Channel Estimation for OFDM in Fast Fading Channels

by

Ping Wan

B.Eng., Zhengzhou Institute of Technology, 1990

M.Sc., Chinese Aeronautical Establishment, 1993

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Channel Estimation for OFDM in Fast Fading Channels

by

Ping Wan

B.Eng., Zhengzhou Institute of Technology, 1990

M.Sc., Chinese Aeronautical Establishment, 1993

Supervisory Committee

Dr. Michael McGuire, Co-Supervisor

(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, Co-Supervisor

(Department of Electrical and Computer Engineering)

Dr. Aaron Gulliver, Departmental Member

(Department of Electrical and Computer Engineering)

Dr. Mihai Sima, Departmental Member

(Department of Electrical and Computer Engineering)

Dr. Kui Wu, Outside Member

(Department of Computer Science)

Supervisory Committee

Dr. Michael McGuire, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Aaron Gulliver, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Mihai Sima, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Kui Wu, Outside Member
(Department of Computer Science)

ABSTRACT

The increasing demand for high data rate transmission over broadband radio channels has imposed significant challenges in wireless communications. Accurate channel estimation has a major impact on the whole system performance. Specifically, reliable estimate of the channel state information (CSI) is more challenging for orthogonal frequency division multiplexing (OFDM) systems in doubly selective fading channels than for the slower fading channels over which OFDM has been deployed traditionally. With the help of a basis expansion model (BEM), a novel multivariate autoregressive

(AR) process is developed to model the time evolution of the fast fading channel. Relying on pilot symbol aided modulation (PSAM), a novel Kalman smoothing algorithm based on a second-order dynamic model is exploited, where the mean square error (MSE) of the channel estimator is near to that of the optimal Wiener filter. To further improve the performance of channel estimation, a novel low-complexity iterative joint channel estimation and symbol detection procedure is developed for fast fading channels with a small number of pilots and low pilot power to achieve the bit error rate (BER) performance close to when the CSI is known perfectly. The new channel estimation symbol detection technique is robust to variations of the radio channel from the design values and applicable to multiple modulation and coding types. By use of the extrinsic information transfer (EXIT) chart, we investigate the convergence behavior of the new algorithm and analyze the modulation, pilot density, and error correction code selection for good system performance for a given power level. The algorithms developed in this thesis improve the performance of the whole system requiring only low ratios of pilot to data for excellent performance in fast fading channels.

Contents

Supervisory Committee ii

Abstract iii

Table of Contents v

List of Abbreviations ix

List of Symbols xii

List of Figures xvi

Acknowledgements xviii

Dedication xix

1 Introduction 1

1.1 Background 1

1.2 History of Channel Estimation 2

1.3 Technical Challenges of Channel Estimation 3

1.4 Contributions 6

1.5 Thesis Outline 7

2 Literature Review of Channel Estimation for OFDM Systems 9

2.1 OFDM 9

2.2 Channel Estimation Techniques 13
2.2.1 Channel Estimation in Slow Fading Channels 15
2.2.2 Channel Estimation in Fast Fading Channels 18
2.2.3 Joint Channel Estimation Data Detection 22

3 System Model **25**
3.1 Time Blocks for Channel Processing 25
3.1.1 Time Domain and Frequency Domain Model for a Single OFDM Block 27
3.1.2 Time Domain Descriptions for a Transmission Block 30
3.1.3 Coding Block 33
3.2 Three Measurement Models Introduction 34
3.2.1 Measurements for Data Given Channel 34
3.2.2 Measurements for Channel Given Pilots 35
3.2.3 Measurements for Channel Given Data 37
3.3 Time Domain Evolution of the Channel 39

4 Second-Order Kalman Filtering Channel Estimation in Fast Fading **43**
4.1 Introduction 43
4.2 Second-Order Kalman Filtering Algorithm 44
4.3 Wiener Bound 48
4.4 Simulation Results 49
4.5 Conclusions 52

5 Joint Channel Estimation and Data Detection **57**
5.1 Introduction 58
5.2 Iterative receivers for OFDM systems in Fast fading 62
5.2.1 Iterative Channel Estimation 63

5.2.2 Symbol Detection 66
5.3 Simulation Results 67
5.3.1 BER Performance 68
5.3.2 Robustness Analysis 69
5.4 Conclusion 77

6 EXIT Chart Analysis **78**
6.1 Introduction 78
6.1.1 Input of Transfer Function 79
6.1.2 Output of Transfer Function 81
6.2 EXIT Chart for detection with channel estimation in Fast Fading channels 82
6.2.1 EXIT Chart of Decoder 82
6.2.2 EXIT Chart of Decision-Directed Channel Estimation with Detector 83
6.2.3 EXIT Chart of Trajectory 86
6.3 Simulation Results 86
6.4 Conclusion 90

7 Conclusions and Future Work **93**
7.1 Summary 93
7.2 Future Work 95

A Basic Time Domain Channel Impulse Response Description **96**

B Basis Expansion Model **99**
B.1 Complex Exponential Basis Expansion Model 100
B.2 Discrete Prolate Spheroidal Basis Expansion Model 102

B.3 Karhunen-Loève Basis Expansion Model 103

C Whittle-Wiggins-Robinson Algorithm 105

D Least Squares and Linear Minimum Mean Square Error Estimation 109

E Basic Kalman Filter Equations 111

Bibliography 113

List of Abbreviations

Acronym	Definition
1-D	one-dimensional
2-D	two-dimensional
APP	*a posteriori* probability
AR	autoregressive
AWGN	additive white Gaussian noise
BCH	Bose-Chaudhuri-Hocquenghem
BCJR	Bahl-Cocke-Jelinek-Raviv
BEM	basis expansion model
BER	bit error rate
BICM	bit-interleaved coded modulation
CDMA	code division multiple access
CE	channel estimation
CE-BEM	complex-exponential basis expansion model
CFO	carrier frequency offset
CIR	channel impulse response
CSI	channel state information
CP	cyclic prefix
CRB	Cramér-Rao Bound

DAB	digital audio broadcasting
DFE	decision-feedback equalizer
DFT	discrete Fourier transform
DPS	discrete prolate spheroidal
DSL	digital subscriber line
DVB	digital video broadcasting
EXIT	extrinsic information transfer
FFT	fast Fourier transform
HIPERLAN/2	high performance local area network type 2
ICI	intercarrier interference
LDPC	low-density parity-check
IDFT	inverse discrete Fourier transform
IFFT	inverse fast Fourier transform
ISI	intersymbol interference
KL	Karhunen-Loève
LMMSE	linear minimum mean square error
LS	least squares
LLR	log-likelihood ratio
LTE	long-term extension
MAP	maximum *a posteriori* probability
MC	multicarrier
MIMO	multiple-input multiple-output
ML	maximum-likelihood
MMSE	minimum mean square error
MSE	mean square error
OB	OFDM block

OFDM	orthogonal frequency division multiplexing
PAN	personal area network
PAPR	peak-to-average-power ratio
PB	pilot block
PDF	probability density function
PN	pseudo-noise
P/S	parallel-to-serial
PSAM	pilot symbol aided modulation
PSWF	prolate spheroidal wave function
QPSK	quadrature phase-shift keying
QAM	quadrature amplitude modulation
RS	Reed-Solomon
SC	single-carrier
SER	symbol error rate
SISO	soft input soft output
SNR	signal-to-noise ratio
S/P	serial-to-parallel
SVD	singular value decomposition
TB	transmission block
UWB	ultra-wideband
WLAN	wireless local area network
WMAN	wireless metropolitan area network
WSSUS	wide-sense stationary uncorrelated scattering
WWRA	Whittle-Wiggins-Robinson Algorithm
ZP	zero padding

List of Symbols

Acronym	Definition
$\underline{\boldsymbol{A}}$	multivariate AR coefficient matrix related to $\underline{\boldsymbol{x}}[n]$
$\boldsymbol{A}_l$	multivariate AR coefficient matrix related to $\boldsymbol{x}_l[n]$
$\underline{\boldsymbol{A}}_l$	multivariate AR coefficient matrix related to $\underline{\boldsymbol{x}}_l[n]$
c_k	the kth bit
$\boldsymbol{C}_k$	transition matrix between $\boldsymbol{y}_{tk}$ and $\boldsymbol{x}$
$\boldsymbol{C}^P$	transition matrix between $\boldsymbol{y}_t^P$ and $\boldsymbol{x}$
$\boldsymbol{d}$	pilot vector in the time domain
$\boldsymbol{d}_p$	the pth pilot block in the time domain
$\mathcal{D}(\boldsymbol{z})$	diagonal matrix with vector $\boldsymbol{z}$ on its diagonal
$E[x]$	expected value of the random variable x
$\boldsymbol{E}$	basis function matrix for the transmission block
$\boldsymbol{E}_k$	basis function matrix for the kth OFDM block
E_b/N_0	bit-to-noise ratio
$E_n(q)$	the qth basis function value at sample n
f	frequency (hertz)
f_c	carrier frequency
f_d	maximum Doppler frequency
$\boldsymbol{F}$	fast Fourier transform matrix

$\boldsymbol{F}^H$	inverse fast Fourier transform matrix
$\boldsymbol{G}[n]$	measurement matrix
$h(n,l)$	channel impulse response in the discrete time
$h(t,\tau)$	channel impulse response
$\boldsymbol{h}_l$	channel gain vector of the lth path
$\boldsymbol{h}_l^k$	channel gain vector of the lth path for the kth OFDM block
H	Hermitian transpose
$\boldsymbol{H}_k$	channel matrix in the frequency domain for the kth OFDM block
$\boldsymbol{H}_t^k$	channel matrix in the time domain for the kth OFDM block
$J_0(\cdot)$	the zeroth-order Bessel function of the first kind
K	order of multivariate AR process
K_b	number of data blocks
K_p	number of pilot blocks
L	channel order
L_A	*a priori* log-likelihood ratio
L_D	*a posteriori* log-likelihood ratio
L_E	extrinsic log-likelihood ratio
M	number of samples in a transmission block
$\boldsymbol{\mathcal{M}}_D$	transition matrix between $\boldsymbol{y}^D$ and $\boldsymbol{x}$
$\boldsymbol{\mathcal{M}}_k$	transition matrix between $\boldsymbol{y}_k$ and $\boldsymbol{x}$
N	number of subcarriers
N_{cp}	length of cyclic prefix
N_s	length of OFDM block
N_0	power spectrum density of AWGN
$\boldsymbol{P}[n]$	error covariance matrix of the state vector $\boldsymbol{x}[n]$
$P(x)$	probability of x

Q	BEM order
$\boldsymbol{Q}$	covariance matrix of noise
$\boldsymbol{Q}_{lf}$	covariance matrix of noise vector $\boldsymbol{w}_l$
$r_t(\Delta t)$	spaced-time autocorrelation function
$\boldsymbol{R}$	correlation matrix of the measurement noise
$\underline{\underline{\boldsymbol{R}}}$	covariance matrix of $\underline{\boldsymbol{x}}$
$\boldsymbol{R}_{\boldsymbol{x}}$	covariance matrix of $\boldsymbol{x}_l$
$\boldsymbol{R}_{\boldsymbol{h}}$	covariance matrix of $\boldsymbol{h}_l$
$\boldsymbol{s}_k$	transmitted signal vector in the frequency domain for the kth OFDM block
$S_H(f)$	Doppler power spectrum of the channel
$S_v(f)$	power spectrum of the noise
$\boldsymbol{S}_l$	circularly shifting identity matrix $\boldsymbol{I}_N$ down with l samples
t	time variable
T	OFDM symbol duration
T	transpose
T_{cp}	guard interval
T_s	sampling period
$\boldsymbol{u}$	transmitted signal vector in the time domain
$\boldsymbol{u}_k$	transmitted signal vector in the time domain for the kth OFDM block
$\boldsymbol{u}_p$	the pth data block in the time domain
$\tilde{\boldsymbol{u}}$	transmitted signal vector of the transmission block in the time domain
v	vehicle speed
$v_t(n)$	AWGN in the time domain
$\boldsymbol{v}_k$	AWGN vector in the frequency domain for the kth OFDM block
$\boldsymbol{v}[n]$	state noise vector
$\tilde{\boldsymbol{v}}_t$	AWGN noise vector in the transmission block

$\boldsymbol{v}_t^D$	AWGN vector at the data position in the time domain
$\boldsymbol{v}_t^k$	AWGN vector in the time domain for the kth OFDM block
$\boldsymbol{v}_t^P$	AWGN vector at the pilot position in the time domain
W	bandwidth
$\boldsymbol{w}[n]$	measurement vector
$\boldsymbol{x}$	basis coefficient vector
$x_l(q)$	the basis coefficient for path l corresponding to the qth basis function
$\boldsymbol{x}[n]$	state vector
$\boldsymbol{x}_l$	basis coefficient vector of the lth path
$\underline{\boldsymbol{x}}_l[n]$	containing $\boldsymbol{x}_l$ for blocks $n-K+1$ to n
$\underline{\boldsymbol{x}}[n]$	containing $\boldsymbol{x}$ for blocks $n-K+1$ to n
$y_t(n)$	received signal in the time domain at sample n
$\boldsymbol{y}^D$	received signal vector corresponding to data in the frequency domain
$\boldsymbol{y}[n]$	measurement vector at 'time' n
$\tilde{\boldsymbol{y}}_t$	received signal vector of the transmission block in the time domain
$\boldsymbol{y}_{tk}$	received signal vector of the kth OFDM block in the time domain
$\boldsymbol{y}_t^D$	received signal vector corresponding to data in the time domain
$\boldsymbol{y}_t^P$	received signal vector corresponding to pilots in the time domain
$\boldsymbol{y}_{tp}^d$	received signal vector corresponding to $\boldsymbol{u}_p$
$\boldsymbol{y}_{tp}^p$	received signal vector corresponding to $\boldsymbol{d}_p$
Δf	subcarrier spacing
$\boldsymbol{\Phi}$	state transition matrix
σ_v^2	variance of noise
τ	propagation delay
τ_{max}	multipath delay spread

List of Figures

Figure 2.1 A discrete-time baseband OFDM system 12
Figure 2.2 Structure of training and pilots. (a) Training symbols. (b) Pilot symbols. 16
Figure 2.3 Structure of input and output of the channel in the time domain. 20
Figure 2.4 Structure of input and output in the frequency domain 21

Figure 3.1 Block structure with $\kappa = 1$ 27
Figure 3.2 Structure of $\boldsymbol{H}_k$ matrix 30
Figure 3.3 Discrete-time structure in OFDM systems 30
Figure 3.4 Relationship between the transmitted data and the received signal in the time domain for single transmission block 32

Figure 4.1 Channel estimation MSE for $f_dT = 0.05$ 51
Figure 4.2 Channel estimation MSE for $f_dT = 0.1$ 52
Figure 4.3 Channel estimation BER for $f_dT = 0.05$ 53
Figure 4.4 BER resulting from estimated channel state for $f_dT = 0.1$ . . 54
Figure 4.5 Channel MSE for $SNR = 10$ dB and designed $f_dT = 0.05$ . . 55
Figure 4.6 Channel MSE for $f_dT = 0.05$ and designed $SNR = 15$ dB . . 56

Figure 5.1 Block diagram of the transmitter 59
Figure 5.2 Block diagram of the receiver 59
Figure 5.3 Coded OFDM transmitter 62

Figure 5.4 Coded OFDM system 62
Figure 5.5 MSE resulting from estimated CSI for $f_dT = 0.1$ with QPSK . 70
Figure 5.6 BER resulting from estimated CSI for $f_dT = 0.1$ with QPSK . 71
Figure 5.7 BER resulting from estimated CSI for $f_dT = 0.15$ with QPSK 72
Figure 5.8 BER resulting from estimated CSI for $f_dT = 0.1$ with 16-QAM 73
Figure 5.9 BER resulting from estimated CSI with $K_p = 10$ for $f_dT = 0.1$ with 64-QAM 74
Figure 5.10 BER resulting for designed $f_dT = 0.1$ with QPSK 75
Figure 5.11 BER resulting for $f_dT = 0.1$ and true $L = 3$ with QPSK . . . 76

Figure 6.1 Model for the EXIT chart analysis 79
Figure 6.2 Model for the EXIT chart of detector 83
Figure 6.3 EXIT chart 84
Figure 6.4 EXIT chart using $K_p = 10$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_o = 7dB$ 88
Figure 6.5 EXIT chart using $K_p = 5$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_0 = 7dB$ 89
Figure 6.6 EXIT chart using $K_p = 5$ and $K_t = 100$ for $f_d = 0.1$ at $E_b/N_0 = 7dB$ 90
Figure 6.7 EXIT chart using $K_p = 5$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_0 = 8dB$ 91
Figure 6.8 EXIT chart using $K_p = 10$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_o = 7dB$ 92

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, Dr. Michael McGuire, for his valuable guidance, continuous encouragement and insightful technical advice throughout my study. Dr. McGuire and my co-supervisor Dr. Dong have offered great assistance to help me navigate through the bumps in the road. This thesis could not have been completed without their support and help.

I would also like to thank Dr. Aaron Gulliver, Dr. Mihai Sima and Dr. Kui Wu for the valuable suggestions on revising my thesis.

Thanks to many of my colleagues and friends at University of Victoria for being so nice and helpful, which makes my stay a great pleasure. While I cannot write about them all individually, each of them has been important to me at various stages throughout my time at University of Victoria.

Special thanks to Steve, Kevin, Erik, Vicky, Moneca and Mary-Anne for their patience and constant help.

Lastly, I would like to thank my family, for their boundless love, understanding, and constant support in all that I do. This thesis would certainly not have existed without them.

DEDICATION

To my family

Chapter 1

Introduction

1.1 Background

The increasing demand for high data rate transmission over broadband radio channels has imposed significant challenges in wireless communications. High data rate transmission and high mobility of transmitters and/or receivers result in frequency-selective and time-selective, i.e., doubly selective, fading channels for future mobile broadband wireless systems. Mitigating such doubly selective fading effects is critical for efficient data transmission. Moreover, perfect channel state information (CSI) is not available at the receiver. Thus accurate estimate of the CSI has a major impact on the whole system performance [1]. This motivates an extensive channel estimation study of doubly selective fading channels for future wireless communication systems.

To extend the applications of orthogonal frequency division multiplexing (OFDM) in the future mobile broadband wireless communication systems such as mobile WiMAX and long-term extension (LTE), channel estimation techniques for OFDM systems in doubly selective channels are the topic in this thesis. For high data rate transmission systems, employing OFDM converts a wide bandwidth channel into sev-

eral narrow band subchannels (subcarriers). Due to its high bandwidth efficiency, its simple implementation and its robustness over frequency-selective channels, OFDM has been widely applied in wireless communication systems.

For conventional coherent detection, accurate CSI is needed for the receiver processing [2]. Although channel estimation can be avoided by using differential modulation techniques, these techniques will fail catastrophically in the fast fading channel, where the channel impulse response (CIR) varies significantly within the symbol duration [3]. In fact, differential modulation techniques assume that the channel is stationary over the period of two OFDM symbols which is not true for the fast fading channels of greatest interest in this thesis [4]. Thus channel estimation is an integral part of the receiver for fast fading channels. In such a case, the receiver needs to perform channel estimation for each OFDM symbol. Moreover, in fast fading channels, the impulse response of the channel for each propagation path changes from the beginning to the end of each OFDM symbol. The orthogonality among the subcarriers is destroyed and intercarrier interference (ICI) is created, which, if left uncompensated, can cause high bit error rates (BERs). Generally, the compensation for the ICI due to the fast fading channel is based on more complex equalizers such as minimum mean-square error (MMSE) equalizers [5, 6], which need not only the individual subcarrier frequency responses but also the interference among subcarriers in each OFDM symbol. Hence, channel estimation is more challenging for OFDM systems in fast fading channels than in slow fading systems.

1.2 History of Channel Estimation

Of different classes channel estimation techniques that have been developed in the literature, one class of techniques is based on pilot symbols which are known a priori

to the receiver. In this case, a standard approach is pilot symbol aided modulation (PSAM) [7] or so-called pilot assisted transmission [8], where known pilot signals multiplexed with information symbols are sent through the channel at regular intervals so the radio receiver can make direct measurements of the channel variations created by the propagation environment and terminal mobility. The CSI corresponding to pilot symbols is estimated first and the CSI corresponding to data symbols is then obtained by interpolation. In OFDM systems, the existing standards add pilot symbols in the frequency domain.

Another class of channel estimation techniques is blind channel estimation which depends only on the received symbols without inserting pilots [9, 10, 11]. In this case, the CSI is obtained from the received signal by using high-order statistical methods which require a large amount of data and high computational complexity [2].

1.3 Technical Challenges of Channel Estimation

As the case of the fast fading channel, the channel gain at each time sample of one OFDM symbol block is needed for data detection, and as a consequence, the number of unknown channel parameters is larger than the number of measurements if standard channel modelling is employed. In such a case, an underdetermined system occurs as the number of parameters exceeds the number of available measurements created by the pilot symbols [12]. To reduce the number of parameters to be estimated, an alternative approach is modelling the channel using a basis expansion model (BEM) over a time period, called a transmission block, where the channel state is expressed as a superposition of known basis functions weighted by unknown basis coefficients. Thus, the channel estimation problem is converted to the estimation of a limited number of basis coefficients.

Due to Gibbs phenomenon and the bandlimited nature of the fading process, the previous BEM-based channel modelling was not accurate when a short OFDM period was used as the BEM period. To solve this problem, we choose several OFDM blocks as the BEM period or transmission block length. With the use of a short OFDM symbol block, a low cost data detection is available. Therefore, higher order modulation can be used for fast fading channels without an exponential increase in data detection cost.

Due to a lack of modelling of the time evolution of the fading channel, the previous BEM-based channel estimation techniques did not exploit the information from OFDM blocks other than one currently being processed, and hence, such estimation methods did not achieve optimal mean square error (MSE) performance. To model the time evolution over a transmission block consisting of multiple OFDM symbols, a multivariate autoregressive (AR) Gauss-Markov model is exploited to characterize the dependency of the channel state between OFDM blocks for fast fading channels for the purposes of improving the channel estimation ability.

To simplify the channel estimation algorithm for coded OFDM systems, time blocks of three different durations are used at the receiver proposed in this thesis: the OFDM block is used as the basis for detection; the transmission block is used for channel estimation; and the interleaving block, which contains several transmission blocks, is used for decoding. To develop an iterative scheme at the receiver, we derive different measurement models, which include measurement model for data in terms of channel for detection, measurement model for channel in terms of data symbols for channel estimation, and measurement model for channel in terms of pilot symbols for channel estimation.

In the case of pilot-aided channel estimation, the time evolution of basis coefficients is described as a second-order multivariate AR model, which can character-

ize the channel more accurately than a first-order AR model [13]. Subsequently, a Kalman filter is developed to track the basis coefficients from the pilot measurements. A Kalman smoother, which incorporates the following measurements, gives a more refined estimate than the previously presented channel estimation methods. It will be shown that the MSE of the channel estimation is not significantly reduced when a higher order (≥ 3) Kalman filter is used. To reduce the complexity of channel estimation in fast fading channels, new pilot aided channel estimation in the presence of ICI with low complexity is proposed in this thesis.

As the fading rate increases, the severity of the ICI also increases, and thus, more channel state values in each OFDM symbol are needed. Under this condition, more pilot symbols are required to obtain sufficiently accurate channel estimates for reliable data detection. Even with the use of BEMs, the number of channel parameters required to model the channel increases which, in turn, increases the needed pilot to data ratio to achieve acceptable system performance. This reduces the overall data rate during fast fading to unacceptable levels for many applications. The combination of high fading rate and low pilot to data ratio necessitates the channel estimator in OFDM systems and a good tradeoff is needed between the performance of the channel estimation and the pilot to data ratio.

To further improve the performance of channel estimation with low pilot to data ratio, a novel low-complexity iterative joint channel estimation and symbol detection procedure is proposed. The time evolution of the basis coefficients between each transmission block is modelled as a first-order multivariate AR process. An initial channel estimate is obtained based on pilot-aided channel estimation and used for an initial data detection. The detected data is then used to re-estimate the channel by a decision-directed method and the data is detected again. The detected data is used as 'virtual' pilot signal which is dense in the time domain. The error of the channel

estimation is reduced with more 'virtual' pilots being available. As a result, the BER performance of the proposed joint channel estimation data detection technique is much better than that of the conventional PSAM approach. It will be shown that the BER performance is close to that when ideal CSI is available.

1.4 Contributions

The main contributions of the thesis for OFDM system modelling :

- The introduction of a novel multivariate AR Gauss-Markov model for basis coefficients over transmission blocks to model the time evolution of the fast fading channel.
- The introduction of time blocks of three different durations for channel processing in coded OFDM systems to both reduce channel estimation and detection computation cost and error levels.
- The introduction of different measurement models for an iterative receiver in fast fading channels.

The main contributions of the thesis to the subject of channel estimation based on pilots are :

- The introduction of a novel Kalman smoothing algorithm based on a second-order dynamic model with a MSE near that of the optimal Wiener filter and a computational cost on the same order as the previously proposed algorithms for channel estimation and improves the BER performance of wireless communications systems.
- The demonstration of the robustness of the new technique when the Doppler

frequency or signal-to-noise ratio (SNR) is varied from the value used to design the channel estimation filter.

The main contributions of the thesis in the areas of joint channel estimation and data detection are :

- The introduction of a novel joint channel estimation/symbol detection algorithm for fast fading channels requiring only a small number of pilots and low pilot power to achieve performance close to when the CSI is known perfectly.
- The demonstration that this new channel estimation symbol detection technique is robust to variations of the radio channel from the design values.
- The demonstration that the new approach is efficient when applied to different modulation and coding types.
- The demonstration of how the convergence behavior of the new algorithm is analyzed by the extrinsic information transfer (EXIT) chart technique.
- The introduction of the EXIT chart is used to select the modulation, pilot density, and error correction code for good system performance for a given power level.

1.5 Thesis Outline

The structure of the thesis is described as follows. In Chapter 2, we present a survey of the existing techniques for estimating the CSI with pilot symbols. In Chapter 3, first, we introduce three time blocks at the receiver; then, we derive different measurement models for the proposed iterative receiver; finally, we develop a multivariate AR model to characterize the time domain evolution of the channel. Assuming access

to pilot symbols, we propose a second-order Kalman filter algorithm using BEMs of the channel gains to accurately estimate fast fading channels in Chapter 4. A joint channel estimation and data detection approach is discussed in Chapter 5, which reduces the required number of pilot symbols for acceptable error performance in fast fading channels to the levels needed by previous channel estimation algorithms with slow fading channels. Using EXIT chart to investigate how the proposed iterative receiver works well is discussed in Chapter 6. Conclusions and future work will be present in Chapter 7.

Chapter 2

Literature Review of Channel Estimation for OFDM Systems

This chapter presents a literature survey of channel estimation methods for orthogonal frequency division multiplexing (OFDM) system in fast fading channels. As described in Chapter 1, we will focus on pilot-aided channel estimation schemes. In general, two aspects affect the performance of the pilot-aided channel estimation techniques: the design of pilot symbols and the estimation scheme. We begin with the description of OFDM system. Different pilot-aided channel estimation schemes are treated in Section 2.2.

2.1 OFDM

The use of OFDM techniques can be traced back to the late 1950's and early 1960's for military high frequency radio systems. With the availability of simple and cheap implementations of the discrete Fourier transform (DFT) and the inverse DFT (IDFT), the DFT can generate data signal in parallel form and OFDM became popular [14]. In particular, coded OFDM has been adopted by standards and major manufacturers

for a wide range of applications. Now, it has been used in digital audio broadcasting (DAB) and digital video broadcasting (DVB) systems in Europe [15], digital subscriber line (DSL) standards, and wireless local area network (WLAN) standards such as IEEE Std.802.11a/g (WiFi) [16] or high performance local area network type 2 (HIPERLAN/2) in Europe [17], and wireless metropolitan area network (WMAN) standards such as IEEE Std.802.16 (WiMAX) [18], also in ultra-wideband (UWB) personal area network (PAN) (IEEE 802.15.3a). OFDM is also considered in IEEE Std.802.11n in multiple-input multiple-output (MIMO) systems. For future mobile broadband wireless communications, mobile WiMAX and long-term extension (LTE) use OFDM-based modulation too.

A well-known approach to efficiently utilize available channel bandwidth is the multicarrier (MC) transmission scheme first proposed by Chang [19]. OFDM is a type of MC modulation [20], which converts a wide bandwidth W into N narrowband subchannels (subcarriers) over which data is transmitted in parallel [21, 22]. To obtain high spectral efficiency, these subchannels are set to be overlapping and orthogonal under ideal propagation conditions. The subcarrier spacing, denoted as $\Delta f = W/N$, provides the minimum frequency separation required to maintain orthogonality between subcarriers. Consequently, the following equation should be satisfied over the OFDM symbol duration T, i.e., $\int_0^T e^{j\frac{2\pi}{T}(m-n)t}\,\mathrm{d}t = 0$, for different subcarriers m and n. In other words, the OFDM symbol duration is $T = NT_s = 1/\Delta f$, where T_s is the sampling period.

One problem created by multipath propagation is inter-symbol interference (ISI). To manage the ISI, OFDM symbols use either a cyclic prefix (CP) or a zero padding (ZP) guard interval between OFDM blocks which is longer than the delay spread of the channel which is the difference between the maximum and minimum propagation delay of the radio channel. A CP is the repetition of the last N_{cp} samples of the

transmitted signal which are inserted at the beginning of each OFDM symbol. With the used of CP, the transmitted signal is cyclically extended in the guard interval $T_{cp} = N_{cp}T_s$. At the receiver, after removing the first N_{cp} signals corrupted by ISI, the ISI is completely discarded. However, due to adding the CP at the transmitter, the proportional loss of the useful transmission energy is $\frac{T_{cp}}{T+T_{cp}}$. Instead of using the CP, the ZP with N_{cp} zeros can be inserted at the end of each OFDM symbol without energy loss. Unfortunately, this method increases the receiver complexity compared to the traditional CP-based OFDM (CP-OFDM) [23]. Therefore, in this and the following chapters we consider CP-OFDM systems.

The major advantage of OFDM lies in processing frequency-selective channels as multiple flat-fading sub-channels. If the channel is time invariant over the period of an OFDM symbol block, a condition known as slow fading, the orthogonality property is maintained between the subcarriers. In such a case, channel estimation or data detection is simple since each subcarrier is equalized with a single-tap equalizer. However, when the channel is time-varying over one OFDM symbol period, the orthogonality among subcarriers is destroyed, resulting in ICI, which degrades the bit error rate (BER) performance compared to the slow fading channels if it is not properly compensated for. The ICI may occur due to the presence of the fast fading channel or the presence of a carrier frequency offset (CFO) between the transmitter and receiver caused by imperfect synchronization. CFO can be estimated by using a maximum likelihood (ML) estimation algorithm [24, 25, 26]. The potential performance degradation of OFDM caused by fading channels is a function of the fading rate, with faster fading channels requiring more significant mitigation methods to achieve the same error performance as slow fading channels. Furthermore, in the presence of ICI due to fast fading, the channel estimation is more challenging since both the individual subcarrier and the interference created by each subcarrier to its

neighbouring subcarriers need to be estimated. Therefore, more sophisticated detection procedures or channel estimation schemes are required [27, 4, 28]. To further advance the applications of OFDM in future standards, robust channel estimation schemes for fast fading channels are discussed in this thesis.

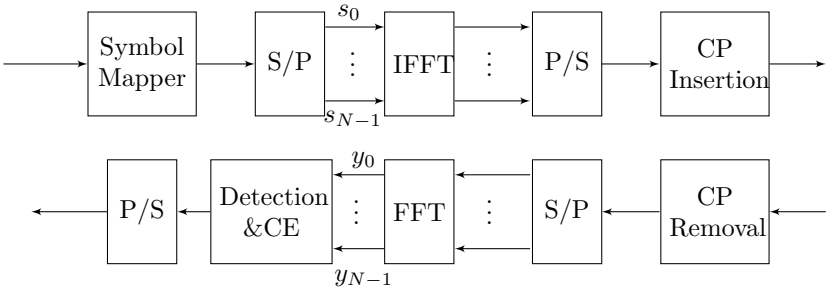

Figure 2.1: A discrete-time baseband OFDM system

The block diagram of discrete-time baseband OFDM system is depicted in Figure 2.1. In discrete-time OFDM systems, the input data is first modulated and then passed through a serial-to-parallel (S/P) converter, whose output in the frequency domain is expressed in vector form as $\boldsymbol{s} = [s_0, ..., s_{N-1}]^T$, where N is the number of subcarriers. Using the inverse fast Fourier transform (IFFT) algorithm, a discrete-time baseband OFDM symbol is converted into time-domain samples as

$$u(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k e^{j2\pi \frac{nk}{N}}, \quad 0 \leq n \leq N-1 \tag{2.1}$$

After inserting CP, the resulting transmitted signal becomes $\mathbf{u} = \{u(N-N_{cp}), ..., u(N-2), u(N-1), u(0), u(1), ..., u(N-1)\}$. This signal is then serially transmitted through a multipath radio propagation channel which is subject to additive white Gaussian noise (AWGN) $v(t)$ with variance $\sigma_v^2 = N_0/2$, where N_0 is power spectral density.

At the receiver, after removing the CP and processing by a parallel-to-serial (P/S) converter, the received signal in the time domain is then converted back to the frequency domain, which is implemented by using the fast Fourier transform (FFT)

algorithm. Due to the use of CP, the interblock interference between contiguous OFDM blocks in the frequency domain is eliminated so each OFDM block can be processed independently, provided that the length of the CP is equal to or larger than the delay spread of the channel.

The main drawbacks of OFDM systems are high peak-to-average-power ratio (PAPR) and high sensitivity to CFO. Moreover, OFDM does not obtain frequency diversity. If a deep fade occurs close to the frequency of a subcarrier, reliable data detection carried by these faded subcarriers becomes difficult [29, 30]. To solve this problem, an alternative method is to employ error-control codes. Therefore, the diversity loss can be circumvented by incorporating error-control coding in conjunction with interleaving. The typical codes are block codes (e.g., Reed-Solomon (RS) or Bose-Chaudhuri-Hocquenghem (BCH)), convolutional codes, trellis codes, turbo codes, and low-density parity-check (LDPC) codes. Another one is combining multi-carrier coding and code division multiple access (CDMA) techniques, where MC-CDMA splits a wide band signal into narrowband signals and also exploits multipath diversity [31]. The new channel estimation techniques which will be discussed later in this thesis can be extended to MC-CDMA.

2.2 Channel Estimation Techniques

When the channel is unknown *a priori* to the receiver, pilot symbol aided modulation (PSAM), where known pilot signals are periodically sent during the transmission, can simplify the channel estimation. In general, the performance of channel estimation depends on the number, the location, and the power of pilot symbols inserted into OFDM blocks. Consider a fading multipath channel with the multipath delay spread τ_{max} and the maximum Doppler frequency f_d. To recover the channel state informa-

tion (CSI), the spaces between pilot symbols in the time and frequency domain must satisfy two-dimensional (2-D) sampling theorem, that is,

$$f_d T d_t \leq 1/2 \tag{2.2}$$

and

$$\tau_{max} \Delta f d_f \leq 1 \tag{2.3}$$

where T is the OFDM block duration, Δf is the subcarrier spacing; d_t and d_f are the numbers of samples between pilot symbols in the time domain and frequency domain, respectively [32]. Within the OFDM symbol duration, the number of pilot symbols in the frequency domain is related to the delay spread; on the other hand, the number of pilot symbols in the time domain is related to the normalized Doppler frequency $f_d T$. Based on 2-D arrangement of pilot symbols, 2-D channel estimators are too complex in practice [33]. Therefore, channel estimation is exploited in one-dimension (1-D) for OFDM systems in general.

In practice, for OFDM systems, the channel estimation techniques can be performed by using either frequency domain samples or time domain samples. Thus, pilot symbols can be added either in the frequency domain where the pilot symbols are frequency multiplexed with the data, or in the time domain where the pilot symbols are time multiplexed with the data. Since single-carrier (SC) modulation, in which data symbols are transmitted in serial fashion, is a form of generalized MC transmission [34, 35], methods for SC channel estimation will also be examined for their applicability to OFDM channel estimation. In the following subsections, we discuss different channel estimation techniques employing pilot schemes in slow and fast fading channels, respectively.

2.2.1 Channel Estimation in Slow Fading Channels

When the channel slowly changes over a number of OFDM symbol blocks, channel estimation can be based on pilot symbols, which are inserted into all subcarriers of OFDM symbol blocks within a specific period, as shown in Figure 2.2 (a), where such pilot symbols are usually called training symbols. Then a batch of OFDM symbols follows the training symbols. For the channel estimation based on training symbols, the CSI corresponding to training symbols are first estimated, the CSI corresponding to the subsequent data symbols can be tracked and further improved by decision directed channel estimation [28]. If the channel varies slowly over OFDM blocks, the estimated CSI based on previous training symbols are generally reliable so such estimated channel state may be used in data detection. As the channel varies fast over time, the training symbols must be sent more frequently to get reliable channel estimates, and hence, the overall system efficiency is reduced [4]. In such a case, channel estimation can be based on pilot symbols, which are periodically inserted into different subcarriers for each OFDM symbol block, as shown in Figure 2.2 (b), and will be discussed later.

In addition, to improve bandwidth efficiency, the superimposed pilot scheme was proposed for flat-selective fading channels [36], where a pseudo-noise (PN) sequence is synchronously added to the information symbols at the transmitter prior to modulation transmission. Based on the first-order statistics method, this superimposed pilot approach for channel estimation was discussed for frequency selective fading channels [37, 38, 39], and for doubly selective fading channels [40, 41]. Although the use of superimposed pilots can improve the bandwidth efficiency, the performance based on superimposed pilot aided channel estimation is worse than that of traditional PSAM [41, 42]. Therefore, we focus our attention on channel estimation based on traditional PSAM in this thesis.

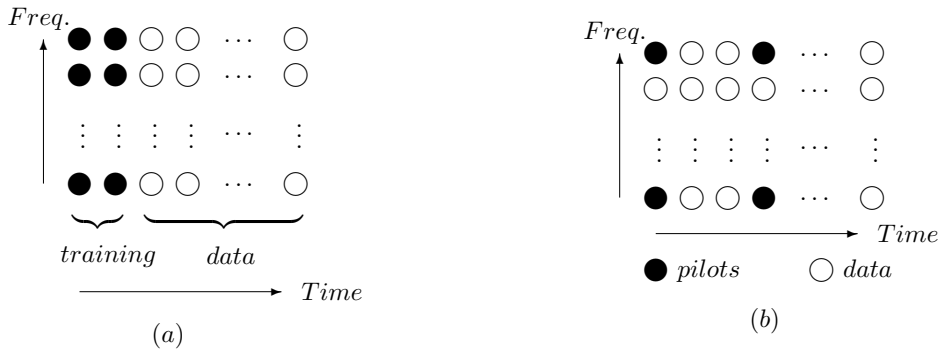

Figure 2.2: Structure of training and pilots. (a) Training symbols. (b) Pilot symbols.

As in the case of the channel varying from one OFDM block to another, pilot symbols are often inserted in every OFDM symbol in general. The pilot spacing in the frequency domain is the main criteria for the pilot placement. When the channel is only frequency selective, the channel gain is different for different frequency components [43]. In order to capture the variation of the channel in the frequency domain, the pilot spacing in the frequency domain should satisfy the sampling criteria (2.3). For slow fading channels, different pilot designs have been discussed in [44, 45, 46, 47, 48, 49, 50], where optimally the pilots are equally spaced in the frequency domain. Consider a fading multipath channel with $L+1$ paths, where L is called as channel order and $\tau_{max} = LT_s$. To minimize the mean-square error (MSE) of the channel estimation, the $L+1$ pilots are equally spaced in each OFDM symbol in the frequency domain in order to estimate the channel condition. If $N/(L+1)$ is not an integer, the design of optimal pilot symbols is discussed in [48], where the $L+1$ pilots are uniformly distributed across subcarriers to minimize symbol error rate (SER). The existing standards such as WiMAX choose frequency multiplexed pilots scheme since these schemes operate only over slow-fading channels. However, more recent innovations allow OFDM to operate over fast-fading channels if sufficiently accurate channel state information is available. Therefore, the majority of this thesis is to dedicate channel estimation in OFDM for fast fading channels.

From the description given above, we now turn our attention to the channel estimation. Several types of estimation techniques are exploited by using least squares (LS), ML, minimum mean square error (MMSE) or linear minimum mean square error (LMMSE) methods. Under the condition that the channel impulse response (CIR) is deterministic and the channel statistics is unknown at the receiver, if the additive noise is unknown at the receiver, LS approach can be used for pilot-aided channel estimation; if the noise is AWGN, ML channel estimation will be optimal and achieves the Cramér-Rao Bound (CRB), and furthermore, LS estimation is equivalent to ML estimation [51]. Whereas in the case that the CIR is random and channel statistics and signal-to-noise ratio (SNR) are known *a priori*, MMSE or LMMSE channel estimation is exploited to find minimize the MSE of the channel for better performance with higher computational complexity compared with LS or ML method. Moreover, if the measurement model is linear without Gaussian assumption, the LMMSE channel estimation can be used. On the other hand, if the measurement model is linear with Gaussian assumption, the LMMSE technique is same as the MMSE method. In practice, LS method is used to get initial channel estimates at the pilot symbols, and further improvement is based on MMSE or LMMSE method [4].

For slow fading channels as discussed above, depending on the pilot arrangement, estimation schemes are performed in the frequency domain using either training symbols or pilot symbols. When the channel is estimated based on training symbols or detected symbols, the CSI may be estimated by using LS algorithm or LMMSE algorithm [52, 53, 54]. When the channel is estimated based on pilot symbols which are usually inserted in the frequency domain, the CSI of the frequency domain at the known pilot subcarriers is first estimated by using LS algorithm [55] or ML algorithm [56], then interpolation is performed between these channel estimates to get the CSI at the data subcarriers via different methods such as the piecewise constant and linear

interpolation methods [22, 57] or high-order polynomial interpolation [58, 59]. When the information of SNR and channel statistics is obtained, MMSE(LMMSE) algorithm is exploited for channel estimation with pilot symbols [60, 61]. Compared with the LS algorithm, since the LMMSE algorithm exploiting channel statistics, LMMSE estimation is more accurate than LS estimation; however, LMMSE estimation has higher computational complexity.

Since data detection ignores ICI for slow fading channels, simple channel estimation techniques are required for the receiver. However, in order to compensate for the ICI in fast fading channels, more sophisticated channel estimation algorithms are required for data detection techniques, which is discussed below.

2.2.2 Channel Estimation in Fast Fading Channels

When the channel varies significantly over one OFDM symbol block, the orthogonality among the OFDM subcarriers is lost and the ICI is created. The severity of ICI depends on the normalized Doppler frequency f_dT [5]. In the presence of ICI, the amount of channel states that need to be estimated for reliable data detection increases. Not only the individual subcarrier frequency responses but also the interference among subcarriers in each OFDM symbol are to be estimated. In such a case, an underdetermined system occurs if standard channel modelling is employed since the number of unknowns are more than the number of measurements (pilot symbols). In order to reduce the number of unknown channel parameters, simplification approaches are exploited for channel estimation. One approach is that the channel is approximated to a piece-wise linear model over one or two adjacent OFDM blocks [62, 63]. However, this modelling approach degrades the performance of the channel estimation at high normalized Doppler frequencies such as 10% [64]. Another approach is to model the channel by a basis expansion model (BEM), where the sam-

ples of the channel state are characterized as a linear combination of a finite number of known basis functions weighted by unknown basis coefficients, which is discussed in detail in Appendix B, as well as Appendix A for channel impulse response description.

Investigations into channel estimation for SC and OFDM modulation in [1, 65, 66], have shown that pilot symbols placed in the time domain maximize the lower bound on channel capacity and minimize the channel estimation mean square error for multi-carrier communication systems such as OFDM. In order to mitigate interference between the pilot block and data block, which is introduced due to multipath propagation of the channel, each pilot block has the same form as $[\mathbf{0}_{1\times L}\ 1\ \mathbf{0}_{1\times L}]$, where the size of $\mathbf{0}_{1\times L}$ is 1 by L. In the case of the fixed data power and pilot power, based on lower bound on the average channel capacity with the LMMSE channel estimation, the optimal pilot symbols are equi-powered and equi-spaced in the time domain [1]. In addition, in order to minimize the MSE of the channel estimator, the optimal number of pilot (data) blocks should be equal to the number of basis coefficients for a given transmission block [1]. From a performance viewpoint, to minimize the total MSE of the estimator which includes the BEM modelling error, the number of pilot (data) blocks may be larger than the number of basis coefficients [65].

Similar to the time-domain pilot scheme, the pilot block in the form of $[\mathbf{0}_{1\times Q}\ 1\ \mathbf{0}_{1\times Q}]$ is equally placed between data blocks in the frequency domain, where Q is the number of basis coefficients, defined as $Q \geq 2\lceil f_d N T_s \rceil$ and $\lceil\ \rceil$ denotes the integer ceiling [67, 68]. The affect of ICI caused by each subcarrier is spread into its neighbouring subcarriers, where the length of ICI caused by each subcarrier is equal to $2Q$. In this case, the pilot symbols are equi-powered and equi-spaced in the frequency domain, and the optimal number of pilot blocks is equal to $L+1$.

Figure 2.3 (a) illustrates the transmitted signal in the time domain, where $\boldsymbol{u}_p$ and $\boldsymbol{d}_p$ are data block and pilot block for $p \in [0, K_p - 1]$, respectively, and K_p is the

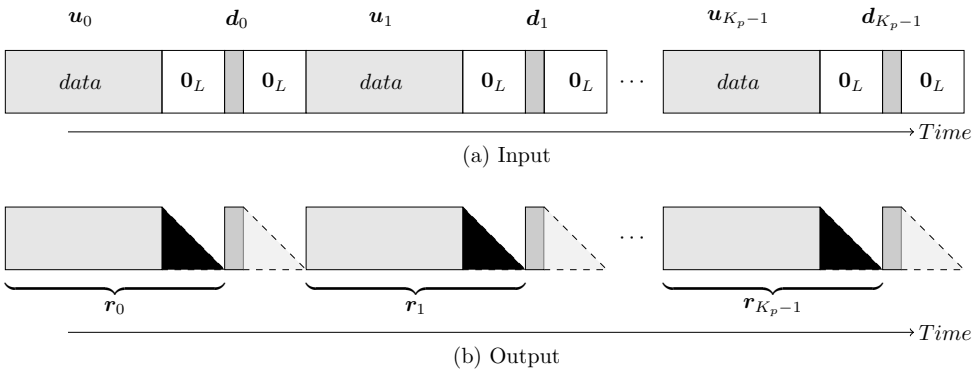

Figure 2.3: Structure of input and output of the channel in the time domain.

number of pilot blocks. Figure 2.3 (b) illustrates the received signal in the the time domain without noise. Due to the multipath spread, the output vector of the channel, $\boldsymbol{r}_p$, corresponding to data block $\boldsymbol{u}_p$ can be spilt into two parts. The rectangular part represents the desired output of the channel. The black triangular part is the interference overlapping into the adjacent pilot block from the preceding data block, where the length of the interference block is equal to L. Similarly, the output vector of the channel corresponding to the pilot block has two parts. From Figure 2.3, we can see that, as long as the pilot block has the form as described above, there is no inter-block interference between pilot block and data block, and hence linear channel estimation techniques can be used in the time domain. Therefore, time multiplexed pilots scheme can temporally separate pilot blocks from data blocks, and also be used for the purpose of timing synchronization.

Figure 2.4 (a) illustrates the transmitted signal in the frequency domain, where $\boldsymbol{s}_m$ and $\boldsymbol{b}_m$ are data block and pilot block for $m \in [0, \mathcal{M}-1]$, respectively, and $\mathcal{M}$ is the number of pilot blocks. Figure 2.4 (b) illustrates the received signal in the frequency domain without noise. Due to the Doppler spread, the output vector of the channel, $\boldsymbol{r}_m^p$, corresponding to pilot block $\boldsymbol{b}_m^p$ can be spilt into three parts, the rectangular part and the two triangular parts. As shown in Figure 2.4 (b), there

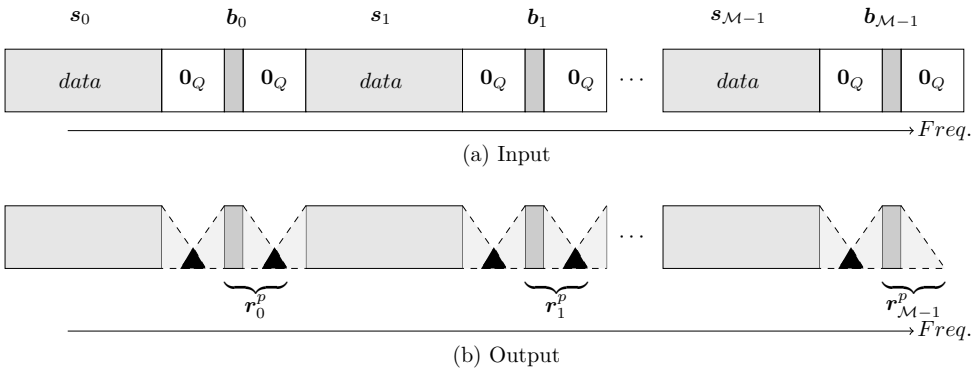

Figure 2.4: Structure of input and output in the frequency domain

is the inter-block interference (black portions) between $\boldsymbol{r}_m^p$ and the adjacent output corresponding to data block [67]. In this case, time multiplexed pilots scheme has more advantage than the frequency multiplexed pilots scheme for channel estimation. Therefore, in the following chapters, we focus our attention on channel estimation based on time multiplexed pilots method.

From the use of a BEM, the channel gains for a finite duration block are represented by vectors of basis coefficients. The advantage of the use of basis coefficients as the state vector is that the length of the basis coefficient vectors for a block is less than the length of the channel gain vectors so the number of unknowns to characterize the channel is reduced. Based on PSAM, basis coefficients can be estimated by using LS method [69, 70, 71], MMSE method [72, 73] or LMMSE method [1, 74]

Since single block BEM channel estimation techniques described above do not use information contained within blocks other than the current block, the previous schemes do not achieve optimal MSE performance. By taking the advantage of autoregressive (AR) model that can capture the channel dynamics [75], a new dynamic model is derived to describe the time evolution of the basis coefficients as a multivariate AR process in Chapter 3. Relying on a second-order multivariate AR model, a Kalman smoother is developed to track basis coefficients from the pilot measurements

in Chapter 4.

As the fading rate increases, more pilot symbols are required to obtain sufficiently accurate channel estimates for reliable detection. This reduces the overall data rate during fast fading to unacceptable levels for many applications. To improve the performance in communication systems, according to turbo principle which was originally proposed for channel decoding [76, 77], the iterative processing has been received considerable attention for slow fading channels, and furthermore it is being taken into account for fast fading channels. We will discuss this technique in detail below.

2.2.3 Joint Channel Estimation Data Detection

By using the turbo processing principle, the performance advantage of the turbo equalizer has motivated a significant amount of research on methods of the iterative receiver for slow fading channels, where the CSI is unknown a priori to the receiver. To improve the performance for frequency-selective fading channels, an alternative method is to use iterative channel estimation and symbol detection in the presence of ISI. Channel estimation can be performed with detection simultaneously based on blind methods [78, 79, 80] or based on pilot symbols [81]. For the trellis-based equalizer, those methods have high computational complexities. To reduce the complexity of equalizer, channel estimation is separated from data detection in general, where PSAM is used to obtain an initial channel estimate.

The early work on the design of a reduced complexity iterative receiver focused on methods that add pilot symbols in both the transmitter and the receiver, the detected data and pilot symbols at the receiver are then used to re-estimate the channel and the data is detected again in SC systems for flat fading channels [82, 83, 84], for frequency-selective fading channels [85], for time-varying frequency-selective fading channels [86], for block-fading channels [87], and for MC-CDMA in [88].

To simplify the receiver structure in slow fading channels discussed above, only detected data values are then used to re-estimate the channel in decision directed estimation approach and the data is detected again in practice. With good initial channel estimates, joint channel estimation and symbol detection method has been shown to perform well with OFDM in slow fading channels, where the channel state is time invariant over an OFDM symbol period [89, 90, 91, 92]. In order to reduce error propagation due to decision directed channel estimation, a threshold is exploited for the detected data, and hence, only accurate detected data is used for channel estimation in [90, 91]. For doubly selective channels, this technique has been proposed for SC in [93, 94, 95, 96] with BER performance worse than the case of the ideal CSI.

Using pilots inserted in the frequency domain for OFDM systems, joint channel estimation and data detection method was proposed in [68] with high complexity of detection and long symbol duration, or [71] with high pilot to data ratio and performance degradation. Within a single OFDM symbol period, LS or MMSE algorithm was discussed for channel estimation in [71, 68]. In order to detect data efficiently in the presence of interference between the sub-carriers, more costly detection algorithms such as maximum *a posteriori* probability (MAP) are employed, which is implemented using Bahl-Cocke-Jelinek-Raviv (BCJR) technique [97]. By choosing the transmission block of the CE-BEM exactly as the OFDM symbol period in [68], this CE-BEM has high modelling error due to Gibb's phenomenon [98]. Since the sampling signal over the transmission block is equivalent to a signal OFDM block which is truncated by a rectangular window, spectral leakage exists [98]. To increase spectral resolution, the pre-existing iterative channel estimation/joint detection methods extend the OFDM symbol period to a longer period.

However, the main cost of detection is mitigating ICI, which increases with the normalized Doppler frequency. For a given Doppler frequency, the ICI is related to

OFDM symbol duration. Although short OFDM symbol blocks have little ICI, the channel estimation does not work well because of the truncation problem. On the other hand, long OFDM symbol blocks require expensive detection as the channel variation over the time period of the OFDM block causes the compensation for ICI to become more difficult to handle.

Furthermore, the cost of detection algorithms increases exponentially with the order of modulation, making the previous algorithms unsuitable for OFDM systems using higher order modulation in fast fading channels, limiting the practical use of the previous iterative channel estimation/joint detection algorithms in fast fading to lower order modulation such as quaternary phase-shift keying (QPSK). Additionally, these methods also require a high pilot to data ratio for effective operation based on a single OFDM block.

To overcome the drawbacks described above, it is necessary to devise the channel estimation based on a number of OFDM symbol blocks. This motivates a new low-complexity joint channel estimation/symbol detection scheme for OFDM subject to fast fading with a low pilot-to-data ratio, where the BEM channel coefficients are found for transmission block consisting of multiple OFDM symbols. Moreover, the new method can be used for coded OFDM systems which are widely applied in standards. The detailed discussion will be described in Chapter 5.

Chapter 3

System Model

In this chapter, we focus our attention on system model for fast fading channels. Unlike the previous time blocks for channel processing, three time blocks are discussed in this chapter, which are OFDM block, transmission block and interleaving block. To simplify the channel estimation, we derive different measurement models, which include measurement model for data in terms of channel for detection, measurement model for channel in terms of data symbols for channel estimation, and measurement model for channel in terms of pilot symbols for channel estimation. To model the time evolution over the transmission block, the unknown parameters of the channel are characterized as a multivariate autoregressive (AR) Gauss-Markov model.

3.1 Time Blocks for Channel Processing

The proposed iterative scheme includes three parts, which are data detection, channel estimation and decoding. This motivates the use of three durations of time blocks at the receiver, as shown in Figure 3.1: the OFDM block (OB), which contains N_{cp} cyclic prefix (CP) and N subcarriers, is used as the basis for detection; the transmission block (TB), which includes K_b OFDM blocks and K_p pilot blocks (PBs), is used

for channel estimation; and the interleaving block, which contains K_t transmission blocks, is used for decoding.

For fast fading channels, the channel is modelled as the basis expansion model (BEM), where the channel gains are characterized as a linear combination of known basis functions weighted by unknown basis coefficients. Due to Gibbs phenomenon, when a short duration OFDM period is used as the BEM period, the channel will not be modelled accurately. To solve this problem, we propose to make the BEM period consisting of multiple OFDM symbols. In such a case, a short OFDM symbol block can still be used with a longer duration BEM period and hence, a low cost data detection method is required. As a consequence, higher order modulation can used for fast fading channels without an exponential increase in data detection cost.

The computational complexity of the channel estimator for a transmission block of M samples is related to the number of the BEM coefficients, $Q + 1$, which is bounded by $Q \geq 2\lceil f_d M T_s \rceil$, where f_d is the maximum Doppler frequency, T_s is the sampling interval that equals to the data symbol period [1]. Therefore, for a fixed Doppler frequency, the longer the transmission block, the larger the number of the BEM coefficients and hence the larger the cost of the channel estimator. Because of truncation effects, the shorter the transmission block, the more spectral leakage will be seen as the bandlimited fading process is imperfectly modelled for the short time duration of the transmission block creating significant modelling errors. Therefore, we propose a suitable size for transmission block in channel estimation. In order to deal with bursts of errors due to temporary deep fading, a large interleaver is used to enable the convolutional code to correct spread data bits over both times when the channel is in a good condition and in a bad condition [99].

In the following sections, we describe these three types of blocks, their applications, and how fast fading affects the different time blocks for processing.

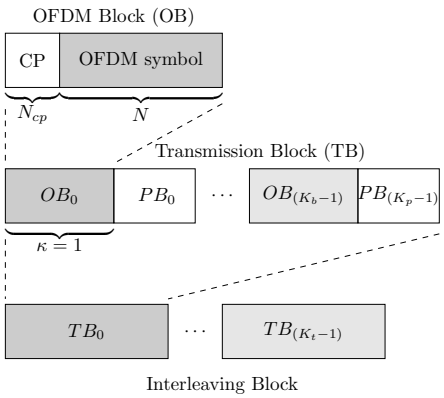

Figure 3.1: Block structure with $\kappa = 1$

3.1.1 Time Domain and Frequency Domain Model for a Single OFDM Block

The fading channel in wireless communications is often modelled as a wide-sense stationary uncorrelated scattering (WSSUS) process, which is discussed in detail in Appendix A. Given a sampling interval T_s, the impulse response of the radio channel in the discrete time is denoted as $h(n, l)$ for propagation path l at sample n and $h(n, l) = 0$ for $l < 0$ or $l > L$, where L is the channel order, the received signal in the time domain at sample n is given by

$$y_t(n) = \sum_{l=0}^{L} h(n, l)u(n - l) + v_t(n) \tag{3.1}$$

where $y_t(n)$ is the received signal at sample n, $u(n)$ is the transmitted data at the nth sample, $v_t(n)$ is an additive white Gaussian noise (AWGN) with zero mean and variance σ_v^2.

Consider an OFDM block with the length of $N_s = N + N_{cp}$. For notational convenience, we choose n to denote the sample index and k to denote the index of the OFDM block so that $n = kN_s + N_{cp} + i$ for $i \in [0, N - 1]$ being the index of the sample within the considered OFDM block. In order to express the received

signal in matrix form, we define a number of vectors and matrices that are needed in this development. First, the sequences for the kth OFDM block are denoted as $y_t^k(i) = y(kN_s + N_{cp} + i)$, $u^k(i) = u(kN_s + N_{cp} + i)$, $h^k(i,l) = h(kN_s + N_{cp} + i, l)$ and $v_t^k(i) = v(kN_s + N_{cp} + i)$. Then, we define $\boldsymbol{y}_{tk}$, $\boldsymbol{u}_k$, $\boldsymbol{h}_l^k$ and $\boldsymbol{v}_{tk}$ as the received vector, the data vector, the channel gain vector and the noise vector of the lth path for block k, respectively, where $\boldsymbol{h}_l^k = [h^k(0,l), ..., h^k(N-1,l)]^T$, $\boldsymbol{y}_{tk} = [y_t^k(0), ..., y_t^k(N-1)]^T$, $\boldsymbol{u}_k = [u^k(0), ..., u^k(N-1)]^T$ and $\boldsymbol{v}_{tk} = [v_t^k(0), ..., v_t^k(N-1)]^T$. After removing CP, the received signal can be rewritten in matrix-vector form as

$$\boldsymbol{y}_{tk} = \boldsymbol{H}_t^k \boldsymbol{u}_k + \boldsymbol{v}_{tk} \tag{3.2}$$

where

$$\boldsymbol{H}_t^k = \begin{bmatrix} h^k(0,0) & 0 & \cdots & 0 & h^k(0,L) & \cdots & h^k(0,1) \\ h^k(1,1) & h^k(1,0) & 0 & \cdots & 0 & \cdots & h^k(1,2) \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h^k(L,L) & h^k(L,L-1) & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & h^k(N-1,L) & \cdots & \cdots & \cdots & h^k(N-1,0) \end{bmatrix} \tag{3.3}$$

or

$$\boldsymbol{H}_t^k = \sum_{l=0}^{L} \boldsymbol{\mathcal{D}}(\boldsymbol{h}_l^k) \boldsymbol{S}_l \tag{3.4}$$

where $\mathcal{D}(\boldsymbol{z})$ denotes a diagonal matrix with vector $\boldsymbol{z}$ on its diagonal and all other entries being zero, $\boldsymbol{S}_l$ is given by circularly shifting identity matrix $\boldsymbol{I}_N$ down with the delay of l samples. A circular matrix shift is the operation of rearranging the matrix, circularly shifting rows down (upper) or column left (right). The size of $\boldsymbol{H}_t^k$ is N by N. When the channel is time invariant over the period of an OFDM symbol, and the

length of CP is larger than the channel order L, then all of $h^k(i,l)$'s are not a function of i for all subcarriers i where $i \in [0, N-1]$, i.e., $h^k(i,l)$ is same for all i's, resulting in $\boldsymbol{H}_t^k$ being a circulant matrix. However, when the channel is time varying over one OFDM block, $h^k(i,l)$ has different value for each i and $\boldsymbol{H}_t^k$ is not a circulant matrix.

In the frequency domain, the received signal at the kth OFDM block is then expressed as

$$\begin{aligned} \boldsymbol{y}_k &= \boldsymbol{F}\boldsymbol{H}_t^k\boldsymbol{F}^H\boldsymbol{s}_k + \boldsymbol{F}\boldsymbol{v}_{tk} \\ &= \boldsymbol{H}_k\boldsymbol{s}_k + \boldsymbol{v}_k \end{aligned} \tag{3.5}$$

where $\boldsymbol{F}$ is the fast Fourier transform (FFT) matrix with the size of N by N and the (m,n)th entry $F_{m,n} = 1/\sqrt{N}e^{-j2\pi(m-1)(n-1)/N}$, $\boldsymbol{F}^H$ is the inverse fast Fourier transform (IFFT) matrix. $\boldsymbol{y}_k$ is an N by 1 vector defined as $\boldsymbol{y}_k = \boldsymbol{F}\boldsymbol{y}_{tk}$, $\boldsymbol{H}_k = \boldsymbol{F}\boldsymbol{H}_t^k\boldsymbol{F}^H$ and $\boldsymbol{v}_k = \boldsymbol{F}\boldsymbol{v}_{tk}$. $\boldsymbol{s}_k$ is the transmitted signal vector in the frequency domain at the kth OFDM block, the size is N by 1.

In the special case that $\boldsymbol{H}_t^k$ is a circulant matrix, the matrix $\boldsymbol{H}_k$ is diagonal, and hence, no interference between subcarriers exist. Therefore, no intercarrier interface (ICI) occurs in slow fading channels. However, when the channel is time varying over the OFDM symbol duration, the sample of the channel gain for each propagation path changes from the beginning to the end of each OFDM symbol. Therefore, the matrix $\boldsymbol{H}_k$ is not diagonal and becomes a banded matrix, the gray part as shown in Figure 3.2. Cross-terms indicating interference between subcarriers are created and result in ICI. In this case, the channel gain at each time sample of the OFDM block is needed for data detection. In fact, the ICI on a subcarrier mainly comes from several neighboring subcarriers. If the Doppler frequency increases, more symbol energy leaks to neighboring subcarriers as shown in Figure 3.2.

In the case that the channel impulse response (CIR) is known, low-complexity de-

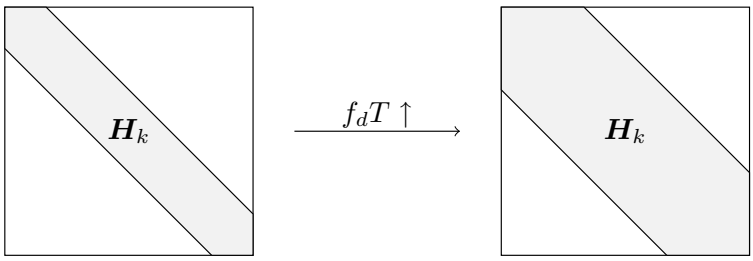

Figure 3.2: Structure of $\boldsymbol{H}_k$ matrix

tection methods have been proposed to mitigate ICI. Whereas, in the case of initially unknown channels, we propose joint channel estimation data detection which will be discussed in detail in Chapter 5.

3.1.2 Time Domain Descriptions for a Transmission Block

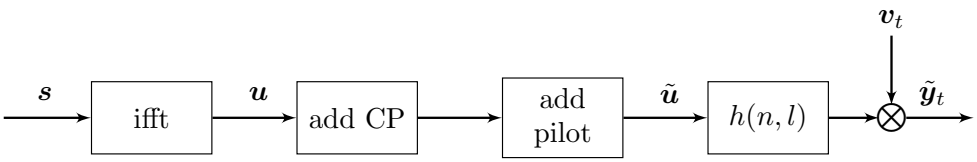

Figure 3.3: Discrete-time structure in OFDM systems

When modelling a fast fading channel, we recall that the transmission block consists of K_b OFDM blocks and K_p pilot blocks. On the one hand, short transmission block will increase the error of the channel model, on the other hand, a long transmission block will increase the complexity of the channel estimation. Therefore, there is a tradeoff between the complexity and the channel modelling error. The choices of K_b and K_p for the transmission block will be discussed in detail in Chapter 6.

As described in Chapter 2, the time multiplexed pilots scheme is preferable to channel estimation in fast fading channels due to the advantage of temporal separation from the data symbols. With a pilot block in the form of $[\mathbf{0}_{1\times L}\ 1\ \mathbf{0}_{1\times L}]$, as shown in Figure 2.3, inter-block interference between pilot block and OFDM symbol block can

be avoided. Within a transmission block, K_p pilot blocks are uniformly inserted in K_b OFDM blocks, where $K_b = \kappa K_p$, κ is the number of OFDM blocks between two adjacent pilot blocks and $\kappa \geq 1$ is an integer. To minimize the total channel of mean square error (MSE), we will show that the optimal κ is 1 in Chapter 5, i.e., 1 pilot block for every OFDM block, and hence, the transmitted signal $\boldsymbol{u}$ in time domain is expressed as

$$\tilde{\boldsymbol{u}} = \begin{bmatrix} \boldsymbol{u}_0^T & \boldsymbol{d}_0^T & \cdots & \boldsymbol{u}_{K_p-1}^T & \boldsymbol{d}_{K_p-1}^T \end{bmatrix}^T \tag{3.6}$$

where $\boldsymbol{u}_p$ and $\boldsymbol{d}_p$ for $p = 0, ..., K_p - 1$, are data and pilot block, respectively; $\boldsymbol{u} = [\boldsymbol{u}_0^T, ..., \boldsymbol{u}_{K_p-1}^T]^T$, and $\boldsymbol{d} = [\boldsymbol{d}_0^T, ..., \boldsymbol{d}_{K_p-1}^T]^T$.

After adding pilot symbols, the received signal vector for the transmission block in the time domain is then expressed as

$$\tilde{\boldsymbol{y}}_t = \begin{bmatrix} \boldsymbol{y}_{t0}^d \\ \boldsymbol{y}_{t0}^p \\ \vdots \\ \boldsymbol{y}_{t(K_p-1)}^d \\ \boldsymbol{y}_{t(K_p-1)}^p \end{bmatrix} \tag{3.7}$$

where $\boldsymbol{y}_{tp}^d$ and $\boldsymbol{y}_{tp}^p$ for $p = 0, ..., K_p - 1$, are the received signal vector corresponding to $\boldsymbol{u}_p$ and $\boldsymbol{d}_p$, respectively.

Similar to (3.2), we express $\tilde{\boldsymbol{y}}_t$ as

$$\tilde{\boldsymbol{y}}_t = \tilde{\boldsymbol{H}}_t \tilde{\boldsymbol{u}} + \tilde{\boldsymbol{v}}_t \tag{3.8}$$

where $\tilde{\boldsymbol{H}}_t$ is the channel matrix of transmission block, $\tilde{\boldsymbol{v}}_t$ is the AWGN vector in the transmission block.

The relationship between $\tilde{\boldsymbol{H}}_t$ and the received signal $\tilde{\boldsymbol{y}}_t$ is illustrated in Figure 3.4,

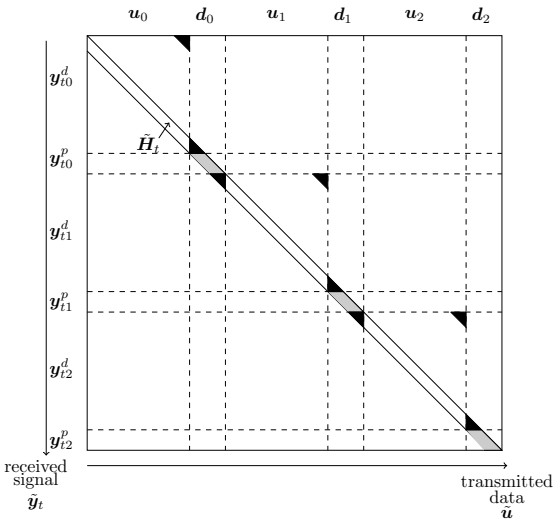

Figure 3.4: Relationship between the transmitted data and the received signal in the time domain for single transmission block

with the transmitted index on the X-axis and the received index on the Y-axis. Since $\tilde{\boldsymbol{H}}_t$ is a banded matrix, $\tilde{\boldsymbol{H}}_t$ is split into three parts: $\boldsymbol{H}_t$, $\boldsymbol{H}_t^P$ and $\boldsymbol{H}_{tsd}$. In the figure, $\boldsymbol{H}_t$ is the white central diagonal part, corresponding to the relationship between the transmitted signal $\boldsymbol{u}_i$ and the received signal $\boldsymbol{y}^d_{ti}$; $\boldsymbol{H}_t^P$ is the gray part, corresponding to the relationship between the transmitted signal $\boldsymbol{d}_i$ and the received signal $\boldsymbol{y}^p_{ti}$; and $\boldsymbol{H}_{tsd}$ is the black triangular part, corresponding to the interblock interference that emerges due to the channel delay spread. As a result, the received signal, $\tilde{\boldsymbol{y}}_t$, in the transmission block can be separated into two portions, which are the received signal at the data position, $\boldsymbol{y}_t^D$, and the received signal at the pilot position, $\boldsymbol{y}_t^P$. Therefore, Equation (3.7) can be separated as

$$\boldsymbol{y}_t^D = \begin{bmatrix} \boldsymbol{y}^d_{t0} \\ \vdots \\ \boldsymbol{y}^d_{t(K_p-1)} \end{bmatrix} = \boldsymbol{H}_t\boldsymbol{u} + \boldsymbol{H}_{tsd}\tilde{\boldsymbol{d}} + \boldsymbol{v}_t^D \tag{3.9}$$

$$\boldsymbol{y}_t^P = \begin{bmatrix} \boldsymbol{y}_{t0}^p \\ \vdots \\ \boldsymbol{y}_{t(K_p-1)}^p \end{bmatrix} = \boldsymbol{H}_t^P \boldsymbol{d} + \boldsymbol{v}_t^P \tag{3.10}$$

where $\tilde{\boldsymbol{d}} = [\tilde{\boldsymbol{d}}_0^T, ..., \tilde{\boldsymbol{d}}_{K_p-1}^T]^T$ and $\tilde{\boldsymbol{d}}_p$ contains the first and last L entries of $\boldsymbol{d}_p$ for $p = 0, ..., K_p - 1$. $\boldsymbol{v}_t^D$ and $\boldsymbol{v}_t^P$ are the corresponding noise vectors. The term $\boldsymbol{H}_{tsd}\tilde{\boldsymbol{d}}$ corresponds to the interference of the pilot blocks to their adjacent data blocks. By use of the pilot form discussed above, $\tilde{\boldsymbol{d}}_p = \boldsymbol{0}$ for $p = 0, ..., K_p - 1$, and hence, the interference part can be mitigated. In such a case, Equation (3.9) becomes

$$\boldsymbol{y}_t^D = \begin{bmatrix} \boldsymbol{y}_{t0}^d \\ \vdots \\ \boldsymbol{y}_{t(K_p-1)}^d \end{bmatrix} = \boldsymbol{H}_t \boldsymbol{u} + \boldsymbol{v}_t^D \tag{3.11}$$

$\boldsymbol{y}_t^D$ and $\boldsymbol{y}_t^P$ will be discussed in detail in Section 3.2.1 and Section 3.2.2, respectively.

3.1.3 Coding Block

Since the channel is time-variant and frequency-selective, the channel exhibits bursty error characteristics [2, 99]. In order to deal with temporary deep fades, an effective method is to exploit coding with a interleaver. After interleaving, errors are spread out so that error events for two adjacent bits in the output data stream are independent [2]. The major drawback with large interleavers is decoding delay and the computational complexity in the iterative decoding scheme.

The size of the interleaver is chosen based on the Doppler frequency. Lower Doppler frequency requires a longer interleaver to obtain the same time diversity as a higher Doppler frequency and vice versa. On the other hand, in order to keep the complexity of the channel estimation low, the channel is estimated over a transmis-

sion block and interleaving is exploited over a coding block which contains several transmission blocks. In Chapter 6, we will show how to use the extrinsic information transfer (EXIT) chart to select the interleaver size for good performance.

3.2 Three Measurement Models Introduction

As mentioned before, the proposed iterative receiver includes data detection, channel estimation and decoding. This motivates the need for three different measurement models. Due to the separation between the received signal $\tilde{\boldsymbol{y}}_t$, we can develop the three measurements models: measurements for data given channel, measurements for channel given pilots, and measurements for channel given data, which will be discussed in this section.

3.2.1 Measurements for Data Given Channel

For OFDM systems, if a transmission block contains K_p pilot blocks,where each pilot block has the form of $[\mathbf{0}_{1\times L}\ 1\ \mathbf{0}_{1\times L}]$, and K_p OFDM blocks, where each OFDM block contains $N_s = N + N_{cp}$ samples, the index n in Equations (3.2) and (3.3) becomes $n = kN_s + N_{cp} + p(2L+1) + i$ for $k \in [0, K_b - 1]$ $p \in [0, K_p - 1]$ and $i \in [0, N-1]$, the size of $\boldsymbol{y}_t^D$ in Equation (3.11) is K_pN by 1 and $\boldsymbol{H}_t$ is given by

$$\boldsymbol{H}_t = \begin{bmatrix} \boldsymbol{H}_t^0 & & \\ & \ddots & \\ & & \boldsymbol{H}_t^{(K_p-1)} \end{bmatrix} \tag{3.12}$$

or $\boldsymbol{H}_t = \text{diag}\{\boldsymbol{H}_t^0, ..., \boldsymbol{H}_t^{(K_p-1)}\} = \oplus_{k=0}^{K_p-1} \boldsymbol{H}_t^k$, and $\boldsymbol{H}_t^k$ is the channel matrix in the time domain of the kth OFDM block, expressed as Equation (3.3). Therefore, the

received signal in the frequency domain is expressed as

$$\boldsymbol{y}^D = \begin{bmatrix} \boldsymbol{y}_0^d \\ \vdots \\ \boldsymbol{y}_{(K_p-1)}^d \end{bmatrix} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}^D \tag{3.13}$$

where $\boldsymbol{y}^D = \oplus_{k=0}^{K_p-1} \boldsymbol{F}\boldsymbol{y}_t^D$, $\boldsymbol{H} = \text{diag}\{\boldsymbol{H}_0, ..., \boldsymbol{H}_{(K_p-1)}\} = \oplus_{k=0}^{K_p-1} \boldsymbol{H}_k$, $\boldsymbol{s}$ is the transmitted data in the frequency domain, $\boldsymbol{s} = [\boldsymbol{s}_0^T, ..., \boldsymbol{s}_{K_p-1}^T]^T$; $\boldsymbol{v}^D$ is AWGN vector in the frequency domain.

For single-carrier (SC) systems, if a transmission block contains K_p pilot blocks,where each pilot block has the form of $[\mathbf{0}_L\ 1\ \mathbf{0}_L]$, and K_b data blocks, where each data block contains $\mathcal{N}$ samples, the index n in Equation (3.1) becomes $n = k\mathcal{N} + p(2L+1) + i$ for $k \in [0, K_b-1]$ $p \in [0, K_p-1]$ and $i \in [0, \mathcal{N}-1]$. Based on the description of Equation (3.1), for the nth sample of a transmission block, the received signal is defined as $y_t(n) = \boldsymbol{h}_{tn}^T \mathcal{U}_n + v_t(n)$ where $\boldsymbol{h}_{tn} = [h(n,0), ..., h(n,L)]^T$, $\mathcal{U}_n = [u(n), ..., u(n-L+1)]^T$.

3.2.2 Measurements for Channel Given Pilots

As indicated in Section 3.1.1, when the channel varies significantly over one OFDM duration, we need $N(L+1)$ channel gains to obtain the channel matrix $\boldsymbol{H}_t^k$ in Equation (3.2). Unfortunately, the number of unknowns is larger than that of measurements. To overcome this difficulty, the estimation of basis coefficients for a basis expansion model (BEM), which is discussed in detail in Appendix B, is instead. Therefore, the channel gain for the transmission block of the lth path is $\boldsymbol{h}_l = \boldsymbol{E}\boldsymbol{x}_l$, where $\boldsymbol{E}$ is the basis matrix and $\boldsymbol{x}_l$ specifies the basis coefficients for path l for the transmission block, defined as $\boldsymbol{x}_l = [x_l(0), ..., x_l(Q)]^T$, Q is the number of BEM order denoted by $Q \geq 2\lceil f_d M T_s \rceil$, where T_s is the sampling time and M is the number of samples in each transmission block; $x_l(q)$ is the coefficient for path l of the qth basis functions,

all defined in Appendix B. With the number of pilot symbols $2L+1$ for each pilot block, the required number of pilot blocks K_p for each transmission block should be satisfied $K_p \geq Q+1$.

Define $\boldsymbol{x}$ as the basis coefficient vector,

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_0 \\ \vdots \\ \boldsymbol{x}_L \end{bmatrix} = \begin{bmatrix} x_0(0) \\ \vdots \\ x_0(Q) \\ \vdots \\ x_L(0) \\ \vdots \\ x_L(Q) \end{bmatrix}. \tag{3.14}$$

Based on the discussion of Section 3.1.2, the received signal corresponding to pilot symbols is obtained as

$$\begin{aligned} \boldsymbol{y}_t^P &= \begin{bmatrix} \boldsymbol{y}_{t0}^p \\ \vdots \\ \boldsymbol{y}_{t(K_p-1)}^p \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_0^p \\ \vdots \\ \boldsymbol{C}_{K_p-1}^p \end{bmatrix} \boldsymbol{x} + \boldsymbol{v}_t^P \\ &= \boldsymbol{C}^P \boldsymbol{x} + \boldsymbol{v}_t^P, \end{aligned} \tag{3.15}$$

where $\boldsymbol{v}_t^P$ is the AWGN vector, the length of $\boldsymbol{y}_t^P$ is $K_p(L+1)$, $\boldsymbol{C}^P = [\boldsymbol{C}_0^{pT}, ..., \boldsymbol{C}_{K_p-1}^{p\ T}]^T$, and

$$\boldsymbol{C}_p^p = \begin{bmatrix} \boldsymbol{e}_{p_p} & \boldsymbol{0}_{Q+1} & \cdots & \boldsymbol{0}_{Q+1} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0}_{Q+1} & \cdots & \cdots & \boldsymbol{e}_{p_p+L} \end{bmatrix},$$

where p_p is the first index of $\boldsymbol{y}_{tp}^p$, defined as $p_p = (k+1)N_s + k(2L+1)$ for $k \in [0, K_p-1]$; $\boldsymbol{e}_i = [E_i(0) \cdots E_i(Q)]$ is a basis function vector at sample i with the size of 1 by $(Q+1)$,

and $\boldsymbol{C}_p^p$ is a block matrix, where the size of $\boldsymbol{C}_p^p$ is $(L+1)$ by $(L+1)(Q+1)$. In the case that the number of pilot blocks K_p is larger than the number of OFDM blocks K_b, as long as the pilot form is the same as the described in Section 3.1.2, we can still use Equation (3.15) with K_b instead of K_p. As a consequence, $p_p = (k+1)N_s + p(2L+1)$ for $k \in [0, K_b-1]$ and $p \in [0, K_p-1]$. Hence, Equation (3.15) denotes the measurement model for channel in terms of pilot symbols for channel estimation in the time domain.

3.2.3 Measurements for Channel Given Data

In iterative joint channel estimation/data detection schemes, previously detected values for the transmitted data are used in conjunction with the measurements to estimate the channel state. This requires a model relating the channel state to the measurements given the detected data values. This model is developed in this section. After the initial channel estimation with pilot symbols, the measurements in the frequency domain $\boldsymbol{y}^D$ taken from the detected symbols are used to calculate the BEM coefficients. In this section, first we turn our attention to develop the relationship between the received signal and the BEM coefficients for each OFDM block. Then we exploit the measurement signals for channel estimation based on the detected data over a transmission block.

With the use of a BEM, the channel gain vector of the kth OFDM block in a transmission block for the lth path are then expressed as $\boldsymbol{h}_l^k = \boldsymbol{E}_k \boldsymbol{x}_l$, where $\boldsymbol{E}_k = \boldsymbol{E}[kN_s + N_{cp} + p(2L+1) : kN_s + N_{cp} + p(2L+1) + N - 1, :)]$, $k \in [0, K_b - 1]$ and

$p \in [0, K_p - 1]$. In such a case, the time domain receiver for the kth OFDM block is

$$\begin{aligned} \boldsymbol{y}_{tk} &= \sum_{l=0}^{L} \boldsymbol{\mathcal{D}}\left(\boldsymbol{E}_k \boldsymbol{x}_l\right) \boldsymbol{S}_l \boldsymbol{F}^H \boldsymbol{s}_k + \boldsymbol{v}_{tk} \\ &= \sum_{l=0}^{L} \boldsymbol{\mathcal{D}}\left(\boldsymbol{S}_l \boldsymbol{F}^H \boldsymbol{s}_k\right) \boldsymbol{E}_k \boldsymbol{x}_l + \boldsymbol{v}_{tk} \\ &= \boldsymbol{C}_k \boldsymbol{x} + \boldsymbol{v}_{tk} \end{aligned} \tag{3.16}$$

where $\boldsymbol{x}$ is basis coefficient vector defined in Equation (3.14), and the matrix $\boldsymbol{C}_k$ is a block matrix, expressed as

$$\boldsymbol{C}_k = \begin{bmatrix} \boldsymbol{C}_{0k} & \cdots & \boldsymbol{C}_{Lk} \end{bmatrix} \tag{3.17}$$

with $\boldsymbol{C}_{lk} = \boldsymbol{\mathcal{D}}\left(\boldsymbol{S}_l \boldsymbol{F}^H \boldsymbol{s}_k\right) \boldsymbol{E}_k$.

The received signal in the frequency domain is then expressed as

$$\begin{aligned} \boldsymbol{y}_k &= \boldsymbol{F} \boldsymbol{C}_k \boldsymbol{x} + \boldsymbol{v}_k \\ &= \boldsymbol{\mathcal{M}}_k \boldsymbol{x} + \boldsymbol{v}_k \end{aligned} \tag{3.18}$$

where $\boldsymbol{\mathcal{M}}_k = [\boldsymbol{M}_{0k} \cdots \boldsymbol{M}_{Lk}]$ with $\boldsymbol{M}_{lk} = \boldsymbol{F}\boldsymbol{C}_{lk} = \boldsymbol{F}\boldsymbol{\mathcal{D}}(\boldsymbol{S}_l \boldsymbol{F}^H \boldsymbol{s}_k)\boldsymbol{E}_k$. When the transmitted data is given, Equation (3.18) can be used for channel estimation based on the decision-directed method.

For each transmission block, the measurement in the frequency domain $\boldsymbol{y}^D$ of

Equation (3.13) is rewritten as

$$\begin{aligned}\boldsymbol{y}^D &= \begin{bmatrix} \boldsymbol{y}_0^d \\ \vdots \\ \boldsymbol{y}_{(K_p-1)}^d \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{M}}_0 \\ \vdots \\ \boldsymbol{\mathcal{M}}_{K_p-1} \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \boldsymbol{v}_0 \\ \vdots \\ \boldsymbol{v}_{K_p-1} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{M}_{00} & \cdots & \boldsymbol{M}_{L0} \\ \vdots & \cdots & \vdots \\ \boldsymbol{M}_{0(K_p-1)} & \cdots & \boldsymbol{M}_{L(K_p-1)} \end{bmatrix} \boldsymbol{x} + \boldsymbol{v}^D \\ &= \boldsymbol{\mathcal{M}}_D \boldsymbol{x} + \boldsymbol{v}^D, \end{aligned} \tag{3.19}$$

When the data in the transmission block is detected, Equation (3.19) is then used to estimate the basis coefficients for data-aided channel estimation, which will be discussed in detail in Chapter 5.

In summary, we have provided different system models that find applications in fast fading channels. Of the models described, Equation (3.5) is used for data detection when the channel gain is available. As the case of pilot aided channel estimation, Equation (3.15) is used for the time domain channel estimation. As the case of decision-directed channel estimation, Equation (3.19) is used for channel estimation based on the detected data.

3.3 Time Domain Evolution of the Channel

Of the different channel models, an autoregressive (AR) Gauss-Markov model is exploited to describe the time-varying channel in general [75]. Since the channel is bandlimited, the evolution of the BEM coefficients for each propagation path in a fast fading channel between transmission blocks can be modelled as a multivariate AR process. This AR model allows the estimated BEM coefficients for the previous

transmission block to be used to estimate the BEM coefficients for the current transmission block, reducing the need for pilot symbols and improving channel estimation accuracy for the initial iteration of the channel estimation/data detection systems.

In order to exploit channel information contained within transmission blocks other than current transmission block, a novel channel model is derived, where the unknown basis coefficients are characterized by a multivariate AR Gauss-Markov model, discussed in detail below.

When the channel is modelled by a BEM for each transmission block, we assume that the basis coefficient vector for each propagation path is a multivariate AR process of order K, the basis coefficient vectors for each propagation path at transmission block n are given by

$$\boldsymbol{x}_l[n] = -\sum_{i=1}^{K} \boldsymbol{A}_l[i]\boldsymbol{x}_l[n-i] + \boldsymbol{w}_l[n-1] \tag{3.20}$$

where the $\boldsymbol{A}_l[i]$'s are multivariate AR coefficient matrices with the size of $(Q+1)$ by $(Q+1)$, and $\boldsymbol{w}_l[n]$ is an AWGN vector process with zero mean and a covariance of $\boldsymbol{Q}_{lf}$. The bandlimited nature of the fading process makes the actual $\boldsymbol{x}_l[n]$ process of infinite order. However, in practice, the process is well modelled as an AR process of finite order K with the approximation becoming more accurate with increasing K. To calculate the multivariate AR coefficients and the covariance $\boldsymbol{Q}_{lf}$, it is necessary to create a state vector $\underline{\boldsymbol{x}}_l[n]$ to contain basis coefficients for blocks $n-K+1$ to n:

$$\underline{\boldsymbol{x}}_l[n] = \begin{bmatrix} \boldsymbol{x}_l[n] \\ \vdots \\ \boldsymbol{x}_l[n-K+1] \end{bmatrix}. \tag{3.21}$$

Equation (3.20) can be rewritten as

$$\underline{\boldsymbol{x}}_l[n] = \begin{bmatrix} -\boldsymbol{A}_l[1] & -\boldsymbol{A}_l[2] & \cdots & -\boldsymbol{A}_l[K] \\ \boldsymbol{I} & \boldsymbol{I} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \\ \boldsymbol{0} & \cdots & \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \underline{\boldsymbol{x}}_l[n-1] + \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{w}_l[n-1] \tag{3.22}$$

where $\boldsymbol{I}$ and $\boldsymbol{0}$ are $(Q+1)$ by $(Q+1)$ identity matrices and $(Q+1)$ by $(Q+1)$ zero matrices, respectively. The values of the AR coefficient matrices satisfy the multichannel Yule-Walker equation [100, 101]:

$$\underline{\boldsymbol{A}}_l \, \underline{\boldsymbol{R}}_l = \begin{bmatrix} \boldsymbol{Q}_{lf} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix} \tag{3.23}$$

where

$$\underline{\boldsymbol{A}}_l = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}_l[1] & \cdots & \boldsymbol{A}_l[K] \end{bmatrix}, \tag{3.24}$$

and

$$\underline{\boldsymbol{R}}_l = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{x}}(0) & \cdots & \boldsymbol{R}_{\boldsymbol{x}}(K) \\ \boldsymbol{R}_{\boldsymbol{x}}(-1) & \cdots & \boldsymbol{R}_{\boldsymbol{x}}(K-1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{R}_{\boldsymbol{x}}(-K) & \cdots & \boldsymbol{R}_{\boldsymbol{x}}(0) \end{bmatrix}. \tag{3.25}$$

These matrices can be efficiently calculated by the use of Whittle-Wiggins-Robinson Algorithm (WWRA) [102, 103], which is discussed in Appendix C. The matrix $\underline{\boldsymbol{R}}_l$ has a Toeplitz-block structure with $\boldsymbol{R}_{\boldsymbol{x}}(i) = E[\boldsymbol{x}_l[n+i]\boldsymbol{x}_l[n]^H]$, and $\boldsymbol{R}_{\boldsymbol{x}}(-i) = \boldsymbol{R}_{\boldsymbol{x}}^H(i)$.

To calculate $\boldsymbol{R_x}(i)$, we get the covariance matrix of each channel path as

$$\begin{aligned}\boldsymbol{R_h}(i) &= E[\boldsymbol{h}_l[n+i]\boldsymbol{h}_l[n]^H] \\ &= E[(\boldsymbol{E}\boldsymbol{x}_l[n+i])(\boldsymbol{E}\boldsymbol{x}_l[n])^H] \\ &= \boldsymbol{E}\boldsymbol{R_x}(i)\boldsymbol{E}^H.\end{aligned} \tag{3.26}$$

Based on Jakes' model, the correlation function of each channel path is expressed as $r_l[k] = \sigma_l^2 J_0(2\pi f_d T_s k)$, where J_0 is the zeroth order Bessel function[104] and σ_l^2 is the variance of the lth path. The average power of the channel is normalized to one, i.e., $\sum_{l=0}^{L} \sigma_l^2 = 1$. So the cross-covariance matrix $\boldsymbol{R_h}(i)$ is given by

$$\boldsymbol{R_h}(i) = \begin{bmatrix} r_l[Mi] & \cdots & r_l[M(i-1)+1] \\ r_l[Mi+1] & \cdots & r_l[M(i-1)+2] \\ \vdots & \ddots & \vdots \\ r_l[M(i+1)-1] & \cdots & r_l[Mi] \end{bmatrix}. \tag{3.27}$$

Since $\boldsymbol{E}^H\boldsymbol{E} = \boldsymbol{I}$, then we calculate $\boldsymbol{R_x}(i)$ from Equations (3.26) and (3.27) as $\boldsymbol{R_x}(i) = \boldsymbol{E}^H\boldsymbol{R_h}(i)\boldsymbol{E}$. A difficulty encountered when computing the coefficients for higher order models is that $\underline{\boldsymbol{R}}_l$ is ill-conditioned due to the band-limited process of the channel making the ratio of the magnitude of the smallest to largest singular value of $\boldsymbol{R}_l$ very large. This problem can be overcome by adding small positive values to the diagonal entries of $\boldsymbol{R_x}(0)$ or by using other techniques [105, 106]. If $\boldsymbol{A}_l$ and $\boldsymbol{Q}_{lf}$ are stationary, which means that the Doppler frequency does not vary over the time, we can calculate these matrices off-line. However, when the Doppler frequency changes over the time, the calculation of $\boldsymbol{A}_l$ and $\boldsymbol{Q}_{lf}$ should be on-line.

Chapter 4

Second-Order Kalman Filtering Channel Estimation in Fast Fading

4.1 Introduction

In this chapter, we consider the problem of channel estimation in the presence of fast fading. The channel state is characterized by basis coefficients over a fixed period of time. Unlike the previous Kalman filter based channel estimation methods which focus on the time variation within one OFDM block, the proposed algorithm takes the time variation within each transmission block into account.

By use of a basis expansion model (BEM), a second-order multivariate autoregressive (AR) model is exploited to characterize the unknown channel state over each transmission block. In this chapter a Kalman smoother is proposed, which uses more measurements and is more accurate compared to the Kalman filtering [100]. It will be shown that the proposed technique provides channel mean square error (MSE) is near to that of the optimal minimum mean square error (MMSE) Wiener filter. The computational cost of this new algorithm is on the same order as existing channel

estimation algorithms. In addition, the proposed approach is robust to variations of the radio channel parameters from the design values. For data detection, a low-complexity MMSE and decision-feedback equalizer (DFE) algorithm [5] is used to suppress intercarrier interference (ICI).

4.2 Second-Order Kalman Filtering Algorithm

Unlike the previous Kalman filter methods proposed for fast fading channel estimation [13, 107], where a Kalman filter was developed in single-carrier (SC) systems, the time evolution of basis coefficients over each transmission block is described as a second-order multivariate AR model for OFDM systems. The second-order Kalman filter will be shown to track the basis coefficients from the pilot measurements requiring only a low pilot to data ratio. Moreover, a Kalman smoother, which incorporates the following measurements, gives a more refined estimate than the previously presented channel estimation methods. The channel state information (CSI) is estimated based on pilot symbols which are multiplexed in the time domain. The basis coefficients are first estimated using pilot symbols, then the CSI is calculated from basis matrices.

For Kalman filtering algorithms, the state space model is used for estimation, which includes the state model and the measurement model. As discussed in Chapter 3, each transmission block includes K_b OFDM blocks (OBs) and K_p pilot blocks, which are uniformly inserted in between the OFDM blocks. To exploit the state model of the basis coefficients over the transmission block, we consider that a multipath channel is modelled as a BEM with $Q+1$ basis coefficients over the transmission block and has $L+1$ paths. From the description of the state model given in Chapter

3, the state vector of the second-order, $K = 2$, is given by

$$\underline{\boldsymbol{x}}[n] = \begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}[n-1] \end{bmatrix} \tag{4.1}$$

where $\boldsymbol{x}[n]$ is the basis coefficient vector with the size of $(L+1)(Q+1)$ by 1 defined in Equation (3.14) at the nth transmission block, that is, $\boldsymbol{x}_l[n]$ is the vector of the Q basis coefficients for the lth channel path for the nth transmission block. Consider an OFDM symbol with N subcarriers, N_{cp} cyclic prefix (CP) samples and each pilot block with the form $[\mathbf{0}_{1\times L}\ 1\ \mathbf{0}_{1\times L}]$, the nth transmission block will include $M = (N+N_{cp})K_b + (2L+1)K_p$ samples for each channel path, defined in vector form as

$$\boldsymbol{h}_l[n] = \begin{bmatrix} h(n,l) \\ \vdots \\ h(n+M-1,l) \end{bmatrix} = \boldsymbol{E}\boldsymbol{x}_l[n] \tag{4.2}$$

where $\boldsymbol{E}$ is the basis function matrix defined in Equation B.2.

The state model is described by

$$\begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}[n-1] \end{bmatrix} = \boldsymbol{\Phi} \begin{bmatrix} \boldsymbol{x}[n-1] \\ \boldsymbol{x}[n-2] \end{bmatrix} + \begin{bmatrix} \boldsymbol{I} \\ \mathbf{0} \end{bmatrix} \boldsymbol{w}[n-1] \tag{4.3}$$

where $\boldsymbol{\Phi}$ is the state transition matrix given by

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12} \\ \boldsymbol{\Phi}_{21} & \boldsymbol{\Phi}_{22} \end{bmatrix}. \tag{4.4}$$

where $\boldsymbol{\Phi}_{ij}$ for $i,j = 1,2$ is a block matrix. To simplify the equation, we use $\oplus$ to specify the direct sum of matrices: $\oplus_{l=0}^{L}\boldsymbol{A}_l = \text{diag}\{\boldsymbol{A}_0, ..., \boldsymbol{A}_L\}$ where diag is the

block diagonal operator [108]. Therefore, $\boldsymbol{\Phi}_{11} = -\oplus_{l=0}^{L} \boldsymbol{A}_l[1]$, $\boldsymbol{\Phi}_{12} = -\oplus_{l=0}^{L} \boldsymbol{A}_l[2]$, $\boldsymbol{\Phi}_{21} = -\oplus_{l=0}^{L} \boldsymbol{I}$, and $\boldsymbol{\Phi}_{22} = -\oplus_{l=0}^{L} \boldsymbol{0}$, where $\boldsymbol{I}$ is $(Q+1)\times(Q+1)$ identity matrix, and the $\boldsymbol{A}_l[i]$'s are multivariate AR coefficient matrices of the propagation path l with the size of $(Q+1)$ by $(Q+1)$, respectively. $\boldsymbol{w}[n]$ is an AWGN process with a covariance matrix $\boldsymbol{Q}$ given by

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \tag{4.5}$$

where $\boldsymbol{Q}_1 = \oplus_{l=0}^{L} \boldsymbol{Q}_{lf}$, $\boldsymbol{Q}_{lf}$ is the covariance matrix of the state vector for path l.

Let σ_l^2 be the variance of the lth path. The average power of the channel is normalized to one, i.e., $\sum_{l=0}^{L} \sigma_l^2 = 1$. In Appendix C, it is shown how to solve for the $\boldsymbol{A}[i]$ and $\boldsymbol{Q}_f$ matrices using the Whittle-Wiggins-Robinson Algorithm (WWRA) [102, 103]. From calculations of $\boldsymbol{A}[i]$ and $\boldsymbol{Q}_f$, we obtain $\boldsymbol{A}_l[i] = \boldsymbol{A}[i]$ and $\boldsymbol{Q}_l = \sigma_l^2 \boldsymbol{Q}_{lf}$.

Now we turn our attention to the measurement model over the nth transmission block, which is expressed as

$$\boldsymbol{y}[n] = \boldsymbol{G}[n]\underline{\boldsymbol{x}}[n] + \boldsymbol{v}[n] \tag{4.6}$$

where $\boldsymbol{G}[n]$ is the measurement matrix and $\boldsymbol{v}[n]$ is the measurement noise with each covariance $\sigma_v^2 = N_0/2$, where N_0 is power spectral density. As described in Chapter 3, the CSI is estimated based on pilot symbols which are multiplexed in the time domain. For time multiplexed pilots, the measurement model of a transmission block is given by Equation (3.15), i.e., $\boldsymbol{y}_t^P = \boldsymbol{C}^P \boldsymbol{x} + \boldsymbol{v}_t^P$, where $\boldsymbol{y}_t^P$ is the received signal vector corresponding to pilot symbols, $\boldsymbol{v}_t^P$ is the AWGN vector. Therefore, for the nth transmission block, we obtain $\boldsymbol{y}[n] = \boldsymbol{y}_t^P[n]$, $\boldsymbol{v}[n] = \boldsymbol{v}_t^P[n]$ and $\boldsymbol{G}[n]$ expressed as

$$\boldsymbol{G}[n] = \begin{bmatrix} \boldsymbol{C}^P[n] & \boldsymbol{0} \end{bmatrix} \tag{4.7}$$

The measurement noise process $\boldsymbol{v}[n]$ is an AWGN vector process independent of $\boldsymbol{w}[n]$, and the variance matrix $\boldsymbol{R}[n]$ is a diagonal matrix with all entries on the main diagonal given by variance σ_v^2.

The Kalman filter equations for a single block of the channel measurements are given in Appendix E. The estimate of $\underline{\boldsymbol{x}}[n]$ given all measurements $\boldsymbol{y}[n]$ up to block m is denoted by $\underline{\hat{\boldsymbol{x}}}[n|m]$ with the corresponding error covariance matrix denoted as $\boldsymbol{P}[n|m]$. $\hat{\boldsymbol{y}}[n|n-1]$ is the predicted measurement for block n, $\boldsymbol{z}[n]$ is the so-called innovation sequence with a covariance matrix given by $\boldsymbol{M}[n|n-1]$, and $\boldsymbol{K}[n]$ is the Kalman gain for block n. Given an initial estimate of the state $\underline{\hat{\boldsymbol{x}}}[0] = \boldsymbol{0}$ with error covariance $\boldsymbol{P}[0] = \oplus_{l=0}^{L} \boldsymbol{P}_l[0]$, where

$$\boldsymbol{P}_l[0] = E\left(\begin{bmatrix} \boldsymbol{x}_l[0] \\ \boldsymbol{x}_l[-1] \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_l[0] \\ \boldsymbol{x}_l[-1] \end{bmatrix}^H \right) = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{x}}[0] & \boldsymbol{R}_{\boldsymbol{x}}[1] \\ \boldsymbol{R}_{\boldsymbol{x}}[-1] & \boldsymbol{R}_{\boldsymbol{x}}[0] \end{bmatrix},$$

and $\boldsymbol{R}_{\boldsymbol{x}}[i]$ is calculated based on Equation (3.26), the Kalman filter, which is calculated based on Appendix E, estimates the channel state at block n given measurements up to $\boldsymbol{y}[n]$.

The Kalman filter is easily modified into a Kalman smoother [100]. The parameter estimation x_k of the Kalman filter only takes into account the information in $[0, k]$. However, by incorporating the future measurements relative to x_k, that is, the measurements between $[0, n]$ for $n > k$, we can obtain a more refined estimate. If the radio receiver can store the channel samples and wait until after the reception of block n before decoding the data symbols for block $n-1$, a substantial improvement in channel estimation MSE and, thus, bit error rate (BER) is obtained. An estimate of the channel BEM coefficients $\hat{\boldsymbol{x}}\,[n-1|n]$ can be obtained from a sub-vector of the state vector $\underline{\hat{\boldsymbol{x}}}\,[n|n]$ from a standard Kalman filter. If the Kalman filter uses an AR model of order $K \geq 2$, this does not require any more computation and the

MSE of $\hat{\boldsymbol{x}}[n-1|n]$ is less than that $\hat{\boldsymbol{x}}[n-1|n-1]$ [100]. The addtional delay is a transmission block length, which is tolerate compared to the interleaver block length.

Now we analyze the computational cost of channel estimation. For a given SNR and f_d, Kalman filter calculation requires the inversion of the square covariance matrix of order $(Q+1)(L+1)$ requiring $\mathcal{O}([(Q+1)(L+1)]^3)$ operations. Comparing with LMMSE estimation discussed in Appendix D, it can be seen that Kalman filter calculation has the same cost as the computation of the LMMSE coefficient estimation within one OFDM symbol block. By taking the advantage of second-order Kalman smoother, it will be seen in Section 4.4 that the Kalman filter provides superior channel estimation error to LMMSE estimation.

4.3 Wiener Bound

For any estimation procedure, it is useful to know what the performance of the best possible estimation procedure of that type is in the sense of MSE. This can be difficult to derive for discrete time estimation procedures, in particular for bandlimited processes because the optimal filter is neither causal or of finite memory [109]. However, it is fairly simple to calculate a bound on the estimation error for the continuous time channel estimation problem via standard Wiener filter theory. Based on sampling theory, estimation bounds for continuous time can be converted to the estimation bounds for discrete time. In this section, we derive the MSE for the continuous time Wiener filter to estimate the radio channel. Since the channel fading process is bandlimited, it satisfies the Nyquist sampling theorem perfectly and the channel can be perfectly reconstructed from its samples. Therefore, there exists an infinite memory discrete time system that can perfectly reproduce the results of the continuous time channel estimation system.

In the continuous time domain, the estimated channel error is $\varepsilon(t) = h(t) - \hat{h}(t)$. According to [109], the MSE of the optimal filter is

$$E[\varepsilon^2(t)]_{min} = \int_{-\infty}^{\infty} \frac{S_H(f)S_v(f)}{S_H(f) + S_v(f)} df \tag{4.8}$$

where $S_H(f)$ is the Doppler power spectrum, given by Equation (A.4) in Appendix A. $S_v(f)$ is the power spectrum of the noise, which is $\bar{N}_0$. Due to insertion of pilot symbols, $\bar{N}_0$ becomes

$$\bar{N}_0 = N_0 \frac{f_s}{f_p} \tag{4.9}$$

where N_0 is the power density of AWGN at the input of the receiver, f_s is the sample rate, and f_p is the pilot rate. Therefore, the noise density $\bar{N}_0$ is proportional to f_s. It is not possible to reduce $\bar{N}_0$ by adding a low pass filter to prevent aliasing prior to the channel estimation filter since this will destroy the time separability of the data signals and pilot signals. If N_0 is fixed and more pilot signals are sent, increasing f_p, the Wiener bound is lowered indicating better channel estimation is possible. The drawback is that the bandwidth efficiency of the radio channel decreases when more pilot signals are used for channel estimations.

Since the Wiener filter is the optimal filter in terms of MSE, the Wiener filter error gives us the lower bound on the MSE for all possible pilot-based channel estimation procedures for a given pilot symbol density and SNR. We will show in the next section that our new Kalman smoother comes close to Wiener filter MSE.

4.4 Simulation Results

We consider an OFDM system with $N = 128$ subcarriers and $N_{cp} = N/8$ CP for each OFDM symbol. Each transmission block includes $K_b = 10$ OFDM symbols. We

choose $Q = 4$, and $P = 5$ pilot symbols are inserted into each transmission block. We simulate a one-path Rayleigh fading channel with the carrier frequency $f_c = 3.5GHz$, the sampling frequency $f_s = 1.4MHz$, and the symbol duration $T = 91.4\mu s$ [18]. The system is designed for velocities $v = 168.8km/h$ and $v = 337.6km/h$, which result in $f_dT = 0.05$ and $f_dT = 0.1$, respectively. The 'DPS-BEM' and 'Oversampled CE-BEM' are the results from using LS estimators to do channel estimation based on the discrete prolate spheroidal BEM (DPS-BEM) [72] and oversampled complex exponential BEM (CE-BEM) [110], respectively. The 'KL-BEM' is the results where we use MMSE estimator to do channel estimation based on Karhunen-Loève BEM (KL-BEM) [69]. The MSE results of the channel estimation are shown in Figure 4.1 for $f_dT = 0.05$ and Figure 4.2 for $f_dT = 0.1$. The Wiener bound is calculated from Equation (4.8). We see that Kalman channel estimation improves MSE performance. The small gap between the Wiener bound and the Kalman smoother shows that the Kalman smoother improves the MSE performance significantly and provides near optimal performance. It was found in simulation studies that for these Doppler frequencies, increasing the order of the Kalman filter to $K > 2$ did not significantly reduce the channel estimation MSE.

For our BER simulations, we use a DFE with MMSE described in [5], where the ICI power on a subcarrier comes from several neighboring subcarriers. In the simulation ,we assume that the ICI power for each subcarrier focuses on 2 neighboring subcarriers. Quaternary phase-shift keying (QPSK) modulation is used for BER simulation. The BER results for different channel estimations are shown in Figure 4.3 for $f_dT = 0.05$ and Figure 4.4 for $f_dT = 0.1$. The 'Perfect CSI' is the resulting BER for the known CSI at the receiver, and the simulated BERs for different channel estimators are shown in these figures. The results show that both Kalman filter and smoother can improve the BER performance compared to previous channel estimation

methods. It can be seen that for $f_dT = 0.1$ and E_b/N_0 around 30 dB the Kalman smoother gives an improvement of about 1 dB for BER performance than the use of BEM channel estimation alone.

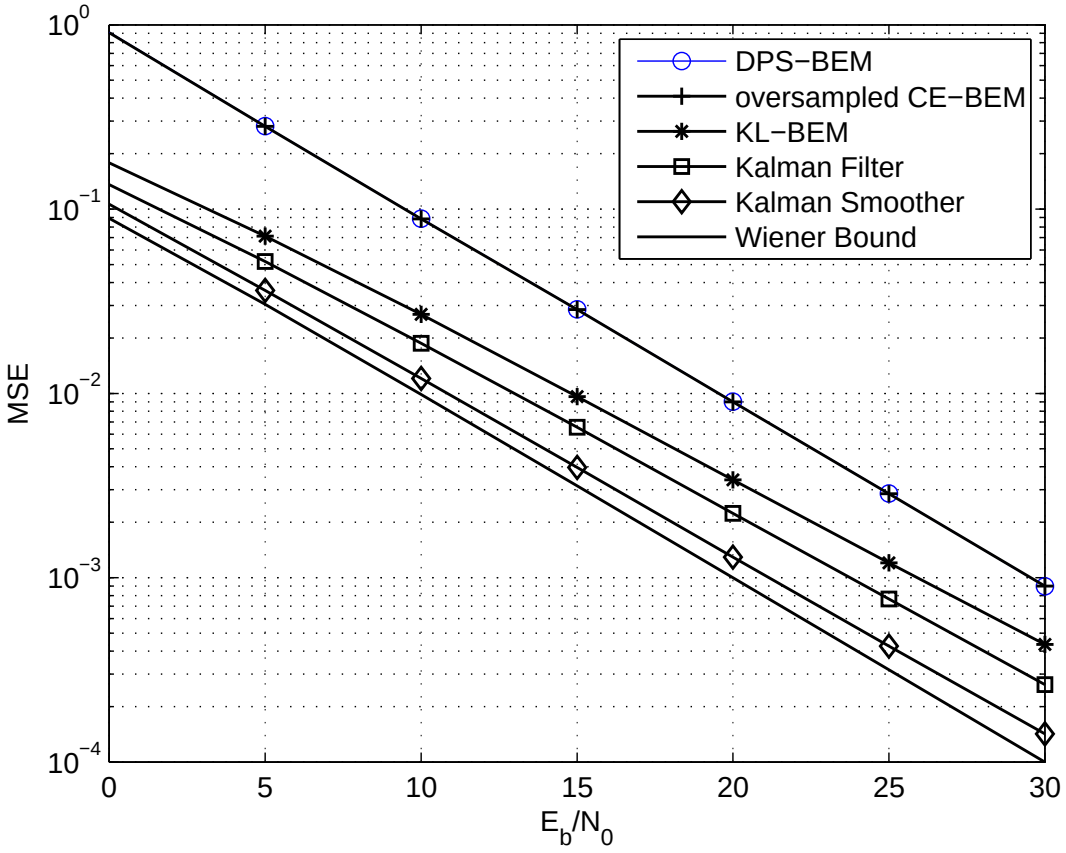

Figure 4.1: Channel estimation MSE for $f_dT = 0.05$

The new Kalman channel estimation is robust when the designed f_dT or SNR used to create the dynamic model and measurement model equations for Kalman filter equations are not same as the true values. Simulation results are shown in Figure 4.5 for the designed $f_dT = 0.05$ when $SNR = 10dB$ and Figure 4.6 for the designed $SNR = 15dB$ when $f_dT = 0.05$. We can see that when the assumed f_dT is lower than the true value, Kalman channel estimation still works well. When the assumed f_dT is higher than the true value, the proposed method will fail. If this

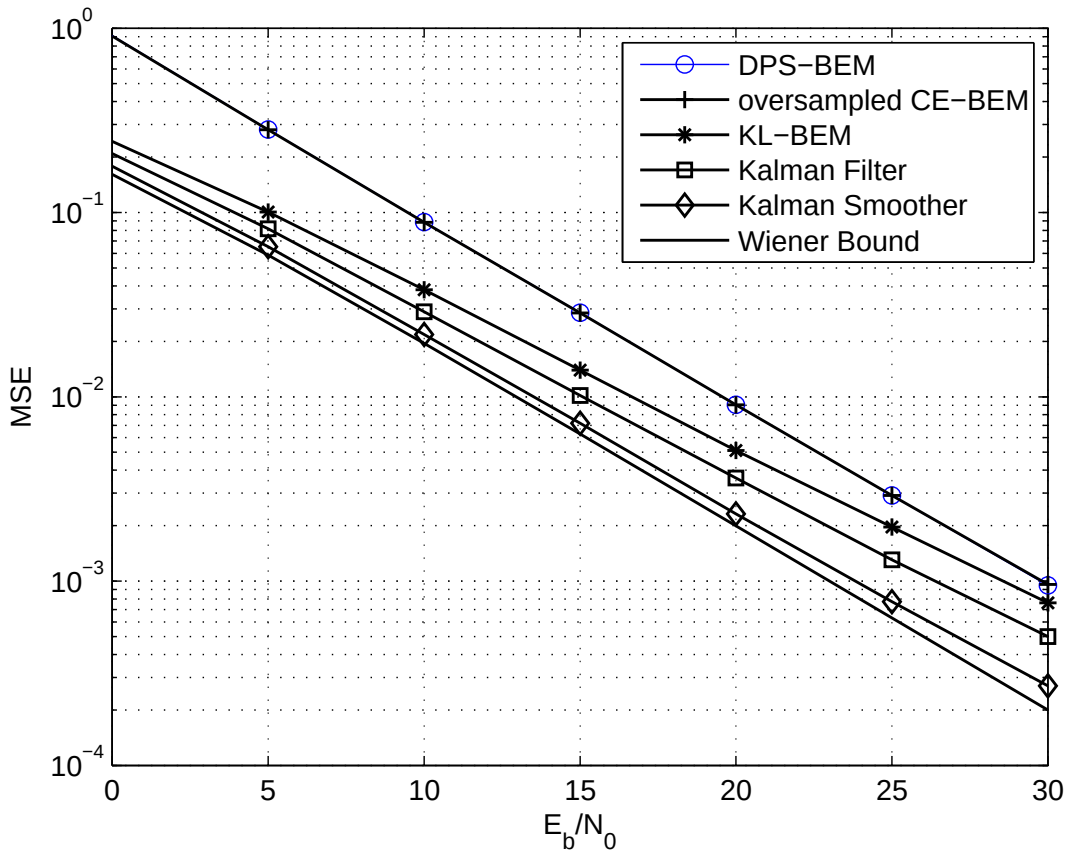

Figure 4.2: Channel estimation MSE for $f_d T = 0.1$

algorithm is used in a radio receiver, a conservative design choice would be to set the assumed Doppler frequency of the Kalman smoother to the highest expected level.

4.5 Conclusions

In this chapter, a time evolution model for basis coefficients in fast fading radio channels is introduced. This model allows channel state information measured in adjacent time intervals to be combined together to improve the overall channel estimation accuracy. A Kalman filter algorithm based on the derived dynamic model for basis coefficients was introduced in this chapter. Simulation results show that Kalman fil-

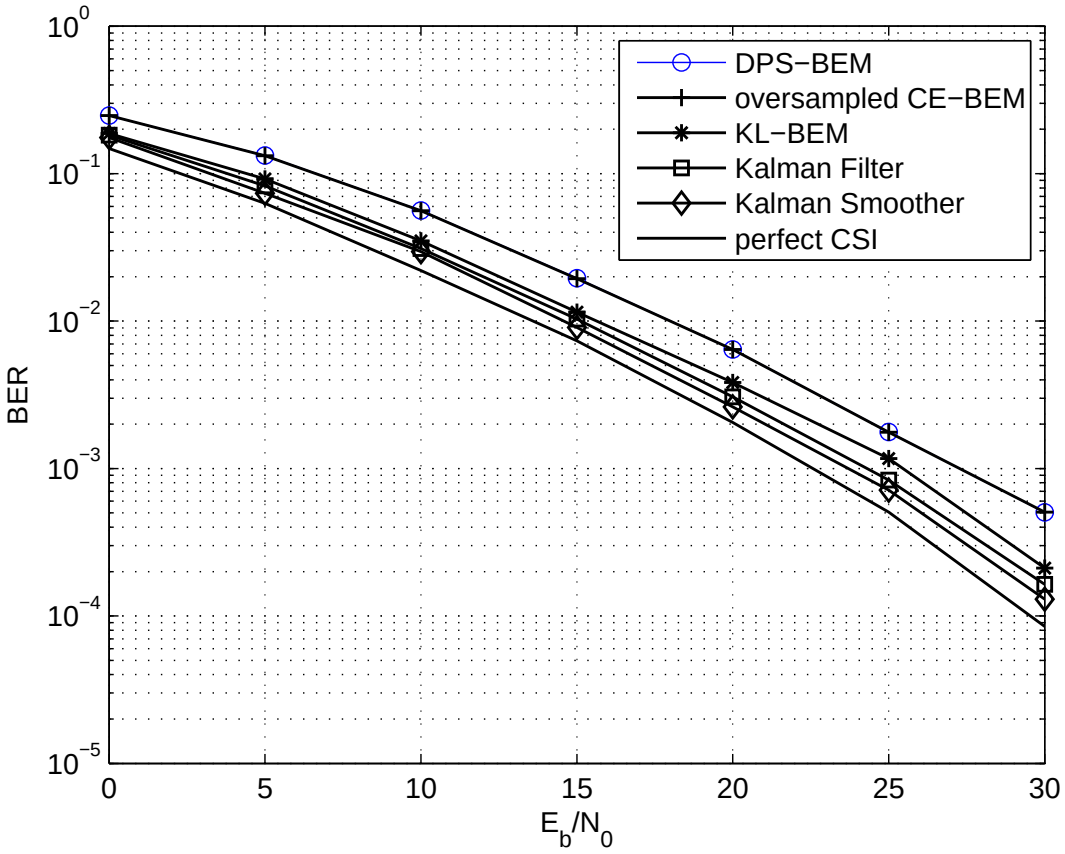

Figure 4.3: Channel estimation BER for $f_dT = 0.05$

ter/smoother improves the MSE and BER performance and is robust to changes of Doppler frequency and SNR levels from the design values.

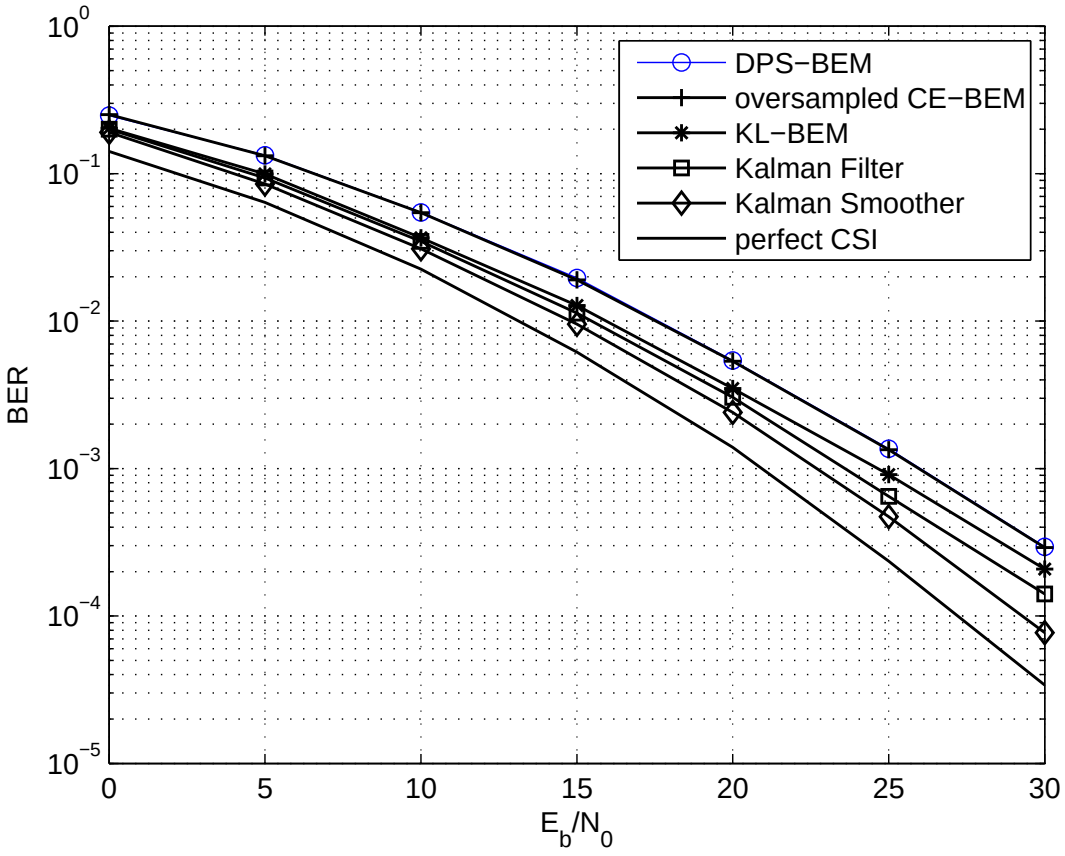

Figure 4.4: BER resulting from estimated channel state for $f_dT = 0.1$

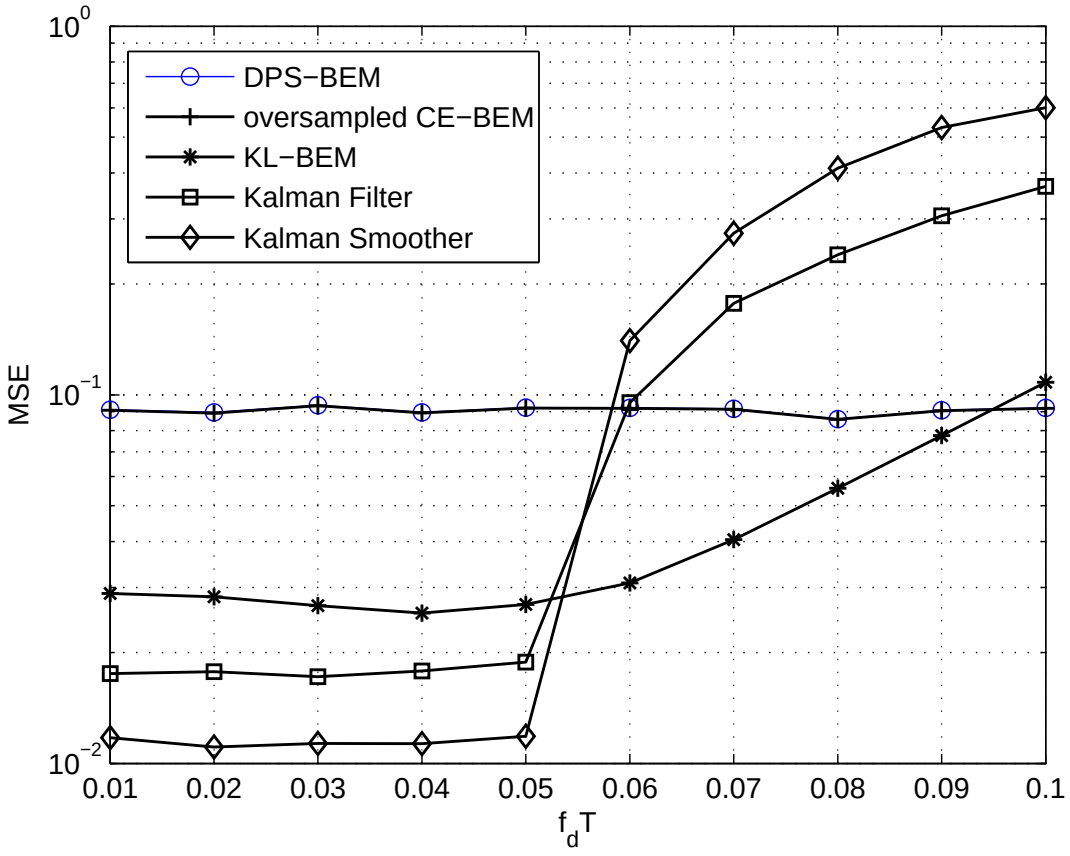

Figure 4.5: Channel MSE for $SNR = 10$ dB and designed $f_dT = 0.05$

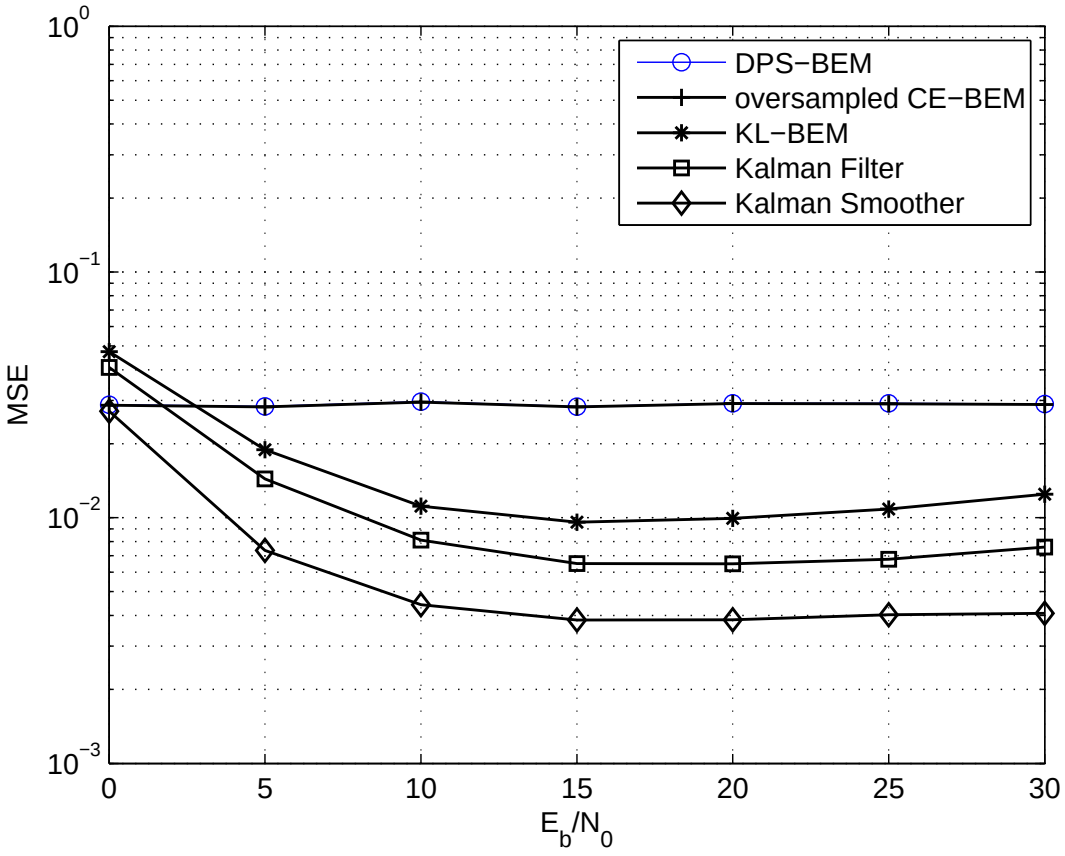

Figure 4.6: Channel MSE for $f_dT = 0.05$ and designed $SNR = 15$ dB

Chapter 5

Joint Channel Estimation and Data Detection

In Chapter 4, the design of a second-order Kalman filtering method for fast fading channels based on pilot symbol aided modulation (PSAM) was introduced. However, as the fading rate increases, more pilot symbols are required to obtain sufficiently accurate channel estimates for reliable data detection. This reduces the overall data rate during fast fading to unacceptable levels for many applications. To improve the performance of PSAM, it is necessary to exploit iterative channel estimation and symbol detection approach. In this method, the detected data is used to estimate the channel state information (CSI) and remove the need for dense pilot signalling for accurate channel estimates.

An iterative receiver combining joint iterative channel estimation with symbol detection for coded orthogonal frequency division multiplexing (OFDM) systems in fast fading channels with Doppler frequencies up to 15% of the OFDM symbol rate is considered in this chapter. The initial rough estimate of the channel using the pilot signals provides a low enough error rate for the detected data from the first

iteration to be used as a 'virtual' pilot signal which is dense in the time domain. Data is detected and decoded before the channel is estimated again based on the detected data. Traditional decision directed channel estimation is sensitive to errors in previously detected data which can cause elevated error levels in data detected at later times, this is known as error propagation. In order to reduce error propagation due to decision directed channel estimation, the initial values for Kalman filtering used for the decision directed channel estimation for each transmission block are only from the channel estimation values obtained from pilot based channel estimation of the previous transmission blocks. The receiver exchanges information between the channel estimator and detector in an iterative fashion to obtain accurate estimates of the CSI, and furthermore, the performance of the detection is improved. Therefore, the proposed approach does not suffer from error propagation and the complexity of decision directed channel estimation is comparable to that of PSAM while reducing the pilot-to-data ratio.

We begin with the design of iterative receivers with perfect CSI. Then, we treat the subject of iterative receiver design for coded OFDM systems in fast fading channels.

5.1 Introduction

In practice, communication systems utilize error correction coding to provide reliable delivery on fading channels [2]. When the coded signal is transmitted, the soft information, typically represented in terms of the log-likelihood ratio (LLR), is exchanged between the detector and the decoder in an iterative fashion [111]. The iterative processing principle, called turbo-decoding or detection, was originally proposed for channel decoding [76, 77]. The iterative processing has received considerable attention due to the improvement of the receiver's performance. The use of turbo-processing

for equalization and detection in fast fading channels with known CSI has been shown to provide good performance [112, 113]. To help with later discussion, we focus our attention on the conventional iterative structure for time-varying known channels in this section.

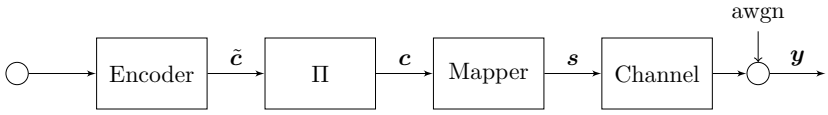

Figure 5.1: Block diagram of the transmitter

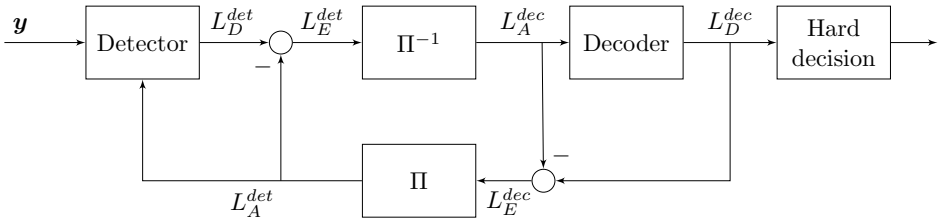

Figure 5.2: Block diagram of the receiver

We consider a coded communication system depicted in Figure 5.1 for the transmitter structure and Figure 5.2 for the iterative receiver structure, where Π and Π^{-1} are denoted the interleaver and the deinterleaver, respectively. At the transmitter, the information bits are first encoded as $\tilde{\boldsymbol{c}}$ and then interleaved into coded bits $\boldsymbol{c} = \Pi(\tilde{\boldsymbol{c}})$, which are mapped to $M-$ary complexity-valued code symbols $\boldsymbol{s}$. The encoded data $\tilde{\boldsymbol{c}}$ is reordered by the interleaver and transmitted over the channel. At the receiver, the deinterleaver puts the LLR in proper sequence and passes it to the decoder.

The *a priori* LLR for the bit c_k is given by

$$L_A(c_k) = \ln \frac{P(c_k = +1)}{P(c_k = -1)}. \tag{5.1}$$

where $P(c_k = +1)$ and $P(c_k = -1)$ are probabilities for bit $c_k = +1$ and $c_k = -1$, respectively. Given the received signal $\boldsymbol{y}$, the *a posteriori* LLR of the coded bit c_k is

expressed as

$$L_D(c_k) = \ln \frac{P(c_k = +1|\boldsymbol{y})}{P(c_k = -1|\boldsymbol{y})}. \tag{5.2}$$

With perfect CSI, the general iterative receiver consists of a soft-input soft-output (SISO) detector, a SISO decoder, a bit interleaver and a deinterleaver as shown in Figure 5.2. The SISO detector takes the received signal $\boldsymbol{y}$ and the *a priori* LLR, L_A^{det}, and outputs the *a posteriori* LLR, L_D^{det}. In most cases, for the initial detection step, no the *a priori* information is available and hence $L_A^{det} = 0$. The extrinsic information of the detector $L_E^{det} = L_D^{det} - L_A^{det}$ is passed through the bit-deinterleaver to become the *a priori* LLR L_A^{dec} of the decoder. The output of decoder, the *a posteriori* LLR L_D^{dec}, is then passed through the hard decision device and the estimated bits are obtained. The extrinsic information $L_E^{dec} = L_D^{dec} - L_A^{dec}$ is fed back to the detector and becomes the *a priori* input after bit-interleaving (Π). The detector uses this soft information to obtain more accurate soft output L_D^{det}, which is then passed as L_E^{det} to decoder for further iterations. This cycle is repeated and more reliable values are exchanged between the detector and decoder [113]. Therefore, the BER performance of the receiver is improved.

The maximum *a posteriori* (MAP) algorithm proposed by Bahl *et al.* [114], called the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm, is used to estimate the *a posteriori* probabilities (APPs) for each bit. With Bayes' rule, the *a posteriori* LLR of the coded bit c_k in Equation (5.2) is rewritten as

$$\begin{aligned} L_D(c_k|\boldsymbol{y}) &= \ln \frac{p(\boldsymbol{y}|c_k = +1)P(c_k = +1)/p(\boldsymbol{y})}{p(\boldsymbol{y}|c_k = -1)P(c_k = -1)/p(\boldsymbol{y})} \\ &= \ln \frac{P(c_k = +1)}{P(c_k = -1)} + \ln \frac{p(\boldsymbol{y}|c_k = +1)}{p(\boldsymbol{y}|c_k = -1)} \\ &= L_A(c_k) + \ln \frac{p(\boldsymbol{y}|c_k = +1)}{p(\boldsymbol{y}|c_k = -1)}. \end{aligned} \tag{5.3}$$

To obtain the $p(\boldsymbol{y}|c)$, we sum over all symbols with the specified bit value: $\mathbb{C}_{k,\pm 1} =$

$\{\boldsymbol{c}|c_k = \pm 1\}$,

$$p\left(\boldsymbol{y}|c_k = \pm 1\right) = \sum_{\mathbb{C}_{k,\pm 1}} p(\boldsymbol{y}|\boldsymbol{c}) P(\boldsymbol{c}|c_k). \tag{5.4}$$

Equation (5.3) is modified as

$$L_D(c_k|\boldsymbol{y}) = L_A(c_k) + \ln \frac{\sum_{\mathbb{C}_{k,+1}} p(\boldsymbol{y}|\boldsymbol{c}) P(\boldsymbol{c}|c_k)}{\sum_{\mathbb{C}_{k,-1}} p(\boldsymbol{y}|\boldsymbol{c}) P(\boldsymbol{c}|c_k)}. \tag{5.5}$$

Using the max-log-MAP approximation in [115], that is,

$$\ln(e^{\delta_1} + \cdots + e^{\delta_n}) \approx \max(\delta_1, ..., \delta_n) \tag{5.6}$$

Equation (5.5) is then simplified as

$$\begin{aligned} L_D(c_k|\boldsymbol{y}) \approx L_A(c_k) &+ \max_{\mathbb{C}_{k,+1}} \left\{\ln p(\boldsymbol{y}|\boldsymbol{c}) + \ln P(\boldsymbol{c}|c_k)\right\} \\ &- \max_{\mathbb{C}_{k,-1}} \left\{\ln p(\boldsymbol{y}|\boldsymbol{c}) + \ln P(\boldsymbol{c}|c_k)\right\} \end{aligned} \tag{5.7}$$

Since (5.7) is derived based on (5.6), only the maximum likelihood is considered when calculating this probability. Therefore, the value $L_D(c_k|\boldsymbol{y})$ actually gives the probability of the most likely path through the trellis [113].

In general, the complexity of the optimal BCJR algorithm is proportional to M^L, and grows exponentially with the channel order, L. In order to reduce the complexity of detection, we propose a low-complexity detection below.

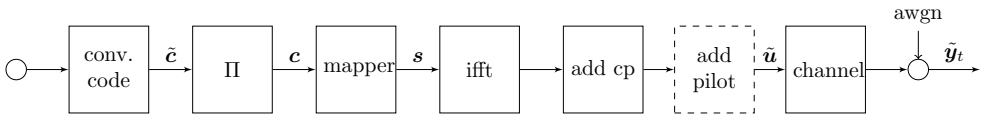

Figure 5.3: Coded OFDM transmitter

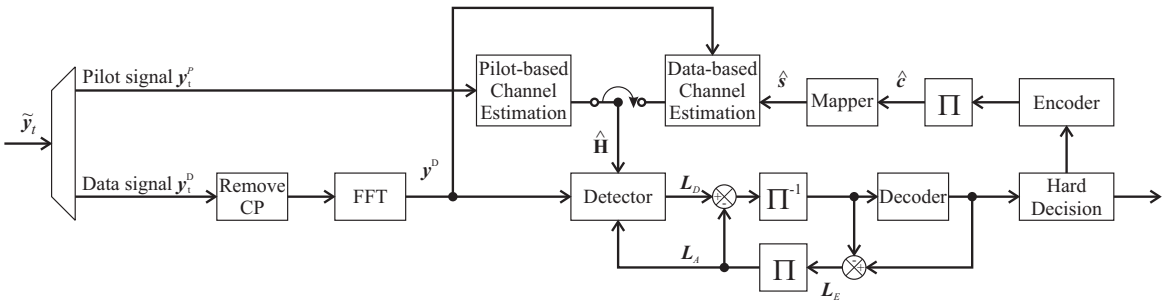

Figure 5.4: Coded OFDM system

5.2 Iterative receivers for OFDM systems in Fast fading

Consider the coded OFDM system depicted in Figures 5.3-5.4. At the receiver, using the initial channel estimates from pilot symbols, data is detected and then decoded. The decoded data is then fed back into the data-based channel estimation block for the estimation of the channel coefficients in the next iteration. The channel estimation, data detection, and data decoding are repeated until convergence is reached or the maximum allowable number of iterations has been performed.

The proposed iterative receiver uses three different time intervals as shown in Fig. 3.1: the OFDM block (OB) is used as the basis for detection, the transmission block (TB) for channel estimation, and the interleaving block for decoding. The channel estimation for an interleaving block is first performed by using the pilot symbols and then the detected data in subsequent iterations using a first order Kalman filter iterating over all transmission blocks, which is described in detail in Section 5.2.1. The data detection (based on OFDM blocks) and decoder (based on interleaving

blocks) are described in Section 5.2.2.

5.2.1 Iterative Channel Estimation

Channel estimation is based on a transmission block, which includes K_b OFDM blocks. The state model of basis coefficients is characterized as a first-order AR model. From the description in Section 3.3, the dynamic model of the first-order for the evolution of the basis coefficient vector from transmission block $n-1$ to n is given by

$$\boldsymbol{x}[n] = \boldsymbol{\Phi}\boldsymbol{x}[n-1] + \boldsymbol{w}[n-1] \tag{5.8}$$

where $\boldsymbol{x}[n]$ is the state vector at transmission block n, the state transition matrix is given as $\boldsymbol{\Phi} = -\oplus_{l=0}^{L} \boldsymbol{A}_l$ where $\oplus$ is the direct sum of matrices, and $\boldsymbol{w}[n]$ is an AWGN vector with the covariance matrix $\boldsymbol{Q} = \oplus_{l=0}^{L} \boldsymbol{Q}_{lf}$, $\boldsymbol{Q}_{lf}$ is the covariance matrix of the state vector for the lth path.

For the initial channel estimation of a transmission block, the pilot measurements in the time domain $\boldsymbol{y}^P$ are used to estimate the basis coefficients. Similar to Equation (4.6), we express the measurement model at transmission block n as

$$\boldsymbol{y}_t^P[n] = \boldsymbol{C}^P[n]\boldsymbol{x}[n] + \boldsymbol{v}_t^P[n] \tag{5.9}$$

where $\boldsymbol{y}_t^P[n]$ is the received signal corresponding to pilot symbols at transmission block n, $\boldsymbol{C}^P[n]$ is the matrix corresponding to basis functions corresponding to pilot symbols at transmission block n, $\boldsymbol{v}_t^P[n]$ is an AWGN vector corresponding to pilot symbols at transmission block n.

Since the size of unknown state vector $\boldsymbol{x}[n]$ is $(Q+1)(L+1)$ by 1, we choose $K_p \geq (Q+1)$ to be able to get a minimum of one measurement per channel basis coefficient. The tradeoff between the complexity and the performance exists in the

iterative scheme. From the point of minimizing the total channel of mean square error (MSE), we select $K_p > (Q+1)$. In such a case, (5.9) is overdetermined, and the modified equation is given by

$$\tilde{\boldsymbol{y}}_P[n] = \tilde{\boldsymbol{C}}_P[n]\boldsymbol{x}[n] + \tilde{\boldsymbol{v}}_P[n], \tag{5.10}$$

where $\tilde{\boldsymbol{y}}_P = (\boldsymbol{C}^P)^H \boldsymbol{y}^P$, $\tilde{\boldsymbol{C}}_P = (\boldsymbol{C}^P)^H \boldsymbol{C}^P$, $\tilde{\boldsymbol{v}}_P = (\boldsymbol{C}^P)^H \boldsymbol{v}_t^P$ with the variance matrix $\boldsymbol{R}_P = \sigma_v^2 (\boldsymbol{C}^P)^H \boldsymbol{C}^P$. Using Equations (5.8) and (5.10), the basis coefficients are estimated with a Kalman filter for the initial iteration.

The initial conditions of the state vector, $\hat{\boldsymbol{x}}[0]$ and error covariance matrix $\boldsymbol{P}[0]$ are taken from the estimated state vector and error covariance matrix for the last transmission block of the preceding interleaving block estimated using the pilot signals only. Using the pilot signals for initial conditions of the Kalman filters prevents error propagation from the decision feedback propagating beyond a single interleaving block. At system initialization, $\hat{\boldsymbol{x}}[0] = \boldsymbol{0}$ and $\boldsymbol{P}[0] = \oplus_{i=0}^{L} \sigma_l^2 \boldsymbol{R}_{\boldsymbol{x}}[0]$, where $\boldsymbol{R}_{\boldsymbol{x}}[i] = E[\boldsymbol{x}_l[n+i]\boldsymbol{x}_l^H[n]]$ and $\boldsymbol{x}_l[n]$ is the basis vector for the lth path defined in Appendix B. From the estimated channel coefficients for a given transmission block, $\hat{\boldsymbol{x}}[n]$, the estimated channel matrix for the kth OFDM symbol is calculated as

$$\hat{\boldsymbol{H}}_k = \boldsymbol{F} \sum_{l=0}^{L} \boldsymbol{\mathcal{D}}(\boldsymbol{E}_k \hat{\boldsymbol{x}}_l) \boldsymbol{S}_l \boldsymbol{F}^H, \quad k \in [0, K_b - 1]. \tag{5.11}$$

where $\mathcal{D}(\boldsymbol{E}_k \hat{\boldsymbol{x}}_l)$ denotes a diagonal matrix with vector $\boldsymbol{E}_k \hat{\boldsymbol{x}}_l$ on its diagonal and all other entries being zero, $\boldsymbol{S}_l$ is given by circularly shifting identity matrix $\boldsymbol{I}_N$ down with the delay of l samples. A circular shift of a matrix is the operation of rearranging the matrix, circularly shifting rows down (upper) or column left (right). $\boldsymbol{F}$ is the Fourier transform matrix.

After the initial channel estimation, the measurements in the frequency domain

$\tilde{\boldsymbol{y}}^D$, expressed as Equation (3.19), taken from the detected symbols are used to calculate basis coefficients. Hence, the measurement model at transmission block n is

$$\tilde{\boldsymbol{y}}^D[n] = \hat{\boldsymbol{\mathcal{M}}}_D[n]\boldsymbol{x}[n] + \boldsymbol{v}^D[n] \tag{5.12}$$

where $\tilde{\boldsymbol{y}}^D[n]$ is the received signal vector at transmission block n corresponding to detected data, $\hat{\boldsymbol{\mathcal{M}}}_D[n]$ is the block matrix corresponding to detected data at transmission block n denoted as $\hat{\boldsymbol{\mathcal{M}}}_D = [\hat{\boldsymbol{\mathcal{M}}}_0^T, ..., \hat{\boldsymbol{\mathcal{M}}}_{K_p-1}^T]^T$, where $\hat{\boldsymbol{\mathcal{M}}}_k = [\hat{\boldsymbol{M}}_{0k} \cdots \hat{\boldsymbol{M}}_{Lk}]$, $\hat{\boldsymbol{M}}_{lk} = \boldsymbol{F}\boldsymbol{\mathcal{E}}_{lk}$, where $\boldsymbol{\mathcal{E}}_{lk} = \boldsymbol{\mathcal{D}}(\boldsymbol{S}_l\boldsymbol{F}^H\hat{\boldsymbol{s}})\boldsymbol{E}_k$ and $\hat{\boldsymbol{s}}$ is the detected OFDM symbol; $\boldsymbol{v}^D$ is an AWGN vector.

Similarly, when (5.12) is overdetermined, a modified measurement equation, given below, is used:

$$\tilde{\boldsymbol{y}}_D[n] = \tilde{\boldsymbol{\mathcal{M}}}_D\boldsymbol{x}[n] + \tilde{\boldsymbol{v}}_D[n] \tag{5.13}$$

where $\tilde{\boldsymbol{y}}_D = \hat{\boldsymbol{\mathcal{M}}}_D^H\tilde{\boldsymbol{y}}^D$, $\tilde{\boldsymbol{\mathcal{M}}}_D = \hat{\boldsymbol{\mathcal{M}}}_D^H\hat{\boldsymbol{\mathcal{M}}}_D$, $\tilde{\boldsymbol{v}}_D = \hat{\boldsymbol{\mathcal{M}}}_D^H\boldsymbol{v}^D$ with the variance matrix $\boldsymbol{R}_D = \sigma_v^2\hat{\boldsymbol{\mathcal{M}}}_D^H\hat{\boldsymbol{\mathcal{M}}}_D$. For the second and following iterations, the Kalman filter is rerun over all transmission blocks and basis coefficients are re-estimated using Equations (5.8) and (5.13).

We now turn our attention to computational complexity calculations. The most expensive portions of the channel estimation algorithm is the computation of the measurement matrices and the matrix inversion required in the Kalman filter iteration calculations. For the initial channel estimation from pilots, the Kalman filtering calculation requires the inversion of the square covariance matrix of order $(Q+1)(L+1)$ requiring $\mathcal{O}([(Q+1)(L+1)]^3)$ operations. This is of the same order or operations as the standard MMSE coefficient estimation within one transmission block.

For the iterative data-based channel estimation, the major computation is the calculation of the measurement matrix $\hat{\boldsymbol{\mathcal{M}}}_D$, where the operation is evaluated from the

number of element-element multiplications required for the matrix multiplications as well as the FFT operation counts. To support this, calculation of $\boldsymbol{\mathcal{E}}_{lk}$ requires $\mathcal{O}(N(Q+1))$ operations and FFT of each column of $\boldsymbol{\mathcal{E}}_{lk}$ requires $\mathcal{O}(N\log_2(N))$ operations, hence $\hat{\boldsymbol{M}}_{lk}$ requires $\mathcal{O}(N(Q+1)\log_2(N))$ operations. Since calculation of $\hat{\boldsymbol{M}}^H_{l1k1}\hat{\boldsymbol{M}}_{l2k2}$ requires $\mathcal{O}(N(Q+1)^2)$, calculation $\tilde{\boldsymbol{\mathcal{M}}}_D$ for each OFDM block requires $\mathcal{O}(N(Q+1)^2(L+1))$. Therefore, the total computational complexity for data-based channel estimation in one OFDM block is $\mathcal{O}(N(Q+1)^2(L+1))$. In the next subsection, it will be shown that this is of the same order as detection, indicating that this channel estimation method, while more expensive than purely pilot-based approaches, will not require an increase in the order of the number of operations for the channel estimation and symbol detection.

5.2.2 Symbol Detection

As mentioned before, the complexity of the optimal detection is very high. In order to reduce the complexity of data detection, we use a suboptimal detection scheme in this section.

We express the estimated frequency matrix $\hat{\boldsymbol{H}}_k$ of Equation (5.11) for a single OFDM block as $\hat{\boldsymbol{H}}_k = \hat{\boldsymbol{H}}_{k1} + \hat{\boldsymbol{H}}_{k2}$, where $\hat{\boldsymbol{H}}_{k1}$ is the diagonal matrix which has its only non-zero values being the main diagonal elements of $\hat{\boldsymbol{H}}_k$, $\hat{\boldsymbol{H}}_{k2}$ is equal to $\hat{\boldsymbol{H}}_k$ with the exception of its main diagonal elements being zero [116].

For the initial detection, symbols are detected by ignoring ICI, using only the elements on the main diagonal of the estimated channel transfer matrix, $\hat{\boldsymbol{H}}^1_{k1}$, and the received signal $\boldsymbol{y}_k$. To suppress ICI efficiently, we estimate the ICI from the estimated symbol $\hat{\boldsymbol{s}}^{i-1}_k$ and remove the estimated ICI from the received signal $\boldsymbol{y}_k$, that is,

$$\boldsymbol{y}^i_k = \boldsymbol{y}_k - \hat{\boldsymbol{H}}^i_{k2}\hat{\boldsymbol{s}}^{i-1}_k \tag{5.14}$$

where $\hat{\boldsymbol{s}}_k^i$ is the symbol estimate of the ith iteration, $\hat{\boldsymbol{H}}_{k2}^i$ is the estimated $\boldsymbol{H}_{k2}$ of the ith iteration, and $\hat{\boldsymbol{s}}_k^0 = \mathbf{0}$ for the first iteration. The values $\boldsymbol{y}_k^i$ and $\hat{\boldsymbol{H}}_{k1}^i$ are used to estimate the symbol $\boldsymbol{s}_k^i$ for the ith iteration.

The detector and the decoder exchange the extrinsic information iteratively in each detection scheme. In each detection loop, the detector uses the estimated channel $\hat{\boldsymbol{H}}_k^i$, the received signal $\boldsymbol{y}_k^i$ and the *a priori* LLR, $L_A^{det}(c)$, and then the *a posteriori* LLR, $L_D^{det}(c|\boldsymbol{y}^i)$ is obtained by

$$L_D^{det}(c_k|\boldsymbol{y}^i) = \ln \frac{P[c_k = +1|\boldsymbol{y}^i, \hat{\boldsymbol{H}}^i]}{P[c_k = -1|\boldsymbol{y}^i, \hat{\boldsymbol{H}}^i]} \tag{5.15}$$

where $\hat{\boldsymbol{H}}^i = \oplus_{k=0}^{K_b K_t - 1} \boldsymbol{H}_k$ is the estimated channel in the frequency domain over an interleaving block, where K_b is the number of OFDM blocks for each transmission block and K_t is the number of transmission blocks for each interleaver block.

In each detection loop, the detector uses the estimated channel $\hat{\boldsymbol{H}}^1$, the received signal $\boldsymbol{y}^i$ and the *a priori* LLR, $L_A^{det}(c_k)$ supplied by a SISO decoder, to obtain the *a posteriori* LLR [117, 118]. The extrinsic LLR obtained from the decoder, $L_E^{dec}(\tilde{c}_k) = L_D^{dec}(\tilde{c}_k|\boldsymbol{y}^i) - L_A^{dec}(\tilde{c}_k)$, is interleaved and then sent back to the detector as the *a priori* LLR, L_A^{det}. After several iterations, the LLRs pass through a hard decision to get the estimated information bits.

For a single OFDM block, the cost of the SISO detection/decoding is $\mathcal{O}(N \log_2 N)$, with the interference calculation in Equation (5.14) requiring $\mathcal{O}(N^2)$ operations making the overall cost $\mathcal{O}(N^2)$ per OFDM block.

5.3 Simulation Results

In this section, the performance of the proposed algorithm is tested in a simulated wireless OFDM system transmitted over a radio channel. We simulate a 4-path

WSSUS channel with mean propagation powers of 8/15, 4/15, 2/15 and 1/15 for propagation delays of 0 to 3 samples respectively. Each path is subject to independent Rayleigh fading generated using the method from [119] with Jakes' model specifying the autocorrelation function. We consider an OFDM system with $N = 128$ samples for each OFDM symbol. The length of CP is $N_{cp} = N/8$. Gray-coded QPSK and 16 or 64-quadrature amplitude modulation (QAM) constellation are used in simulations. The normalized Doppler frequency is set to $f_dT = 0.1$ and 0.15 respectively. For a WiMAX system operating at the carrier frequency of 5.0 GHz and the subcarrier space is approximately 10 kHz, these Doppler frequencies map to radio terminal velocities of 216 km/h and 324 km/h respectively. Based on the existing standards, the simulations employ the convolutional code with 1/2 rate, constraint length 7 and the generator polynomials $G = [133\ 171]$ in octal form.

5.3.1 BER Performance

The channel is estimated using a first-order Kalman filter. Each transmission frame consists of $K_b = 10$ OFDM blocks, so the length of the frame $M = 1510$ samples. Each interleaving block consists of $K_t = 10$ transmission blocks; each transmission block consists of $K_b = 10$ OFDM blocks and $K_p = 10$ pilot blocks. The pilot to data ratio is low to 7/144 and the percentage of total power used on pilots is 1/145, respectively. We choose $Q = 4$ basis functions for $f_dT = 0.1$ and $Q = 6$ basis functions for $f_dT = 0.15$. Simulation results of coded OFDM are shown in Figures 5.5-5.9 for different normalized Doppler frequencies respectively. 'CE ith' shown in figures represents 'channel estimation ith iteration'. The mean square error (MSE) for the channel estimation is shown in Figure 5.5. In Figure 5.6, the BER results of the proposed channel estimation scheme are compared with [68], where the channel is estimated using MMSE algorithm in the frequency domain and the pilot-to-data

ratio is 20/128. The basis functions used in [68] are Fourier basis functions, which is subject to Gibbs phenomenon resulting in channel modelling error. Specifically, the algorithm in [68] uses a BEM over a single OFDM symbol whereas the new algorithms extends the model over several OFDM symbols. It can be seen in Figure 5.6 that for the new method when $f_dT = 0.1$, the BER results using the estimated channel approach the case of the ideal CSI after the third iteration. For $f_dT = 0.15$, BER results are shown in Figure 5.7.

The proposed algorithm is also tested with the higher order modulation constellations of 16 and 64-QAM as shown in Figures 5.8 and 5.9, respectively. In these cases, it can be seen that performance close to the ideal CSI case is obtained for higher E_b/N_0 values. Moreover, the cheap detection caused by short OFDM blocks allows us to scale up to even higher order modulation if necessary.

5.3.2 Robustness Analysis

For the approach to be useful in field implementations, it is necessary for it to be robust to variations of the true channel parameters from the design values. Robustness to variations of Doppler frequency, number of propagation paths, and mean path power will be demonstrated here. The simulation BER results shown in Fig. 5.10 are for $E_b/N_0 = 8$ dB with the designed normalized Doppler frequency of $f_dT = 0.1$. When the designed f_dT is lower than the true value, the proposed method will diverge. If this algorithm is used in a radio receiver, a conservative design choice would be to set the designed normalized Doppler frequency of the Kalman filter to the highest expected level. In order to reduce the complexity cost of the channel estimation, the designed Doppler frequency should be chosen close to the true values.

The approach is also robust to variations of the number of paths in the channel estimator from the true L. The simulation BER results shown in Fig. 5.11 are for

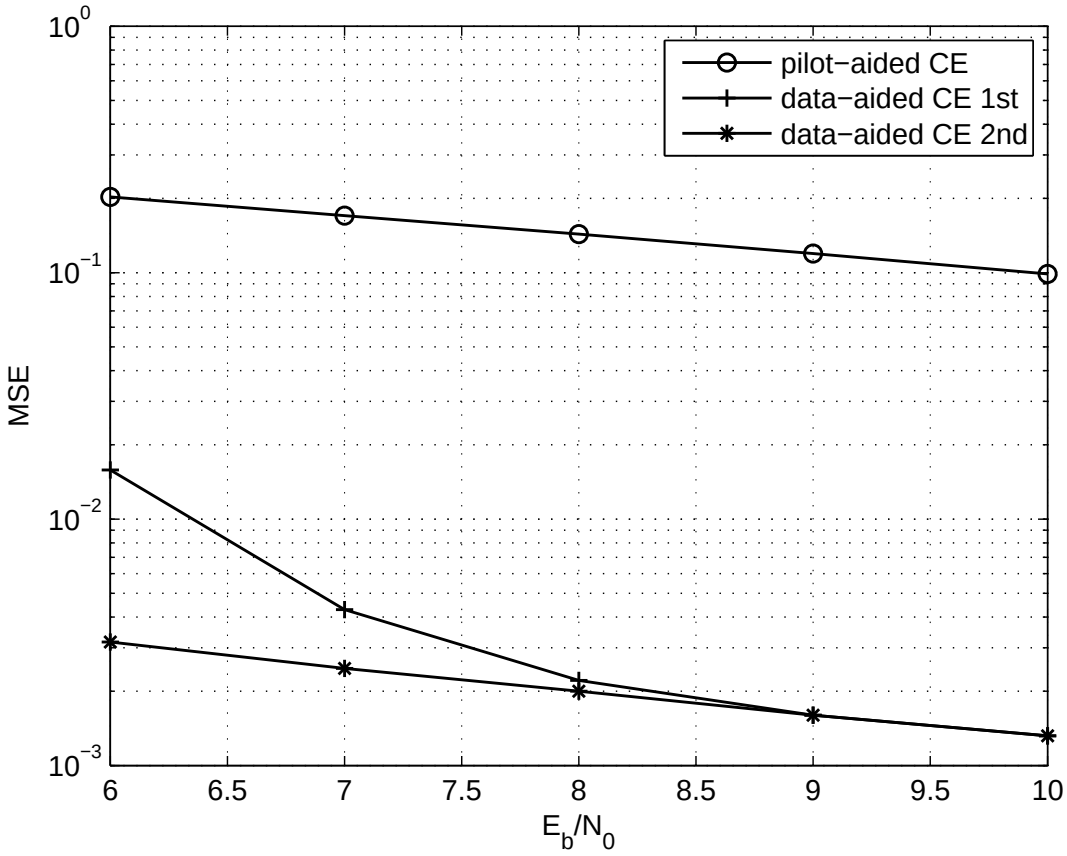

Figure 5.5: MSE resulting from estimated CSI for $f_dT = 0.1$ with QPSK

$E_b/N_0 = 8$ dB with the true $L = 3$. The designed mean power for each path is chosen as 1. When the designed L is lower than the true value, the proposed method will diverge. If this algorithm is used in the receiver, a conservative design choice would be to set the designed number of paths of the Kalman filter to the highest expected level. The channel estimator does not require knowledge of the mean power level of each propagation path. Only the relative delays and number of propagation paths are needed to obtain good channel estimation and data detection/decoding results.

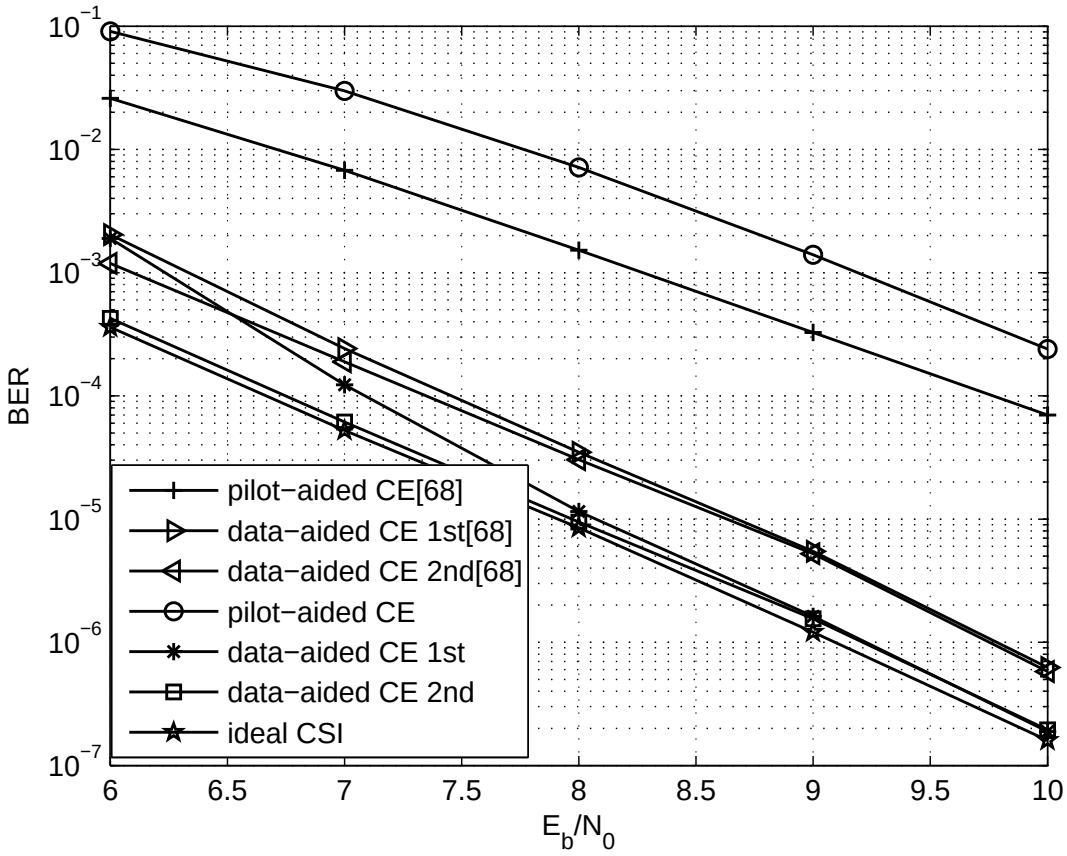

Figure 5.6: BER resulting from estimated CSI for $f_dT = 0.1$ with QPSK

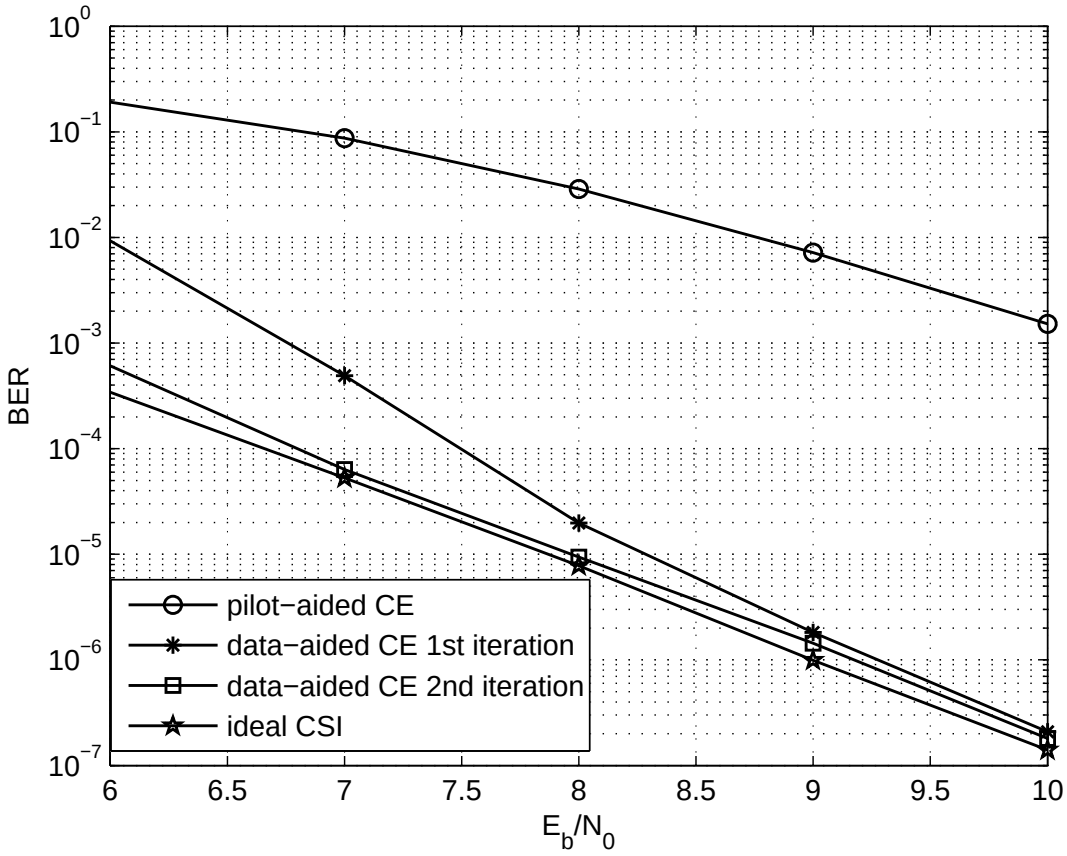

Figure 5.7: BER resulting from estimated CSI for $f_dT = 0.15$ with QPSK

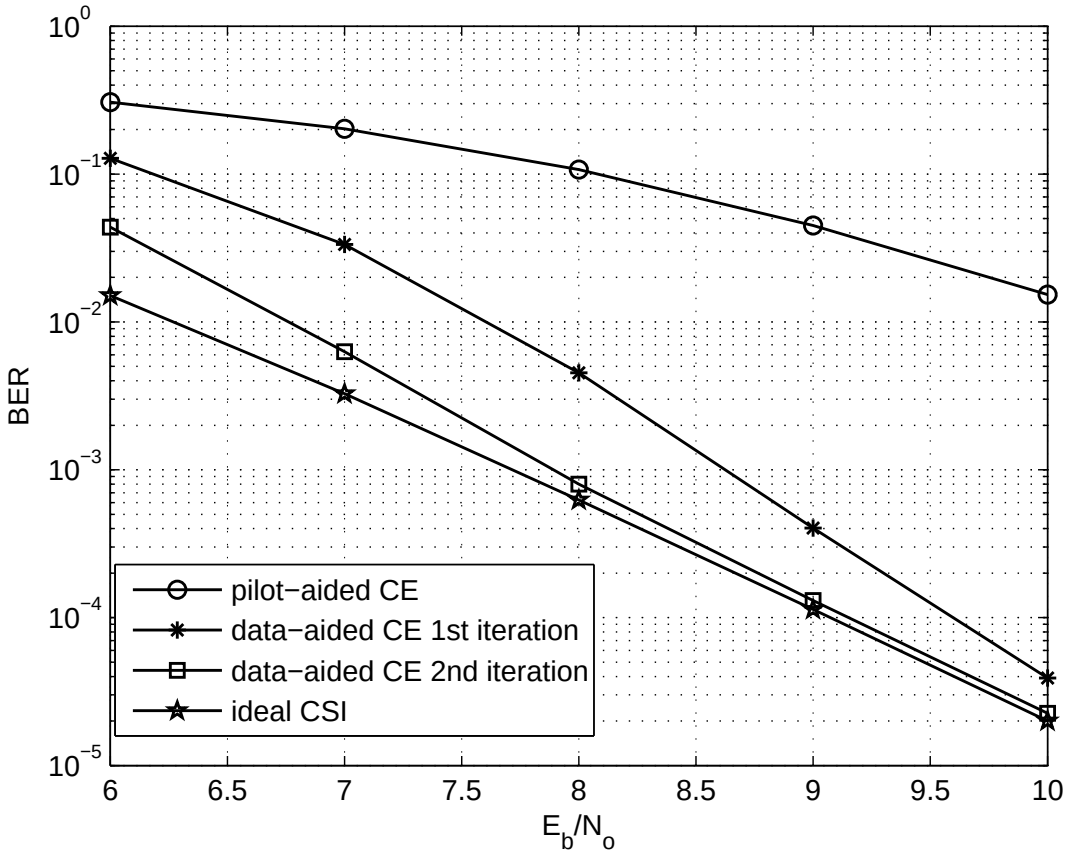

Figure 5.8: BER resulting from estimated CSI for $f_dT = 0.1$ with 16-QAM

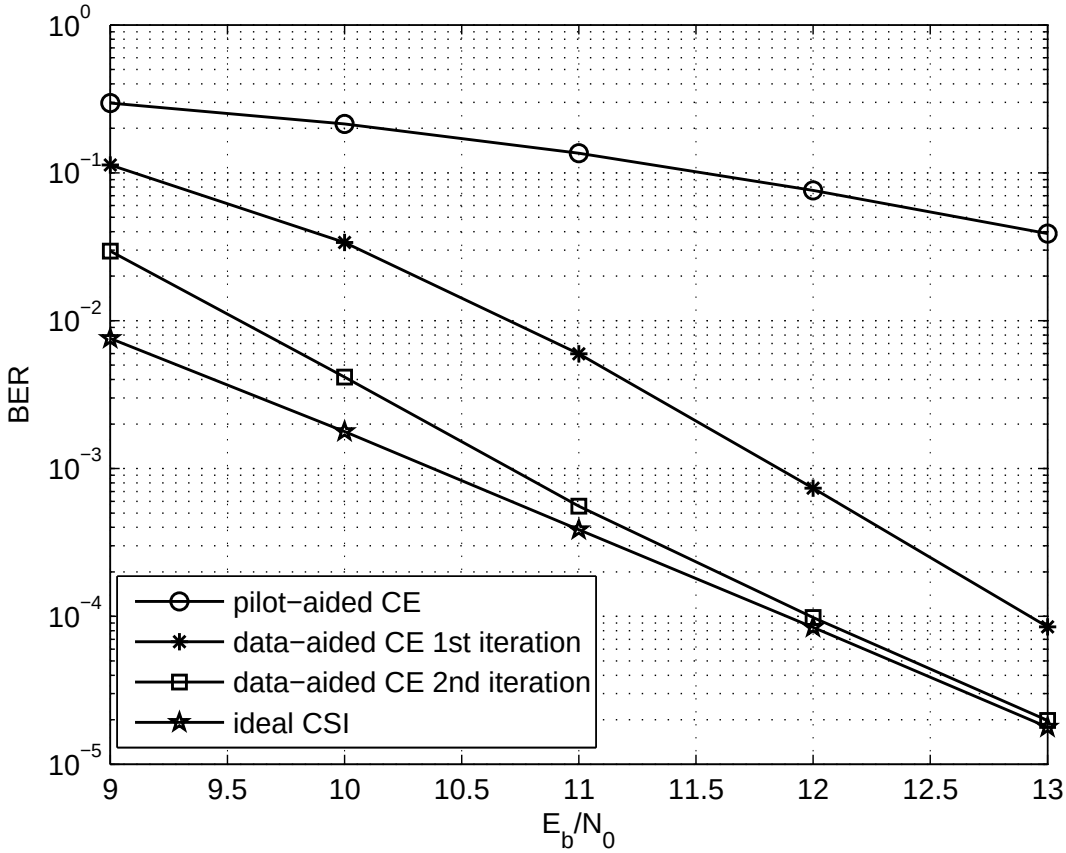

Figure 5.9: BER resulting from estimated CSI with $K_p = 10$ for $f_dT = 0.1$ with 64-QAM

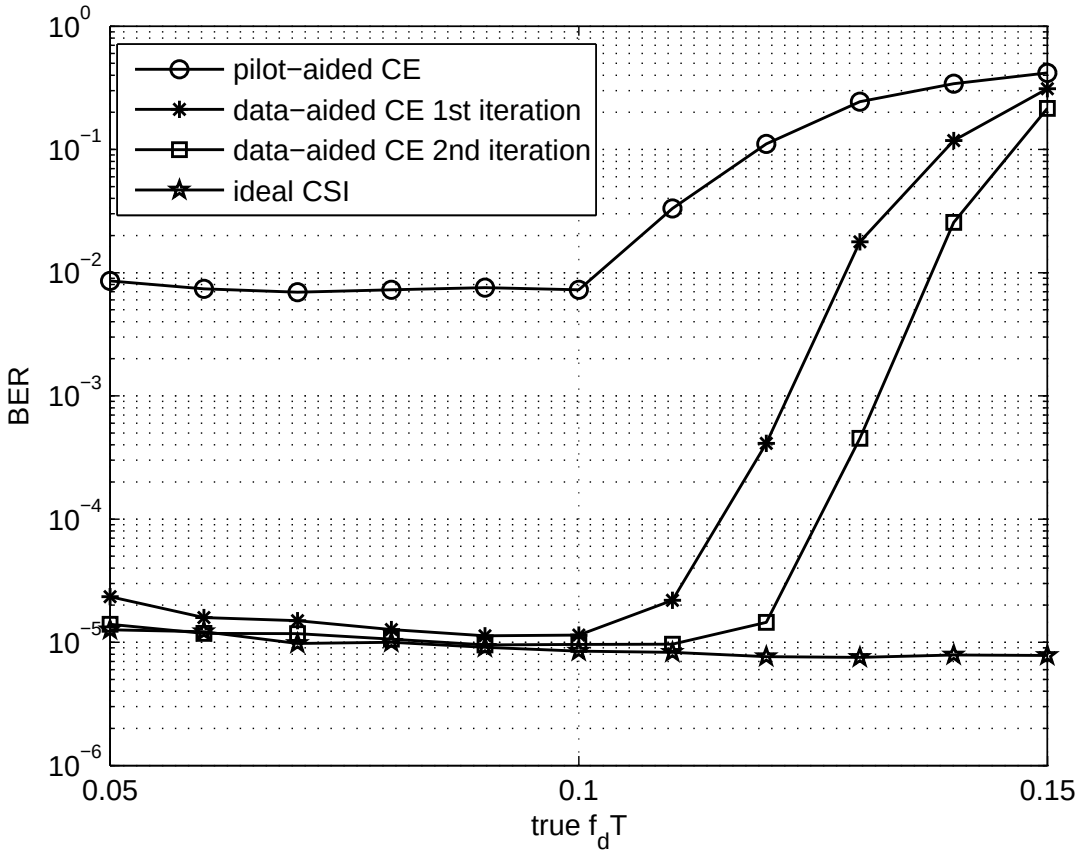

Figure 5.10: BER resulting for designed $f_dT = 0.1$ with QPSK

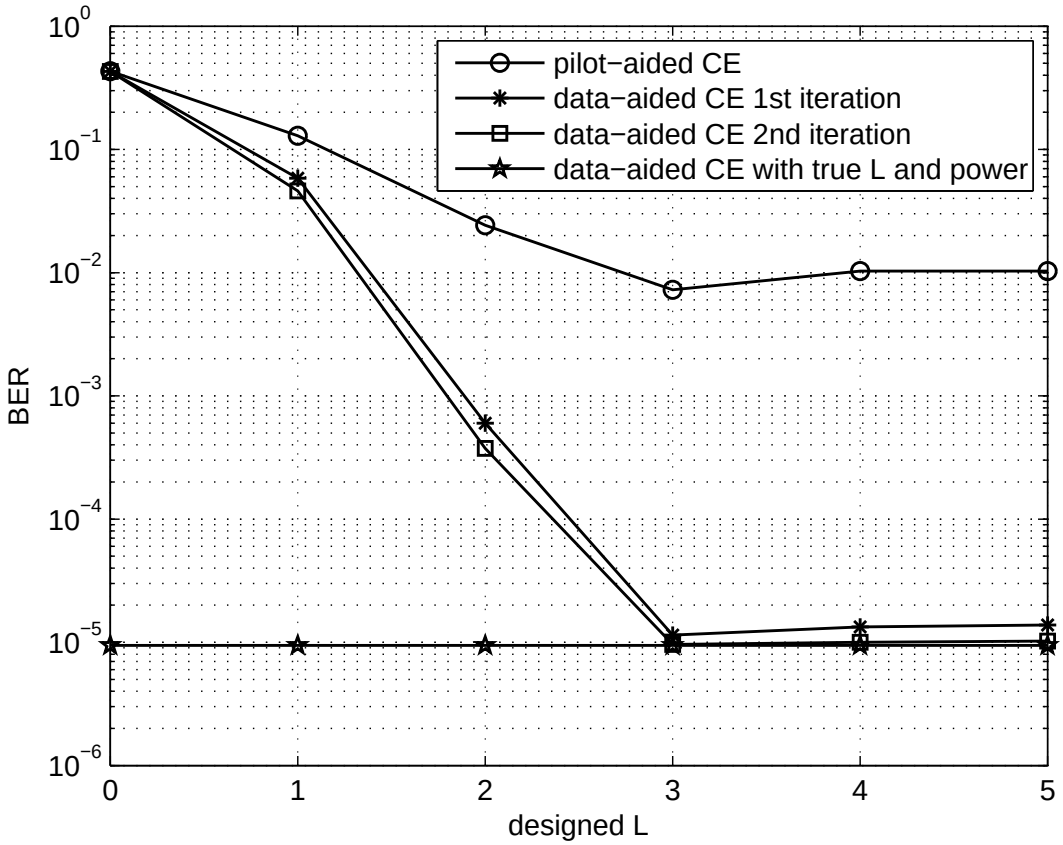

Figure 5.11: BER resulting for $f_dT = 0.1$ and true $L = 3$ with QPSK

5.4 Conclusion

We propose a new channel estimation/symbol detection methodology for fast fading radio channels. The pilot-to-data ratio is low to 7/144. This method requires low pilot overhead and computational complexity. The new method provides BER results nearly as good as the detection/decoding with perfect knowledge of the channel state information. This technique is robust to variations of the channel conditions from the designed values.

Chapter 6

EXIT Chart Analysis

This chapter deals with the convergence behavior of the iterative receiver with estimated channel state information (CSI) in fast fading channels. By using the Extrinsic Information Transfer (EXIT) chart, we can analyze and predict the performance of the previous chapters without simulating the whole iterative receiver. The EXIT chart analysis assumes that all log-likehood ratios (LLRs) are independent. Therefore, it is an asymptotic analysis of the convergence behavior of iterative receivers. The EXIT chart demonstrates that the proposed technique to achieve a bit error rate (BER) nearly as good as when ideal CSI is available. Moreover, the EXIT chart can be used to select the modulation, pilot density, and error correction code for good performance of the proposed iterative receiver.

6.1 Introduction

In order to exploit the EXIT chart in fast fading channels, we give a brief introduction in this section. The EXIT chart is a semi-analytical technique to investigate the convergence behavior of iterative receiver schemes. It was first introduced by Stephan ten Brink [120, 121] to analyze the behavior of turbo codes. The EXIT chart shows

the mapping of the mutual information between the input and output of each system component and the transfer characteristics are plotted into a single diagram which is referred to as the EXIT chart. Now, the EXIT chart has been applied to investigate iterative receivers such as turbo equalizer in multipath channels with time-invariant impulse response [122] and time-variant impulse response [123], optimal mapping in bit-interleaved coded modulation (BICM) [124] in AWGN channels, noncoherent turbo detection [125] in fast fading channels, or [126] in multicarrier interleave division multiple access.

The EXIT chart has also been proposed in [127] for adaptive turbo equalization, [128, 129] for OFDM systems and [130] for BICM. In this chapter, we will show how the EXIT chart can be used to predict the performance of the iterative receiver with the imperfect CSI for OFDM systems in fast fading channels. Figure 6.1 illustrates the analysis of the iterative processing, where the black line is used in the full system, and the dotted line is used for EXIT chart analysis.

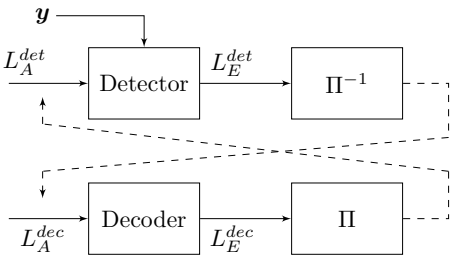

Figure 6.1: Model for the EXIT chart analysis [122]

6.1.1 Input of Transfer Function

The assumption of using a large interleaver assures that the *a priori* LLRs L_A^{dec} of the decoder and L_A^{det} of the detector are modelled as independent Gaussian random

variables, respectively. For simplicity, they are denoted as L_A, expresses as [121].

$$L_A = \mu_A \cdot c + n_A \tag{6.1}$$

where c is the transmitted bit and $c \in \pm 1$. n_A is an additive white Gaussian noise (AWGN) with the variance σ_A^2. $\mu_A = \sigma_A^2/2$. The conditional probability density function (pdf) of L_A is

$$p_A(l|\boldsymbol{c} = c) = \frac{e^{-((l-(\sigma_A^2/2)c)^2)/(2\sigma_A^2)}}{\sqrt{2\pi}\sigma_A^2} \tag{6.2}$$

The mutual information, $I_A = I(\boldsymbol{c}, L_A)$, between the transmitted bits $\boldsymbol{c}$ and the *a priori* LLR L_A, is expressed as [121]

$$\begin{aligned} I_A = \frac{1}{2} \sum_{c=-1,1} \int_{-\infty}^{\infty} & p_A(l|\boldsymbol{c} = c) \\ & \cdot \log_2 \frac{2p_A(l|\boldsymbol{c} = c)}{p_A(l|\boldsymbol{c} = -1) + p_A(l|\boldsymbol{c} = 1)} \mathrm{d}l, \quad 0 \le I_A \le 1 \end{aligned} \tag{6.3}$$

or

$$I_A = 1 - \int_{-\infty}^{\infty} \frac{e^{-((l-(\sigma_A^2/2)c)^2)/(2\sigma_A^2)}}{\sqrt{2\pi}\sigma_A^2} \log_2[1 + e^{-l}]dl \tag{6.4}$$

Equation (6.4) is one-to-one relationship between σ_A and I_A, that is, I_A is a monotonic function of σ_A, denoted as $I_A = J(\sigma_A)$ with $\lim_{\sigma_A \to 0} I_A = 0$ and $\lim_{\sigma_A \to \infty} I_A = 1$ for $\sigma_A > 0$, which correspond to no and perfect *a priori* LLRs respectively.

For a given mutual information I_A, the variance is obtained from $\sigma_A = J^{-1}(I_A)$ [121]. The corresponding *a priori* LLR $L_A^{det}(c_k)$ or $L_A^{dec}(\tilde{c}_k)$ is then generated according to Equation (6.4).

6.1.2 Output of Transfer Function

Similar to the *a priori* LLR L_A, the extrinsic LLRs of the detector and decoder are denoted as L_E for simplicity. The extrinsic information transfer function is then defined as

$$I_E = T(I_A, E_b/N_0) \tag{6.5}$$

where N_0 is the power spectral density of the AWGN, and E_b is the bit energy.

Similar to Subsection 6.1.1, mutual information of the extrinsic information is also denoted as $I_E = I(\boldsymbol{c}, L_E)$, expressed as[121]

$$\begin{aligned} I_E = \frac{1}{2} \sum_{c=-1,1} \int_{-\infty}^{\infty} & p_E(l|\boldsymbol{c}=c) \\ & \cdot log_2 \frac{2p_E(l|\boldsymbol{c}=c)}{p_E(l|\boldsymbol{c}=-1)+p_E(l|\boldsymbol{c}=1)} \mathrm{d}l, \quad 0 \le I_E \le 1 \end{aligned} \tag{6.6}$$

where $p_E(l|\boldsymbol{c})$ is the conditional pdf of L_E. Since no Gaussian distribution is assumed to $p_E(l|\boldsymbol{c})$, $p_E(l|\boldsymbol{c})$ is simulated using Monte Carlo method (histogram method) [121]. $p_E(l|\boldsymbol{c})$ is obtained into two parts corresponding to $c = -1$ and $c = +1$, respectively. I_E is then obtained using the histogram method according to Equation (6.6). In practice, $p_E(l|\boldsymbol{c})$ of the detectors cannot be approximated as Gaussian for fading channels [123, 129].

For the EXIT chart of trajectory, the initial value of the mutual information of the detector starts at the intersection of the detector curve with the $I_A^{det} = 0$ line. The system mutual information moves horizontally to intersect the decoder EXIT curve. In each iteration, the trajectory moves vertically to intersect the detector line and then horizontally to intersect the decoder line.

6.2 EXIT Chart for detection with channel estimation in Fast Fading channels

In a standard EXIT chart for detector/decoder systems, two curves are plotted. One curve is plotted for the detector with the prior mutual information on the X-axis and output extrinsic mutual information on the Y-axis. Another curve is plotted for the decoder, with the *a priori* mutual information on the Y-axis and output extrinsic mutual information on the X-axis. The performance of the system is a function of the distance between these two curves and where these two curves intersect [121].

For all following iterations, the standard EXIT curve drawing rules are employed based on the EXIT curves of the detector with data-aided channel estimation and the decoder; the trajectory moving vertically to the EXIT curve of the detector with data-aided channel estimation and horizontally to the EXIT curve of the decoder.

6.2.1 EXIT Chart of Decoder

For the decoder, transfer function (6.5) is only dependent on the *a priori* input I_A^{dec} values, that is $I_E^{dec} = T^{dec}(I_A^{dec})$. Figure 6.3 depicts EXIT charts of the decoder. The decoder uses a convolutional code with 1/2 rate, constraint length 7 and the generator polynomials $G = [133\ 171]$ in octal form or constraint length 3 and $G = [7\ 5]$, respectively.

To obtain the EXIT curve, 25600 randomly equiprobable coded bits $\tilde{\boldsymbol{c}}$ are used for simulations. Figure 6.3 shows that the stronger the code the steeper the transfer function. When $I_A^{dec} < 0.5$, the code $G = [7\ 5]$ requires a lower I_A^{dec} than the code $G = [133\ 171]$ for the same I_E^{dec}. However, when $I_A^{dec} \geq 0.5$, the code $G = [133\ 171]$ requires a lower I_A^{dec} than the code $G = [7\ 5]$ to approach $I_E^{dec} = 1$, indicating that the stronger the code, the lower error floor after convergence and the faster the

convergence rate.

6.2.2 EXIT Chart of Decision-Directed Channel Estimation with Detector

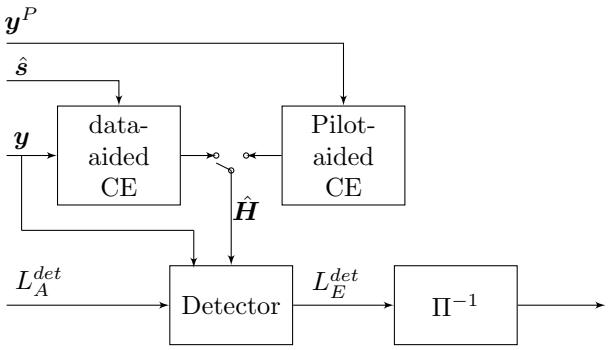

Figure 6.2: Model for the EXIT chart of detector

For the detector, transfer function (6.5) is not only dependent on the *a priori* input I_A^{det} values but also E_b/N_0, CSI and modulation, etc; that is $I_E^{det} = T^{det}(I_A^{det}, E_b/N_0, \boldsymbol{H})$. The model of the detector for the EXIT chart analysis is depicted in Figure 6.2.

For the channel estimation (CE) stage, a first-order Kalman filter is used for two different iterative CE methods discussed in Chapter 4. The initial channel estimate is provided solely from the measurements of the pilots $\boldsymbol{y}^P$. The detected data $\hat{\boldsymbol{s}}$ is then fed back into the data-aided channel estimation block for the estimation of the channel coefficients in the next iteration.

The EXIT chart of the detector is characterized by three EXIT curves: the EXIT curve of the detector using channel estimates obtained using pilot-signals, the EXIT curve of the detector using data-aided channel estimation, the EXIT curve of the detector using perfect CSI. Figure 6.3 shows these curves for a typical fast fading channel with gray-coded quadrature phase-shift keying (QPSK) modulation of gray mapping and the normalized Doppler frequency is set to $f_dT = 0.1$, where T is OFDM

symbol period. We choose a 4-path wide-sense stationary uncorrelated scattering (WSSUS) channel with mean propagation powers of 8/15, 4/15, 2/15 and 1/15 for propagation delays of 0 to 3 samples respectively.

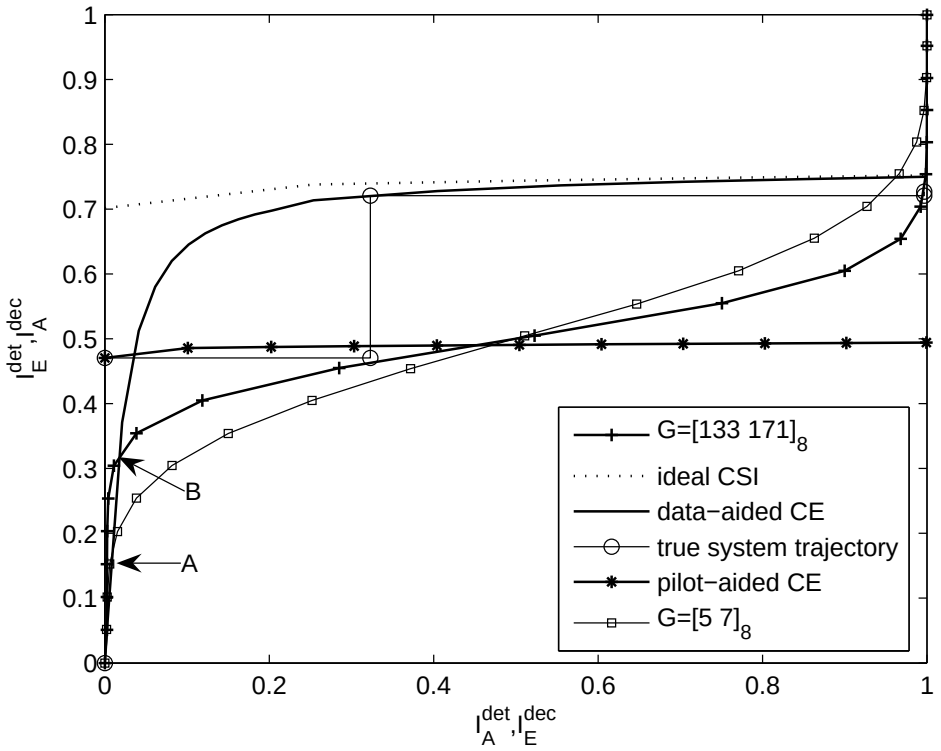

Figure 6.3: EXIT chart

25600 randomly chosen equiprobable coded bits are generated for simulations. For the ideal CSI case, the proposed detector only needs two iterations to reach $I_E^{dec} \approx 1$. In the detector curves for imperfect CSI, the EXIT curve of the detector based on the data-aided CE is very close to that with the ideal CSI when the perfect *a priori* mutual information $I_A^{det} = 1$. Therefore, the degradation based on the data-estimated channel estimation is very low and the BER performance is close to that of the ideal CSI when using a large interleaver.

For the EXIT chart of detector based on data-aided CE, when I_A^{det} is low, the iterative receiver based on data-aided CE cannot approach to a large value I_E^{dec} with

the code $G = [133\ 177]_8$ or $G = [5\ 7]_8$ and gets stuck at the intersection A or B respectively, shown in Figure 6.3. After this bottleneck area, the detector based on data-estimated case can improve the performance for a large I_A^{det}. When the EXIT curve of detector based on data-aided CE with the code $G = [133\ 177]_8$ gets stuck, there is still a path through the bottleneck using the code $G = [5\ 7]_8$. However, the EXIT curve of detector based on data-aided CE with the code $G = [133\ 177]_8$ approaches $I_E^{dec} = 1$ faster than the code $G = [5\ 7]_8$, indicating that a lower error floor is obtained after convergence. Therefore, it is better to construct a decoder to obtain values I_E^{dec} as high as possible after several iterations.

On the other hand, when I_A^{det} is low, the value I_E^{det} based on data-aided CE is lower than that based on pilot-aided CE, indicating that an unreliable detection is obtained comparing to the pilot-estimated case. However, the EXIT curve of the detector based on pilot case is more flat than that of detector based on data-estimated case, it will stop improving the performance and get stuck at the intersection of the EXIT curves between the detector and the decoder, indicating that the performance based on pilot-estimated channel cannot converge and has a high BER.

Therefore, the EXIT chart depicted in Figure 6.3 indicates that it is better to combine the detector based on pilot-estimated case and data-estimated case together, which is discussed in Chapter 5. With this joint channel estimation scheme, we combine the good starting value I_E^{det} of the pilot-based case and the good ending value I_E^{det} of the data-based case. For the initial iteration, the detector based on pilot-estimated case can overcome the poor channel estimation based on the data-estimated case; then it is switched to data-estimated case for the successive iterations which can approach the performance of the ideal CSI.

6.2.3 EXIT Chart of Trajectory

The mutual information trajectory of iterative receiver proposed in Chapter 5 is also shown in Figure 6.3. The detector in the initial iteration uses the pilot-aided channel estimator so the point of I_E^{det} for $I_A^{det} = 0$ is for the pilot-aided channel estimation. The trajectory then moves horizontally to the decoders EXIT curve.

The proposed detector will only provide good final BER performance if the EXIT curve for the detector with data-aided channel estimation is above the curve for the detector using pilot-aided channel estimation at the I_E^{dec} value of the intersection of the EXIT curves of the decoder curve and the detector using pilot-aided channel estimation. That is, there must be enough information provided by the initial detection using pilot-aided channel estimation to allow the error correction code to provide some correction capability. This condition will always be satisfied if the E_b/N_0 value is large, as higher E_b/N_0 values raise the EXIT curves of all detectors without changing the EXIT curve of the decoder. The required E_b/N_0 value is a function of the modulation and error correction code used. For this reason, it recommended that proper operation be verified with EXIT chart analysis. The proposed method will fail at low E_b/N_0 values or if the slope of the error correction codes EXIT curve is very steep for I_A^{dec} values lower than the initial I_E^{det} provided by the detector. In these cases, it is possible to get some BER improvement by running the detector with pilot-aided channel estimation for more than one iteration so the interference cancellation can provide some information improvement.

6.3 Simulation Results

In this section, the performance of the proposed algorithm is tested in a simulated wireless OFDM system with a simulated multipath fast-fading radio channel. We

simulate a 4-path WSSUS channel with mean propagation powers of 8/15, 4/15, 2/15 and 1/15 for propagation delays of 0 to 3 samples respectively. Each path is subject to independent Rayleigh fading generated using the method from [119] with Jakes' model specifying the autocorrelation function. We select an OFDM system with $N = 128$ samples for each OFDM symbol. The length of CP is $N_{cp} = N/8$. The convolutional code used in simulations is with 1/2 rate, constraint length 7 and the generator polynomials $G = [133\ 171]$ in octal form.

The channel is estimated using a first-order Kalman filter. Each transmission frame consists of $K_b = 10$ OFDM blocks. A length of $K_t = 10$ transmission frames is used for the interleaving block. The joint channel estimation, data detection and decoding method described in Chapter 5 is used. Initial channel estimation is performed using $K_p = 35$ or $K_p = 70$ pilots per each transmission block.

We analyze EXIT charts of the decoder and the detector with different estimated channel schemes, where gray-coded QPSK constellation is used in simulations. The normalized Doppler frequency is set to $f_dT = 0.1$, where T is OFDM symbol period. We choose $Q = 4$ basis functions, the corresponding number of pilot blocks for each transmission is $K_p \geq 5$.

To test the predicted performance, the true trajectory is simulated by calculating the mutual information of the extrinsic LLR for the detector and decoder after each iteration. The EXIT charts of the iterative process are shown in Figures 6.4-6.8 for $f_dT = 0.1$ at $E_b/N_0 = 7$dB and $E_b/N_0 = 8$dB.

Figures 6.4-6.6 depict the EXIT charts with $K_t = 10$ for each interleaving block at $E_b/N_0 = 7$dB for different number of pilot blocks. Figure 6.4 shows that the true trajectory needs only three iterations to reach $I_E^{dec} \approx 1$ when $K_p = 10$. However, for a reduced value $K_p = 5$, Figure 6.6 shows that the trajectory gets stuck at $I_E^{dec} \approx 0.96$ after three iterations. The receiver's EXIT trajectory depicted in Figure 6.6 doesn't

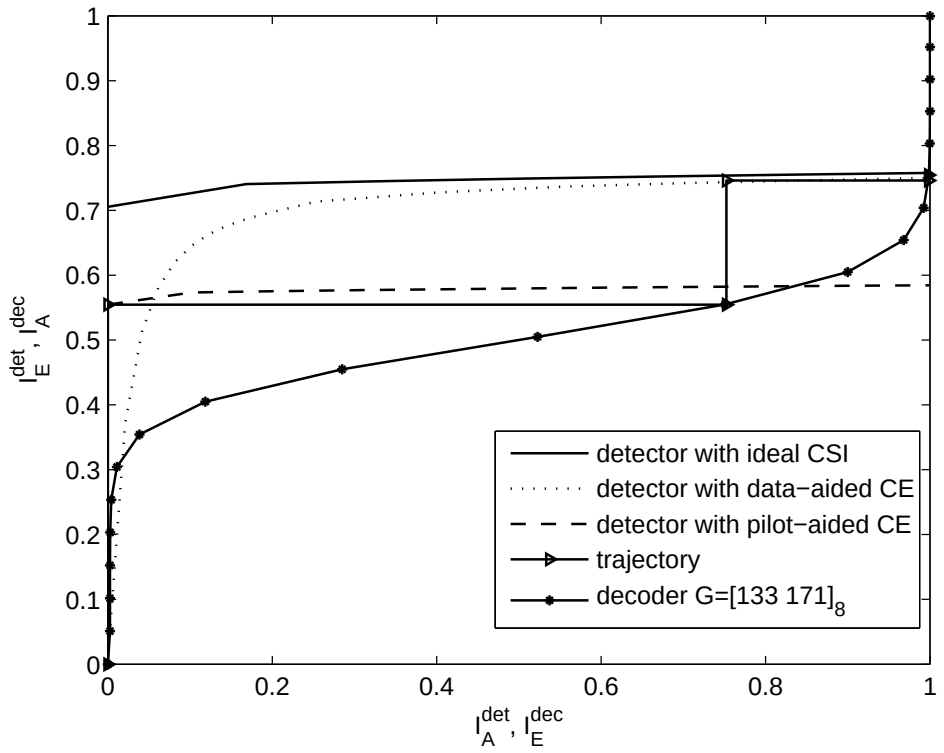

Figure 6.4: EXIT chart using $K_p = 10$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_o = 7dB$

match the transfer characteristics predicted by the EXIT chart due to dependencies of the LLR between the detector and the decoder. Increasing the interleaver length to $K_t = 100$ decreases these dependencies and the simulated trajectory matches the transfer characteristics as can be seen in Figure 6.6. Enlarging K_t increases the computational complexity and memory requirements of the channel estimation. Therefore, finding the optimal number of pilot blocks is a tradeoff between the performance of the error floor and the complexity cost at low SNRs. At this system, we choose $K_p = 10$ for each transmission frame and $K_t = 10$ since the increase in pilot power for $K_p = 10$ over $K_p = 5$ is only a small percentage of total power and the memory requirement in the receiver is reduced by an order of magnitude. It was noted that when $E_b/N_0 > 7$dB, $K_p = 5$ provided good BER performance for $K_t = 10$.

The EXIT curves using different pilot schemes are shown in Figures 6.7 and 6.8

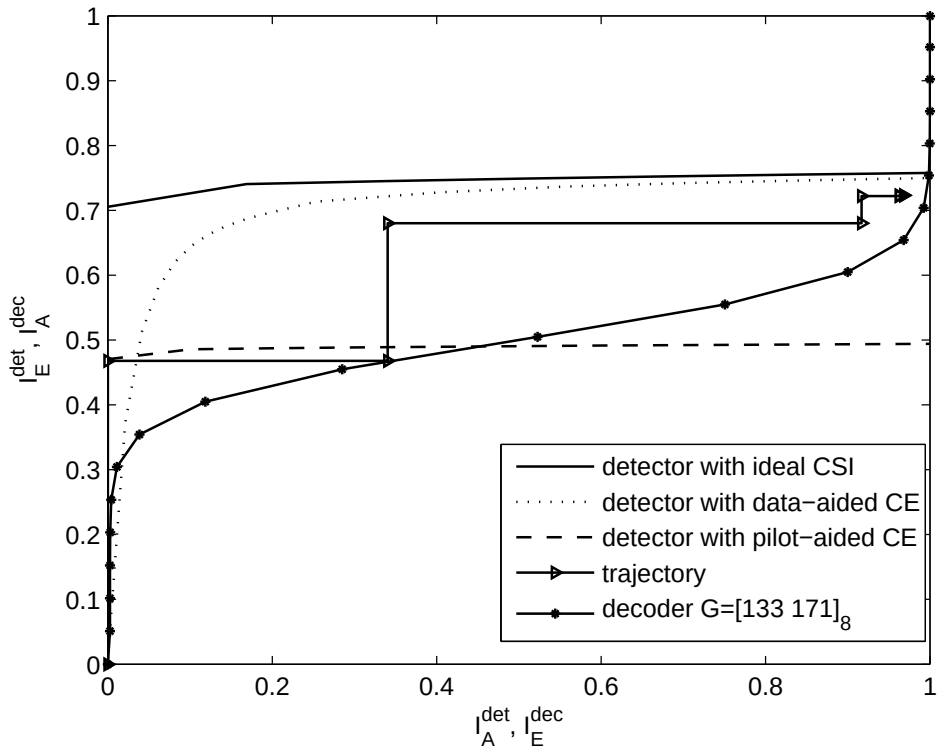

Figure 6.5: EXIT chart using $K_p = 5$ and $K_t = 10$ for $f_dT = 0.1$ at $E_b/N_0 = 7dB$

with $K_t = 10$ at $E_b/N_o = 8$dB. The simulated trajectories are shown that after two iterations based on data-estimated case, the trajectories are close to the EXIT curves of the ideal CSI for different pilot schemes. Therefore, the whole system using $K_p = 5$ for initial detection has the same error floor as that using $K_p = 10$ at a high SNR. The EXIT chart predicts that using $K_p = 5$ is enough for a high SNR.

From the analysis above, the EXIT chart predicts that the detector needs two iterations after the initial detection and the detection based on data-aided channel estimation comes very close to the case of perfect CSI. Therefore, the degradation of the system performance due to channel estimation is very low. In order to make the system converge at low SNR and have a low complexity, we $K_t = 10$ and $K_p = 10$ for the whole iterative system.

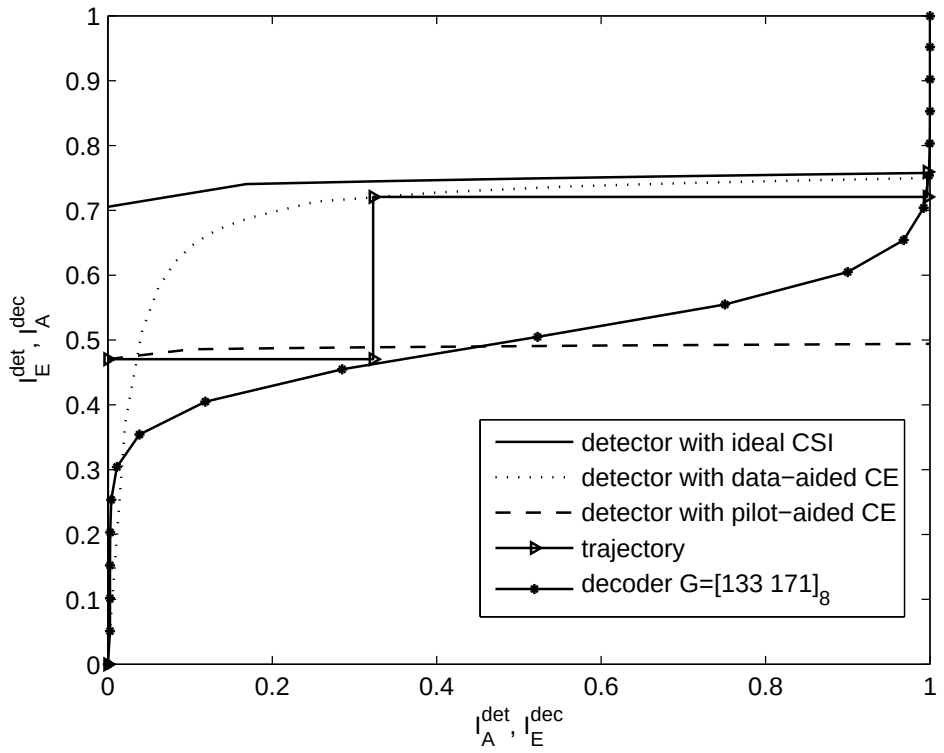

Figure 6.6: EXIT chart using $K_p = 5$ and $K_t = 100$ for $f_d = 0.1$ at $E_b/N_0 = 7dB$

6.4 Conclusion

The EXIT chart technique predicts the convergence behavior of the proposed iterative scheme. EXIT chart analysis can be used to assist with the selection of pilot signals, modulation type, and error correction code as well as the required E_b/N_0 value for the algorithm to converge.

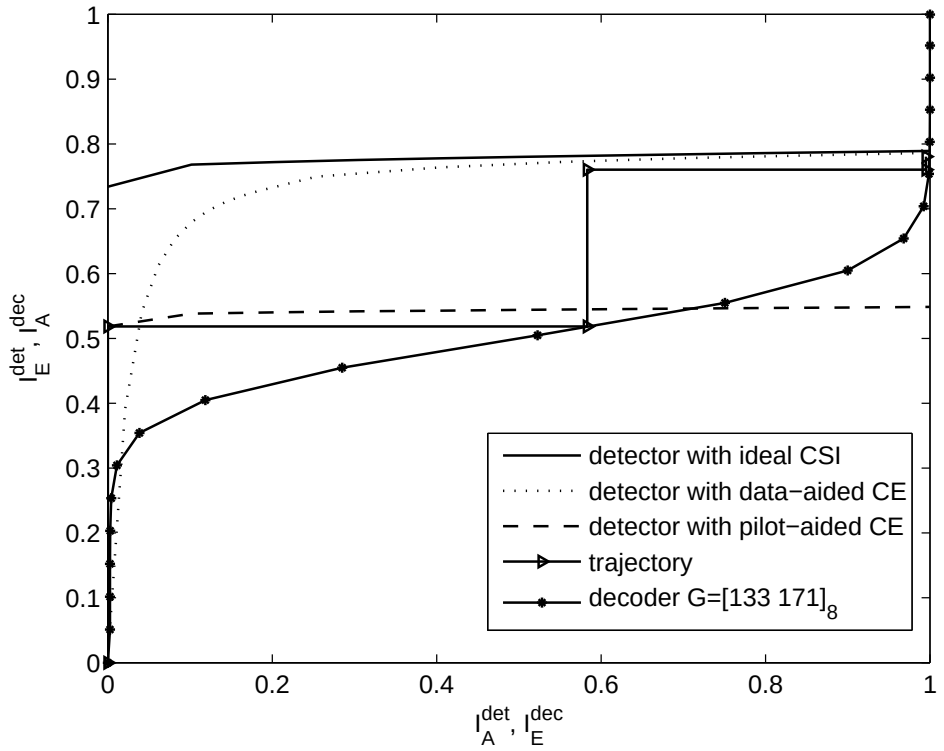

Figure 6.7: EXIT chart using $K_p = 5$ and $K_t = 10$ for $f_d T = 0.1$ at $E_b/N_0 = 8dB$

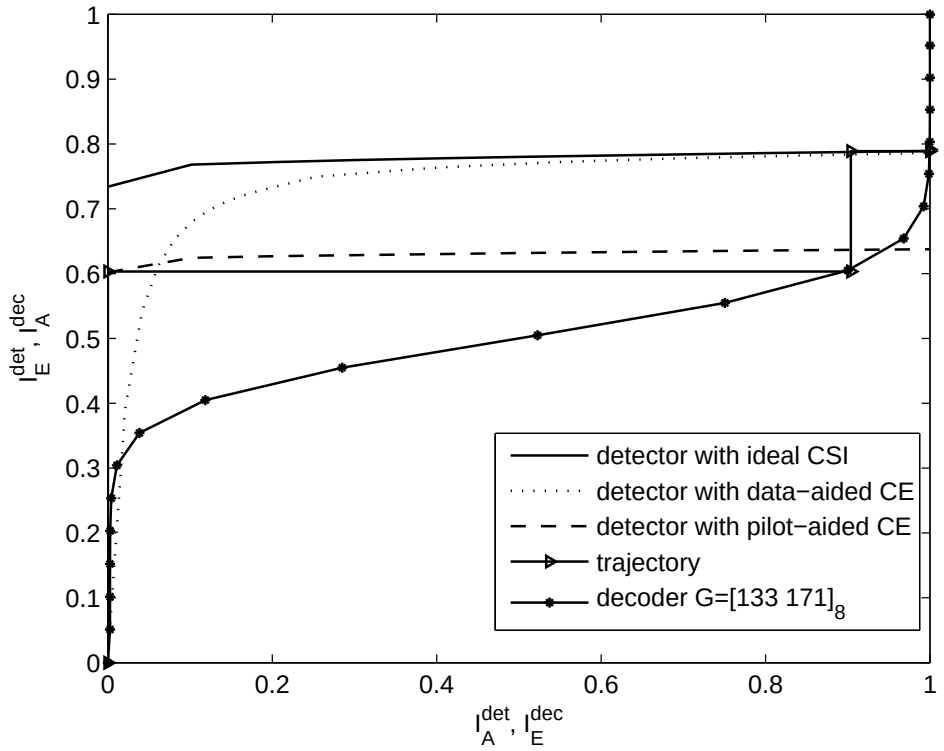

Figure 6.8: EXIT chart using $K_p = 10$ and $K_t = 10$ for $f_d T = 0.1$ at $E_b/N_o = 7dB$

Chapter 7

Conclusions and Future Work

7.1 Summary

Reliable estimate of the channel state information (CSI) for orthogonal frequency division multiplexing (OFDM) systems in doubly selective channels has been studied in this thesis. The algorithms developed in this thesis have improved the performance of the whole system requiring only low ratios of pilot to data for excellent performance in fast fading channels.

The main contributions of the thesis are :

- To enable low cost detection and decoding of OFDM in fast fading, a new structure for channel processing has been used at the receiver in coded OFDM systems, where the OFDM block has been used as the basis for data detection; the transmission block which contains several OFDM blocks has been used for channel estimation; and the interleaving block which contains multiple transmission blocks has been used for decoding. To capture the channel dynamics, a novel multivariate autoregressive (AR) process over transmission blocks has been developed to model the time evolution of the fast fading channel with the

help of a basis expansion model (BEM). To develop the iterative scheme at the receiver, three measurement models have been discussed in this thesis, i.e., the measurement model for data detection based on the estimated or known CSI; the measurement model for channel estimation based on pilot symbols and the measurement model for channel estimation based on the detected data.

- A novel Kalman smoothing algorithm based on a second-order dynamic model has been exploited, where the mean square error (MSE) of the channel is near to that of the optimal Wiener filter and the computational complexity is same as the previous algorithms. Moreover, the proposed method can improve the bit error rate (BER) performance of wireless communication systems. Also, this new technique is robust when the Doppler frequency or signal-to-noise ratio (SNR) is varied from the value used to design the channel estimation filter.

- A novel joint channel estimation data detection algorithm has been developed for fast fading channels with a small number of pilots and low pilot power to achieve the BER performance close to when the CSI is known perfectly. The new channel estimation symbol detection technique is robust to variations of the radio channel from the design values and applicable to multiple modulation and coding types.

- A useful analysis technique has been investigated to analyze the convergence behavior of the new algorithm and select the modulation, pilot density, and error correction code for good system performance for a given power level by the extrinsic information transfer (EXIT) chart technique.

7.2 Future Work

Future work should be done to integrate efficient data detection algorithm into joint channel estimation and data detection techniques for coded OFDM systems in fast fading channels. In particular, the new data detection method should be investigated to see if alternate methods can be found to improve data detection without greatly increasing the complexity.

It will also be interesting to investigate the performance of channel estimation when used for higher Doppler frequencies. Currently, each transmission block contains 10 OFDM symbol blocks. When Doppler frequency increases, the structure of time blocks for channel processing will need to be adjusted to maintain efficient performance for low pilot to data ratios. The required form of the signalling will be investigated in future research.

In addition, the multivariate AR model can be applied to other systems such as single-carrier (SC) systems or multicarrier-code division multiple access (MC-CDMA) systems for fast fading channels.

Future work in the context of channel estimation and symbol detection as well as decoding relies on EXIT chart analysis to achieve lower BER performance over fast fading channels at a low possible complexity.

Appendix A

Basic Time Domain Channel Impulse Response Description

In wireless communications, the radio channel is often assumed to be wide-sense stationary uncorrelated scattering (WSSUS), where the time-variant impulse response $h(t, \tau)$ is modelled as a zero-mean complex Gaussian process for a fixed delay τ, and uncorrelated for different delays [131, 2]. In such a case, the autocorrelation function of $h(t, \tau)$ is given by

$$E[h(t, \tau_1)h^*(t + \Delta t, \tau_2)] = r_h(\Delta t, \tau_1)\delta(\tau_1 - \tau_2) \tag{A.1}$$

If the arrival angle of the received waveform is a uniformly distributed random variable over $[-\pi, \pi]$, $r_h(\Delta t; \tau)$ can be separated in time and delay [104], i.e.,

$$r_h(\Delta t; \tau) = r_t(\Delta t)r_\tau(\tau) \tag{A.2}$$

where $r_\tau(\tau)$ is the multipath intensity profile, and the Fourier transform of the auto-correlation function $r_t(\Delta t)$ is

$$r_t(\Delta t) = J_0(2\pi f_d \Delta t) \tag{A.3}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind and $f_d = v f_c/c$ is the maximum Doppler frequency, where v is the vehicle speed in meters per second (m/s), f_c is the carrier frequency, and c is the speed of light (3×10^8m/s) [2]. In the case of outdoor radio channels, a vertical receive antenna with constant azimuthal gain, the angles of arrival of the radio waves are uniformly distributed over $(-\pi, \pi)$, the normalized Jakes' Doppler power spectrum, $S_H(f)$, is given by [99]

$$S_H(f) = \begin{cases} \frac{1}{\pi f_d} \frac{1}{\sqrt{1-(f/f_d)^2}}, & \text{if } |f| \le f_d \\ 0, & \text{otherwise} \end{cases} \tag{A.4}$$

Suppose the multipath delay spread is τ_{max}, the tapped delay line model for the channel can be truncated at $L = \lfloor \tau_{max} W \rfloor$ taps, where W is bandwidth, $\lfloor \cdot \rfloor$ is the integer floor function, the impulse response is expressed as

$$h(t, \tau) = \sum_{l=0}^{L} h(t, \tau_l) \delta(\tau - \tau_l) \tag{A.5}$$

where $\tau_l = l/W$ and $\tau_0 < \tau_1 < \cdots < \tau_L$ for $l = 0, 1, ..., L$. The $h(t, \tau_l)$'s represent the tap weights corresponding to the L different delays τ_l. $\delta(t)$ is Dirac's delta function. Since $h(t, \tau_l)$ is characterized as WSSUS, $h(t, \tau_l)$'s are mutually uncorrelated with variance $\sigma_l^2 = r_\tau(\tau_l)$. In such a case, when the $h(t, \tau_l)$'s are Gaussian random process, they are statistically independent. Usually, the total variances of the channel taps are assumed to be normalized to unity, i.e., $\sum_{l=0}^{L} \sigma_l^2 = 1$. In deed, power is often

normalized in communication systems [99]. Moreover, if the envelope $|h(t;\tau)|$ at time t is Rayleigh-distributed, the channel is a Rayleigh fading channel.

Appendix B

Basis Expansion Model

This appendix provides a summary of basis expansion model (BEM). Beginning with the papers by Giannakis *et al.* [132], Borah *et al.* [133] and Martone [134], BEMs have been widely developed to model doubly selective channels in wireless communication, where the time variant impulse responses $h(n, l)$ are expressed by the sum of time-varying basis functions, weighted by time-invariant channel coefficients within a fixed time duration, which is called the transmission block. A general BEM expansion of $h(n, l)$ can be written as

$$h(n, l) = \sum_{q=0}^{Q} E_n(q) x_l(q) \tag{B.1}$$

where $E_n(q)$ gives the value at time n of the qth basis function and $x_l(q)$ is the coefficient for path l corresponding to the qth basis function. Due to the bandlimited nature of Equation (A.4), Q is BEM order denoted by $Q = 2\lceil f_d M T_s \rceil$, where T_s is the sampling period and M is the number of samples in each transmission block [1], and $Q + 1$ is the number of basis coefficients.

It is convenient to express the BEM in matrix form. Hence, we define the basis

function matrix as

$$\boldsymbol{E} = \begin{bmatrix} E_0(0) & \cdots & E_0(Q) \\ \vdots & \ddots & \vdots \\ E_{M-1}(0) & \cdots & E_{M-1}(Q) \end{bmatrix} \tag{B.2}$$

Let $\boldsymbol{x}_l$ denote the coefficient vector for the lth path for the current transmission block, i.e.,

$$\boldsymbol{x}_l = \begin{bmatrix} x_l(0) \\ \vdots \\ x_l(Q) \end{bmatrix} \tag{B.3}$$

Thus, the channel gain in each transmission block for the lth path is expressed as

$$\boldsymbol{h}_l = \boldsymbol{E}\boldsymbol{x}_l. \tag{B.4}$$

For a given basis function $\boldsymbol{E}$, $\boldsymbol{x}_l$ vector gives all the necessary information about the path l for a given block. Different BEM matrices have been performed on the criterion for optimizing the BEM with various robustness properties [72, 110, 69]. Below, we discuss these matrices.

B.1 Complex Exponential Basis Expansion Model

The complex exponential basis expansion model (CE-BEM) or Fourier BEM comes from [132]. As Q increases, the CE-BEM approximates the random fading coefficient model in mobile communication channels [132].

In [1], the Fourier basis functions over a window of M samples are expressed as

$$E_n(q) = e^{j\omega_q n}, \quad q \in [0, Q] \tag{B.5}$$

where $\omega_q = 2\pi(q - Q/2)/M$. In this case, a discrete-time baseband equivalent matrix-vector channel model is given as

$$\boldsymbol{E} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\frac{2\pi Q/2}{M}} & \cdots & e^{j\frac{2\pi Q/2}{M}} \\ \vdots & \cdots & \vdots \\ e^{-j\frac{2\pi Q/2(M-1)}{M}} & \cdots & e^{j\frac{2\pi Q/2(M-1)}{M}} \end{bmatrix} \tag{B.6}$$

This model approximates doubly selective channels as a linear combination of a finite number of complex exponentials. Since this sampling period of the Doppler spectrum is equal to the symbol period T_s, sampling signal over a finite transmission block is equivalent to the signal truncated by a rectangular window. Then the truncated spectrum by using the discrete Fourier transform (DFT) is spread out over the whole frequency range, which is so-called spectral leakage [98]. As a result, the modelling error between the true channel and the CE-BEM is quite high.

The basic limitation of the CE-BEM to cope with modelling error has motivated a considerable amount of research into other BEMs. As an alternative, we may oversample the Doppler spectrum, which is discussed in [110], where ω_q in equation(B.5) becomes $\omega_q = 2\pi(q - Q/2)/(kM)$ with $Q = 2\lceil f_d kMT_s \rceil$ and $k = 2$. In fact, this choice of Q may be increased compared to critical sampling case.

Instead of rectangular window, other windowing functions can be used in order to reduce leakage. However, the mainlobe levels of the general windowing functions such as Hamming, Hanning, etc. are wider than the rectangular window, but the sidelobe levels are lower [98]. The optimal windowing function is the time limited sequence whose energy is most concentrated in a finite frequency interval, and hence, is related to the zeroth discrete prolate spheroidal (DPS) sequence [135, 136, 137], which is described in detail in the following section.

B.2 Discrete Prolate Spheroidal Basis Expansion Model

Another approach is to use prolate spheroidal basis functions, which come from the the prolate spheroidal wave functions (PSWFs). The PSWFs provide an orthogonal set of the bandlimited real functions, with each concentration being each eigenvalue, and are most concentrated in the bandwidth [138, 139]. In the discrete time, these basis functions are the DPS sequences [72]. For a given transmission block with M samples and the Doppler frequency f_d, the CSI, $h(n, l)$ of each path l, can be stated in terms of PSWFs, DPS sequences and their associated eigenvalues λ's.

The DPS-BEM functions for channel modelling are bandlimited to the Doppler frequency and are concentrated on the time interval $[0,\ M-1]$. Since the extension of the finite sequence for the time duration has least energy [138], it can overcome the drawback of windowing in the CE-BEM.

For a given transmission block with M samples and the Doppler frequency f_d, the DPS sequences $e_m(q)$'s can be concentrated on M samples and bandlimited to $f_d T_s$, where T_s is the sampling period, [138]

$$\sum_{m=0}^{M-1} \frac{\sin[2\pi f_d T(i-m)]}{\pi(i-m)} e_m(q) = \lambda_q e_i(q), \quad i \in [-\infty, \infty] \tag{B.7}$$

where λ_q is the qth eigenvalue, which measures the energy concentration of the qth DPS sequence.

The DPS sequences are orthogonal on both the infinite interval and the finite

interval $[0, M-1]$, that is [138]

$$\begin{aligned}\sum_{i=0}^{M-1} e_m(i)e_m(j) &= \lambda_i \sum_{-\infty}^{\infty} e_m(i)e_m(j) \\ &= \delta_{ij}, \quad i,j \in [0, M-1]\end{aligned} \tag{B.8}$$

The eigenvalue λ_i is close to 1 for $i \leq \lceil 2f_dT_sM \rceil$ while λ_i is close to 0 for $i > \lceil 2f_dT_sM \rceil$ [138]. Hence, for a large M, a small number of $\lceil 2f_dT_sM \rceil$ is sufficient to describe the DPS sequence [138].

Let $\boldsymbol{A}$ be the $M \times M$ matrix with the element

$$A_{m,n} = \frac{sin[2\pi f_dT_s(m-n)]}{\pi(m-n)}, \quad m,n \in [0, ..., M-1]. \tag{B.9}$$

Equation (B.7) is rewritten as

$$\boldsymbol{A}\boldsymbol{e}(q) = \lambda_q \boldsymbol{e}(q), \quad q \in [0, ..., M-1]. \tag{B.10}$$

Therefore, the DPS sequences $\boldsymbol{e}(q)$ is the eigenvector of $\boldsymbol{A}$ corresponding to the eigenvalue λ_q.

In general, for a given interval M and Doppler frequency f_d, define the basis expansion number $Q = \lceil 2f_dT_skM \rceil$, where $k = 1$ or $k = 2$ [72, 67]. In practice, when the true Doppler frequency is unavailable, a conservative design would be to set the assumed Doppler frequency the highest expected level.

B.3 Karhunen-Loève Basis Expansion Model

In the MSE criterion, Karhunen-Loève (KL) expansion model, which was proposed in [140, 141, 142, 69], is optimal in minimizing the channel modeling MSE. For a given

length M, the correlation matrix of $\boldsymbol{h}_l$ from Equation (B.4) is

$$\boldsymbol{R}_l = E[\boldsymbol{h}_l \boldsymbol{h}_l^H] = E[\boldsymbol{E}\boldsymbol{x}_l(\boldsymbol{E}\boldsymbol{x}_l)^H] = \boldsymbol{E}E[\boldsymbol{x}_l \boldsymbol{x}_l^H]\boldsymbol{E}^H, \quad \text{for } l \in [0, L] \tag{B.11}$$

where $\boldsymbol{R}_l$ is the size of M by M. Using singular value decomposition (SVD), $\boldsymbol{R}_l$ is expressed as

$$\boldsymbol{R}_l = \boldsymbol{U}_l \boldsymbol{V}_l \boldsymbol{U}_l^H \tag{B.12}$$

where $\boldsymbol{U}_l$ is a unitary matrix with the size of M by M, $\boldsymbol{V}_l$ is a diagonal matrix with M nonnegative real values called eigenvalues in non-increasing form.

Let us choose Q most significant eigenvalues. Corresponding to $\boldsymbol{V}_l$ with Q diagonal entries, $\boldsymbol{U}_l$ becomes M by Q. In order to minimize the channel modeling error, KL basis expansion functions $\boldsymbol{E}$ choose $\boldsymbol{U}_l$ with the size of M by Q.

Based on Jakes' model, the correlation function of the channel impulse response $h(n, l)$ is expressed as

$$R_l(m, n) = E[h(m, l)h^H(n, l)] = \sigma_l^2 J_0(2\pi f_d T_s |m - n|), \quad \text{for } m, n \in [0, M] \tag{B.13}$$

where J_0 is the zeroth order Bessel function [104], σ_l^2 is the variance of the lth path.

Let $\boldsymbol{R}_l = \sigma_l^2 \boldsymbol{R}$, where $\boldsymbol{R}$ is the normalized correlation matrix. Similarly, We apply SVD on $\boldsymbol{R}\boldsymbol{U}\boldsymbol{V}\boldsymbol{U}^H$, where $\boldsymbol{U} = \boldsymbol{U}_l$, $\boldsymbol{V}_l = \sigma_l^2 \boldsymbol{V}$. Therefore, the KL basis functions are chosen as $\boldsymbol{E} = \boldsymbol{U}$.

Appendix C

Whittle-Wiggins-Robinson Algorithm

In this appendix, we provide the calculation of the values for a multivariate autoregressive (AR) process based on Whittle-Wiggins-Robinson Algorithm (WWRA) [102, 103].

Consider an unknown vector of size q for time index n expressed as

$$\boldsymbol{x}[n] = \begin{bmatrix} x_1[n] & x_2[n] & \cdots & x_q[n] \end{bmatrix}^T \tag{C.1}$$

with the covariance matrix defined as

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{x}}[j] &= E\{\boldsymbol{x}[n+j]\boldsymbol{x}[n]^H\} \\ &= \begin{bmatrix} R_{x_1x_1}[j] & R_{x_1x_2}[j] & \cdots & R_{x_1x_q}[j] \\ R_{x_2x_1}[j] & R_{x_2x_2}[j] & \cdots & R_{x_2x_q}[j] \\ \vdots & \vdots & \ddots & \vdots \\ R_{x_qx_1}[j] & R_{x_qx_2}[j] & \cdots & R_{x_qx_q}[j] \end{bmatrix} \end{aligned} \tag{C.2}$$

A Kth-order multivariate AR process is modelled as

$$\boldsymbol{x}[n] = -\sum_{i=1}^{K} \boldsymbol{A}[i]\boldsymbol{x}[n-i] + \boldsymbol{e}_f[n] \tag{C.3}$$

where $\boldsymbol{A}[i],\ i = 1, 2, ..., K$ are q by q matrices, and $\boldsymbol{e}_f[n]$ is an additive white Gaussian noise (AWGN) vector process with zero mean and a covariance of $\boldsymbol{Q}^f = E\{\boldsymbol{e}_f[n]\boldsymbol{e}_f[n]^H\}$. High-order AR model can provide a more accurate model with high complexity. However, low-order AR model is more practically used to approximate discrete-time random process [75, 143].

Define

$$\underline{\boldsymbol{x}}[n] = \begin{bmatrix} \boldsymbol{x}^T[n] & \cdots & \boldsymbol{x}^T(n-K+1) \end{bmatrix}^T \tag{C.4}$$

The covariance matrix of $\underline{\boldsymbol{x}}[n]$ is a positive semi-definite block matrix defined as

$$\begin{aligned}
\underline{\boldsymbol{R}}_k &= E\{\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}[n]^H\} \\
&= \begin{bmatrix} E\{\boldsymbol{x}[n]\boldsymbol{x}[n]^H\} & E\{\boldsymbol{x}[n]\boldsymbol{x}[n-1]^H\} & \cdots & E\{\boldsymbol{x}[n]\boldsymbol{x}[n-K]^H\} \\ E\{\boldsymbol{x}[n-1]\boldsymbol{x}[n]^H\} & E\{\boldsymbol{x}[n-1]\boldsymbol{x}[n-1]^H\} & \cdots & E\{\boldsymbol{x}[n-1]\boldsymbol{x}[n-K]^H\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{\boldsymbol{x}[n-K]\boldsymbol{x}[n]^H\} & E\{\boldsymbol{x}[n-K]\boldsymbol{x}[n-1]^H\} & \cdots & E\{\boldsymbol{x}[n-K]\boldsymbol{x}[n-K]^H\} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{x}}[0] & \boldsymbol{R}_{\boldsymbol{x}}[1] & \cdots & \boldsymbol{R}_{\boldsymbol{x}}[K] \\ \boldsymbol{R}_{\boldsymbol{x}}[-1] & \boldsymbol{R}_{\boldsymbol{x}}[0] & \cdots & \boldsymbol{R}_{\boldsymbol{x}}[1-K] \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}_{\boldsymbol{x}}[-K] & \boldsymbol{R}_{\boldsymbol{x}}[1-K] & \cdots & \boldsymbol{R}_{\boldsymbol{x}}[0] \end{bmatrix}
\end{aligned} \tag{C.5}$$

where $\underline{\boldsymbol{R}}_k$ has both a Hermitian and a block-Toeplitz structure, that is $\boldsymbol{R}_{\boldsymbol{x}}[n] = \boldsymbol{R}_{\boldsymbol{x}}[-n]^H$, but the $\boldsymbol{R}_{\boldsymbol{x}}[n]$ are not Hermitian ($\boldsymbol{R}_{\boldsymbol{x}}[n] \neq \boldsymbol{R}_{\boldsymbol{x}}^H[n]$).

Equation (C.3) is then expressed as

$$\underline{\boldsymbol{A}}_k[n]\underline{\boldsymbol{x}}[n] = \boldsymbol{e}_f[n] \tag{C.6}$$

where

$$\underline{\boldsymbol{A}}_k = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}[1] & \cdots & \boldsymbol{A}[K] \end{bmatrix}, \tag{C.7}$$

The values of the AR coefficient matrices satisfy the multichannel Yule-Walker equation [100, 101]:

$$\underline{\boldsymbol{A}}_k \ \underline{\boldsymbol{R}}_k = \begin{bmatrix} \boldsymbol{Q}^f & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix} \tag{C.8}$$

This is a multivariate forward linear prediction filter with order K. The multivariate backward linear filter process is expressed as

$$\underline{\boldsymbol{B}}_k[n]\underline{\boldsymbol{x}}[n] = \boldsymbol{e}_b[n] \tag{C.9}$$

where

$$\underline{\boldsymbol{B}}_k = \begin{bmatrix} \boldsymbol{B}[1] & \cdots & \boldsymbol{B}[K] & \boldsymbol{I} \end{bmatrix}, \tag{C.10}$$

The multichannel Yule-Walker equations for the backward coefficients of a stationary AR process are written by

$$\underline{\boldsymbol{B}}_k \ \underline{\boldsymbol{R}}_k = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{Q}^b \end{bmatrix} \tag{C.11}$$

where $\boldsymbol{Q}^b = E\{\boldsymbol{e}_b[n]\boldsymbol{e}_b[n]^H\}$ is the covariance matrix of the noise process for the backward AR process.

The values for the multivariate AR model can be solved efficiently with the WWRA [102, 103]. The WWRA is the extension to the multivariate case of the Levinson-Durbin algorithm. The forward and backward coefficient matrices for the

$(k+1)$th order are calculated by

$$\mathbf{A}_{k+1}(k+1) = -\Delta_{k+1}(\mathbf{Q}_k^b)^{-1} \tag{C.12}$$

$$\mathbf{B}_{k+1}(k+1) = -\nabla_{k+1}(\mathbf{Q}_k^f)^{-1} \tag{C.13}$$

where

$$\Delta_{k+1} = \sum_{i=0}^{K} \mathbf{A}_k(i)\mathbf{R}(k+1-i) \tag{C.14}$$

$$\nabla_{k+1} = \Delta_{k+1}^H \tag{C.15}$$

The update recursion are

$$\underline{\mathbf{A}}_{k+1} = \begin{bmatrix} \underline{\mathbf{A}}_k & \mathbf{0} \end{bmatrix} + \mathbf{A}_{k+1}(k+1) \begin{bmatrix} \mathbf{0} & \underline{\mathbf{B}}_k \end{bmatrix} \tag{C.16}$$

$$\underline{\mathbf{B}}_{k+1} = \begin{bmatrix} \mathbf{0} & \underline{\mathbf{B}}_k \end{bmatrix} + \mathbf{B}_{k+1}(k+1) \begin{bmatrix} \underline{\mathbf{A}}_k & \mathbf{0} \end{bmatrix} \tag{C.17}$$

The update error covariance matrices are

$$\mathbf{Q}_{k+1}^f = \mathbf{Q}_k^f + \mathbf{A}_{k+1}(k+1)\nabla_{k+1} \tag{C.18}$$

$$\mathbf{Q}_{k+1}^b = \mathbf{Q}_k^b + \mathbf{B}_{k+1}(k+1)\Delta_{k+1} \tag{C.19}$$

The initial conditions are

$$\mathbf{Q}_0^f = \mathbf{Q}_0^b = \mathbf{R}(0) \tag{C.20}$$

$$\underline{\mathbf{A}}_0 = \underline{\mathbf{B}}_0 = \mathbf{I} \tag{C.21}$$

Appendix D

Least Squares and Linear Minimum Mean Square Error Estimation

Consider a linear system mode as

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{v} \tag{D.1}$$

where $\boldsymbol{y}$ is a known vector, $\boldsymbol{x}$ is an unknown vector, $\boldsymbol{C}$ is a known transition matrix, and $\boldsymbol{v}$ is an additive white Gaussian noise (AWGN) with variance σ_v^2.

Least squares (LS) estimation is [51]

$$\hat{\boldsymbol{x}} = (\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{x} \tag{D.2}$$

Linear minimum mean square error (LMMSE) approach is optimal in minimizing the MSE of the channel estimates in the presence of AWGN. LMMSE estimation exploits the channel statistics and the signal-to-noise ratio (SNR) values. LMMSE of

the variable $\boldsymbol{x}$ is given by [1]

$$\hat{\boldsymbol{x}} = \frac{1}{\sigma_v^2}[\boldsymbol{R}_x^{-1} + \frac{1}{\sigma_v^2}(\boldsymbol{C})^H\boldsymbol{C}]^{-1}(\boldsymbol{C})^H\boldsymbol{y} \tag{D.3}$$

where $\boldsymbol{R}_x = E[\boldsymbol{x}\boldsymbol{x}^H]$, and σ_v^2 are known to the receiver.

Appendix E

Basic Kalman Filter Equations

Consider a state space model expressed as

$$\boldsymbol{x}[n] = \boldsymbol{\Phi}\boldsymbol{x}[n-1] + \boldsymbol{w}[n-1] \tag{E.1}$$

$$\boldsymbol{y}[n] = \boldsymbol{G}[n]\boldsymbol{x}[n] + \boldsymbol{v}[n] \tag{E.2}$$

where $\boldsymbol{x}[n]$ is the state vector at time n with the covariance matrix $\boldsymbol{P}[n]$, $\boldsymbol{y}[n]$ is the measurement vector; $\boldsymbol{\Phi}$ and $\boldsymbol{G}[n]$ are known transition matrix; $\boldsymbol{w}[n]$ is the state noise with zero mean and covariance $\boldsymbol{Q}[n]$, and $\boldsymbol{v}[n]$ is the measurement noise with zero mean and covariance $\boldsymbol{R}[n]$. The purpose of the Kalman filter is to use measurements and the states to estimate the values which is better than the estimate obtained by using measurements alone.

Let us choose the initial values at time $n = 0$ for Equation (E.1). From the practical point of view, the true value of the random constant has a standard normal probability distribution, so we assume the initial state $\boldsymbol{x}[0] = \boldsymbol{0}$. Similarly, we choose the initial value of the covariance of the state vector, $\boldsymbol{P}[0] = E([\boldsymbol{x}[0] - \overline{\boldsymbol{x}}[0]][\boldsymbol{x}[0] - \overline{\boldsymbol{x}}[0]]^H])$, where $\overline{\boldsymbol{x}}[0] = E[\boldsymbol{x}[0]]$. Given the initial values $\boldsymbol{x}[0]$ and $\boldsymbol{P}[0]$, basic Kalman filter equations for one iteration are [100]:

- Predict:

$$\hat{\boldsymbol{x}}\left[n|n-1\right] = \boldsymbol{\Phi}\hat{\boldsymbol{x}}\left[n-1|n-1\right] \tag{E.3}$$

$$\boldsymbol{P}\left[n|n-1\right] = \boldsymbol{\Phi}\boldsymbol{P}\left[n-1|n-1\right]\boldsymbol{\Phi}^{H} + \boldsymbol{Q} \tag{E.4}$$

- Update:

$$\hat{\boldsymbol{y}}\left[n|n-1\right] = \boldsymbol{G}[n]\hat{\boldsymbol{x}}\left[n|n-1\right] \tag{E.5}$$

$$\boldsymbol{M}\left[n|n-1\right] = \boldsymbol{G}[n]\boldsymbol{P}\left[n|n-1\right]\boldsymbol{G}^{H}[n] + \boldsymbol{R}\left[n\right] \tag{E.6}$$

$$\boldsymbol{z}\left[n\right] = \boldsymbol{y}\left[n\right] - \hat{\boldsymbol{y}}\left[n|n-1\right] \tag{E.7}$$

$$\boldsymbol{K}\left[n\right] = \boldsymbol{P}\left[n|n-1\right]\boldsymbol{G}^{H}[n]\left(\boldsymbol{M}\left[n|n-1\right]\right)^{-1} \tag{E.8}$$

$$\underline{\hat{\boldsymbol{x}}}\left[n|n\right] = \underline{\hat{\boldsymbol{x}}}\left[n|n-1\right] + \boldsymbol{K}\left[n\right]\boldsymbol{z}\left[n\right] \tag{E.9}$$

$$\boldsymbol{P}\left[n|n\right] = \left(\boldsymbol{I} - \boldsymbol{K}\left[n\right]\boldsymbol{G}[n]\right)\boldsymbol{P}\left[n|n-1\right] \tag{E.10}$$

Bibliography

[1] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmission over doubly selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[2] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[3] S. Haykin and M. Moher, *Modern Wireless Communication*. Upper Saddle River, N.J.: Prentice Hall, 2005.

[4] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," in *IEEE Commun. Surveys Tutorials*, vol. 9, no. 2, 2nd Quarter 2007, pp. 18–48.

[5] X. Cai and G. B. Giannakis, "Bounding performance and suppressing intercarrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2047–2056, Dec. 2003.

[6] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1002–1011, April 2004.

[7] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[8] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, Nov. 2004.

[9] C. Tepedelenlioğlu and G. B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Processing*, vol. 48, no. 7, pp. 2029–2043, July 2000.

[10] H. Bölcskei, J. R. W. Heath, and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna ssytems," *IEEE Trans. Signal Processing*, vol. 50, no. 1, pp. 99–109, Jan. 2002.

[11] T. Cui and C. Tellambura, "Blind receiver design for OFDM systems over doubly selective channels," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 906–917, May 2007.

[12] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, N.J.: Prentice Hall, 2002.

[13] M. McGuire and P. Wan, "Analysis of joint channel estimation and equalization using a Kalman filter," in *Proc. CCECE'06 conf.*, 2006, pp. 912–915.

[14] S. B. Weinstein, "The history of orthogonal frequency-division multiplexing," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 26–35, Nov. 2009.

[15] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[16] "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High speed physical layer in the 5 GHz brand," *IEEE Std 802.11, 1999 Edition (R2003).*

[17] "Broadband radio access networks (BRAN); HIPERLAN type 2; physical (PHY) layer, HIPERLAN *II*," 2001.

[18] "Local and metropolitan area networks part 16: Air interface for broadband wireless access systems," *IEEE Std 802.16-2004*.

[19] R. W. Chang, "Synthesis of band limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 45, pp. 1775–1797, Dec. 1966.

[20] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[21] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," no. 5, pp. 628–634, 1971.

[22] L. J. C. Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Commun. Mag.*, no. 7, pp. 665–675, 1985.

[23] B. Muquest, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zeros padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.

[24] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.

[25] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[26] Y. Li, H. Minn, N. Al-Dhahir, and A. R. Calderbank, "Pilot designs for consistent frequency-offset estimation in OFDM systems," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 864–877, May 2007.

[27] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications," *Pro.IEEE*, vol. 88, no. 5, pp. 611–640, May 2000.

[28] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[29] Z. Wang and G. B. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 707–720, Mar. 2003.

[30] X. Cai, S. Zhou, and G. B. Giannakis, "Group-orthogonal multicarrier CDMA," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 90–99, Jan. 2004.

[31] M. McGuire and M. Sima, "Parallel detection of MC-CDMA in fast fading," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5916–5927, Dec. 2009.

[32] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filering," in *Proc. ICASSP97, Munich, German*, vol. 3, Apr. 1997, pp. 1845–1848.

[33] Y. Shen and E. Martinez, "Channel estimation in OFDM systems," *Freescale Semiconductor*, pp. 1–15, 2006.

[34] Z. Wang and G. B. Giannakis, "Wireless multicarrier transmissions: Where Fourier meets Shannon," *IEEE Signal Processing Mag.*, vol. 17, pp. 29–48, May 2000.

[35] N. Benvenuto, R. Dinis, D. Falconer, and S. Tomasin, "Single carrier modulation with nonlinear frequency domain equalization: an idea whose time has come - again," *Proceedings of the IEEE*, vol. 98, no. 1, pp. 68–96, Jan. 2010.

[36] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *Proc. Globalcom99, Rio de Janeiro, Brazil*, vol. 4, 1999, pp. 2162–2166.

[37] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Processing Lett.*, vol. 10, no. 6, pp. 57–60, Mar. 2003.

[38] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Processing Lett.*, vol. 12, no. 3, pp. 226–229, Mar. 2005.

[39] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 413–415, 2003.

[40] J. K. Tugnait, X. Meng, and S. He, "Doubly selective channel estimation using superimposed training and exponential bases models," *EURASIP Journal on Applied Signal Processing*, pp. 1–11, 2006.

[41] S. He and J. K. Tugnait, "On doubly slective channels estimation using superimposed training and discrete prolate shperoidal sequences," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 3214–3228, July 2008.

[42] W.-C. Huang, C.-P. Li, and H.-J. Li, "On the power allocation and system capacity of OFDM sytems using superimposed training schemes," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1731–1740, May 2009.

[43] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, N.J: Prentice Hall PTR, 2002.

[44] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[45] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.

[46] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Processing*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[47] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138–2145, Sept. 2004.

[48] X. Cai and G. B. Giannakis, "Error probability minimizing pilots for OFDM with M-PSK modulation over Rayleigh-fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 146–155, Jan. 2004.

[49] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1403–1418, May 2004.

[50] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1615–1634, June 2003.

[51] S. M. Kay, *Fundamentals of statistical signal processing: Estimation Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1993.

[52] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. VTC*, Sep. 1995.

[53] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, 1998.

[54] M. R. Raghavendra and K. Giridhar, "Improving chcannel estimation in OFDM systems for sparse multipath channels," *IEEE Signal Processing Lett.*, vol. 12, no. 1, pp. 52–55, Jan. 2005.

[55] W. G. Jeon, K. H. Paik, and Y. S. Cho, "An efficient channel estimation technique for OFDM systems with tranmitter diversity," in *Proc. IEEE PIMRC 2000*, Sept. 2000, pp. 1246–1250.

[56] M. Morelli and U. Mengali, "A comparison of pilot-aided chanel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[57] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[58] Y. Zhao and A. Huang, "A novel channel estimation method for OFDM mobile communication systems based on pilot signals and transform-domain processing," in *Proc. IEEE VTC*, vol. 3, May 1997, pp. 2089–2093.

[59] X. Wang and K. J. R. Liu, "An adaptive channel estimation algorithm using time-frequency polynomial model for OFDM with fading multipath channels," *EURASIP Journal on Applied Signal Processing*, no. 8, pp. 818–830, Aug. 2002.

[60] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, July 1998.

[61] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, July 2000.

[62] S. Chen and T. Yao, "Intercarrier interference suppression and channel estimation for OFDM systems in time-varying frequency-selective fading channels," *IEEE Trans. Consumer Electron.*, vol. 50, no. 2, pp. 429–435, May 2004.

[63] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, 2005.

[64] T. Y. Al-Naffouri, K. M. Z. Islam, N. Al-Dhahir, and S. Lu, "A model reduction approach for OFDM channel estimation under high mobility conditions," *IEEE Trans. Signal Processing*, vol. 58, no. 4, pp. 2181–2193, Apr. 2010.

[65] T. Whitworth, M. Ghogho, and D. McLernon, "Optimal training and basis expansion model parameters for doubly-selective channel estimation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1490–1498, 2009.

[66] A. P. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block tranmissions for doubly selective channels," *IEEE Trans. Signal Processing*, vol. 56, no. 3, pp. 1148–1160, 2008.

[67] Z. Tang, R. Cannizzaro, G. Leus, and P. Baneli, "Pilot-assisted time-varying chnnel estimation for OFDM systems," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 2226–2238, May 2007.

[68] I. Barhumi and M. Moonen, "MLSE and MAP equalization for transmission over doubly selective channels," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4120–4128, Oct. 2009.

[69] K. A. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *IEEE GLOBECOM*, Dec. 2005, pp. 3503–3507.

[70] W. Qin and Q.-C. Peng, "Time-varying channel estimation and symbol detection using superimposed training in OFDM systems," in *Wireless Pers Commun*, vol. 47, Oct. 2008, pp. 293–301.

[71] H. Hijazi and L. Ros, "Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 140–151, Jan. 2009.

[72] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Processing*, vol. 53, no. 9, pp. 3597–3607, Sept. 2005.

[73] I. Barhumi, G. Leus, and M. Moonen, "Estimation and direct equalization of doubly selective channels," *EURASIP Jour. on Applied Signal Processing*, pp. 1–15, 2006.

[74] T. Cui, C. Tellambura, and Y. Wu, "Low-complexity pilot-aided channel estimation for OFDM sywtems over doubly-selective channels," in *Proc. ICC'05 Conf.*, vol. 3, May 2005, pp. 1980–1984.

[75] C. Komninakis, C. Fragouli, A. Sayed, and R. Wesel, "Mulit-input multi-output fading channel tracking snd equalization using kalman estimation," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 1065–1076, May 2002.

[76] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo codes," in *Proc. ICC'93 Conf.*, 1993, pp. 1064–1070.

[77] J. Hagenauer, E. Offer, and L. Papke, "Iterative deocding of binary block and convolutional codes," *IEEE Trans. Inform. Theory.*

[78] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 3686–3691, Oct. 2000.

[79] J. Garcia-Frias and J. D. Villasenor, "Combined turbo detection and decoding for unknown ISI channels," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 79–85, Jan. 2003.

[80] A. Hansson and T. Aulin, "Generalized APP detection of continuous phase modulation over unknown ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 10, pp. 1615–1619, Oct. 2005.

[81] L. Davis, I. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, Oct. 2001.

[82] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assited turbo codes over flat-fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1697 – 1705, Sept. 2001.

[83] B. Mielczarek and A. Svensson, "Improved iterative channel estimation and turbo decoding over flat-fading channels," in *Proc. IEEE VTC2002-fall*, vol. 2, 2002, pp. 975–980.

[84] ——, "Modeling fading channel-estimation errors in pilot-symbol-assisted systems, with application to turbo codes," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1822 – 1832, Nov. 2005.

[85] R. Otnes and M. Tücheal, "Iterative channel estimation for turbo eqalization of time-varying frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1918–1923, Nov. 2004.

[86] M. F. Flanagan and A. D. Fagan, "Iterative channel estimation, equalization, and decoding for pilot-symbol assited modulation over frequency selective fast fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1661 – 1670, July 2007.

[87] M. Nicoli, S. Ferrara, and U. Spagnolini, "Soft-iterative channel estimation: methods and performance analysis," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2993 – 3006, June 2007.

[88] J. W. Thomas Zeman, Christoph F. Mecklenbra and R. R.Muller, "Iterative joint time-variant channel estimation and multi-user detection for MC-CDMA," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1469–1478, June 2006.

[89] F. Sanzi, S. Jelting, and J. Speidel, "A comparative study iterative channel estimations for mobile OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 849–859, Sept. 2003.

[90] G. Auer and J. Bonnet, "Threshold controlled iterative channel estimation for coded OFDM," in *Proc. VTC-Spring Conf.*, 22-25 April 2007, pp. 1737–1741.

[91] I. Nevat and J. Yun, "Error propagation mitigation for iterative channel tracking, detection and decoding of BICM-OFDM systems," in *Proc. ISWCS Conf.*, 17-19 Oct. 2007, pp. 75–80.

[92] S. Stefanatos and A. k. Katsaggelos, "Joint data detection and channel tracking for OFDM systems with phase noise," *IEEE Trans. Signal Processing*, vol. 56, no. 9, pp. 4230–4243, Sept. 2008.

[93] M. Qaisrani and S. Lambotharan, "An iterative (turbo) channel estimation and symbol detection technique for doubly selective channels," in *Proc. VTC-2007 Spring Conf.*, 22-25 April 2007, pp. 2253–2256.

[94] X. Li and T. F. Wong, "Turbo equalization with nolinear Kalman filtering for time-varying frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 60, no. 2, pp. 691–700, Feb. 2007.

[95] Q. Guo, L. Ping, and D. D. Huang, "A low-complexity iterative channel estimation and detection technique for double selective channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4340–4349, Aug. 2009.

[96] H. Kim and J. K. Tugnait, "Turbo equalization for doubly-selective fading channels using nolinear Kalman filtering and basis expansion models," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2076–2087, June 2010.

[97] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[98] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, 4th ed. New Jersey: Pearson Prentice Hall, 2006.

[99] B. Sklar, *Digital Communications Fundamentals and applications*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, 2000.

[100] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

[101] S. L. Marple Jr., *Digital Spectral Analysis with Applications*. Englewood Cliffs, N.J.: Prentice Hall, 1987.

[102] P. Whittle, "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix," *Biometrika*, vol. 50(1-2), pp. 129–134, 1963.

[103] R. A. Wiggins and E. A. Robinson, "Recursive solution to the multichannel fitering problem," *J.Geophys.Res.*, vol. 70, pp. 1885–1891, April 1965.

[104] W. C. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE Press, 1994.

[105] K. E. Baddour and N. C. Beaulieu, "Accurate simulation of multiple cross-correlated rician fading channels," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1980–1987, Nov. 2004.

[106] M. McGuire and M. Sima, "Low-order Kalman filters for channel estimation," in *PACRIM 2005*, Aug. 2005, pp. 165–168.

[107] Y. Li, H. Minn, N. Al-Dhahir, and A. R. Calderbank, "Pilot designs for consistent frequency-offset estimation in OFDM systems," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 864–877, May 2007.

[108] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[109] P. Z. Peebles Jr., *Probability, Random Variables, and Random Signal Principles*, 4th ed. McGraw Hill, 2000.

[110] I. Barhumi, G. Leus, and M. Moonen, "Time-domain and frequency-domain per-tone equalization for OFDM over doubly selective channels," *Signal Processing*, vol. 84, no. 11, pp. 2055–2066, Nov. 2004.

[111] J. Choi, *Adaptive and iterative signal processing in communicstions.* Cambridge University Press, 2006.

[112] C. Douillard, A. Picart, M. JéZéquel, P. Didier, C. Berrou, and A. Glavieux, "Iterative correction of itersymbol interference: turbo-equalization," *Eur. Trans. Commun.*, vol. 6, pp. 507–511, 1995.

[113] L. Hanzo, J. P. Woodard, and P. Robertson, "Turbo decoding and detection for wireless applications," *Pro. IEEE*, vol. 95, no. 6, pp. 1178–1200, June 2007.

[114] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimising symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 3, pp. 284–287, Mar. 1974.

[115] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimalMAP decoding algorithms operating in the log domain," in *Proc. ICC'95 Conf.*, vol. 2, Jun. 1995, pp. 1009–1013.

[116] L. Zou, Q. Chang, C. Xiu, and Q. Zhang, "Channel estimation and ICI cancellation for OFDM systems in fast time-varying environments," *IEICE Trans.Commun.*, vol. E91-B, no. 4, pp. 1203–1206, April 2008.

[117] J. Choi, *Adaptive and iterative signal processing in communicstions.* Cambridge University Press, 2006.

[118] *Coded Modulation Library.* http://www.iterativesolutions.com.

[119] D. J. Young and N. C.Beaulieu, "The generation of correlated Rayleigh random variates by inverse discrete Fourier trnasform," *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 1114–1127, July 2000.

[120] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," in *AEÜ Int. J. Electron.Commun.*, vol. 54, no. 6, Nov. 2000, pp. 389–398.

[121] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[122] R. K. Michael Tüchler and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–76, May 2002.

[123] K. Li and X. Wang, "EXIT chart analysis of turbo multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 300–311, Jan. 2005.

[124] F. Schreckenbach, N. Görtz, and J. Hagenauer, "Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[125] M. Wetz, D. Huber, W. G.Teich, and J. Lindner, "Robust transmission over frequency selective fast fading channels with noncoherent turbo detection," in *VTC2009-Fall*, Sept. 2009.

[126] R. Zhang and L. Hanzo, "Three design aspects of multicarrier interleave division multiple access," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3607–3617, Nov. 2008.

[127] R. Otnes and M. Tüchler, "EXIT chart analysis applied to adaptive turbo equalization," in *Proc. Nordic Signal Processing Symposium*, 2002.

[128] F. Sanzi, S. Jelting, and J. Speidel, "A comparative study iterative channel estimations for mobile OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 849–859, Sept. 2003.

[129] S. Sand, C. Mensing, and A. Dammann, "Transfer chart analysis of iterative OFDM receivers with data aided channel esitamtion," in *Processings* 3^{rd} *COST389 workshop - enabling technologies for B3G systems*, July 2006, pp. 1009–1013.

[130] Y. Huang and J. A. Ritcey, "Joint iterative channel estimation and decoding for bit-interleaved coded modulation over correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2549–2558, Sept. 2005.

[131] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[132] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Pro. IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.

[133] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commn.*, vol. 47, no. 6, pp. 862–873, June 1999.

[134] M. Martone, "Wavelet-based separating kernels for sequence estimation with unknown rapidly time-varying channels," *IEEE Trans. Lett.*, vol. 3, no. 3, pp. 78–80, March 1999.

[135] T. Verma, S. Bilbao, and T. H. Y. Meng, "The digital prolate spheroidal window," in *ICASSP*, vol. 3, May 1996, pp. 1351– 1354.

[136] Z. Lin, R. W. McCallum, and H. Wang, "Computation and performance of the prolate-spheroidal wave function window in spectral estimation," in *ICASSP*, vol. 5, May 1996, pp. 2976– 2978.

[137] D. V. D. Ville, W. Philips, and I. Lemahieu, "On the n-dimensional extension of the discrete prolate spheroidal window," *IEEE Signal Processing Lett.*, vol. 9, no. 3, pp. 89–91, Mar. 2002.

[138] D. Slepian, "Prolate spheriodal wave functions, Fourier analysis and uncertainty–v: The discrete case," *Bell Syst. Tech. J.*, vol. 57, pp. 1371–1430, May-Aug. 1978.

[139] ——, "Some comments on fourier analysis, uncertainty and modeling," *JSTOR: SIAM Review*, vol. 25, no. 3, pp. 379–393, July 1983.

[140] M. Visintin, "Karhunen-loève expansion of a fast Rayleigh fading process."

[141] K.-W. Yip and T.-S. Ng, "Karhunen-loève expansion of the WSSUS channel output and its application to efficient simulation," *IEEE J. Select. Areas Commun.*, vol. 15, no. 4, pp. 640–646, May 1997.

[142] H. Senol, H. A. Cirpan, and E. Panayirci, "A low-complexity KL expansion-based channel estimation for OFDM systems," *EURASIP Journal on Wireless Commun. Netw.*, no. 2, pp. 163–174, Feb. 2005.

[143] H. S. Wang and P.-C. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, May 1996.