ldots, [(\sigma^2)^{(i)} + (\psi_{n_{\mathrm{R}} n_{\mathrm{T}}})^{(i)}]^{-1}\right\} \quad (5.37a)$$

$$[\boldsymbol{h}_{\mathrm{SC},k}^{(i)}]_n = \frac{(\psi_n)^{(i)}}{(\sigma^2)^{(i)} + (\psi_n)^{(i)}} \cdot [\boldsymbol{z}_k - A_{\mathrm{LOS},k}\, \boldsymbol{x}^{(i)}]_n \quad (5.37b)$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$, and the conditional maximization (CM) steps in (5.36a) and(5.36b) simplify accordingly. In this case, the CRB expressions are also simplified (see Appendix D.5.2):

$$\mathrm{CRB}_{\sigma^2,\sigma^2} = \frac{\sigma^4}{n_{\mathrm{R}}(N - n_{\mathrm{T}})K} \quad (5.38a)$$

$$[\mathrm{CRB}_{\psi,\psi}]_{n,n} = \mathrm{CRB}_{\psi_n,\psi_n}\frac{(\sigma^2 + \psi_n)^2}{K} + \frac{\sigma^4}{n_{\mathrm{R}}(N - n_{\mathrm{T}})K} \quad (5.38b)$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$. Here, $\mathrm{CRB}_{\psi_n,\psi_n}$ is an increasing function of both σ^2 and ψ_n. As expected, both $\mathrm{CRB}_{\sigma^2,\sigma^2}$ and $\mathrm{CRB}_{\psi_n,\psi_n}$ decrease proportionally to $1/K$ as the number of coherent intervals K grows.

5.3.3.2 ECME algorithm for REML estimation

The ECME REML algorithm for independent fading follows using arguments similar to those in Appendix D.3.2. It iterates between (5.17a)–(5.17c) for $k = 1, 2, \ldots, K$ and

$$(\sigma^2)^{(i+1)} = \frac{K n_{\mathrm{R}} N}{K n_{\mathrm{R}} N - r} \cdot \widetilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{K n_{\mathrm{R}} N - r} \cdot \sum_{k=1}^{K} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)})^H W_k^{(i)} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)}) \quad (5.39a)$$

$$(\psi_n)^{(i+1)} = (\psi_n)^{(i)} + \frac{1}{K} \cdot \sum_{k=1}^{K}\left\{|[\boldsymbol{h}_{\mathrm{SC},k}^{(i)}]_n|^2 - [(\psi_n)^{(i)}]^2 \cdot \Big[Z_k{}^H Z_k \left(W_k^{(i)}\right.\right.$$
$$\left.\left. - W_k^{(i)} \mathcal{A}_{\mathrm{LOS},k} \cdot \Big(\sum_{l=1}^{K} \mathcal{A}_{\mathrm{LOS},l}^H W_l^{(i)} \mathcal{A}_{\mathrm{LOS},l}\Big)^{-1} \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)}\right) Z_k{}^H Z_k\Big]_{n,n}\right\}. \quad (5.39b)$$

where $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$. For $\Phi_k \Phi_k^H = I_{n_{\mathrm{T}}}$, we can use (5.37) to simplify (5.39a) and (5.39b).

As in the correlated fading case, the above ECME algorithms converge to variance estimates that are *always* in the parameter space. They can be initialized using the moment estimators in (5.22) and diagonal elements of (5.21). For the unstructured LOS array response model, we can remove (5.17b) from the ML and REML ECME iterations and use the closed-form expression for the ML estimate of $\boldsymbol{x}$ in (5.29a).

5.4 Numerical examples

We evaluate the estimation accuracy and computational efficiency of the ML and REML methods in Section 5.3. Our performance metric is MSE of an estimator, calculated using 5000 independent trials. Numerical simulations were performed using both block- and continuous-fading scenarios. Throughout this section, we employed the Alamouti transmission scheme for a $n_{\mathrm{R}} \times n_{\mathrm{T}} = 2 \times 2$ MIMO system with $N = 30$ QPSK symbols per coherent interval (normalized so that $\Phi_k \Phi_k^H = I_2$) and generated additive white complex Gaussian noise $\boldsymbol{e}_k(t)$ with variance $\sigma^2 = 0.01$.

5.4.1 Block-fading scenario

In the block-fading case, we generated the simulated data using the measurement model in Section 5.2. The LOS component was generated using (5.7) with $x = 1$, $\boldsymbol{a}_{\mathrm{T,LOS}} = [1, \exp(-j\pi/6)]^T$, $\boldsymbol{a}_{\mathrm{R,LOS}} = [1, \exp(-j\pi/3)]^T$, and $\omega_{\mathrm{D,LOS}} = \pi/2 \cdot 10^{-2}$.

In this section, we consider the *correlated block-fading scenario* and apply the ML and REML algorithms in Section 5.3.1 using the unstructured and structured LOS array response models in (5.6) and (5.7). The spatial fading covariance matrix was:

$$\Psi = \sigma^2 \cdot \begin{bmatrix} 1 & 0.2+0.1j & 0.4-0.5j & 0 \\ 0.2-0.1j & 2 & 0 & 0.1+0.1j \\ 0.4+0.5j & 0 & 4 & 0.3-0.3j \\ 0 & 0.1-0.1j & 0.3+0.3j & 8 \end{bmatrix}. \tag{5.40}$$

In Figs. 5.1 and 5.2, we present the MSEs and corresponding CRBs for the ML and REML estimates of selected parameters as functions of the number of coherent intervals K. The ML

estimates of ρ were computed using the closed-form expressions in (5.26) and (5.29) for the unstructured LOS model and the alternating-projection ML algorithm (5.27)–(5.28) for the structured LOS model. For $K \geq 10$, the alternating-projection algorithm converged in less than five iterations. In the cases where (5.29b) and (5.28) were not positive semidefinite, we ran the ECME ML algorithm described in Section 5.3.1.1. In terms of CPU time, the alternating-projection ML algorithm was five to seven times faster than the ECME ML algorithm. The REML estimation was performed using the ECME algorithm in Section 5.3.1.2 which converged in less than seven iterations.

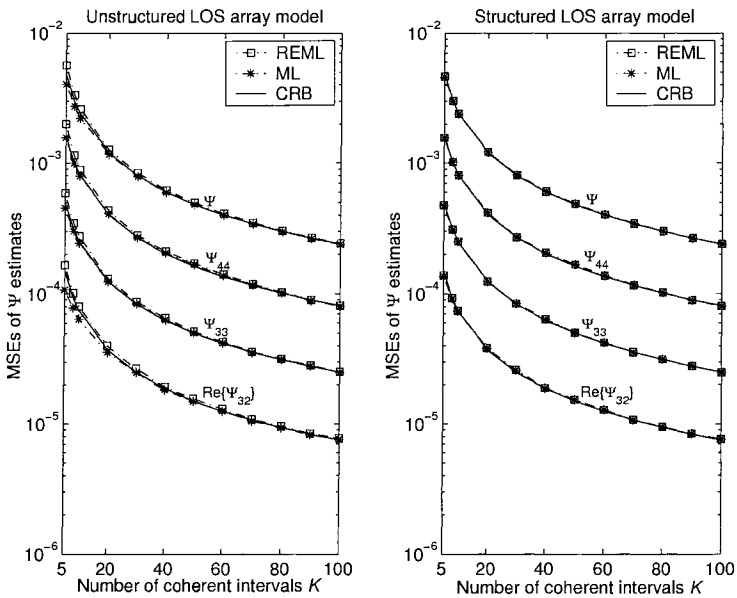

Figure 5.1 MSEs and CRBs of ML and REML estimates of some variance components under the unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario.

In Table 5.1, we show the percentages of trials in which the estimates of Ψ in (5.28) and (5.29b) were not positive semidefinite, as functions of K. These percentages decay rapidly with K; however, they are high for small K, underlining the importance of the ECME approach which handles parameter constraints automatically.

In Fig. 5.1, the MSEs and CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, and sum of all elements of ψ are shown as functions of K for the unstructured (left) and structured (right) LOS array response models. Due to the CRB decoupling between the mean and

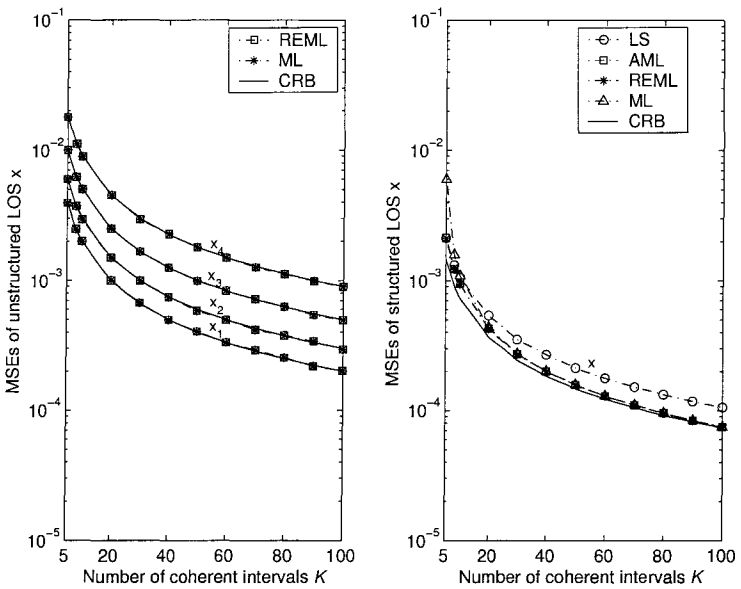

Figure 5.2 MSEs and CRBs of ML and REML estimates of LOS coefficients under the unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario.

variance-component parameters (see (D.25) in Appendix D.5), the CRBs for the variance components are the same regardless of the LOS array response parametrization; the corresponding MSEs are also approximately equal.

In Fig.5.2, we present the MSEs and CRBs for the ML and REML estimates of the unstructured LOS array response vector (left) and ML, REML, AML, and LS estimates of the structured-array LOS complex amplitude x (right), as functions of K. Here, the linear LS estimate of x is computed by substituting (5.7) into (5.20). As expected, the MSEs and CRBs

	$K = 10$	$K = 20$	$K = 30$	$K = 40$	$K = 50$	$K = 60$
structured LOS	63.2%	13.7%	3.7%	0.5%	0.1%	0.05%
unstructured LOS	65.3%	18.3%	3.9%	0.8%	0.1%	0.04%

Table 5.1 Percentages of trials in which the alternating-projection and closed-form estimators for the structured and unstructured LOS array response models in (5.28) and (5.29b) were not positive semidefinite, as functions of K.

are smaller for the structured LOS model. For larger K, the (closed-form) AML and (iterative) ML and REML estimates of x achieve similar MSE performances. However, the AML estimator performs poorly when K is small, see also the discussion in Section 5.3.2.

An analytical expression for the MSE of the linear LS estimate of x is given below for the special case of $\Phi_k \Phi_k^H = I_{n_T}$ and assuming the block-fading scenario:

$$\begin{aligned} \mathrm{MSE}_x &= \mathrm{MSE}_{\mathrm{Re}\{x\}} + \mathrm{MSE}_{\mathrm{Im}\{x\}} \\ &= \frac{\sigma^2}{K \cdot \boldsymbol{a}_{\mathrm{T}}^H \boldsymbol{a}_{\mathrm{T}} \cdot \boldsymbol{a}_{\mathrm{R}}^H \boldsymbol{a}_{\mathrm{R}}} + \frac{(\boldsymbol{a}_{\mathrm{T,LOS}}^H \otimes \boldsymbol{a}_{\mathrm{R,LOS}}^H)\, \Psi\, (\boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}})}{K \cdot (\boldsymbol{a}_{\mathrm{T}}^H \boldsymbol{a}_{\mathrm{T}})^2 \cdot (\boldsymbol{a}_{\mathrm{R}}^H \boldsymbol{a}_{\mathrm{R}})^2}. \end{aligned} \tag{5.42}$$

The CRBs were computed using the results in Appendix D.5.1.

Since the MSE performances of the ML and REML estimators are similar, it is of interest to compare their biases as well. Figure 5.3 compares the absolute biases for the ML and REML estimates of the variance components ψunder the unstructured (left) and structured (right) LOS array response models; the biases are shown as functions of K. The obtained results confirm the *bias-correction property* of REML variance-component estimation, see also the discussion in Section 5.3. Compared with ML, the REML approach yields significant bias improvements when the rank r of the deterministic component is large (e.g., unstructured LOS array model) and for small sample sizes K.

5.4.2 Continuous-fading scenario

We now study the performance of the proposed methods in *continuous fading* where the scattering channel coefficients are temporally correlated according to the Jakes' model, see [34, 95] and references therein. First, denote by $\boldsymbol{h}_{\mathrm{SC},k}(t)$ the scattering channel vector at time t in the kth coherent interval. In adjacent coherent intervals, we model the $Kn_{\mathrm{R}}n_{\mathrm{T}}N \times 1$ vector of all scattering coefficients:

$$\boldsymbol{h}_{\mathrm{SC}} = [\boldsymbol{h}_{\mathrm{SC},1}(1)^T \cdots \boldsymbol{h}_{\mathrm{SC},1}(N)^T, \cdots \boldsymbol{h}_{\mathrm{SC},K}(1)^T, \cdots \boldsymbol{h}_{\mathrm{SC},K}(N)^T]^T \tag{5.43}$$

as a zero-mean complex Gaussian vector with covariance matrix [34]:

$$\mathrm{E}[\boldsymbol{h}_{\mathrm{SC}} \boldsymbol{h}_{\mathrm{SC}}^H] = J(\omega_{\mathrm{D}}) \otimes \Psi \tag{5.44}$$

where $\omega_{\mathrm{D}} \in (0, \pi)$ is the *maximum angular Doppler frequency* (corresponding to the Doppler spread of $2\omega_{\mathrm{D}}$). And the (p, q) element of the $KN \times KN$ matrix $J(\omega_{\mathrm{D}})$ is

$$[J(\omega_{\mathrm{D}})]_{p,q} = J_0(\omega_{\mathrm{D}}(p - q)) \tag{5.45}$$

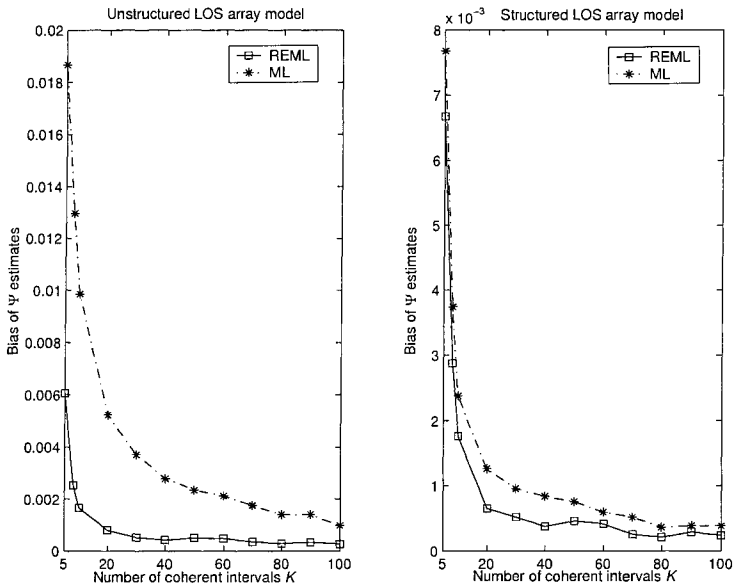

Figure 5.3 Biases for the ML and REML estimates of the sum of all elements of ψ under the correlated block-fading scenario and unstructured (left) and structured (right) LOS array response models in correlated block-fading scenario.

We further assume that the LOS component of the channel response matrix changes with t according to the following model:

$$H_{\mathrm{LOS},k}(t) = \boldsymbol{a}_{\mathrm{R,LOS}}\, \boldsymbol{a}_{\mathrm{T,LOS}}^{T}\, \exp\{j\omega_{\mathrm{D,LOS}} \cdot [(k-0.5)\, N + t]\} \cdot x \tag{5.46}$$

where the LOS angular Doppler shift $\omega_{\mathrm{D,LOS}}$ should be bounded by the maximum Doppler frequency, i.e. $|\omega_{\mathrm{D,LOS}}| \leq \omega_{\mathrm{D}}$. Combining the scattering and LOS channel components, we obtain the following continuous-fading measurement model:

$$\boldsymbol{y}_k(t) = [H_{\mathrm{SC},k}(t) + H_{\mathrm{LOS},k}(t)]\, \boldsymbol{\phi}_k(t) + \boldsymbol{e}_k(t) \tag{5.47}$$

for $t = 1, \ldots, N$, $k = 1, \ldots, K$, where $\boldsymbol{h}_{\mathrm{SC},k}(t) = \mathrm{vec}\,\{H_{\mathrm{SC},k}(t)\}$. The above model accounts for correlations among the coherent intervals and time variations of the scattering and LOS channel coefficients within a coherent interval.

In the following examples, we consider the correlated fading scenario with Ψ given in (5.40), maximum Doppler frequency $\omega_{\mathrm{D}} = 2\pi \cdot 10^{-2}$ [6], and the LOS parameters $x = 1$, $\boldsymbol{a}_{\mathrm{T,LOS}} =$

[6] This is consistent with the mobile speed of 100 mi/h for the carrier frequency 1.9 GHz and a symbol rate of 30 kHz, see [53].

$[1, \exp(-j\pi/6)]^T$, $\boldsymbol{a}_{\mathrm{R,LOS}} = [1, \exp(-j\pi/3)]^T$, and $\omega_{\mathrm{D,LOS}} = \pi/2 \cdot 10^{-2}$. We first assume that the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is known, and then consider the case where $\omega_{\mathrm{D,LOS}}$ is unknown.

Known $\omega_{\mathrm{D,LOS}}$: We computed the ML and REML estimates of $\boldsymbol{\rho}$ using the methods in Section 5.3.1 where the coherent-interval length was chosen as $N = 30$. Figs. 5.4 and 5.5 show the MSEs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, sum of all elements of $\boldsymbol{\psi}$, and the LOS coefficients, as functions of K; Fig. 5.5 (right) shows the MSEs for the AML and LS estimates of the structured-array LOS complex amplitude x. We also compare these MSEs with the corresponding block-fading CRBs, thus quantifying the *performance loss* that each method incurs due to the temporal correlation. Interestingly, for small K the MSEs of the ML and REML estimates are close to the block-fading CRBs. Under the continuous-fading scenario, the proposed variance-component estimates are mostly affected by correlations among the coherent intervals and time variations of the scattering component within a coherent interval, whereas the LOS-coefficient estimates are mostly affected by time variations of the LOS component.

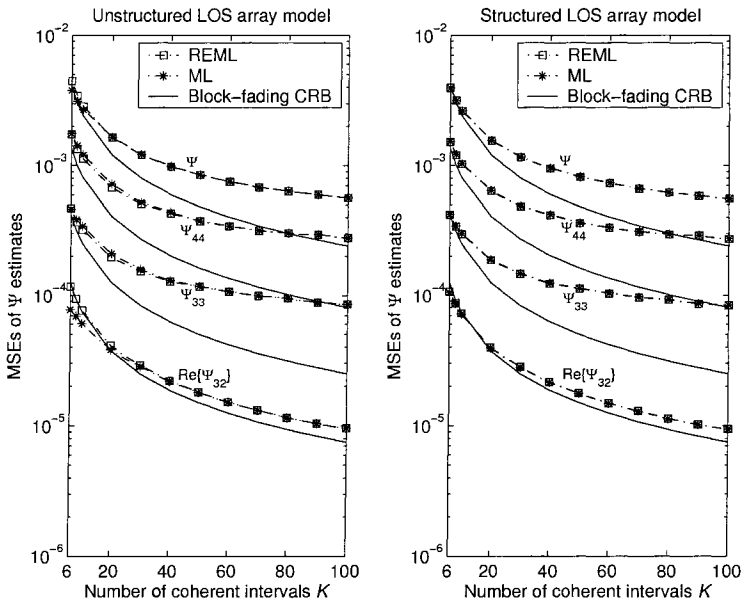

Figure 5.4 MSEs and block-fading CRBs of ML and REML estimates of some variance components under correlated Rician continuous–fading scenario and unstructured (left) and structured (right) LOS array response models.

Unknown $\omega_{\mathrm{D,LOS}}$: We now consider the scenario where the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is

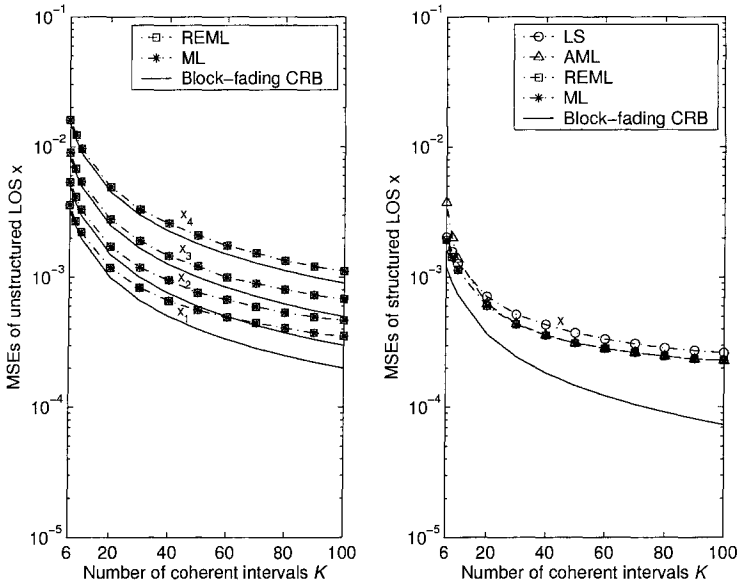

Figure 5.5 MSEs and block-fading CRBs of ML, REML, AML, and LS estimates of the LOS coefficients under correlated Rician continuous-fading scenario and unstructured (left) and structured (right) LOS array response models.

unknown and *estimated* by maximizing (5.31). Here, we selected the coherent interval length $N = 30$ as in the previous example. Following the *estimated likelihood approach* in [81, ch. 10.7], we treat the obtained estimate of $\omega_{\mathrm{D,LOS}}$ as a known constant and apply the ML and REML methods in Section 5.3.1. Figs. 5.6 and 5.7 show the MSEs and block-fading CRBs for the (estimated) ML, REML, AML, and LS estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, sum of all elements of $\boldsymbol{\psi}$, and LOS coefficients, as functions of K. For the unstructured LOS array response model, the ML estimates of $\boldsymbol{x}$ outperform the corresponding REML estimates.

5.5 Summary

In this chapter, we have developed ML and REML methods for estimating the mean and covariance parameters of MIMO fading channels under correlated and independent block-fading scenarios. For unitary space-time codes and orthogonal designs in correlated fading, we obtained closed-form expressions of exact and approximate ML estimates of the unknown parameters. We also evaluated the performance of the proposed methods via numerical simulations under the block- and continuous-fading scenarios. Simulation results show that the

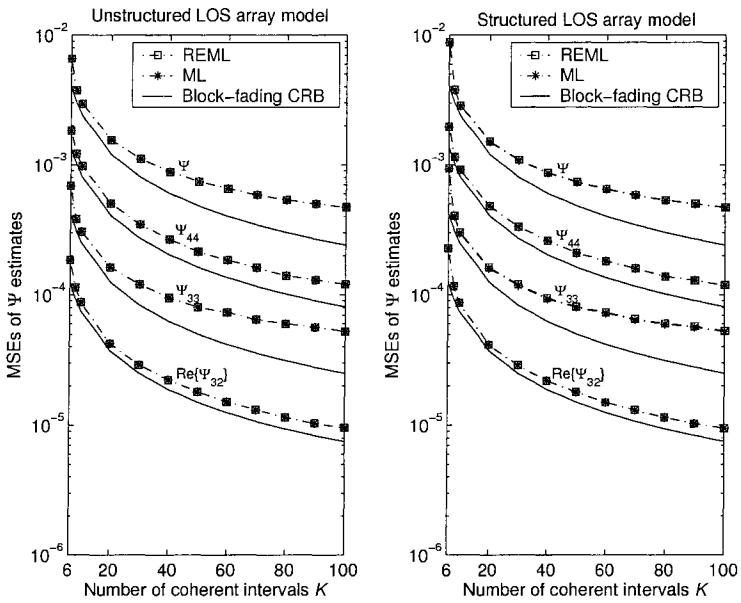

Figure 5.6 MSEs and block-fading CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, and sum of all elements of ψ under the correlated Rician continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and unstructured (left) and structured (right) LOS array response models, as functions of K.

proposed estimators are almost efficient under the block-fading scenario, having mean-square errors close to the corresponding CRBs.

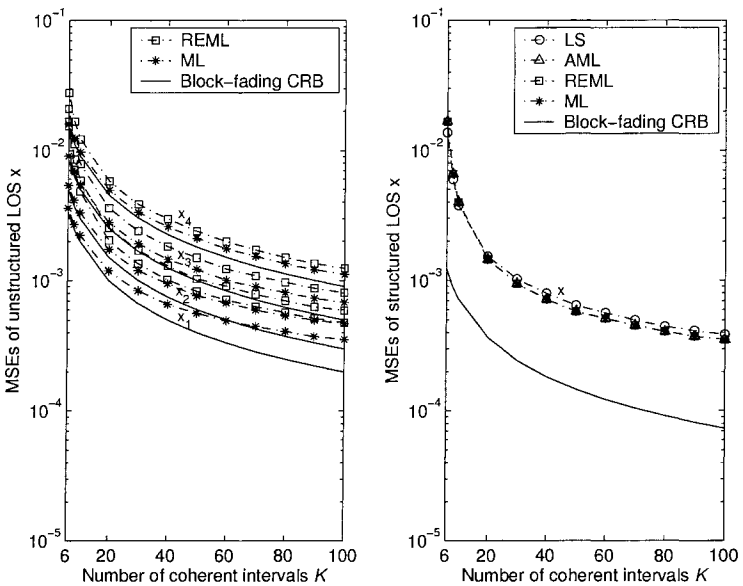

Figure 5.7 MSEs and block-fading CRBs for the ML, REML, AML, and LS estimates of the LOS coefficients under the correlated continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and unstructured (left) and structured (right) LOS array response models, as functions of K.

CHAPTER 6. CONCLUSIONS AND FUTURE WORK

In this thesis, we studied two important techniques that may be used to improve the communication performance, data rate, or bandwidth efficiency, for wireless communication systems that suffer from possibly severe multipath fading, shadow fading, time variations, etc. The first technique that we concentrated on is EGC or HS/EGC equipped with QAM. EGC or HS/EGC offers enhanced BER performance and increased data rate, and QAM ensures high bandwidth efficiency. We proposed a receiver design and a decision variable, provided a unified approach based on MGF/CHF to efficiently evaluate the BER performance for square and rectangular M-QAM over a variety of fading channels, and characterized the performance loss due to ICE for the Rayleigh and Nakagami fading channels. Numerical simulations verified our theoretic results. Our results offer wireless communication systems designers new alternatives to design bandwidth efficient systems with desired tradeoff between diversity gain and complexity. The second technique is the estimation of channel statistic properties, such as the mean and correlations in fading and fading-shadowing channels. We proposed several ML based recursive algorithms to estimate these quantities, and derived the CRBs for these estimation problems. We demonstrated that our algorithms are efficient in the sense of achieving the CRBs asymptotically. This technique makes it possible for wireless communication systems designers to take advantage of the channel statistic properties in their analysis and design.

Our future work includes the following.

HS/EGC performance loss due to ICE. In Chapter 2, we quantified the performance loss due to ICE in the EGC framework. In Chapter 3, we studied HS/EGC systems that are designed using perfect CSI. Clearly, the quantification of the performance loss due to ICE for HS/EGC systems elude our study. Due to the practical importance of HS/EGC QAM, it is of interest to investigate its ICE issue, which will also be a natural extension of this thesis. The approach that has been successfully applied in Chapter 2, namely the analysis based on the

"effective channel" or "effective SNR" may be also found useful in this future work. The study on HS/EGC with ICE will make the framework of studying the ICE effect on the available diversity receivers more complete and provide us with deeper understanding of the sensitivity of different diversity systems to ICE.

Design and Performance evaluation of systems incorporating EGC QAM or HS/EGC QAM. There exist studies of systems that are designed based on MRC QAM. For example, in a time-division multiple-access (TDMA) digital cellular system, a trellis-coded 16-QAM with MRC was proposed, see [3, 57]. It would be practically appealing to study the performance of trellis-coded EGC QAM or HS/EGC QAM. To analyze the overall, end-to-end communication performance of such systems, it is necessary to completely characterize the underlying EGC QAM or HS/EGC QAM. Therefore, after the resolution of the performance analysis of the underlying EGC QAM or HS/EGC QAM techniques, we are in a better position to proceed to the performance analysis of the overall systems.

Other natural extensions of performance analysis. We are also interested in quantifying the BER performance for channels with temporal correlation (we allowed temporal correlation only in the Rayleigh case), and studying multidimensional diversity techniques that involve the combination of two or more conventional means of realizing diversity (e.g. space, multipath, and frequency). Due to the potentially broad applications of these studies, they have received a great deal of attentions recently. The thorough study performed in this thesis gives us advantages to study these extensions.

Effects of channel statistic estimation errors. In the future, we would like to incorporate the proposed estimators into the design of (optimal) transmitters and study the effects of the channel statistic estimation errors on the performance. Note that there exists extensive study of transmitter design based on channel statistic properties (e.g. [44, 108]), i.e. mean and correlation. The benefit of such an approach is that it requires much less frequent side information updates at the transmitter and hence reduces the complexity and overhead. However, in the existing references, it is widely assumed that the transmitter obtains side information from an estimator at the receiver side, and the estimator uses methods based on Method of Moments. It is well known that, though easy to implement, Method of Moments does not guarantee any optimality [58], and hence it undermines the optimality of designed transmitter. An alternative and perhaps better way is to employ our proposed ML estimator at the receiver side to

generate the estimated channel correlations, which are accessed by the transmitter through a feedback link. We would like to see how much improvement we may obtain by implement the transmitter design based on our ML estimation. It would then be interesting to study the tradeoff between the training sequence length and the overall data rate/performance. In other words, a larger training sequence length can give us a more accurate estimation of correlation and improve the data rate/performance, but we would have a shorter period of time to transmit our messages. Such a result would be instructive in practical design.

MIMO composite gamma-lognormal channel estimation. Also as an extension of our work in Chapter 4, we would like to study the MIMO composite gamma-lognormal channel estimation. This study is meaningful since in applications such as distributed antenna systems, there exist multiple transmitters and receivers, and the channels are modelled as composite gamma-lognormal fading channels.

Performance evaluation of schemes with beamforming based on channel estimation and with diversity combining. Finally, we remark that another direction of future work is an integrating piece of the two parts in this thesis. More specifically, for a fading channel, we may employ EGC QAM or HS/EGC QAM technique at the receiver, in addition to a beamformer at the transmitter. The beamformer uses the channel statistic properties provided by the channel estimation algorithms proposed in this thesis. The performance evaluation of this scheme will be challenging but potentially useful due to broad presence of EGC receivers and beamformers in many wireless communication applications.

APPENDIX A. DERIVATION OF $\hat{m}$

To facilitate BER analysis for EGC QAM in Nakagami fading channels, the Nakagami-m parameter for $\hat{c}_l(i)$, denoted by $\hat{m}$, has to be derived. This is especially important for BER analysis for small-to-medium average SNRs. We derive $\hat{m}$ based on the concept of the amount of fading.

Consider a slow Nakagami fading channel. The channel estimate $\hat{c}_l$ given by (2.5) is re-written as

$$\hat{c}_l = c_{w_l} + n_{w_l}, \tag{A.1}$$

where $c_{w_l} = \mathbf{w}_l^H \mathbf{c}_{l,\mathrm{PS}}$ and $n_{w_l} = \mathbf{w}_l^H \mathbf{n}_{l,\mathrm{PS}}$. Then, $\hat{c}_l$ has zero mean, variance $\sigma_{\hat{c}}^2$ and Nakagami-parameter $\hat{m}$. The AF corresponding to the SNR $\gamma_l = c_l^2/N_0$ is given by $\mathrm{AF} = 1/m = E[\gamma_l^2]/E[\gamma_l]^2 - 1$ [94, eqs. (2.5) and (2.24)]. It follows that

$$\frac{1}{\hat{m}} = \frac{E[\hat{c}_l^4]}{E[\hat{c}_l^2]^2} - 1 \tag{A.2}$$

Below we drop the branch index l when no confusion arises. Using (A.1) we obtain $|\hat{c}|^2 = |c_w|^2 + |n_w|^2 + 2\mathrm{Re}(c_w^* n_w)$, which leads to

$$E[|\hat{c}|^2] = E[|c_w|^2] + E[|n_w|^2] + 2E[\mathrm{Re}(c_w^* n_w)] = \sigma_{c_w}^2 + \sigma_{n_w}^2. \tag{A.3}$$

where $\sigma_{c_w}^2 = E[|c_w|^2]$ and $\sigma_{n_w}^2 = E[|n_w|^2]$ are the variances of the signal and the noise, respectively. Here, $E[\mathrm{Re}(c_w^* n_w)] = 0$ because c_w and n_w are independent of each other and both have zero mean. Equation (A.1) further leads to

$$E[|\hat{c}|^4] = E[|c_w|^4] + E[|n_w|^4] + 4E[\{\mathrm{Re}(c_w^* n_w)\}^2] + 2E[|c_w|^2]E[|n_w|^2]. \tag{A.4}$$

Using eq. (2.23) of [94] (by setting $k = 2$ and $\bar{\gamma} = \sigma_{c_w}^2$ therein), we obtain that

$$E[|c_w|^4] = \frac{\Gamma(m+2)}{\Gamma(m)m^2}\sigma_{c_w}^4 = (1 + \frac{1}{m})\sigma_{c_w}^4$$

Using a similar procedure, we have $E[|n_w|^4] = \frac{\Gamma(1+2)}{\Gamma(1)}\sigma_{n_w}^4 = 2\sigma_{n_w}^4$. Since c_w and n_w are independent and both are circularly-symmetric random variables, we obtain $E\{[\mathrm{Re}(c_w^* n_w)]^2\} = \frac{1}{2}E[|c_w|^2]E[|n_w|^2] = \frac{1}{2}\sigma_{c_w}^2\sigma_{n_w}^2$.

Substituting these results into (A.4) leads to

$$E[|\hat{c}|^4] = (1+\frac{1}{m})\sigma_{c_w}^4 + 2\sigma_{n_w}^4 + 4\sigma_{c_w}^2\sigma_{n_w}^2$$

Therefore, (A.2) is equivalent to $1/\hat{m} = E[|\hat{c}|^4]/(\sigma_{c_w}^2 + \sigma_{n_w}^2)^2 - 1 = \frac{\sigma_{c_w}^4/m + \sigma_{n_w}^4 + 2\sigma_{c_w}^2\sigma_{n_w}^2}{(\sigma_{c_w}^2+\sigma_{n_w}^2)^2}$, or that

$$\hat{m} = \frac{(\sigma_{c_w}^2 + \sigma_{n_w}^2)^2}{\sigma_{c_w}^4/m + \sigma_{n_w}^4 + 2\sigma_{c_w}^2\sigma_{n_w}^2} \tag{A.5}$$

where $\sigma_{c_w}^2 = E[|c_w|^2] = \mathbf{w}_l^H \mathbb{R}_{c,\mathrm{PS}} \mathbf{w}_l$ and $\sigma_{n_w}^2 = E[|n_w|^2] = \mathbf{w}_l^H \mathbb{R}_{n,\mathrm{PS}} \mathbf{w}_l$. Here, $\mathbb{R}_{c,\mathrm{PS}} = E[\mathbf{c}_{l,\mathrm{PS}}\mathbf{c}_{l,\mathrm{PS}}^H]$ and $\mathbb{R}_{n,\mathrm{PS}} = E[\mathbf{n}_{l,\mathrm{PS}}\mathbf{n}_{l,\mathrm{PS}}^H]$ are the correlation matrices of the channel gain vector and noise vector, respectively. For a slow Nakagami fading channel with white noise, we have $\mathbb{R}_{c,\mathrm{PS}} = \sigma_c^2 \mathbf{1}_{F\times F}$ and $\mathbb{R}_{n,\mathrm{PS}} = N_0 I_F$. Thus,

$$\sigma_{c_w}^2 = \sigma_c^2 \mathbf{w}_l^H \mathbf{1}_{F\times F} \mathbf{w}_l \qquad \text{and} \qquad \sigma_{n_w}^2 = N_0 \mathbf{w}_l^H \mathbf{w}_l \tag{A.6}$$

By combining (A.6) and (A.5) we obtain

$$\hat{m} = \frac{[\bar{\gamma}(\mathbf{w}_l^H \mathbf{1}_{F\times F}\mathbf{w}_l) + (\mathbf{w}_l^H \mathbf{w}_l)]^2}{\bar{\gamma}^2(\mathbf{w}_l^H \mathbf{1}_{F\times F}\mathbf{w}_l)^2/m + (\mathbf{w}_l^H \mathbf{w}_l)^2 + 2\bar{\gamma}\mathbf{w}_l^H \mathbf{1}_{F\times F}\mathbf{w}_l(\mathbf{w}_l^H \mathbf{w}_l)} \tag{A.7}$$

As a simple check, (A.7) shows that when $\bar{\gamma}$ is large, $\hat{m} \simeq m$; and when $\bar{\gamma}$ is small, $\hat{m} \simeq 1$. For Rayleigh fading channels ($m = 1$), (A.7) shows that $\hat{m} = m = 1$ for whatever the ASNR $\bar{\gamma}$, as expected.

For the MMSE-CE in slow Nakagami fading channels, we observe that $\mathbf{w}_l = c\mathbf{1}_{F\times 1}$, where c is a constant determined by the system and channel parameters. We obtain

$$\hat{m} = \frac{[\bar{\gamma}F\mathbf{w}_l^H \mathbf{w}_l + \mathbf{w}_l^H \mathbf{w}_l]^2}{\bar{\gamma}^2F^2(\mathbf{w}_l^H \mathbf{w}_l)^2/m + (\mathbf{w}_l^H \mathbf{w}_l)^2 + 2\bar{\gamma}F(\mathbf{w}_l^H \mathbf{w}_l)^2} = \frac{(1+F\bar{\gamma})^2}{1 + 2F\bar{\gamma} + F^2\bar{\gamma}^2/m} \tag{A.8}$$

As F and/or $\bar{\gamma}$ increase, $\hat{m}$ approaches m.

For the ideal MMSE-CE and other channel estimators without frequency offset, ρ is real-valued. For non-MMSE channel estimators (e.g. Sinc- and Gaussian-interpolators [59]) in presence of frequency offset, ρ may be a complex value. Thus, to generalize our results, we assume ρ to be complex unless otherwise stated.

APPENDIX B. CDF AND TMGF FOR FADING AMPLITUDES

B.1 CDF of Nakagami-q random variable

The PDF of Nakagami-q amplitude α_l may be written as [12, 94]

$$f_{\alpha_l}(x) = 2Ax\exp(-Bx^2)I_0(gx^2), \tag{B.1}$$

where $A = \frac{1}{\bar{\gamma}_l\sqrt{1-b_l^2}}$, $B = \frac{1}{\bar{\gamma}_l(1-b_l^2)}$, $g = \frac{b_l}{\bar{\gamma}_l(1-b_l^2)}$, $\bar{\gamma}_l$ is the average SNR at the lth branch given by $\bar{\gamma}_l = E(\alpha_l^2) = (E_d/N_0)E[|c_l|^2] = E[|c_l|^2]/N_0$, and $I_0(z)$ is the zero-order modified Bessel function of the first kind. The CDF of α_l is given by

$$F_{\alpha_l}(x) = 1 - \int_x^\infty f_{\alpha_l}(y)dy = 1 - \int_x^\infty 2Ay\exp(-By^2)I_0(gy^2)dy$$

Note that $I_0(z) = \sum_{k=0}^\infty \frac{(z^2/4)^k}{(k!)^2}$. With a change of variable $\tilde{y} = y^2$, we get

$$\begin{aligned} F_{\alpha_l}(x) &= 1 - A\int_{x^2}^\infty \exp(-B\tilde{y})I_0(g\tilde{y})d\tilde{y} = 1 - A\sum_{k=0}^\infty \int_{x^2}^\infty \frac{g^{2k}\tilde{y}^{2k}}{4^k(k!)^2}\exp(-B\tilde{y})d\tilde{y} \\ &= 1 - A\sum_{k=0}^\infty \frac{g^{2k}}{4^k(k!)^2}\Gamma(2k+1, x^2B)B^{-(2k+1)} \\ &= 1 - \sqrt{1-b^2}\sum_{k=0}^\infty \Gamma\left(2k+1, \frac{x^2}{\bar{\gamma}_l(1-b^2)}\right)\frac{b^{2k}}{4^k(k!)^2}. \end{aligned} \tag{B.2}$$

where $\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}dt$ is the complementary incomplete Gamma function [1].

B.2 Closed-form TMGF for Nakagami-m Fading Channels

In Nakagami-m fading channels, the PDF for amplitude α_l is given by

$$f_{\alpha_l}(x) = \frac{2}{\Gamma(m_l)}\left(\frac{m_l}{\bar{\gamma}_l}\right)^{m_l} x^{2m_l-1}\exp\left(-\frac{m_l}{\bar{\gamma}_l}x^2\right) \tag{B.3}$$

The TMGF may be written as

$$\begin{aligned}\Phi_{\alpha_l}(s,x) &= C\int_x^{\infty} e^{-By^2-sy}y^{2m_l-1}dy \\ &= C\exp\left(\frac{s^2}{4B}\right)\int_x^{\infty}\exp\left\{-\left(\sqrt{B}y+\frac{s}{2\sqrt{B}}\right)^2\right\}y^{2m_l-1}dy\end{aligned}$$

where $C=\frac{2}{\Gamma(m_l)}(\frac{m_l}{\bar{\gamma}_l})^{m_l}$ and $B=\frac{m_l}{\bar{\gamma}_l}$. By defining $z=\sqrt{B}y+\frac{s}{2\sqrt{B}}$, we have

$$\Phi_{\alpha_l}(s,x) = \frac{C\exp\left(\frac{s^2}{4B}\right)}{B^{m_l}}\int_{\sqrt{B}x+\frac{s}{2\sqrt{B}}}^{\infty}\exp(-z^2)\left(z-\frac{s}{2\sqrt{B}}\right)^{2m_l-1}dz$$

Assume m_l is a positive integer or half-integer, such that $2m_l-1$ is an integer. For convenience, we define the equality that

$$L_k(a)=\int_a^{\infty}z^k e^{-z^2}dz=\frac{1}{2}\Gamma(\frac{k+1}{2},a^2). \tag{B.4}$$

Using the binomial expansion of $(z-\frac{s}{2\sqrt{B}})^{2m_l-1}$ we obtain

$$\begin{aligned}\Phi_{\alpha_l}(s,x) &= \frac{C\exp\left(\frac{s^2}{4B}\right)}{2B^{m_l}}\sum_{k=0}^{2m_l-1}\binom{2m_l-1}{k}\left(-\frac{s}{2\sqrt{B}}\right)^{2m_l-1-k} \\ &\quad\times\Gamma\left(\frac{k+1}{2},(\sqrt{B}x+\frac{s}{2\sqrt{B}})^2\right) \\ &= \frac{1}{\Gamma(m_l)}\exp\left(\frac{s^2\bar{\gamma}_l}{4m_l}\right)\sum_{k=0}^{2m_l-1}\binom{2m_l-1}{k}\left(-\frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right)^{2m_l-1-k} \\ &\quad\times\Gamma\left(\frac{k+1}{2},\left(\sqrt{m_l/\bar{\gamma}_l}x+\frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right)^2\right) \\ &= \frac{2}{\Gamma(m_l)}\exp\left(\frac{s^2\bar{\gamma}_l}{4m_l}\right)\sum_{k=0}^{2m_l-1}\binom{2m_l-1}{k}\left(-\frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right)^{2m_l-1-k} \\ &\quad\times L_k\left(\sqrt{m_l/\bar{\gamma}_l}x+\frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right).\end{aligned} \tag{B.5}$$

Below, we provide an efficient recursive procedure to evaluate the complementary incomplete Gamma function $\Gamma(\frac{k+1}{2},z)$ involved in (B.4) and (B.5). Consider a real-valued a first. We have

$$\begin{aligned}L_n(a) &= \int_a^{\infty}z^n e^{-z^2}dz=\int_a^{\infty}e^{-z^2}d\frac{z^{n+1}}{(n+1)} \\ &= -e^{-a^2}\frac{a^{n+1}}{(n+1)}+\int_a^{\infty}\frac{z^{n+1}}{(n+1)}e^{-z^2}2zdz \\ &= -e^{-a^2}\frac{a^{n+1}}{(n+1)}+\frac{2}{(n+1)}L_{n+2}(a)\end{aligned} \tag{B.6}$$

which is equivalent to $L_{n+2}(a) = \frac{n+1}{2} L_n(a) + \frac{1}{2} e^{-a^2} a^{n+1}$, or that

$$L_n(a) = \frac{n-1}{2} L_{n-2}(a) + \frac{1}{2} e^{-a^2} a^{n-1} \tag{B.7}$$

for $n = 2, \ldots, \infty$. For the case of complex-valued a, we need to rewrite $L_n(a)$ as

$$L_n(a) = \int_{a_R + ja_I}^{\infty + ja_I} z^n e^{-z^2} dz \tag{B.8}$$

where $a_R = \mathrm{Re}(a)$ and $a_I = \mathrm{Im}(a)$ are the real and imaginary parts of a, respectively. Equation (B.8) shows that the integral is now along the line parallel to the real axis. Utilizing the equality that $e^{-z^2} \cdot \frac{z^{n+1}}{(n+1)} \Big|_{a_R+ja_I}^{\infty+ja_I} = -e^{-(a_R+ja_I)^2} \frac{(a_R+ja_I)^{n+1}}{(n+1)} = -e^{-a^2} \frac{a^{n+1}}{n+1}$, we can readily show that (B.7) is also valid for the complex-valued a.

To employ (B.7) we still need to determine the explicit expressions for $L_0(a)$ and $L_1(a)$. By definition, we have

$$L_0(a) = \int_a^\infty e^{-z^2} dz = \sqrt{\pi} Q(\sqrt{2}a), \tag{B.9}$$

$$L_1(a) = \int_a^\infty z e^{-z^2} dz = \frac{1}{2} e^{-a^2}. \tag{B.10}$$

For complex-valued a, using (3.22) or (3.23) we derived for the Rayleigh fading case leads to

$$L_0(a) = \frac{\sqrt{\pi}}{2} - a \cdot \phi\left(1, \frac{3}{2}; -a^2\right) = \frac{\sqrt{\pi}}{2} - \frac{1}{2a}\left[\phi\left(1, \frac{1}{2}; a^2\right) - 1\right] \exp(-a^2). \tag{B.11}$$

By using (B.7) with (B.10) and (B.11) (or (B.11)), the closed-form TMGF in (B.5) can now be evaluated.

To calculate the TMGF with better numerical stability than (B.5), we recommend to merge the factor $\exp\left(\frac{s^2 \bar{\gamma}_l}{4m}\right)$ into the iterations for $L_k(a)$, and the modified formulas are given below,

$$\Phi_{\alpha_l}(s, x) = \frac{2}{\Gamma(m_l)} \sum_{k=0}^{2m_l - 1} \binom{2m_l - 1}{k} \left(-\frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right)^{2m_l - 1 - k} \tilde{L}_k(s, x) \tag{B.12}$$

where $\tilde{L}_k(s, x) = \exp\left(\frac{s^2 \bar{\gamma}_l}{4m_l}\right) L_k(a)$ and $a = \left(\sqrt{m_l/\bar{\gamma}_l}\, x + \frac{s}{2\sqrt{m_l/\bar{\gamma}_l}}\right)$. The iterative procedure to calculate $\tilde{L}_n(s, x)$ is given by

$$\tilde{L}_n(s, x) = \frac{n-1}{2} \tilde{L}_{n-2}(s, x) + \frac{1}{2} a^{n-1} \exp\left(-\frac{m_l}{\bar{\gamma}_l} x^2 - sx\right) \tag{B.13}$$

for $n = 2, \ldots, \infty$, with

$$\tilde{L}_0(s, x) = 0.5\sqrt{\pi} \exp\left(\frac{s^2 \bar{\gamma}_l}{4m_l}\right) - \frac{[\phi(1, \frac{1}{2}; a^2) - 1]}{2a} \exp(-\frac{m_l}{\bar{\gamma}_l} x^2 - xs) \tag{B.14}$$

$$\tilde{L}_1(s, x) = 0.5 \exp(-\frac{m_l}{\bar{\gamma}_l} x^2 - xs) \tag{B.15}$$

By using (B.12) with (B.13), (B.14), and (B.15), we can now evaluate the TMGF efficiently and accurately. We have verified that the validity of all the closed-form TMGF expressions derived in Sections 3.4.1 and B.2 by comparing them with numerical integration.

The numerical stability of these closed-form TMGF expressions closely rely on the numerical accuracy of the function $\phi(1, \frac{1}{2}, a^2)$. The series involved in $\phi(1, \frac{1}{2}, a^2)$ yields a high accuracy for small-to-medium $|s|$ and x, but may not always converge very well for large $|s|$ and/or x (e.g. $\omega > 20$, $x > 30$). Thus, for the latter case we suggest to use the numerical integration instead of the closed-form formulas to compute the TMGF.

APPENDIX C. EM ALGORITHM

We derive the EM algorithm presented in Section 4.2.2. Observe that the complete-data log-likelihood can be written as

$$\begin{aligned} L_c(\boldsymbol{y},\boldsymbol{u};\boldsymbol{\theta}) &= \sum_{k=1}^{K} \ln p_u(u_k;\mu,\sigma^2) + \sum_{k=1}^{K}\sum_{t=1}^{N} \ln p_{y|u}(y_k(t)|u_k;m) \\ &= K\cdot\Big\{\ln\Big(\frac{\xi}{\sqrt{2\pi}}\Big) - \tfrac{1}{2}\ln\sigma^2 + Nm\ln m - N\ln[\Gamma(m)] \\ &\quad +(m-1)N\cdot\Big[\frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)\Big] - mN\cdot\Big[\frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\frac{y_k(t)}{u_k}\Big] \\ &\quad -\frac{\xi^2}{2\sigma^2}\cdot\frac{1}{K}\sum_{k=1}^{K}(\ln u_k)^2 + \Big(\frac{\xi\mu}{\sigma^2} - mN - 1\Big)\cdot\Big(\frac{1}{K}\sum_{k=1}^{K}\ln u_k\Big) - \frac{\mu^2}{2\sigma^2}\Big\}. \end{aligned} \tag{C.1}$$

Therefore, the complete-data sufficient statistics are

$$T_1(\boldsymbol{u}) = \frac{1}{K}\cdot\sum_{k=1}^{K}\ln u_k, \tag{C.2a}$$

$$T_2(\boldsymbol{u}) = \frac{1}{K}\cdot\sum_{k=1}^{K}(\ln u_k)^2, \tag{C.2b}$$

$$\begin{aligned} T_3(\boldsymbol{y},\boldsymbol{u}) &= \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\Big[\frac{y_k(t)}{u_k}\Big] - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t) \\ &= \frac{1}{K}\cdot\sum_{k=1}^{K}\Big(\frac{\overline{y}_k}{u_k}\Big) - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t), \end{aligned} \tag{C.2c}$$

where $\overline{y}_k$ was defined in (4.11).

The complete-data log-likelihood (C.1) is easily maximized with respect to μ and σ^2, yielding the following estimates:

$$\widehat{\mu} = \frac{1}{K}\sum_{k=1}^{K}10\log_{10}(u_k) = \xi T_1(\boldsymbol{u}), \tag{C.3a}$$

$$\widehat{\sigma}^2 = \frac{1}{K}\sum_{k=1}^{K}[10\log_{10}(u_k) - \widehat{\mu}]^2 = \xi^2\, T_2(\boldsymbol{u}) - \widehat{\mu}^2. \tag{C.3b}$$

Then, to find the ML estimate of m based on the complete data, we need to maximize

$$\Xi(\boldsymbol{y}, \boldsymbol{u}; m) = m \ln m - \ln[\Gamma(m)] - m\, T_1(\boldsymbol{u}) - m\, T_3(\boldsymbol{y}, \boldsymbol{u}) \tag{C.4}$$

with respect to m. The above expression follows by dividing the concentrated complete-data log-likelihood function $L_c(\boldsymbol{y}, \boldsymbol{u}; [m, \widehat{\mu}, \widehat{\sigma}^2]^T)$ by KN and neglecting terms that are independent of m. It can be maximized using the Newton-Raphson iteration, which requires the first two derivatives of $\Xi(\boldsymbol{y}, \boldsymbol{u}; m)$ with respect to m:

$$\frac{\partial \Xi(\boldsymbol{y}, \boldsymbol{u}; m)}{\partial m} = \ln m - \frac{\Gamma'(m)}{\Gamma(m)} - \varphi(\boldsymbol{y}, \boldsymbol{u}), \tag{C.5a}$$

$$\frac{\partial^2 \Xi(\boldsymbol{y}, \boldsymbol{u}; m)}{\partial m^2} = \frac{1}{m} - \frac{\Gamma(m)\Gamma''(m) - [\Gamma'(m)]^2}{\Gamma(m)^2}, \tag{C.5b}$$

where

$$\varphi(\boldsymbol{y}, \boldsymbol{u}) = T_1(\boldsymbol{u}) + T_3(\boldsymbol{y}, \boldsymbol{u}) - 1. \tag{C.6}$$

The complete-data likelihood belongs to an exponential family of distributions, i.e., the log-likelihood (C.1) is linear in the *natural sufficient statistics* (C.2) [18, ch. 1.6.2] for the definition of the multiparameter exponential family and natural sufficient statistics. Also, the number of parameters is equal to the number of sufficient statistics. In this case, the EM algorithm is easily derived as follows [77, ch. 1.5.3]:

- The E step reduces to computing the conditional expectations of the complete-data natural sufficient statistics in (C.2) given the observed data $\boldsymbol{y}$, see (4.20) [1].
- The M step reduces to replacing the complete-data sufficient statistics (C.2) that occur in the complete-data ML estimate expressions of $\boldsymbol{\theta}$ in (C.3)–(C.6) with their conditional expectations computed in the E step, see (4.22).

[1] Note that $\sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t)/(KN)$ is constant with respect to the conditional expectation, and hence $\mathrm{E}_{u|y}[\sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t)/(KN) | \boldsymbol{y}; \boldsymbol{\theta}] = \sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t)/(KN)$.

APPENDIX D. ECME ALGORITHM

D.1 The Log-likelihood function

We derive expressions for $\Sigma_k(\boldsymbol{\gamma})^{-1}$ and simplify the log-likelihood expression in (5.12). Using the matrix inversion lemma [51], we get:

$$\Sigma_k(\boldsymbol{\gamma})^{-1} = \sigma^{-2}\left[I_{n_{\mathrm{R}}N} - Z_k(Z_k{}^H Z_k)^{-1} Z_k{}^H\right] + Z_k W_k(\boldsymbol{\gamma}) Z_k{}^H \tag{D.1}$$

which further implies

$$\Upsilon_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} = \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) Z_k{}^H. \tag{D.2}$$

Now, the log-likelihood (5.12) can be rewritten as

$$\begin{aligned} L(\boldsymbol{\rho}) \;=\; & -K n_{\mathrm{R}} N \ln(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{k=1}^{K} \boldsymbol{y}_k^H \left[I_{n_{\mathrm{R}}N} - Z_k(Z_k{}^H Z_k)^{-1} Z_k{}^H\right] \boldsymbol{y}_k - \sum_{k=1}^{K} (\boldsymbol{z}_k \\ - \;& \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x})^H W_k(\boldsymbol{\gamma})\,(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}) - \sum_{k=1}^{K} \ln\left|I_{n_{\mathrm{R}}n_{\mathrm{T}}} + (1/\sigma^2)\cdot Z_k{}^H Z_k\, \Psi\right| \end{aligned} \tag{D.4}$$

which follows by using (D.1) and the fact that

$$|\Sigma_k(\boldsymbol{\gamma})| = (\sigma^2)^{n_{\mathrm{R}}N} \cdot \left|I_{n_{\mathrm{R}}n_{\mathrm{T}}} + (1/\sigma^2)\cdot Z_k{}^H Z_k\, \Psi\right| \tag{D.5}$$

see (5.11b) and [51, Theorem 18.1.1].

D.2 Restricted maximum likelihood

We derive a REML log-likelihood expression for the measurement model in Section 5.2. Define a vector of error contrasts $\boldsymbol{u} = B^H \boldsymbol{y}$, where B is a matrix whose columns span the space orthogonal to the column space of Υ. Then, the REML log-likelihood is obtained as the log-likelihood function for the error contrasts. Without loss of generality, we choose

$B^H B = I_{Kn_{\mathrm{R}}N-1}$. Then, (5.16) follows by using the identities

$$\begin{aligned} |B^H \Sigma(\gamma) B| &= |\Upsilon^H \Sigma(\gamma)^{-1} \Upsilon| \cdot |\Sigma(\gamma)| \,/\, |\Upsilon^H \Upsilon| \\ B\,[B^H \Sigma(\gamma) B]^{-1}\, B^H &= \Pi(\gamma) \end{aligned}$$

and neglecting terms that do not depend on γ. Note that (5.16) can be further simplified by using (D.4):

$$\begin{aligned} L_{\mathrm{REML}}(\gamma) \;=\; & -Kn_{\mathrm{R}}N \ln \sigma^2 - \sum_{k=1}^{K} \ln |I_{n_{\mathrm{R}}n_{\mathrm{T}}} + Z_k{}^H Z_k \Psi/\sigma^2| - \frac{1}{\sigma^2} \sum_{k=1}^{K} \boldsymbol{y}_k^H \,[I_{n_{\mathrm{R}}N} \\ & - Z_k (Z_k{}^H Z_k)^{-1} Z_k{}^H]\, \boldsymbol{y}_k - \ln \Big| \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k(\gamma) \mathcal{A}_{\mathrm{LOS},k} \Big| \\ & - \sum_{k=1}^{K} [\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \widehat{\boldsymbol{x}}(\gamma)]^H W_k(\gamma)\, [\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \widehat{\boldsymbol{x}}(\gamma)] \end{aligned} \tag{D.8}$$

where $\widehat{\boldsymbol{x}}(\gamma)$ has been defined in (5.13).

D.3 ECME algorithms for correlated fading

D.3.1 ML estimation

We derive the ECME ML algorithm in Section 5.3.1.1. The first conditional maximization (CM) step in (5.17b) follows by substituting the most recent estimate of γ into (5.13), where (5.13) is the ML estimate of $\boldsymbol{x}$ that maximizes the observed-data likelihood function for fixed γ. The second CM step in (5.18a) follows by reparametrizing the variance components using $\Psi' = \Psi/\sigma^2$ and σ^2 (rather than Ψ and σ^2), which allows us to find the closed-form solution for the ML estimate of σ^2 that maximizes the observed-data likelihood function for fixed Ψ' and $\boldsymbol{x}$:

$$\begin{aligned} \widehat{\sigma}^2(\boldsymbol{x}, \Psi') \;=\; & \frac{1}{Kn_{\mathrm{R}}N} \sum_{k=1}^{K} \{\boldsymbol{y}_k^H \,[I_{n_{\mathrm{R}}N} - Z_k (Z_k{}^H Z_k)^{-1} Z_k{}^H]\, \boldsymbol{y}_k \\ & + (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x})^H\, W_k'(\Psi')\, (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x})\} \end{aligned} \tag{D.10}$$

where

$$W_k'(\Psi') = (Z_k{}^H Z_k + Z_k{}^H Z_k \Psi' Z_k{}^H Z_k)^{-1}. \tag{D.11}$$

Equation (D.9a) follows by maximizing the reparametrized log-likelihood function in (D.4) with $W_k(\gamma)$ replaced by $(1/\sigma^2) \cdot W_k'(\Psi')$.

Then, the second CM step for updating the estimate of σ^2 in (5.18a) is obtained by substituting the most recent estimates of $\boldsymbol{x}$ and $\Psi' = \Psi/\sigma^2$ into (D.9a). Finally, we apply the standard EM algorithm to update Ψ by treating $\boldsymbol{h}_{\mathrm{SC},k}$ as the missing data, where $\boldsymbol{x}$ and σ^2 are *fixed.*

Consequently,

$$\boldsymbol{r}_k = \boldsymbol{y}_k - \Upsilon_k \boldsymbol{x}, \quad k = 1, 2, \ldots, K \tag{D.12}$$

are the *observed data.* If $\boldsymbol{h}_{\mathrm{SC},k}$, $k = 1, 2, \ldots, K$ were known (forming the *complete data* together with $\boldsymbol{r}_k$, $k = 1, 2, \ldots, K$), we could easily find the complete-data ML estimate of Ψ as follows:

$$\widehat{\Psi} = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{h}_{\mathrm{SC},k} \boldsymbol{h}_{\mathrm{SC},k}^H \tag{D.13}$$

where $\widehat{\Psi}$ is also the *natural complete-data sufficient statistic* for estimating Ψ. Then, the third CM step for updating the estimate of Ψ in (5.18b) follows by computing the conditional expectation of (D.13) given the observed data $\boldsymbol{r}_k$, $k = 1, 2, \ldots, K$. We first find the distribution of the missing data $\boldsymbol{h}_{\mathrm{SC},k}$, $k = 1, 2, \ldots, K$ conditional on the observed data $\boldsymbol{r}_k$, $k = 1, 2, \ldots, K$. The joint distribution of $\boldsymbol{r}_k$ and $\boldsymbol{h}_{\mathrm{SC},k}$ is complex Gaussian with mean and covariance

$$\mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k} \end{bmatrix} \right\} = \boldsymbol{0} \tag{D.14a}$$

$$\mathrm{cov}\left\{ \begin{bmatrix} \boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k} \end{bmatrix} \right\} = \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k} \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k} \end{bmatrix}^H \right\} = \begin{bmatrix} \Sigma_k(\boldsymbol{\gamma}) & Z_k \Psi \\ \Psi^H Z_k{}^H & \Psi \end{bmatrix} \tag{D.14b}$$

and then [58, result 7] implies that $\boldsymbol{h}_{\mathrm{SC},k}$ conditional on $\boldsymbol{r}_k$ are complex Gaussian vectors with means and covariances equal to

$$\begin{aligned} \mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}|\boldsymbol{r}_k] &= \Psi Z_k{}^H \Sigma_k(\boldsymbol{\gamma})^{-1} \boldsymbol{r}_k = \Psi Z_k{}^H Z_k W_k(\boldsymbol{\gamma}) Z_k{}^H \boldsymbol{r}_k \\ &= \Psi Z_k{}^H Z_k W_k(\boldsymbol{\gamma})(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}) \qquad \text{(D.15a)} \\ \mathrm{cov}(\boldsymbol{h}_{\mathrm{SC},k}|\boldsymbol{r}_k) &= \Psi - \Psi Z_k{}^H \Sigma_k(\boldsymbol{\gamma})^{-1} Z_k \Psi = \Psi - \Psi Z_k{}^H Z_k W_k(\boldsymbol{\gamma}) Z_k{}^H Z_k \Psi \qquad \text{(D.15b)} \end{aligned}$$

where (D.15a) and (D.15b) follow by using (D.1) and (5.14). Now, (5.17c) follows from (D.15a) and the third CM step in (5.18b) is obtained by substituting the most recent estimates of $\mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}|\boldsymbol{r}_k]$ and Ψ into

$$\mathrm{E}[\widehat{\Psi} \,|\, \boldsymbol{r}_k] = \frac{1}{K} \sum_{k=1}^{K} \left\{ \mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}|\boldsymbol{r}_k] \, \mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}|\boldsymbol{r}_k]^H + \Psi \left[I_{n_{\mathrm{R}} n_{\mathrm{T}}} - Z_k{}^H Z_k W_k(\boldsymbol{\gamma}) Z_k{}^H Z_k \Psi \right] \right\}. \tag{D.16}$$

D.3.2 REML estimation

We derive the ECME REML algorithm in Section 5.3.1.2. The CM step for updating the fading covariance matrix in (5.23b) follows by replacing $\Sigma_k(\boldsymbol{\gamma})^{-1}$ in (D.16) with (see [65, eq. (3.10)] and [78, Section 2.3])

$$\begin{aligned} & \Sigma_k(\boldsymbol{\gamma})^{-1} - \Sigma_k(\boldsymbol{\gamma})^{-1} \Upsilon_k \Big[\sum_{l=1}^{K} \Upsilon_l^H \Sigma_l(\boldsymbol{\gamma})^{-1} \Upsilon_l \Big]^{-1} \Upsilon_k^H \, \Sigma_k(\boldsymbol{\gamma})^{-1} \\ = \; & [(\sigma^2)^{(i)}]^{-1} [I_{n_{\mathrm{R}}N} - Z_k (Z_k{}^H Z_k)^{-1} Z_k{}^H] + Z_k W_k(\boldsymbol{\gamma}) Z_k{}^H \\ & - Z_k W_k(\boldsymbol{\gamma}) \mathcal{A}_{\mathrm{LOS},k} \Big[\sum_{l=1}^{K} \mathcal{A}_{\mathrm{LOS},l}^H W_l(\boldsymbol{\gamma}) \mathcal{A}_{\mathrm{LOS},l} \Big]^{-1} \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) Z_k{}^H \end{aligned} \tag{D.18}$$

and substituting the most recent estimates of $\boldsymbol{x}$ and Ψ. To obtain the right-hand side of (D.18), we used (D.1) and (D.2). The CM step in (5.23a) follows by reparametrizing the variance components using $\Psi' = \Psi/\sigma^2$ and σ^2, which allows us to find the closed-form solution for the REML estimate of σ^2 that maximizes the observed-data restricted likelihood function for fixed Ψ' and $\boldsymbol{x}$:

$$\begin{aligned} \widehat{\sigma}^2(\Psi') \; = \; & \frac{K n_{\mathrm{R}} N}{K n_{\mathrm{R}} N - r} \cdot \widetilde{\sigma}^2 + \frac{1}{K n_{\mathrm{R}} N - r} \cdot \\ & \sum_{k=1}^{K} [\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \, \widehat{\boldsymbol{x}}(\Psi')]^H \, W_k'(\Psi') \, [\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \, \widehat{\boldsymbol{x}}(\Psi')] \end{aligned} \tag{D.20}$$

where

$$\widehat{\boldsymbol{x}}(\Psi') = \Big[\sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k'(\Psi') \mathcal{A}_{\mathrm{LOS},k} \Big]^{-1} \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k'(\Psi') \boldsymbol{z}_k \tag{D.21}$$

and $W_k'(\Psi')$ has been defined in (D.11). Then, the CM step for updating the estimate of σ^2 in (5.23a) is obtained by substituting the most recent estimates of $\widehat{\boldsymbol{x}}(\Psi')$ and $\Psi' = \Psi/\sigma^2$ into (D.19a).

D.4 ML estimation for correlated fading with constant $\Phi_k \Phi_k{}^H$

We derive the ML estimation algorithm in Section 5.3.2, described by equation (5.26) and iteration (5.27)–(5.28). We estimate $\boldsymbol{\rho}$ by iterating between the following two steps:

(i) **LOS component estimation:** fix $\boldsymbol{\gamma}$ and estimate $\boldsymbol{x}$ using (5.13);

(ii) **variance-component estimation:** fix $\boldsymbol{x}$, compute $\boldsymbol{r}_k$ using (D.12), and estimate the variance-component vector $\boldsymbol{\gamma}$ by maximizing the log-likelihood function (5.12):

$$L(\boldsymbol{\gamma};\boldsymbol{x}) = -\sum_{k=1}^{K} [\boldsymbol{r}_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} \boldsymbol{r}_k + \ln|\pi \Sigma_k(\boldsymbol{\gamma})|]. \tag{D.22}$$

Note that the above *alternating-projection* approach is embedded in all the algorithms discussed in this paper. We show that, for constant $\Phi_k \Phi_k^H = \Gamma_\Phi$ and fixed $\boldsymbol{x}$, the ML estimates of Ψ and σ^2 are their method-of-moments estimates:

$$\begin{aligned}
\widehat{\sigma}^2_{\mathrm{ML}} &= \frac{1}{K n_{\mathrm{R}}(N-n_{\mathrm{T}})} \sum_{k=1}^{K} \boldsymbol{r}_k^H \left[I_{n_{\mathrm{R}}N} - Z_k (Z_k^H Z_k)^{-1} Z_k^H\right] \boldsymbol{r}_k \\
&= \frac{1}{K n_{\mathrm{R}}(N-n_{\mathrm{T}})} \sum_{k=1}^{K} \boldsymbol{y}_k^H \left[I_{n_{\mathrm{R}}N} - Z_k C^{-1} Z_k^H\right] \boldsymbol{y}_k \qquad \text{(D.23a)} \\
\widehat{\Psi}_{\mathrm{ML}}(\boldsymbol{x}) &= \frac{1}{K} \sum_{k=1}^{K} \left[C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1}\right] - \widehat{\sigma}^2_{\mathrm{ML}} \cdot C^{-1} \qquad \text{(D.23b)}
\end{aligned}$$

where (D.23a) follows from (D.23a) by using (5.25) and (D.1).

We now prove that the expressions (D.23) indeed maximize (D.22). First, simplify (D.22) using $C = Z_k^H Z_k$ and (D.4):

$$\begin{aligned}
L(\boldsymbol{\gamma};\boldsymbol{x})\Big|_{Z_k^H Z_k = C} &= -\sum_{k=1}^{K} \ln|\pi \Sigma_k(\boldsymbol{\gamma})| - \frac{1}{\sigma^2} \cdot K n_{\mathrm{R}}(N-n_{\mathrm{T}}) \cdot \widehat{\sigma}^2_{\mathrm{ML}} \\
&\quad - \mathrm{tr}\left[\left(\sigma^2 C^{-1} + \Psi\right)^{-1} \sum_{k=1}^{K} C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1}\right]. \qquad \text{(D.24)}
\end{aligned}$$

The above log-likelihood function is the logarithm of a multivariate complex Gaussian pdf which belongs to the multiparameter exponential family of distributions [18, ch. 1.6.2]. This fact can be directly verified by inspecting (D.24), which is a linear function of the following *natural sufficient statistics*: $\widehat{\sigma}^2_{\mathrm{ML}}$ and $(1/K) \cdot \sum_{k=1}^{K} C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1}$. Then, [18, Theorem 2.3.1] implies that the moment estimates of σ^2 and Ψ in (D.23) are also their ML estimates, provided that $\widehat{\Psi}_{\mathrm{ML}}$ in (D.23b) is positive semidefinite.

D.5 Cramer-Rao bound

We derive the CRB for the vector of unknown parameters $\boldsymbol{\rho}$. Using Fisher information matrix [58, p. 525], the CRB for $\boldsymbol{\rho}$ is

$$\mathrm{CRB} = \begin{bmatrix} \mathrm{CRB}_{x,x} & 0 \\ 0 & \mathrm{CRB}_{\gamma,\gamma} \end{bmatrix} = \begin{bmatrix} \mathcal{I}_{x,x}^{-1} & 0 \\ 0 & \mathcal{I}_{\gamma,\gamma}^{-1} \end{bmatrix} \tag{D.25}$$

where

$$\mathcal{I}_{x,x} = \mathcal{I}_{[\mathrm{Re}\{\boldsymbol{x}\}^T,\mathrm{Im}\{\boldsymbol{x}\}^T]^T,[\mathrm{Re}\{\boldsymbol{x}\}^T,\mathrm{Im}\{\boldsymbol{x}\}^T]^T} 2\,\mathrm{Re}\Big\{\sum_{k=1}^{K} D_{x,k}^H \Sigma_k(\boldsymbol{\gamma})^{-1} D_{x,k}\Big\} \tag{D.26a}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{p,q} = \sum_{k=1}^{K} \mathrm{tr}\Big\{\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_p}\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_q}\Big\} \tag{D.26b}$$

for $p, q = 1, 2, \ldots, \dim(\boldsymbol{\gamma})$, and

$$D_{x,k} = \Big[\frac{\partial(\Upsilon_k \boldsymbol{x})}{\partial(\mathrm{Re}\{\boldsymbol{x}\})^T}, \frac{\partial(\Upsilon_k \boldsymbol{x})}{\partial(\mathrm{Im}\{\boldsymbol{x}\})^T}\Big] = [1, j] \otimes \Upsilon_k.$$

Substituting (D.28) into (D.26a) yields

$$\mathcal{I}_{x,x} = \mathrm{Re}\Big\{\begin{bmatrix} 1 & j \\ -j & 1 \end{bmatrix} \otimes P_{x,x}(\boldsymbol{\gamma})\Big\} \tag{D.28}$$

where

$$P_{x,x}(\boldsymbol{\gamma}) = 2 \cdot \sum_{k=1}^{K} \Upsilon_k^H\, \Sigma_k(\boldsymbol{\gamma})^{-1}\, \Upsilon_k = 2 \cdot \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H\, W_k(\boldsymbol{\gamma})\, \mathcal{A}_{\mathrm{LOS},k}. \tag{D.29}$$

It is easy to show that

$$\mathrm{CRB}_{x,x} = \mathcal{I}_{x,x}{}^{-1} = \begin{bmatrix} \mathrm{Re}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} & -\mathrm{Im}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} \\ \mathrm{Im}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} & \mathrm{Re}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} \end{bmatrix}. \tag{D.30}$$

We now use (D.1) to simplify (D.26b). For $p = 1$,

$$\begin{aligned}[\mathcal{I}_{\gamma,\gamma}]_{1,q} = [\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\gamma_q} &= \sum_{k=1}^{K} \mathrm{tr}\Big[\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \sigma^2}\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_q}\Big] \\ &= \begin{cases} n_{\mathrm{R}}(N - n_{\mathrm{T}})K\sigma^{-4} + \sum_{k=1}^{K} \mathrm{tr}\left[{Z_k}^H Z_k W_k(\boldsymbol{\gamma}) {Z_k}^H Z_k W_k(\boldsymbol{\gamma})\right], & q = 1 \\ \sum_{k=1}^{K} \mathrm{tr}\left[{Z_k}^H Z_k W_k(\boldsymbol{\gamma}) {Z_k}^H Z_k W_k(\boldsymbol{\gamma}) {Z_k}^H Z_k \cdot (\partial \Psi / \partial \psi_{q-1})\right], & q = 2, \ldots, \dim(\boldsymbol{\gamma}) \end{cases}\end{aligned} \tag{D.31}$$

whereas for $p, q > 1$ we have

$$\begin{aligned}[\mathcal{I}_{\gamma,\gamma}]_{p,q} &= [\mathcal{I}_{\gamma,\gamma}]_{\psi_{p-1},\psi_{q-1}} [\mathcal{I}_{\psi\psi}]_{p-1,q-1} = \sum_{k=1}^{K} \mathrm{tr}\Big[\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \psi_{p-1}}\Sigma_k(\boldsymbol{\gamma})^{-1}\frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \psi_{q-1}}\Big] \\ &= \sum_{k=1}^{K} \mathrm{tr}\Big[{Z_k}^H Z_k W_k(\boldsymbol{\gamma}) {Z_k}^H Z_k \frac{\partial \Psi}{\partial \psi_{p-1}} {Z_k}^H Z_k W_k(\boldsymbol{\gamma}) {Z_k}^H Z_k \frac{\partial \Psi}{\partial \psi_{q-1}}\Big].\end{aligned} \tag{D.32}$$

We partition $\mathcal{I}_{\gamma,\gamma}$ as

$$\mathcal{I}_{\gamma,\gamma} = \begin{bmatrix} i_{\sigma^2\sigma^2} & \boldsymbol{i}_{\psi,\sigma^2}^T \\ \boldsymbol{i}_{\psi,\sigma^2} & \mathcal{I}_{\psi,\psi} \end{bmatrix} \tag{D.33}$$

where i_{σ^2,σ^2} and $\boldsymbol{i}_{\psi,\sigma^2}$ are computed using (D.31), and $\mathcal{I}_{\psi,\psi}$ is computed using (D.32). We adopt the same block partitioning of $\mathrm{CRB}_{\gamma\gamma} = \mathcal{I}_{\gamma\gamma}^{-1}$. Then,

$$\mathrm{CRB}_{\psi,\psi} = \mathcal{I}_{\psi,\psi}^{-1} + \frac{\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}\boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}}{i_{\sigma^2,\sigma^2} - \boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}} \tag{D.34a}$$

$$\mathrm{CRB}_{\sigma^2\sigma^2} = \frac{1}{i_{\sigma^2,\sigma^2} - \boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}} \tag{D.34b}$$

which follow by using the formula for the inverse of a partitioned matrix [51].

D.5.1 CRB for correlated fading with constant $\Phi_k\Phi_k^H$

We simplify the above CRB expressions with correlated fading and constant $\Phi_k\Phi_k^H = \Gamma_\Phi$. In this case,

$$W_k(\boldsymbol{\gamma}) = W(\boldsymbol{\gamma}) = (\sigma^2 C + C\Psi C)^{-1}$$

where C has been defined in (5.25). For the unstructured LOS array response model in (5.6), (D.30) simplifies to

$$\mathrm{CRB}_{x,x} = \frac{1}{2K}\cdot\begin{bmatrix} \mathrm{Re}\{\sigma^2 C + C\Psi C\} & -\mathrm{Im}\{\sigma^2 C + C\Psi C\} \\ \mathrm{Im}\{\sigma^2 C + C\Psi C\} & \mathrm{Re}\{\sigma^2 C + C\Psi C\} \end{bmatrix}. \tag{D.35}$$

Under the structured LOS array response model in (5.7), (D.30) becomes

$$\mathrm{CRB}_{x,x} = \frac{1}{2K\cdot(\boldsymbol{a}_{\mathrm{T,LOS}}^H\otimes\boldsymbol{a}_{\mathrm{R,LOS}}^H)\,W(\boldsymbol{\gamma})\,(\boldsymbol{a}_{\mathrm{T,LOS}}\otimes\boldsymbol{a}_{\mathrm{R,LOS}})}\cdot I_2. \tag{D.36}$$

The equations in (D.31) simplify to

$$i_{\sigma^2,\sigma^2} = K\cdot\{n_{\mathrm{R}}(N-n_{\mathrm{T}})\sigma^{-4} + \mathrm{tr}[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})]\} \tag{D.37a}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\Psi_{p,p}} = K\cdot\mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p,p}}\right] = K\cdot[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,p} \tag{D.37b}$$

and, for $p > q$,

$$\begin{aligned}[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\mathrm{Re}\{\Psi_{p,q}\}} &= K\cdot\mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C\cdot\frac{\partial\Psi}{\partial\,\mathrm{Re}\{\Psi_{p,q}\}}\right] \\ &= 2K\cdot\mathrm{Re}\{[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,q}\} \end{aligned} \tag{D.38a}$$

$$\begin{aligned}[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\mathrm{Im}\{\Psi_{p,q}\}} &= K\cdot\mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C\cdot\frac{\partial\Psi}{\partial\,\mathrm{Im}\{\Psi_{p,q}\}}\right] \\ &= 2K\cdot\mathrm{Im}\{[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,q}\} \end{aligned} \tag{D.38b}$$

where $\Psi_{p,q}$ denotes the (p,q) element of Ψ, for $p,q = 1,2,\ldots,n_{\mathrm{R}}n_{\mathrm{T}}$. Also, for $p_1 > q_1$ and $p_2 > q_2$, (D.32) becomes

$$\begin{aligned}
&[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_1,q_1}\},\mathrm{Re}\{\Psi_{p_2,q_2}\}} = [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_2,q_2}\},\mathrm{Re}\{\Psi_{p_1,q_1}\}} \\
&= K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Re}\{\Psi_{p_1,q_1}\}}CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Re}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K\,\mathrm{Re}\left\{[CW(\gamma)C]_{q_2,p_1}[CW(\gamma)C]_{q_1,p_2} + [CW(\gamma)C]_{q_2,q_1}[CW(\gamma)C]_{p_1,p_2}\right\} \qquad \text{(D.39a)} \\
&[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_1,q_1}\},\mathrm{Im}\{\Psi_{p_2,q_2}\}} = [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\mathrm{Re}\{\Psi_{p_1,q_1}\}} \\
&= K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Re}\{\Psi_{p_1,q_1}\}}CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= -2K\,\mathrm{Im}\left\{[CW(\gamma)C]_{q_2,p_1}[CW(\gamma)C]_{q_1,p_2} + [CW(\gamma)C]_{q_2,q_1}[CW(\gamma)C]_{p_1,p_2}\right\} \qquad \text{(D.39b)} \\
&[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_1,q_1}\},\mathrm{Im}\{\Psi_{p_2,q_2}\}} = [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\mathrm{Im}\{\Psi_{p_1,q_1}\}} \\
&= K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Im}\{\Psi_{p_1,q_1}\}}CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K\,\mathrm{Re}\left\{-[CW(\gamma)C]_{q_2,p_1}[CW(\gamma)C]_{q_1,p_2} + [CW(\gamma)C]_{q_2,q_1}[CW(\gamma)C]_{p_1,p_2}\right\} \qquad \text{(D.39c)}
\end{aligned}$$

and, for $p_1 = q_1$ and $p_2 > q_2$,

$$\begin{aligned}
&[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\mathrm{Re}\{\Psi_{p_2,q_2}\}} = [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_2,q_2}\},\Psi_{p_1,p_1}} \\
&= K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Re}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K\,\mathrm{Re}\left\{[CW(\gamma)C]_{q_2,p_1}[CW(\gamma)C]_{p_1,p_2}\right\} \qquad \text{(D.40a)} \\
&[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\mathrm{Im}\{\Psi_{p_2,q_2}\}} = [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\Psi_{p_1,p_1}} \\
&= K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\gamma)C\frac{\partial\Psi}{\partial\,\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= -2K\,\mathrm{Im}\left\{[CW(\gamma)C]_{q_2,p_1}[CW(\gamma)C]_{p_1,p_2}\right\} \qquad \text{(D.40b)}
\end{aligned}$$

and, for $p_1 = q_1$ and $p_2 = q_2$,

$$[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\Psi_{p_2,p_2}} = K\,\mathrm{tr}\left[CW(\gamma)C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\gamma)C\frac{\partial\Psi}{\partial\Psi_{p_2,p_2}}\right] = K|[CW(\gamma)C]_{p_1,p_2}|^2. \qquad \text{(D.41)}$$

D.5.2 CRB for independent fading with $\Phi_k\Phi_k^H = I_{n_{\mathrm{T}}}$

For independent fading and unitary space-time codes, i.e., $Z_k^H Z_k = I_{n_{\mathrm{R}}n_{\mathrm{T}}}$, we have:

$$W_k(\gamma) = W(\gamma) = \mathrm{diag}\{(\sigma^2+\psi_1)^{-1},\ldots,(\sigma^2+\psi_{n_{\mathrm{R}}n_{\mathrm{T}}})^{-1}\} \qquad \text{(D.42)}$$

and the CRB expressions (D.31) and (D.32) simplify to:

$$i_{\sigma^2,\sigma^2} = n_{\mathrm{R}}(N-n_{\mathrm{T}})K\cdot\sigma^{-4} + K\cdot\sum_{n=1}^{n_{\mathrm{R}}n_{\mathrm{T}}}(\sigma^2+\psi_n)^{-2} \qquad \text{(D.43a)}$$

$$\boldsymbol{i}_{\psi,\sigma^2} = K \cdot [(\sigma^2+\psi_1)^{-2}, (\sigma^2+\psi_2)^{-2}, \ldots, (\sigma^2+\psi_{n_{\mathrm{R}}n_{\mathrm{T}}})^{-2}]^T \quad \text{(D.43b)}$$

$$\mathcal{I}_{\psi,\psi} = K \cdot \mathrm{diag}\{(\sigma^2+\psi_1)^{-2}, (\sigma^2+\psi_2)^{-2}, \ldots, (\sigma^2+\psi_{n_{\mathrm{R}}n_{\mathrm{T}}})^{-2}\} \quad \text{(D.43c)}$$

which yields (5.38) after applying (D.34). The CRB expressions for the LOS coefficients under the unstructured and structured LOS array response models follow by substituting $C = I_{n_{\mathrm{R}}n_{\mathrm{T}}}$, $\Psi = \mathrm{diag}(\psi_1, \psi_2, \ldots, \psi_{n_{\mathrm{R}}n_{\mathrm{T}}})$ and (D.42) into (D.35) and (D.36).

BIBLIOGRAPHY

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formula, Graphs, and Mathematical Tables.* New York:Dover, 9th edition, 1970.

[2] A. A. Abu-Dayya and N. C. Beaulieu. Microdiversity on Rician fading channels. *IEEE Transactions on Communications*, 42:2258 – 2267, June 1994.

[3] S. A. Al-Semari and T. E. Fuja. Performance analysis of coherent TCM systems with diversity reception in slow Rayleigh fading. *IEEE Transactions on Vehicular Technology*, 48:198–212, 1999.

[4] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal of Selected Areas Communications*, 16:1451–1468, Oct. 1998.

[5] M.-S. Alouini, X. Qi, and Y.-C. Ko. Closed-form analysis of dual-diversity equal-gain combining over Rayleigh fading channels. *IEEE Transactions on Wireless Communications*, 2:1120–1125, Nov. 2003.

[6] M.-S. Alouini and M. K. Simon. An MGF-based performance analysis of generalized selection combining over Rayleigh fading channels. *IEEE Transactions on Communications*, 48:401–415, Mar. 2000.

[7] M.-S. Alouini and M. K. Simon. Performance analysis of coherent equal gain combining over Nakagami-m fading channels. *IEEE Transactions on Vehicular Technology*, 50:1449–1463, Nov. 2001.

[8] M.-S. Alouini and M. K. Simon. Performance analysis of generalized selection combining over Weibull fading channels. *Proc. 58th IEEE Vehicular Technology Conference*, pages 1735–1739, Oct. 2001.

[9] M.-S. Alouini, X. Tang, and A. Goldsmith. An adaptive modulation scheme for simultaneous voice and data transmission over fading channels. *IEEE Journal of Selected Areas Communications*, 17:837–850, May 1999.

[10] A. Annamalai and C. Tellambura. A new approach to performance evaluation of generalized selection diversity receivers in wireless channels. *Proc. 58th IEEE Vehicular Technology Conference*, pages 2309–2313, Fall 2001.

[11] A. Annamalai, C. Tellambura, and V. K. Bhargava. Exact evaluation of maximal-ratio and equal-gain diversity receivers for M-ary QAM on Nakagami fading channels. *IEEE Transactions on Communications*, 47:1335–1344, Sept. 1999.

[12] A. Annamalai, C. Tellambura, and V. K. Bhargava. Equal-gain diversity receiver performance in wireless channels. *IEEE Transactions on Communications*, 48:1732–1745, Oct. 2000.

[13] R. Annavajjala and L. B. Milstein. On the performance of diversity combining schemes on Rayleigh fading channels with noisy channel estimates. *Proc. IEEE Military Communications Conference*, pages 320–325, Oct. 2003.

[14] I. Barhumi, G. Leus, and M. Moonen. Optimal training design for MIMO OFDM systems in mobile wireless channels. *IEEE Transactions on Signal Processing*, 51:1615–1624, June 2003.

[15] N. C. Beaulieu. An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of rayleigh random variables. *IEEE Transactions on Communications*, 26:1463–1474, Sept. 1990.

[16] N. C. Beaulieu and A. A. Abu-Dayya. Analysis of equal gain diversity on Nakagami fading channels. *IEEE Transactions on Communications*, 39:225–234, Feb. 1991.

[17] J. O. Berger, B. Liseo, and R. L. Wolpert. Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science*, 14(1):1–29, 1999.

[18] P. J. Bickel and K. A. Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics.* Upper Saddle River,NJ: Prentice-Hall, 2nd edition, 2000.

[19] D. G. Brennan. Linear diversity combining techniques. *Proceedings of the IEEE*, 91(2):331–356, Feb. 2003.

[20] X. Cai and G. B. Giannakis. Adaptive PSAM accounting for channel estimation and prediction errors. *IEEE Transactions on Wireless Communications*, 4:246–256, Jan. 2005.

[21] L. Cao and N. C. Beaulieu. Exact error rate analysis of diversity 16-QAM with channel estimation error. *IEEE Transactions on Communications*, 52:1019–1029, June 2004.

[22] L. Cao and N. C. Beaulieu. Closed-form BER results for MRC diversity with channel estimation errors in Ricean fading channels. *IEEE Transactions on Wireless Communications*, 4:1440–1447, July 2005.

[23] J. K. Carvers. An analysis of pilot symbol assisted modulation for Rayleigh fading channels. *IEEE Transactions on Vehicular Technology*, 40:686–693, Nov. 1991.

[24] J. M. Chaufray, P. Loubaton, and P. Chevalier. Consistent estimation of Rayleigh fading channel second-order statistics in the context of the wideband CDMA mode of the UMTS. *IEEE Transactions on Signal Processing*, 49:3055–3064, Dec. 2001.

[25] Y. Chen and N. C. Beaulieu. Maximum likelihood estimation of the K factor in Ricean fading channels. *IEEE Communications Letters*, 9(12):1040 – 1042, 2005.

[26] J. Cheng, C. Tellambura, and N. C. Beaulieu. Performance analysis of digital modulations on Weibull fading channels. *Proc. 60th IEEE Vehicular Technology Conference*, 1:236–240, Oct. 2003.

[27] K. Cho and D. Yoon. On the general BER expression of one- and two-dimentional amplitude modulations. *IEEE Transactions on Communications*, 50:1074–1080, July 2002.

[28] G. E. Corazza and F. Vatalaro. A statistical model for land mobile satellite channels and its application to nongeostationary orbit systems. *IEEE Transactions on Vehicular Technology*, 43:738–742, Aug. 1994.

[29] M. Davidian and D. M. Giltinan. *Nonlinear Models for Repeated Measurement Data.* London,UK: Chapman & Hall, 1995.

[30] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, 39:1–38, July 1977.

[31] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[32] A. Dogandvzić. Chernoff bounds on pairwise error probabilities of space-time codes. *IEEE Transactions on Information Theory*, 49:1327–1336, May 2003.

[33] A. Dogandvzić and J. Jin. Estimating statistical properties of MIMO Ricean fading channels. *Proc. 2nd IEEE Sensor Array Multichannel Signal Processing Workshop*, pages 149–153, Aug. 2002.

[34] A. Dogandvzić and B. Zhang. Estimating Jakes' Doppler power spectrum parameters using the Whittle approximation. *IEEE Transactions on Signal Processing*, 53:987–1005, Mar. 2005.

[35] X. Dong and L. Xiao. Symbol error probability of two-dimensional signaling in ricean fading with imperfect channel estimation. *IEEE Transactions on Vehicular Technology*, 54:538–549, Mar. 2005.

[36] A. Duel-Hallen, S. Q. Hu, and H. Hallen. Long-range prediction of fading signals: Enabling adapting transmission for mobile radio channels. *IEEE Signal Processing Magazine*, 17(3):62–75, May 2000.

[37] G. Durisi and S. Benedetto. Performance of coherent and noncoherent receivers for UWB communications. *Proc. IEEE International Conference on Communications*, 6:3429–3433, June 2004.

[38] T. Eng, K. Ning, and L. B. Milstein. Comparison of diversity combining techniques for Rayleigh-fading channels. *IEEE Transactions on Communications*, 44:1117–1129, Dec. 1996.

[39] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela. Link-optimal space-time processing with multiple transmit and receive antennas. *IEEE Communications Letters*, 5:85–87, Mar. 2001.

[40] G. Fuks, J. Goldberg, and H. Messer. Bearing estimation in a Ricean channel — part i: Inherent accuracy limitations. *IEEE Transactions on Signal Processing*, 5:925–937, May 2001.

[41] A. Goldsmith. *Wireless Communications.* Cambridge University Press, 2005.

[42] A. Goldsmith and S.-G. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Transactions on Communications*, 45:1218–1230, Oct. 1997.

[43] A. Goldsmith, L. J. Greenstein, and G. J. Foschini. Error statistics of real-time power measurements in cellular channels with multipath and shadowing. *IEEE Transactions on Vehicular Technology*, 43:439–446, Aug. 1994.

[44] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE Journal of Selected Areas Communications*, 21:684–702, June 2003.

[45] A. J. Goldsmith and P. P. Varaiya. Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, 43:1986–1992, June 1997.

[46] C. H. Gowda, V. Annampedu, and R. Viswanathan. Diversity combining in antenna array base station receiver for DS/CDMA system. *IEEE Communications Letters*, 2:180–182, July 1998.

[47] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* New York: Academic, 5th edition, 1995.

[48] F. Hansen and F. I. Meno. Mobile fading—Rayleigh and lognormal superimposed. *IEEE Transactions on Vehicular Technology*, 26:332–335, Nov. 1977.

[49] L. Hanzo, W. Webb, and T. Keller. *Single- and multi-carrier quadrature amplitude modulation: principles and applications for personal communications, WLANs and broadcasting.* New York: IEEE Press, 2000.

[50] D. A. Harville. Bayesian inference for variance components using only error contrasts. *Biometrika*, 61:383-385, Aug. 1974.

[51] D. A. Harville. *Matrix Algebra From a Statistician's Perspective.* New York: Springer-Verlag, 1997.

[52] C. R. Henderson. Maximum likelihood estimation of variance components. 1973. unpublished manuscript, Dept. Animal Science, Cornell Univ., Ithaca, NY.

[53] B. M. Hochwald and T. L. Marzetta. Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading. *IEEE Transactions on Information Theory*, 46:543–564, Mar. 2000.

[54] C.-D. Iskander. Performance of coherent receivers with hybrid selection/equal-gain combining in Nakagami-m fading. *Proc. 61st IEEE Radio and Wireless Conference*, pages 111–114, Sept. 2004.

[55] S. K. Jayaweera and H. V. Poor. MIMO capacity results for rician fading channels. *Proc. IEEE Global Telecommunications Conference (GlobeCom)*, pages 1806–1810, June 2003.

[56] T. Jiang, N. D. Sidiropoulos, and G. B. Giannakis. Kalman filtering for power estimation in mobile communications. *IEEE Transactions on Wireless Communications*, 2:151–161, Jan. 2003.

[57] Y. Kamio and S. Sampei. Performance of a trellis-coded 16QAM/TDMA system for land mobile communications. *IEEE Transactions on Vehicular Technology*, 43:528–536, 1994.

[58] S. M. Kay. *Fundamentals of Statistical Signal Processing — Estimation Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[59] Y.-S. Kim, C.-J. Kim, G.-Y. Jeong, Y.-J. Bang, H.-K. Park, and S.-S. Choi. New Rayleigh fading channel estimator based on PSAM channel sounding technique. *Proc. IEEE International Conference on Communications*, pages 1518–1520, June 1997.

[60] Y.-C. Ko and M.-S. Alouini. Estimation of Nakagami-m fading channel parameters with application to optimized transmitter diversity systems. *IEEE Transactions on Wireless Communications*, 2(2):250 C 259, March 2003.

[61] Y.-C. Ko and M.-S. Alouini. Estimation of Nakagami-m fading channel parameters with application to optimized transmitter diversity systems. *IEEE Transactions on Wireless Communications*, 2:250–259, Mar. 2003.

[62] Y.-C. Ko, M.-S. Alouini, and M. K. Simon. Outage probability of diversity systems over generalized fading channels. *IEEE Transactions on Communications*, 48:1783–1787, Nov. 2000.

[63] N. Kong and L. B. Milstein. Average SNR of a generalized diversity selection combining scheme. *IEEE Communications Letters*, 3:57–59, Mar. 1999.

[64] J. H. Kotecha and A. M. Sayeed. Transmit signal design for optimal estimation of correlated MIMO channels. *IEEE Transactions on Signal Processing*, 52:546–557, Feb. 2004.

[65] N. Laird, N. Lange, and D. Stram. Maximum likelihood computations with repeated measures: Applications of the EM algorithm. *Journal of America Statistical Association*, 82:97–105, 1987.

[66] C. H. Liu and D. B. Rubin. The ECME algorithm: A simple extension of EM and ECM with fast monotone convergence. *Biometrika*, 81:633–648, Dec. 1994.

[67] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke. The land mobile satellite communication channel—recording, statistics, and channel model. *IEEE Transactions on Vehicular Technology*, 40:375–386, May 1991.

[68] Q. Ma and C. Tepedelenlioglu. Practical multiuser diversity with outdated channel feedback. *IEEE Transactions on Vehicular Technology*, 54:1334–1345, July 2005.

[69] Y. Ma and C. C. Chai. Unified error probability analysis for generalized selection combining in Nakagami fading channels. *IEEE Journal of Selected Areas Communications*, 18:2198–2210, Nov. 2000.

[70] Y. Ma and S. Pasupathy. Efficient performance evaluation for generalized selection combining on generalized fading channels. *IEEE Transactions on Wireless Communications*, 3:29–34, Jan. 2004.

[71] Y. Ma, R. Schober, and S. Pasupathy. Effect of imperfect channel estimation on MRC diversity in Rician fading channels. *IEEE Transactions on Vehicular Technology*, 54:2137–2142, Nov. 2005.

[72] Y. Ma, R. Schober, and S. Pasupathy. Performance of M-PSK with GSC and EGC with Gaussian weighting errors. *IEEE Transactions on Vehicular Technology*, 54:149–162, Jan. 2005.

[73] Y. Ma, R. Schober, and D. Zhang. Exact BER of M-QAM with MRC and imperfect channel estimation in Rician fading channels. *to appear in IEEE Transactions on Wireless Communications*, 2006.

[74] Y. Ma, Z. Wang, and S. Pasupathy. Asymptotic performance of wireless communications with generalized selection combining. *Proc. IEEE Global Telecommunications Conference (GlobeCom)*, pages 1679–1683, Dec. 2003.

[75] T. L. Marzetta. EM algorithm for estimating the parameters of a multivariate complex Rician density for polarimetric SAR. *Proc. IEEE International Conference Acoustic, Speech, Signal Processing*, pages 3651–3654, May 1995.

[76] P. McCullagh and J. A. Nelder. *Generalized Linear Models.* London, UK: Chapman & Hall, 2nd edition, 1989.

[77] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions.* New York: Wiley, 1997.

[78] X.-L. Meng and D. van Dyk. Fast EM-type implementations for mixed effects models. *Journal of the Royal Statistical Society: Series B*, 60:559–578, 1998.

[79] T. J. Moulsley and E. Vilar. Experimental and theoretical statistics of microwave amplitude scintillations on satellite down-links. *IEEE Transactions on Antennas Propagation*, 30:1099–1106, Nov. 1982.

[80] M. K. Ozdemir, E. Arvas, and H. Arslan. Dynamics of spatial correlation and implications on MIMO systems. *IEEE Communications Magazine*, 42:514–519, June 2004.

[81] Y. Pawitan. *In All Likelihood: Statistical Modelling and Inference Using Likelihood.* New York: Oxford University Press, 2001.

[82] M. Peritsky. Statistical estimation of mean signal strength in a Rayleigh-fading environment. *IEEE Transactions on Communications*, 21(11):1207 C 1213, Nov. 1973.

[83] P. Polydorou and P. Ho. Symbol error probability for double selection/equal gain combining of branches in a Rayleigh fading environment without channel statistical information. *Proc. IEEE Global Telecommunications Conference (GlobeCom)*, pages 3343–3347, 2001.

[84] B. Porat. *A Course in Digital Signal Processing.* New York: Wiley, 1997.

[85] W. H. Press, S. A. Teukolsky, W. T. Vettering, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge, UK: Cambirdge Univ. Press, 1992.

[86] Y. Qiao, S. Yu, P. Su, and L. Zhang. Research on an iterative algorithm of LS channel estimation in MIMO OFDM systems. *IEEE Transactions on Broadcasting*, 51:149–153, Mar. 2005.

[87] W. Roh and A. Paulraj. MIMO channel capacity for the distributed antenna systems. *Proc. 56th IEEE Vehicular Technology Conference*, pages 706–709, Sept. 2002.

[88] W. Roh and A. Paulraj. Outage performance of the distributed antenna systems in a composite fading channel. *Proc. 56th IEEE Vehicular Technology Conference*, pages 1520–1524, Sept. 2002.

[89] P. R. Sahu and A. K. Chaturvedi. Performance analysis of predetection EGC receiver in Weibull fading channel. *IEEE Electronics Letters*, 41:85 – 86, Jan. 2005.

[90] S. Sanayei and A. Nosratinia. Antenna selection in MIMO systems. *IEEE Communications Magazine*, 42:68–73, Oct. 2004.

[91] V. Saravanan and T. F. Wong. Equal gain combining for acquisition of UWB signals. *Proc. IEEE Military Communications Conference*, 2:880–885, Oct. 2003.

[92] G. A. F. Seber and C. J. Wild. *Nonlinear Regression.* New York: Wiley, 1989.

[93] M. G. Shayesteh and A. Aghamohammadi. On the error probability of linearly modulated signals on frequency-flat Ricean, Rayleigh and AWGN channels. *IEEE Transactions on Communications*, 43:1454–1466, Feb./Mar./Apr. 1995.

[94] M. K. Simon and M.-S. Alouini. *Digital Communication over Fading Channels.* New York: Wiley, 2000.

[95] G. Stuber. *Principles of Mobile Communication.* Norwell, MA: Kluwer, 3 edition, 2003.

[96] H. Suzuki. A statistical model for urban radio propagation. *IEEE Transactions on Communications*, 25:673–680, July 1977.

[97] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Transactions on Information Theory*, 45:1456–1467, July 1999.

[98] R. A. Thisted. *Elements of Statistical Computing: Numerical Computation.* New York: Chapman & Hall, 1988.

[99] C. van der Plas and J.-P. M. G. Linnartz. Stability of mobile slotted ALOHA network with Rayleigh fading, shadowing, and near-far effect. *IEEE Transactions on Vehicular Technology*, 39:359–366, Nov. 1990.

[100] R. van Nee and R. Prasad. *OFDM wireless multimedia communications.* Boston: Artech House, 2000.

[101] F. Vatalaro, F. Mazzenga, G. De Maio, and A. Forcella. The generalized Rice lognormal channel model—first and second order statistical characterization and simulation. *International Journal Satellite Communications*, 20:29–45, Jan.-Feb. 2002.

[102] E. F. Vonesh and V. M. Chinchilli. *Linear and Nonlinear Models for the Analysis of Repeated Measurements.* New York: Marcel Dekker, 1997.

[103] S. K. Wilson and J. M. Cioffi. Probability density functions for analyzing multi-amplitude constellations in Ryaleigh and Ricean channels. *IEEE Transactions on Communications*, 47:380–386, Mar. 1999.

[104] M. Z. Win and J. H. Winters. Analysis of hybrid selection/maximum-ratio combining in Rayleigh fading. *IEEE Transactions on Communications*, 47:1773–1776, Dec. 1999.

[105] B. Xia and J. Wang. Effect of channel-estimation error on QAM systems with antenna diversity. *IEEE Transactions on Communications*, 53:481–488, Mar. 2005.

[106] A. Zeira and B. Friedlander. Direction finding with time-varying arrays. *IEEE Transactions on Signal Processing*, 43:927–937, Apr. 1995.

[107] Q.T. Zhang. A note on the estimation of Nakagami-m fading parameter. *IEEE Communications Letters*, 6(6):237 C 238, June 2002.

[108] S. Zhou and G. B. Giannakis. Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations. *IEEE Transactions on Information Theory*, 49:1673–1690, July 2003.