Optimal Detection with Imperfect Channel Estimation for Wireless Communications

This thesis is submitted in partial fulfilment of the requirements for

Doctor of Philosophy (Ph.D.)

Junruo Zhang

Communications Research Group

Department of Electronics

University of York

September 2009

Abstract

In communication systems transmitting data through unknown fading channels, traditional detection techniques are based on channel estimation (e.g., by using pilot signals), and then treating the estimates as perfect in a minimum distance detector. In this thesis, we derive and investigate an optimal detector that does not estimate the channel explicitly but jointly processes the received pilot and data symbols to recover the data. This optimal detector outperforms the traditional detectors (*mismatched detectors*). In order to approximate correlated fading channels, such as fast fading channels and frequency-selective fading channels, basis expansion models (BEMs) are used due to high accuracy and low complexity.

There are various BEMs used to represent the time-variant channels, such as Karhunen-Loeve (KL) functions, discrete prolate spheroidal (DPS) functions, generalized complex exponential (GCE) functions, B-splines (BS), and the others. We derive the mean square error (MSE) of a generic BEM-based linear channel estimator with perfect or imperfect knowledge of the Doppler spread in time-variant channels. We compare the performance and complexity of minimum mean square error (MMSE) and maximum likelihood (ML) channel estimators using the four BEMs, for the case with perfect Doppler spread. Although all BEM-based MMSE estimators allow achievement of the optimal performance of the Wiener solution, the complexity of estimators using KL and DPS BEMs is significantly higher than that of estimators using BS and GCE BEMs. We then investigate the sensitivity of BEM-based estimators to the mismatched Doppler spread. All the estimators are sensitive to underestimation of the Doppler spread but may be robust to overestimation. The results show that the traditional way of estimating the fading statistics and generating the KL and DPS basis functions by using the maximum Doppler spread will lead to a degradation of the performance. A better performance can be ob-

tained by using an overestimate of the Doppler spread instead of using the maximum Doppler spread. For this case, due to the highest robustness and the lowest complexity, the best practical choice of BEM is the B-splines.

We derive a general expression for optimal detection for pilot-assisted transmission in Rayleigh fading channels with imperfect channel estimation. The optimal detector is specified for single-input single-output (SISO) Rayleigh fading channels. The slow (time-invariant) fading channels and fast (time-variant) fading channels following Jakes' model are considered. We use the B-splines to approximate the channel gain time variations and compare the detection performance of the optimal detector with that of different mismatched detectors using ML or MMSE channel estimates. Furthermore, we investigate the detection performance of an iterative receiver implementing the optimal detector in the initial iteration and mismatched detectors in following iterations in a system transmitting turbo-encoded data. Simulation results show that the optimal detection outperforms the mismatched detection with ML channel estimation. However, the improvement in the detection performance compared to the mismatched detection with the MMSE channel estimation is modest. We then extend the optimal detector to channels with more unknown parameters, such as spatially correlated MIMO Rayleigh fading channels, and compare the performance of the optimal detector with that of mismatched detectors. Simulation results show that the benefit in detection performance caused by using the optimal detector is not affected by the spatial correlation between antennas, but becomes more significant when the number of antennas increases.

This optimal detector is extended to the case of orthogonal frequency-division multiplexing (OFDM) signals in frequency-selective fading channels. We compare the performance and complexity of this optimal detector with that of mismatched detectors using ML and MMSE channel estimates in SISO and MIMO channels. In SISO systems, the performance of the optimal detector is close to that of the mismatched detector with MMSE channel estimates. However, the optimal detector significantly outperforms the mismatched detectors in MIMO channels.

Acknowledgements

I would firstly like to gratefully acknowledge my supervisor Yuriy Zakharov for his helpful support and constant encouragement. I have learned an incredible amount from him throughout my Ph.D. research. He has been actively interested in my work and has always been available to advise me. I am deeply benefited from his motivation, enthusiasm, preciseness, patience, and immense knowledge in signal processing that, taken together, make him a great mentor.

I would also like to thank all my colleagues in the Communications Research Group for being the surrogate family during the period of my study.

This thesis is dedicated to my parents. I am forever indebted to them for their understanding, endless encouragement and patience. Their unconditional love is the source of my strength which is going to drive and motivate me forever.

Finally, I will dedicate this thesis to my grandfather. I know you are always loving and encouraging me in heaven.

Declaration

Some of the research presented in this thesis has resulted in some publications. These publications are listed at the end of Chapter 1.

All work presented in this thesis as original is so, to the best knowledge of the author. References and acknowledgements to other researchers have been given as appropriate.

Glossary

AWGN   **A**dditional **W**hite **G**aussian **N**oise

BEM   **B**asis **E**xpansion **M**odel

BER   **B**it-**E**rror-**R**ate

BS   **B**-**S**plines

BPSK   **B**inary **P**hase **S**hift **K**eying

CE   **C**omplex **E**xponential

CP   **C**yclic **P**refix

dB   Decibel

DPS   **D**iscrete **P**rolate **S**pheroidal

FEC   **F**orward **E**rror **C**orrection

Hz   **H**ertz

ICI   **I**nter-**C**arrier **I**nterference

ISI   **I**nter-**S**ymbol **I**nterference

IDFT   **I**nverse **D**iscrete **F**ourier **T**ansform

GCE   **G**eneralized **C**omplex **E**xponential

KL   **K**arhunen-**L**oeve

LLR   **L**og-**L**ikelihood **R**atio

MAP   **M**aximum **A** **P**osteriori

MIMO   **M**ultiple-**I**nput **M**ultiple-**O**utput

ML   **M**aximum **L**ikelihood

MLSE   **M**aximum **L**ikelihood **S**equential **E**stimation

MMSE   **M**inimum **M**ean **S**quare **E**rror

MSE   **M**ean **S**quare **E**rror

NSC   **N**on-**S**ystematic **C**onvolutional

OFDM   **O**rthogonal **F**requency **D**ivision **M**ultiplexing

PCI   **P**erfect **C**hannel **I**nformation

PDF   **P**robability **D**ensity **F**unction

PSAM   **P**ilot **S**ymbol **A**ssisted **M**odulation

PSK   **P**hase-**S**hift **K**eying

QAM   **Q**uadrature **A**mplitude **M**odulation

rms   **r**oot-**m**ean-**s**quare

RSC   **R**ecursive **S**ystematic **C**onvolutional

SIHO	**S**oft-**I**nput **H**ard-**O**utput
SISO	**S**oft-**I**nput **S**oft-**O**utput
SISO	**S**ingle-**I**nput **S**ingle-**O**utput
SNR	**S**ignal to **N**oise **R**atio
STBC	**S**pace-**T**ime **B**lock **C**odes
STTC	**S**pace-**T**ime **T**rellis **C**odes
SVD	**S**ingular **V**alue **D**ecomposition

Contents

Bibliography   **1**

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1 Overview

Many wireless communication techniques and components require knowledge of the channel state to achieve their optimal performance. In practice, this knowledge is often acquired by estimation. The estimation can be performed blindly by using only unknown data symbols, but more frequently, it is performed with the aid of pilot symbols which are known at the receiver side. Although occupying transmission bandwidth and energy, pilot-based channel estimation and detection offers reliable performance with a relatively low complexity, especially for time-variant or frequency-selective fading channels. Therefore, pilot symbol assisted modulation (PSAM) is widely proposed to detect data symbols in fading channels by inserting known pilot symbols into data blocks [1–18]. In this thesis, we investigate the channel estimation and data symbol detection techniques

in PSAM systems in Rayleigh fading channels such as time-invariant flat fading channel, time-variant flat fading channel and frequency selective fading channel. Specifically, we define that a time-invariant fading channel is quasi-stationary, which indicates that in each transmission block, the channel coefficients are constant over all symbols but obey Rayleigh fading between different blocks.

Accurately estimating time-variant and/or frequency-selective fading channels is a challenge and the estimation results affect the system performance. In order to approximate the channel coefficients at data positions by using pilot symbols, basis expansion models (BEMs) are widely used, due to their reliable performance and lower complexity than the Wiener filter [19]. For example, with a BEM, estimation of a realization of the random process describing the time-variant channel is transformed into estimation of a few time-invariant expansion coefficients [20]. There are different BEMs, such as complex exponential (CE) model [19, 21–24], generalized complex exponential (GCE) model [25], B-splines (BS) [26–28], discrete prolate spheroidal (DPS) basis functions [20, 29, 30] and Karhunen-Loeve (KL) basis functions [31, 32] to model correlated fading channels. In this thesis, the BEM-based channel estimators are investigated in time-variant Rayleigh fading channels following Jakes' model. We derive mean square error (MSE) of minimum mean square error (MMSE) and maximum likelihood (ML) channel estimators based on different BEMs, and compare their performance and complexity for the case with perfect and/or inaccurate knowledge of the Doppler spread. Based on this comparison, the estimator using B-splines is chosen and applied to approximate the time-variant channel in this thesis.

Due to noise and to the finite number of pilot symbols in a transmission block, the channel estimate is not perfect. In [33, 34], the effects of channel estimation errors on the detection performance of PSAM systems were evaluated. However, most of works in [1–18] consider a traditional minimum distance detector which suffers an extra error on detection performance by treating channel estimates as perfect. In order to achieve better detection performance, optimal detection with imperfect channel estimates in communication systems with PSAM was proposed and investigated in [35, 36]. The optimal detector does not estimate the channel explicitly, but jointly processes received pilot and data symbols to recover the data. The optimal detector in [35] is obtained for communication scenarios in channels with uncorrelated fading and white Gaussian noise, and

its performance is compared with a minimum distance detector (*mismatched detector*) using ML channel estimates. In [36], the performance of optimal and mismatched detectors in single-input single-output (SISO) channels with time-variant Rayleigh fading was investigated. In this thesis, we derive a generic optimal detector and apply it for different scenarios, i.e., time-variant flat channels obeying the Clarke's model, time invariant frequency-selective channels, and spatially correlated multiple-input multiple-output (MIMO) channels, and compare its performance with mismatched detectors using ML, regularized-ML and MMSE channel estimates. We obtain this optimal detector for the case when the channel gain time variations and channel frequency response are approximated by using BEMs.

It is well known that the estimation of time variations in time-variant channels are very challenging at low signal-to-noise ratio (SNR). Our solution is to apply forward error correcting (FEC) channel codes, such as turbo codes and iterative channel estimation/detection schemes by feeding the output information of the FEC decoder back to the channel estimator or detector. In this thesis, we compare the performance of iterative receivers applying ML, regularized-ML and MMSE channel estimation with soft-input hard-output and soft-input and soft-output turbo decoding schemes. We also investigate the iterative receiver implementing the optimal detector, and compare its bit-error-rate (BER) performance with that of iterative receivers applying mismatched detectors.

1.2 Contributions

Major contributions in this thesis can be summed up as follows:

- MSE of a generic BEM-based linear channel estimator for time-variant fading channels has been derived. The MSE performance and complexity of estimators using different BEMs have been compared in cases with perfect and inaccurate knowledge of the Doppler spread. The estimators have been shown to be very sensitive to underestimation of the Doppler spread but may have little sensitivity to overestimation. The estimation using a slight overestimate of the Doppler spread to calculate the fading statistics and generate the basis functions can significantly out-

perform the estimation using the maximum Doppler spread. The B-splines have been shown to be the best practical choice for BEM providing good performance and low complexity.

- The optimal detection has been derived for general correlated fading channels. The optimal detection is shown to outperform mismatched detection with ML and regularized ML channel estimation. In SISO Rayleigh fading channels, when QAM signals are transmitted, the performance of the mismatched detection with MMSE estimation is shown to be close to that of the optimal detection.

- It has been proved that the symbol-by-symbol optimal detection of PSK symbols in spatial uncorrelated SIMO Rayleigh fading channels is equivalent to the mismatched detection with the MMSE channel estimation.

- The optimal detector has been specified for MIMO Rayleigh fading channels. The optimal detector has been shown to significantly outperform mismatched detectors when the number of antennas increases.

- The optimal detection has been specified for orthogonal frequency division multiplexing (OFDM) transmission in SISO and MIMO frequency-selective fading channels. The optimal detector has been shown to significantly outperform mismatched detectors when the number of antennas increases.

- The performance of an iterative receiver incorporating the optimal detector with soft-input soft-output turbo decoder has been investigated. The iterative receiver applying the optimal detector in the initial iteration has been shown to outperform iterative receivers applying mismatched detectors in all iterations.

1.3   Thesis Outline

The rest of the report is separated into following chapters, according to the different systems investigated and analyzed.

- Chapter 2: Fundamental Techniques

In this chapter, fundamental techniques used throughout this thesis are introduced. We firstly compare different simulators of time-variant channels and apply the one whose statistics match to those of the desired reference Clarke's model. We also describe the basic principles of BEMs, which are used to approximate the fading channels. Turbo encoder and decoder are also briefly introduced.

- Chapter 3: Basis expansion model based channel estimation of time-varying channels

  In this chapter, we investigate the pilot assisted channel estimators based on BEMs in time-variant Rayleigh fading channels. We derive the MSE of a generic linear channel estimator with a linearly independent BEM. We also compare the performance and complexity of ML and MMSE estimators using different BEMs, such as KL, DPS, GCE and BS BEMs for the cases with perfect and inaccurate knowledge of the Doppler spread.

- Chapter 4: Optimal and mismatched detection in SISO frequency-flat Rayleigh fading channels with imperfect channel estimation

  This chapter presents the basic principles of the pilot assisted optimal detection which does not require estimating the channel explicitly but jointly processes the received data and pilot symbols to recover the data with minimum error. We derive a generic optimal detector, and compare its performance with that of mismatched detectors in single-input single output (SISO) time-invariant fading channels. We then extend the optimal detector to the case of time-variant channels and use B-splines as basis functions to approximate the time variations of the channel gain. The comparison of bit-error-rate (BER) and MSE performance between iterative receivers applying optimal detector and mismatched detectors is also presented.

- Chapter 5: Optimal and mismatched detection in MIMO frequency-flat Rayleigh fading channels with imperfect channel estimation

  In this chapter, we firstly specify the optimal detector for spatially correlated MIMO time-invariant Rayleigh fading channels and investigate the benefit caused by using the optimal detector. We then extend the optimal detector to MIMO time-variant fading channels with temporal fading correlation following Jakes' model and compare its detection performance with that of mismatched detectors. We also prove that the optimal symbol-by-symbol detector in spatially uncorrelated single-input

multiple-output (SIMO) channels with PSK modulation is equivalent to the mismatched detector with MMSE channel estimates.

- Chapter 6: Optimal and mismatched detection of OFDM signals in MIMO frequency-selective time-invariant fading channels with imperfect channel estimation

In this chapter, we specify the optimal detector for OFDM signals in SISO and MIMO frequency-selective fading channels and compare its performance with that of mismatched detectors. We compare the complexity of different BEMs and investigate their performance of approximating the channel frequency response. We also investigate the performance of iterative receivers incorporating the optimal detector in the initial iteration for turbo coded transmission in SISO channels, and compare the performance of the optimal detector with that of the mismatched detectors.

1.4 Notations

In this thesis, we use capital and small bold fonts to denote matrices and vectors, i.e., $\mathbf{A}$ and $\mathbf{a}$, respectively. Elements of the matrix and vector are denoted as $A_{m,n} = [\mathbf{A}]_{m,n}$ and $a_m = [\mathbf{a}]_m$. The symbol j is an imaginary unit $j = \sqrt{-1}$. We denote $\Re\{\cdot\}$ and $\Im\{\cdot\}$ as the real and imaginary components of a complex number, respectively; $(\cdot)^*$ denotes complex conjugate; $\mathbf{I}_Q$ denotes an $Q \times Q$ identity matrix; $(\cdot)^T$ and $(\cdot)^H$ denote matrix transpose and Hermitian transpose, respectively. $\otimes$ denotes the Kronecker product. $\lceil \cdot \rceil$ denotes the smallest integer. $E\{\cdot\}$ denotes the statistical expectation operator and $\mathrm{tr}\{\cdot\}$ denotes the trace operator.

1.5 Publication List

Journal Papers [37, 38]

1. J. Zhang, Y. V. Zakharov, and V. M. Baronkin, "Optimal detection in MIMO OFDM

systems with imperfect channel estimation", accepted by IET Communications, 2009.

2. Y. V. Zakharov, V. M. Baronkin, and J. Zhang, "Optimal and mismatched detection of QAM signals in fast fading channels with imperfect channel estimation", IEEE Trans. on Wireless Commun., vol. 8, no. 2, pp. 617-621, 2009.

3. J. Zhang, R. N. Khal and Y. V. Zakharov, "The sensitivity of channel estimators using basis expansion models to the mismatched Doppler frequency", under revision in IET Communications, 2009.

Conference Papers [39–48]

1. J. Zhang and Y. V. Zakharov, "Iterative B-spline estimator using superimposed training in doubly-selective fading channels", 41st Asilomar Conf. Signals, Systems and Computers, ACSSC 2007, Pacific Grove, CA, US, Nov., 2007.

2. Y. V. Zakharov, V. M. Baronkin, and J. Zhang, "Optimal detection of QAM signals in fast fading channels with imperfect channel estimation", ICASSP'2008, 3205–3208, Las Vegas, USA, March, 2008.

3. J. Zhang, V. M. Baronkin and Y. V. Zakharov, "Optimal detector of OFDM signals for imperfect channel estimation", EUSIPCO'2008, Lausanne, Switzerland, Aug., 2008.

4. J. Zhang, Y. V. Zakharov, and V. M. Baronkin, "Optimal detection in MIMO Rayleigh fast fading channels with imperfect channel estimation", 42st Asilomar Conf. Signals, Systems, and Computers, ACSSC 2008, Pacific Grove, CA, US, Oct., 2008.

5. R. N. Khal, Y. V. Zakharov, and J. Zhang, "Joint channel and frequency offset estimators for frequency-flat fast fading channels", 42st Asilomar Conf. Signals, Systems, and Computers, ACSSC 2008, Pacific Grove, CA, US, Oct., 2008.

6. J. Zhang, Y. V. Zakharov, and R. N. Khal, "Optimal detection for STBC MIMO systems in spatially correlated Rayleigh fast fading channels with imperfect channel estimation", 43st Asilomar Conf. Signals, Systems, and Computers, ACSSC 2009, Pacific Grove, CA, US, 2009.

7. R. N. Khal, J. Zhang, and Y. V. Zakharov, "Robustness of joint Bayesian frequency offset and channel estimation based on basis expansion models", 43st Asilomar Conf. Signals, Systems, and Computers, ACSSC 2009, Pacific Grove, CA, US, 2009.

8. J. Zhang, and Y. V. Zakharov, "Optimum detection in spatially uncorrelated SIMO Rayleigh fast fading channels with imperfect channel estimation", under revision by ICASSP'2010, Dallas, USA.

9. J. Zhang, R. N. Khal, and Y. V. Zakharov, "Sensitivity of MMSE channel estimator with B-splines to the mismatched Doppler frequency", under revision by ICASSP'2010, Dallas, USA.

10. R. N. Khal, Y. V. Zakharov, and J. Zhang, "B-spline based joint channel and frequency offset estimation in doubly-selective fading channels", under revision by ICASSP'2010, Dallas, USA.

Chapter 2

Fundamental Techniques

Contents

In this chapter, fundamental techniques used throughout this thesis are introduced: simulators of time-variant fading channels, BEMs and turbo codes.

2.1   Simulator of time-variant fading channels

In this thesis, we will investigate the channel estimation and signal detection in time-variant Rayleigh fading channels. Before comparing the performance of different estimation and detection schemes, we should firstly model and simulate the fading channel accurately. This section introduces a simulator of time-variant Rayleigh fading channels, which is used in the subsequent chapters.

After 1960's, Clarke's model [49] and its simplified version by Jakes [50] are widely used to simulate time-variant Rayleigh fading channels. Although the simplicity of the

---

original Jakes' model makes it popular, there are two deficiencies that can not be ig-
nored [51]: the original Jakes' model is a deterministic model and it is difficult to generate
the multiple independent fading channels, such as frequency-selective (multipath) fading
and MIMO channels. Various modifications [52–55] and improvements [51, 56, 57] have
been reported for generating multiple uncorrelated fading waveforms needed for mod-
eling frequency selective fading and MIMO channels, such as Inverse Discrete Fourier
Transform (IDFT) [58] and the autoregressive approach [59]. It is pointed in [60] that
Jakes' simulator is not wide-sense stationary when averaged across the physical ensem-
ble of fading channels. In [60], an improved simulator, named Pop-Beaulieu simulator,
is applied to remove this stationarity problem by introducing random phase shifts in the
low-frequency oscillators. However, it is shown that the Pop-Beaulieu simulator has defi-
ciencies in some of its high-order statistics [57].

Based on the Pop-Beaulieu simulator, novel sum-of-sinusoids statistical simulation
models with small number of sinusoids are proposed for Rayleigh fading channels
in [51, 57]. These modified models improve the original Jakes' model by introducing ran-
dom path gain, random initial phase and random Doppler frequency for sinusoids within
these models [57]. The high-order statistical properties of these novel models, such as
the autocorrelations and cross-correlations of the quadrature components, the autocorre-
lation of the complex envelop, and the probability density functions (PDFs) of the fading
envelop, asymptotically approach the desired ones as the number of sinusoids approaches
infinity [51, 57].

In this section, we introduce the reference Clarke's model mathematically and analyze
the deficiencies of the Jakes' model and the Pop-Beaulieu model. Then, we introduce a
modified model proposed in [51, 57] which provides good convergence of the probability
density functions of the envelope, the level crossing rate, the average fading duration, and
the autocorrelation of the squared fading envelope, even when the number of sinusoids is
as small as 8 [57]. This modified model is used to generate multiple independent time-
variant channels in this thesis.

2.1.1 The reference model and its simplifications

Clarke's model serves as a mathematical reference model for the other sum-of-sinusoid simulation models. This model assumes that the field incident on the wireless receiver consists of a number of azimuthal plane waves with arbitrary carriers phases, arbitrary arrival angles and equal average amplitude [49]. A low-pass fading process can be used to describe a frequency-flat fading channel containing N propagation channels as

$$g(t) = E_0 \sum_{n=1}^{N} C_n \exp\left[j(\omega_d t \cos\alpha_n + \phi_n)\right], \tag{2.1}$$

where E_0 is a constant scaling the fading energy, C_n, α_n and ϕ_n are the random path gain, arrival angle of incoming waves and initial phase corresponding to the n-th propagation channel; $\omega_d = 2\pi\nu$ is the maximum angular Doppler frequency, where ν is the maximum Doppler frequency, which depends on the motion velocity v, the carrier frequency f_C. The Doppler frequency can be calculated by

$$\nu = \frac{v f_C}{c_0} \tag{2.2}$$

where c_0 is the speed of light. For example, we consider a system operating at carrier frequency $f_C = 2\text{GHz}$, with the user moving with velocity $v = 30\text{m/s}$, and symbol duration $10^{-4}s$. Based on these parameters, the normalized Doppler spread is $\nu T_s = 0.02$.

The Doppler frequency of the n-th propagation channel is calculated by

$$\nu_n = \nu \cos\alpha_n. \tag{2.3}$$

Both α_n and ϕ_n are uniformly distributed over $[-\pi, \pi)$ for all n and they are mutually independent.

In complex form, (2.1) can be decomposed as

$$g(t) = g_r(t) + j g_i(t), \tag{2.4}$$

where

$$g_r(t) = \sqrt{E_0} \sum_{n=1}^{N} C_n \cos\left(\omega_d t \cos\alpha_n + \phi_n\right) \tag{2.5}$$

and

$$g_i(t) = \sqrt{E_0} \sum_{n=1}^{N} C_n \sin\left(\omega_d t \cos\alpha_n + \phi_n\right). \tag{2.6}$$

When N is large, $g_r(t)$ and $g_i(t)$ can be modeled as Gaussian random processes according to the central limit theorem [50]. The statistics for fading simulators, such as autocorrelation, cross-correlation functions and are given by

$$
\begin{aligned}
R_{g_r g_r}(\tau) &= E\left\{g_r(t)g_r(t+\tau)\right\} = J_0(\omega_d \tau), \\
R_{g_i g_i}(\tau) &= J_0(\omega_d \tau), \\
R_{g_r g_i}(\tau) &= R_{g_i g_r} = 0, \\
R_{gg}(\tau) &= 2J_0(\omega_d \tau), \\
R_{|g|^2 |g|^2}(\tau) &= 4 + 4J_0^2(\omega_d \tau),
\end{aligned}
\tag{2.7}
$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind. For simplicity, we set $E_0 = \sqrt{2}$ and $\sum_1^N E\{C_n^2\} = 1$. For Clarke's model, the fading envelope $|g(t)|$ is Rayleigh distributed while the phase $\Theta_g(t) = \arctan[g_r(t), g_i(t)]$ is uniformly distributed [49], i.e.

$$
f_{|g|}(x) = x \exp(-\frac{x^2}{2}), \quad x \geq 0
\tag{2.8}
$$

and

$$
f_{\Theta_g}(\theta_g) = \frac{1}{2\pi}, \quad \theta_g \in [-\pi, \pi).
\tag{2.9}
$$

Jakes' model is well known as a simplified model of the Clarke's model. If the phase, amplitude and arrival angle for each incoming propagation channel are fixed, Clarke's model is transformed to Jakes' model. Specifically, the following parameters are set

$$
\begin{aligned}
C_n &= \frac{1}{\sqrt{N}}, \quad n = 1, 2, \ldots, N, \\
\alpha_n &= \frac{2\pi n}{N}, \quad n = 1, 2, \ldots, N, \\
\phi_n &= 0, \quad n = 1, 2, \ldots, N.
\end{aligned}
\tag{2.10}
$$

The normalized low-pass fading processes of this model are given by

$$
\begin{aligned}
\mu(t) &= \mu_r(t) + j\mu_i(t), \\
\mu_r(t) &= \frac{2}{\sqrt{N}} \sum_{n=0}^{M} a_n \cos(\omega_n t), \\
\mu_i(t) &= \frac{2}{\sqrt{N}} \sum_{n=0}^{M} b_n \cos(\omega_n t),
\end{aligned}
\tag{2.11}
$$

where $N = 4M + 2$, and

$$
a_n = \begin{cases} \sqrt{2}\cos\beta_0, & n = 0, \\ 2\cos\beta_n, & n = 1, 2, \ldots, M, \end{cases}
$$

$$
b_n = \begin{cases} \sqrt{2}\sin\beta_0, & n = 0, \\ 2\sin\beta_n, & n = 1, 2, \ldots, M, \end{cases}
$$

$$
\beta_n = \begin{cases} \frac{\pi}{4}, & n = 0, \\ \frac{\pi n}{M}, & n = 1, 2, \ldots, M, \end{cases} \tag{2.12}
$$

$$
\omega_n = \begin{cases} \omega_d, & n = 0, \\ \omega_d \cos\frac{2\pi n}{N}, & n = 1, 2, \ldots, M. \end{cases}
$$

The simplification in (2.10) makes this simulation model deterministic [52, 53]. In [60], it is shown that the statistical variance of the Jakes' simulator fading process is time variant and therefore, Jakes' model averaged across the ensemble of physical fading channels is wide-sense nonstationary. Various approaches are applied to conquer these deficiencies [54, 55, 58–61]. Among these approaches, the Pop-Beaulieu simulator introduced in [60] is wide-sense stationary and widely used as the foundation of further researches on the simulators.

The normalized low-pass fading process of the Pop-Beaulieu simulator is given by

$$
f(t) = f_r(t) + jf_i(t), \tag{2.13}
$$

where

$$
f_r(t) = \frac{2}{\sqrt{N}} \sum_{n=0}^{M} a_n \cos\left(\omega_n t + \phi_n\right) \tag{2.14}
$$

and

$$
f_i(t) = \frac{2}{\sqrt{N}} \sum_{n=0}^{M} b_n \sin\left(\omega_n t + \phi_n\right), \tag{2.15}
$$

where a_n and b_n are the same as those defined in (2.12). It is clear that the Pop-Beaulieu simulator adds ϕ_n, a random phase uniformly distributed on $[-\pi, \pi)$, to the original Jakes' model which assumes that $\phi_n = 0$ for all n. The introduction of the random ϕ_n allows the Pop-Beaulieu simulator becoming wide-sense stationary. However, some problems with high order statistics remain [51].

The autocorrelation and cross-correlation functions of the Pop-Beaulieu simulator are given by [62]

$$R_{f_r f_r}(\tau) = \frac{4}{N} \left[\sum_{n=0}^{M} \frac{a_n^2}{2} \cos(\omega_n \tau) \right],$$

$$R_{f_i f_i}(\tau) = \frac{4}{N} \left[\sum_{n=0}^{M} \frac{b_n^2}{2} \cos(\omega_n \tau) \right],$$

$$R_{f_r f_i}(\tau) = \frac{4}{N} \left[\sum_{n=0}^{M} \frac{a_n b_n}{2} \cos(\omega_n \tau) \right], \tag{2.16}$$

$$R_{f_i f_r}(\tau) = R_{f_r f_i}(\tau),$$

$$R_{ff}(\tau)(\tau) = \frac{4}{N} \left[\sum_{n=0}^{M} 2 \cos(\omega_n \tau) + \cos(\omega_d \tau) \right],$$

$$R_{|f|^2 |f|^2}(\tau) = 4 + 2R_{f_r f_r}^2(\tau) + 4R_{f_r f_i}^2 + \frac{8}{N} J_0(2\omega_d \tau) + \frac{16(N-1)}{N^2}.$$

By comparing (2.16) with (2.7), it is clear that the second-order statistics $[R_{f_r f_r}(\tau), R_{f_i f_i}(\tau), R_{f_r f_i}(\tau), R_{f_i f_r}(\tau)]$ of the Pop-Beaulieu simulator approach those of the desired Clarke's model only if M is infinite. When M is finite, these second-order statistics will significantly deviate from the desired values [51]. Moreover, even if M is infinite, the higher-order statistics $\left[R_{ff}(\tau), R_{|f|^2 |f|^2}(\tau) \right]$ can not match to the desired ones [62].

In order to overcome these deficiencies, an improved simulation model, whose statistical properties can perfectly match the desired Clarke's model, is introduce by Zheng and Xiao in [51, 57], and we will describe this improved model in the next section.

2.1.2 An improved simulation model

An improved simulation model proposed in [51, 57] solves the deficiencies of Jakes' model by reintroducing the randomness of the three variables C_n, α_n and ϕ_n. The normalized low-pass fading process of the model is defined as

$$h(t) = \sqrt{E_0} \sum_{n=1}^{N} \hat{C}_n \exp[j(\omega_d t \cos \hat{\alpha}_n + \phi_n)], \tag{2.17}$$

and

$$\hat{C}_n = \frac{\exp(j\psi_n)}{\sqrt{N}}, \quad n = 1, 2, \ldots, N, \tag{2.18}$$

$$\hat{a}_n = \frac{2\pi n - \pi + \theta}{N}, \quad n = 1, 2, \ldots, N. \tag{2.19}$$

It should be clarified that $N/2$ is an integer, and ψ_n, θ, and ϕ_n are mutually independent random variables uniformly distributed on $[-\pi, \pi)$ [51, 57]. By substituting (2.18) into (2.17), we obtain the improved simulation model as

$$h(t) = \frac{\sqrt{E_0}}{\sqrt{N}} \left\{ \sum_{n=1}^{N/2} e^{j\psi_n} \left[e^{j(\omega_d t \cos \hat{a}_n + \phi_n)} + e^{-j(\omega_d t \cos \hat{a}_n + \phi_n)} \right] \right\}, \tag{2.20}$$

in which $e^{j(\omega_d t \cos \hat{a}_n + \phi_n)}$ represents the waves with Doppler frequencies from the range $[\omega_d \cos(2\pi/N), \omega_d]$ to the range $[-\omega_d \cos(2\pi/N), -\omega_d]$, while $e^{-j(\omega_d t \cos \hat{a}_n + \phi)}$ represents the waves with Doppler frequencies from the range of $[-\omega_d \cos(2\pi/N), -\omega_d]$ to $[\omega_d \cos(2\pi/N), \omega_d]$. The Doppler frequencies are overlapped [51]. Equation (2.20) can be further simplified to be

$$h(t) = \frac{\sqrt{E_0}}{\sqrt{N}} \left\{ \sum_{n=1}^{M} \sqrt{2} e^{j\psi_n} \left[e^{j(\omega_n t + \phi_n)} + e^{-j(\omega_n t + \phi_n)} \right] \right\}, \tag{2.21}$$

where $M = N/4$, and $\omega_n = \omega_d \cos \hat{a}_n$. A new simulation model can be defined based on (2.21) as

$$h(t) = h_r(t) + j h_i(t) \quad,$$

$$h_r(t) = \sqrt{\frac{2}{M}} \sum_{n=1}^{M} \cos(\psi_n) \cos \left[\omega_d t \cos(\alpha_n) + \phi_n \right] \quad, \tag{2.22}$$

$$h_i(t) = \sqrt{\frac{2}{M}} \sum_{n=1}^{M} \sin(\psi_n) \cos \left[\omega_d t \cos(\alpha_n) + \phi_n \right] \quad,$$

where

$$\alpha_n = \frac{2\pi n - \pi + \theta}{4M}, \quad n = 1, \ldots, M, \tag{2.23}$$

and θ, ϕ_n, φ_n are statistically independent and uniformly distributed on $[-\pi, \pi)$. In [51], the value of ϕ_n has been chosen to be the same for all n, which is incorrect. This leads to a mistake on the probability density function of the time-invariant fading envelop where $\omega_d = 0$ [20]. Here we follow the corrected version used in [57] and reintroduce the randomness of ϕ_n. Therefore, ψ_n and ϕ_n can be combined together and (2.21) can be further simplified as

$$h(t) = \frac{\sqrt{E_0}}{\sqrt{N}} \left\{ \sum_{n=1}^{M} \sqrt{2} \left[e^{j(\omega_n t + \chi_n)} + e^{-j(\omega_n t + \chi_n)} \right] \right\}, \tag{2.24}$$

where $\chi_n = (\psi_n + \phi_n)$ and the PDF of χ is the convolution of the density functions of ψ_n and ϕ.

The statistics of this simulation model are derived in [51] as

$$
\begin{aligned}
R_{h_r h_r}(\tau) &= J_0(\omega_d \tau), \\
R_{h_i h_i}(\tau) &= J_0(\omega_d \tau), \\
R_{h_r h_i}(\tau) &= R_{h_i h_r} = 0, \\
R_{hh}(\tau) &= 2J_0(\omega_d \tau), \\
R_{|h|^2 |h|^2}(\tau) &= 4 + 4J_0^2(\omega_d \tau), \text{ if } M \text{ is infinite.}
\end{aligned}
\tag{2.25}
$$

It is clear that except the autocorrelation function of the squared envelop $R_{|h|^2 |h|^2}(\tau)$, the statistics of this improved model do not depend on the value of M, and exactly the same as the desired statistics of Clarke's model described by (2.7). Furthermore, the high-order statistic $R_{|h|^2 |h|^2}(\tau)$ asymptotically approaches the desired autocorrelation $R_{|g|^2 |g|^2}(\tau)$ when M increases. Numerical results in [51] show that a good approximation has been observed when M is as small as 8.

In order to evaluate the improved fading simulator, we compare its simulation performance with analytical results of the corresponding mathematical reference model. We set that $M = 8$, and the normalized Doppler frequency $\nu T_s = 0.02$, where T_s is the duration of a transmitted symbol. The simulation results are based on ensemble averages of 100 and 1000 random trials.

Firstly, we consider the case of a time-variant channel. Fig. 2.1 and Fig. 2.2 show simulation results for autocorrelations of real and imaginary components of the fading, respectively, and Fig. 2.3 shows the cross-correlation of the real and imaginary parts of the fading. The reference is calculated based on (2.7) for the purpose of comparison. Note that $R_{h_r h_i}$ is almost the same as $R_{h_i h_r}$, therefore, only $R_{h_i h_r}$ is shown here.

It is observed that the simulated autocorrelations and cross-correlations match the desired ones closely even when M is as small as 8 and the number of random trials is only 100. A better match can be obtained if more random trials are performed.

Multiple mutual uncorrelated fading channels, which are required for MIMO channels
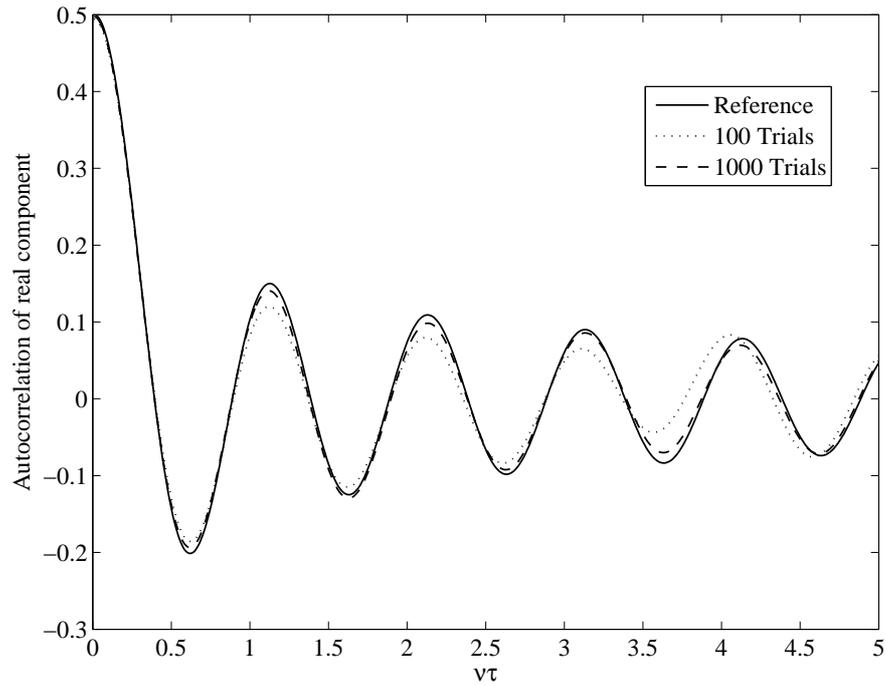
Figure 2.1: Autocorrelation of the simulated real part of the fading, $h_r(t)$ and the reference.
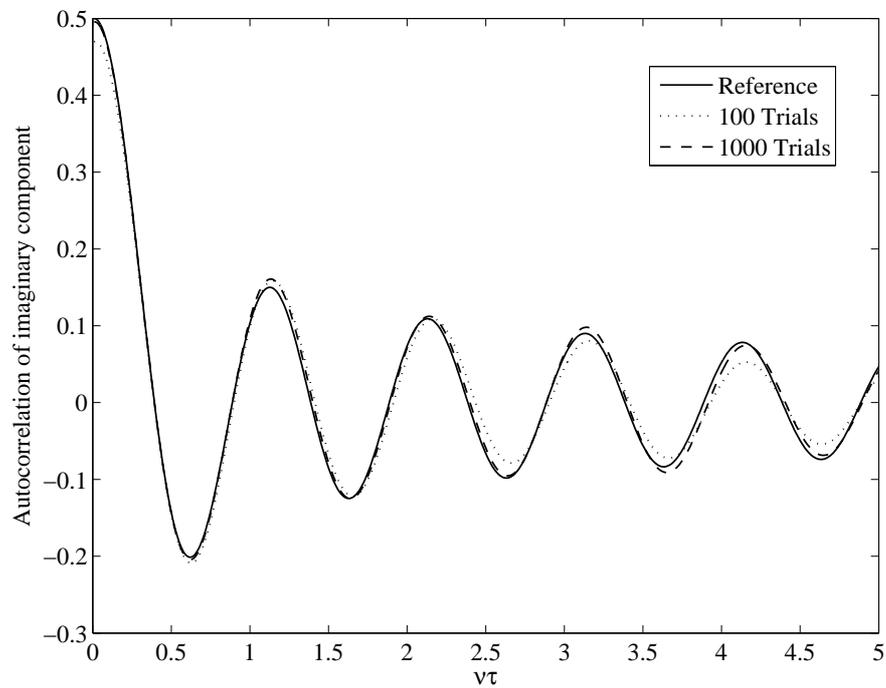
Figure 2.2: Autocorrelation of the simulated imaginary part of the fading, $h_i(t)$ and the reference.
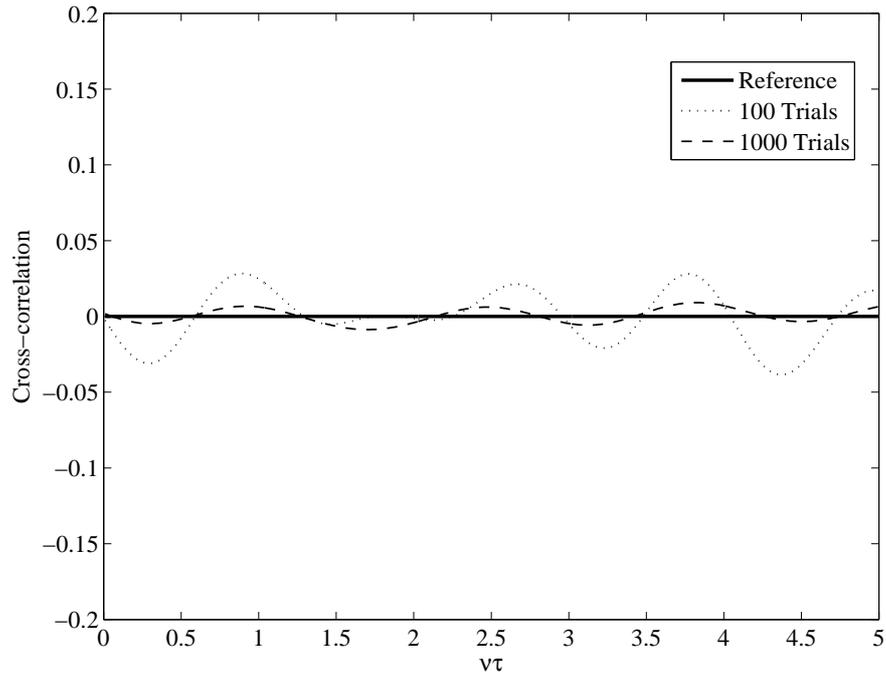
Figure 2.3: Cross-correlation of the simulated real and imaginary parts of the fading, $h(t)$ and the reference.
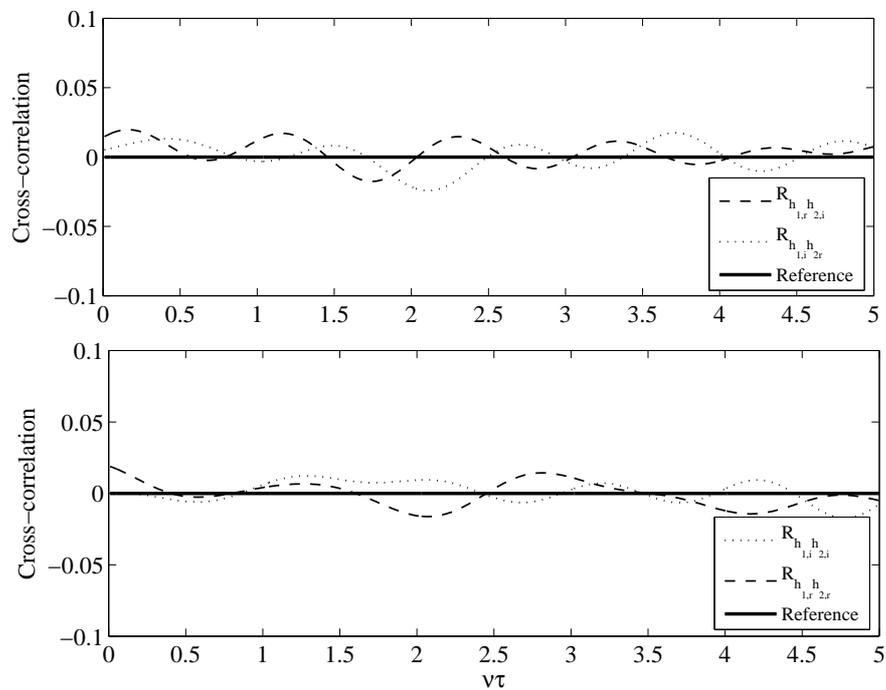
Figure 2.4: Cross-correlation of two independent fading channels $h_1(t)$ and $h_2(t)$ and reference.

or double selective channels, can also be obtained by using this improved simulation model. We can use $h_k(t)$ to denote the k-th Rayleigh fading channel defined by

$$
\begin{aligned}
h_k(t) &= h_r(t) + jh_i(t) \quad, \\
h_{k,r}(t) &= \sqrt{\frac{2}{M}} \sum_{n=1}^{M} \cos(\psi_{n,k}) \cos(\omega_d t \cos(\alpha_{n,k}) + \phi_{n,k}) \quad, \\
h_{k,i}(t) &= \sqrt{\frac{2}{M}} \sum_{n=1}^{M} \sin(\psi_{n,k}) \cos(\omega_d t \cos(\alpha_{n,k}) + \phi_{n,k}) \quad,
\end{aligned} \tag{2.26}
$$

with

$$
\alpha_{n,k} = \frac{2\pi n - \pi + \theta_k}{4M}, \quad n = 1, \ldots, M, \tag{2.27}
$$

where $\psi_{n,k}$, $\phi_{n,k}$ and θ_k are mutually independent and randomly distributed over $[-\pi, \pi)$ for all n and k. Therefore, $h_k(t)$ is uncorrelated with $h_l(t)$ for $k \neq l$. In order to show that the multiple fading channels generated by this model are uncorrelated, we simulate two independent fading channels $h_1(t)$ and $h_2(t)$ and plot the cross-correlation between 4 quadrature components in Fig. 2.4. The simulation results are obtained by performing 1000 random trials. It is seen that the cross-correlations between multipath fading channels are small.

Based on the discussions and simulations above, we find that the improved simulation model can perfectly match the desired Clarke's model, for both single time-variant channel and multiple time-variant channels. Therefore, all of time-variant Rayleigh fading channels used in the following chapters are simulated by this model.

2.2 Basis expansion models

The traditional approach to estimate the time-variant fading channel is based on applying the Wiener filter for tracking time variations of the channel gain [1]. Although a high accuracy of estimation can be achieved [26], the Wiener filter requires a high computational load [63]. In recent years, the basis expansion model (BEM) is widely used for estimating time-variant fading channels, due to its high accuracy and low complexity [19,20,25–32,64,65]. With a BEM, estimation of a realization of the random process describing the time-variant channel is transformed into estimation of a few time-invariant

expansion coefficients [64], and the time-variant channel can be modeled as

$$\bar{\mathbf{h}} = \mathbf{Ba}, \tag{2.28}$$

where the $M \times 1$ vector $\mathbf{a} = [a_1, \ldots, a_M]^T$ contains the expansion coefficients, and the $N \times M$ matrix $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_m, \ldots, \mathbf{b}_M]$ collects M linearly independent columns $\mathbf{b}_m$.

According to the different ways to generate the matrix $\mathbf{B}$, the family of BEMs can be categorized into two categories. The first category applies the basis functions whose generation depends on the physical (e.g. fading rate) or statistical information of the fading channel [20, 29–32, 64], while the second group employs a simple series representation such as complex exponential or polynomial series [19, 25, 27, 28, 64]. In this section, we will introduce two BEMs for each category:

The widely used BEMs in the first category are Karhunen-Loeve (KL) [31, 32] and discrete prolate spheroidal (DPS) [20, 29, 30, 64] BEMs. The generation of KL and DPS basis functions depends on the knowledge of statistical information of fading. The problem though is that if the assumed channel statistics deviate from the true ones, e.g., due to inaccurate information of the maximum velocity of the mobile, the performance of these BEMs may degrade. An alternative approach is to use the second category of BEMs with fixed functions. In this category, the generalized complex exponential (GCE) and B-spline (BS) BEMs are widely used.

KL BEM

The KL BEM provides the best performance among these four BEMs [32, 64], since it assumes that the statistical information of fading is perfectly known at the receiver side. The KL basis functions $v_m(n)$ are eigenvectors of the fading covariance matrix. For example, the covariance matrix of Jakes' fading process is defined as

$$[\boldsymbol{\Upsilon}]_{t_1, t_2} = J_0[2\pi\nu(t_1 - t_2)]. \tag{2.29}$$

We order the eigenvalues λ_m of $\boldsymbol{\Upsilon}$ as: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$, and assume that when m is larger than a fixed value $M << N$, λ_m decreases rapidly and can be neglected [32]. Then, the matrix $\mathbf{B}$ of the KL BEM can be represented as

$$[\mathbf{B}]_{n,m} = v_m(n), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{2.30}$$

DPS BEM

Although the modeling error introduced by the KL BEM is insignificant [31, 32], the covariance matrix of fading is not always available at the receiver side in a practical scenario. Alternatively, a BEM based on DPS functions was proposed in [20]. The DPS BEM corresponds to the discrete KL BEM with a rectangular spectrum [20]. The DPS basis functions are also named Slepian sequences, which are bandlimited to the Doppler spread $[-\nu, \nu]$ and simultaneously most concentrated in the certain time interval of length M [66]. DPS sequences are widely used for channel estimation both in time and frequency domains [20, 30, 67]. Here we will introduce the principle of DPS sequences briefly.

The target is to find the sequences $u[m]$ which maximize the energy concentration in the interval with length N [20]

$$\lambda = \frac{\sum_{n=0}^{N-1} |u[n]|^2}{\sum_{n=-\infty}^{\infty} |u[n]|^2}, \tag{2.31}$$

while being bandlimited to ν; hence

$$u[n] = \int_{-\nu}^{\nu} U(\nu) e^{j2\pi n\nu} d\nu, \tag{2.32}$$

where

$$U(\nu) = \sum_{n=-\infty}^{\infty} u[n] e^{-j2\pi n\nu}, \tag{2.33}$$

and $0 \leq \lambda \leq 1$.

The solution of this constrained maximization problem are the DPS sequences [66], which are the eigenvectors of the following eigenvalue equation

$$\sum_{q=1}^{N} \frac{\sin(2\pi\nu(q-n))}{\pi(q-n)} u_m(q) = \lambda_m u_m(n), \tag{2.34}$$

where $u_m(n)$ is the mth basis function with length N bandlimited to the frequency range $[-\nu, \nu]$, and λ_m is an eigenvalue indicating the fraction of energy contained in the frequency range $[-\nu, \nu]$ of the corresponding eigenvector [67].

The DPS sequence $u_0[n]$ is the unique sequence that is bandlimited and most time-concentrated in a given interval with length N, $u_1[n]$ is the next sequence having maximum energy energy concentration among the DPS sequences orthogonal to $u_0[n]$, and so

on. Thus, the DPS sequences are a set of orthogonal sequences which are bandlimited and high (but not complete) time-concentrated in a certain interval with length N [20]. The eigenvalues λ_m are a measure for this energy-concentration and ordered starting with the maximum one as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$. Therefore, $u_m(n)$ is the mth function corresponding to the mth most maximum eigenvalue; M should be chosen to provide λ_m close to 1 when $m << M$ and close to 0 when $m >> M$ [29]. The option of M is described in [66], as

$$M = 2\lceil \nu N \rceil + 1, \tag{2.35}$$

$\lceil x \rceil$ denotes the smallest integer value larger than or equal to x. The rigorous proof can be found in [68]. Then, the matrix $\mathbf{B}$ containing samples of the DPS basis functions can be represented as

$$[\mathbf{B}]_{n,m} = u_m(n), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{2.36}$$

GCE BEM

The GCE BEM, which is also known as oversampled complex exponential (CE) model [25] or non-critically sampled CE model [69], is a modified model of the the CE BEM. The CE BEM is introduced in [19] to approximate the time variant fading channels. Its basis functions are complex exponentials that have a period equal to the length of the considered interval. Normally, the channel modeled by CE BEM is represented as [23,70]

$$h(n) = \sum_{m=1}^{M} a_m e^{\frac{j2\pi}{N}(n-1)[(m-1)-M/2]}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{2.37}$$

Although the CE BEM is widely used to approximate the time-variant fading channel [21, 23, 71–74], the modeling error of CE BEM is significant. The rectangular window in (2.37), which corresponds to critically sampling the Doppler spectrum, results in spectral leakage, which means, the energy from low frequency CE coefficients leaks to the full frequency range [20]. This results in a floor in the BER performance for time-variant channels with Doppler spread as shown in [75].

Since only a limit Doppler range of windowed channel is considered, the sidelobes might be significantly eliminated and more samples are taken in within that range [25]. An improved modeling performance is obtained by using the GCE BEM, which applies a

set of complex exponentials with the period longer than the window length related to the CE BEM [25, 70]. This corresponds to oversampling the Doppler spectrum of windowed channel. For the GCE BEM, elements of the matrix $\mathbf{B}$ are given by [25, 70]

$$[\mathbf{B}]_{n,m} = e^{\frac{j2\pi}{\kappa N}(n-1)[(m-1)-M/2]}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N, \tag{2.38}$$

where κ is a real number larger than 1; usually, $\kappa = 2$ is used [25].

BS BEM

The B-splines have previously been investigated in application to estimating the Clarke's model [26–28, 65, 76] since its high approximation accuracy and low computational complexity. An optimal spline of order q, approximating the random process $h(t)$ with zero mean and variance σ_h^2, is a spline providing an MSE which is defined as

$$\varepsilon^2 = \frac{1}{\sigma_h^2 T} \int_0^T E\{[h(t) - \hat{h}(t)]^2\}dt, \tag{2.39}$$

where $\hat{h}(t)$ is an approximation of $h(t)$ by applying splines, and T is the sampling interval. An optimal spline of order q can be represented as

$$\hat{h}(t) = \sum_{-\infty}^{m=\infty} a_m b_q(t - mT), \tag{2.40}$$

where $b_q(t)$ is the B-spline of order q, and a_m are spline coefficients. $b_q(t)$ is a $(q + 1)$ fold convolution of the B-spline of zero degree [77]

$$b_0(t) = \begin{cases} 1, & \text{if } |t| < \frac{T}{2} \\ \frac{1}{2}, & \text{if } |t| = \frac{T}{2} \\ 0, & \text{otherwise}, \end{cases} \tag{2.41}$$

where T is the sampling interval. Usually, $b_q(t)$ are described by the Fourier transform [27]

$$B_q(\omega) = \int_{-\infty}^{\infty} b_q(t)e^{-j\omega t}dt = T \left[\frac{\sin(\frac{\omega T}{2})}{\frac{\omega T}{2}} \right]^{q+1}. \tag{2.42}$$

The optimal spline approximation can be described by a "prefilter-sampling-postfilter" scheme which is shown in Fig.2.5 [27] where $G(\omega)$ and $F(\omega)$ are transfer functions of the prefilter and postfilter, and $\delta(t)$ is the Dirac delta function [78]. The postfilter transfer
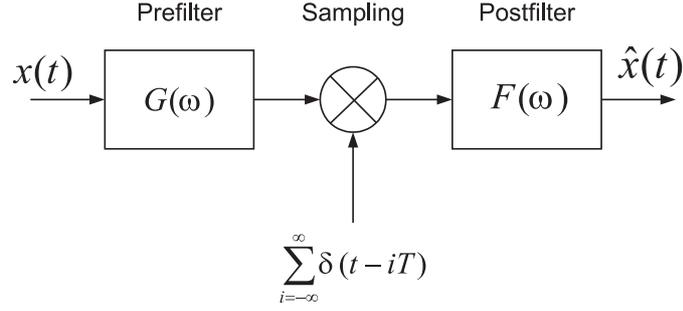
Figure 2.5: Prefilter-sampling-postfilter scheme describing spline approximation of the process $x(t)$,

function $F(\omega)$ is the Fourier transform of the B-splines, $F(\omega) = B_q(\omega)$, while the prefilter has the transform function [27]

$$G(\omega) = \left[(\frac{\omega T}{2}) \sin(\frac{\omega T}{2})\right]^{-q-1} \times \left[\sum_{n=-\infty}^{\infty} \left(\frac{\omega T}{2} + n\pi\right)^{-2q-2}\right]^{-1}. \qquad (2.43)$$

If the the random process $h(t)$ obeys Clarke's model, the MSE of the approximation by applying optimal splines of an arbitrary order q can be calculated by [27]

$$\varepsilon^2 \approx \frac{\pi^{2q+2} B_{2q+2}}{[(q+1)!]^2 \gamma^{2q+2}} + \frac{\pi^{2q+4}(q+1)(2q+3)B_{2q+4}}{[(q+2)!]^2 \gamma^{2q+4}}, \qquad (2.44)$$

where B_n are Bernoulli numbers [79], and the sampling factor $\gamma = 1/(\nu T)$.

To build the basis functions, we use the B-spline of order q [76]

$$B_q(t) = \frac{1}{q!} \sum_{i=0}^{q+1} (-1)^i \binom{q+1}{i} \left(\frac{t}{T} + \frac{q+1}{2} - i\right)_+^q, \qquad (2.45)$$

where $T = (N-1)/(M-q)$ is the sampling interval separating two adjacent BS functions, and $(x)_+ = \max\{0, x\}$. In this case, elements of the basis function matrix are given by

$$[\mathbf{B}]_{n,m} = B_q\left((n-1) - \left(m - \frac{q+1}{2}\right)T\right), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N.$$
$$(2.46)$$

The accuracy and complexity of B-spline approximation depend on the order q of the spline.

As shown above, the KL and DPS BEMs can approximate the time-variant fading channel with insignificant modeling error but require the statistics of fading and have to

suffer extra error caused by inaccurate estimation of these statistics. Although the GCE and BS BEMs do not require the knowledge of the statistical information of fading by using a simple series representation as basis functions, they will introduce higher modeling errors than KL and DPS BEMs. We will compare the performance and complexity of these four BEMs in Chapter 3 and use the one which can provide a good performance and affordable complexity to approximate the fading channels in this thesis.

2.3   Turbo codes

Turbo codes were first introduced by Berrou, Glavieux and Thitimajshima at the International Conference on Communication (ICC) in 1993 [80]. In AWGN channels, the performance of a half rate turbo code is only 0.7 dB away from the Shannon capacity limit at BER= 10^{-5}. The remarkable achievement terminates the conventional thought that the Shannon limit can only be approached by using extraordinarily long codes with extremely complex decoding processes [81]. As one of the most powerful error-control codes, Turbo codes have been developed rapidly and attract substantial attention in wireless communication community due to its outstanding ability of error correction [82–88].

Turbo codes are based on two fundamental concepts, concatenated coding and iterative decoding, the latter of which is the core of the 'turbo principle' since it is the method that allows the outstanding performance of turbo codes. As turbo codes will be used in some chapters of this thesis, we will briefly introduce the structure of the turbo encoder and main turbo decoding algorithms, i.e., the optimal maximum *a posteriori* (MAP) and Log-MAP algorithms, and the suboptimal MAX-Log-MAP algorithm. For more detailed description of turbo codes, readers are referred to [89–91].

2.3.1   Turbo encoder

The structure of the turbo encoder used in this report can be explained by its formal name, parallel concatenated recursive systematic convolutional (RSC) code. Fig. 2.6 gives an example of the structure of a turbo encoder. Two RSC encoders are concatenated and

an interleaver is in between them. Comparing with non-systematic convolutional (NSC) codes, RSC codes apply a feedback loop (recursive part) and set one of the outputs equal to the input data (systematic part). The structure of a RSC encoder and the corresponding NSC encoder are shown in Fig. 2.7 and Fig. 2.8, respectively. For both encoders, the code rate is 1 and the constraint length is 3. The generator polynomials of the feedback and output connectivity in the RSC encoder are $[7, 5]$ in octal notations, respectively.

The working principles of the turbo encoder are described here. A length N data sequence $\mathbf{d} = [d[1], \ldots, d[N]]$ is encoded by the first RSC encoder, the output of which is a length N coded sequence $\mathbf{x}_p^1 = [x_p^1[1], \ldots, x_p^1[N]]$. Then, the original data sequence is interleaved and encoded by the seconde RSC encoder to generate another length N coded sequence $\mathbf{x}_p^2 = [x_p^2[1], \ldots, x_p^2[N]]$. Finally, $\mathbf{d}$, $\mathbf{x}_p^1$ and $\mathbf{x}_p^2$ are multiplexed together to generate the final turbo coded sequence. Without puncturing, this results in a code rate of $1/3$. Higher code rates can be obtained by applying a puncturing scheme.
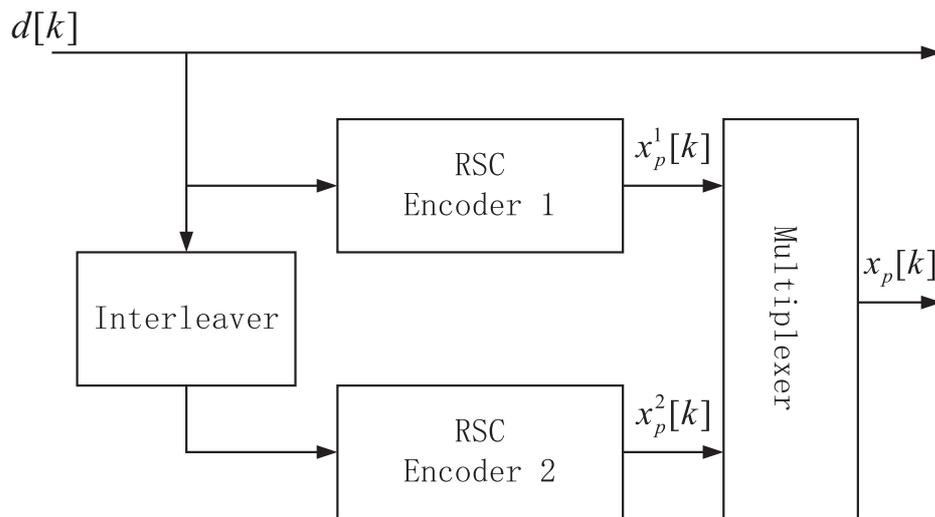
Figure 2.6: Structure of a Turbo encoder.

The interleaver is a device that simply reorders the input data sequence, while an deinterleaver, which will be used in the decoder to recover the original order of the data sequence. It is the joint influence of the interleaver and RSC encoder leading to a high code weight composite codeword for most of the time which is critical to the performance of turbo code [92]. There are numerous interleavers that can be used in the turbo encoders, i.e. pseudo-random [93], block [94], and s-random interleavers [95–98]. In this report, we apply the s-random interleaver due to its superior performance [90]. The output pattern
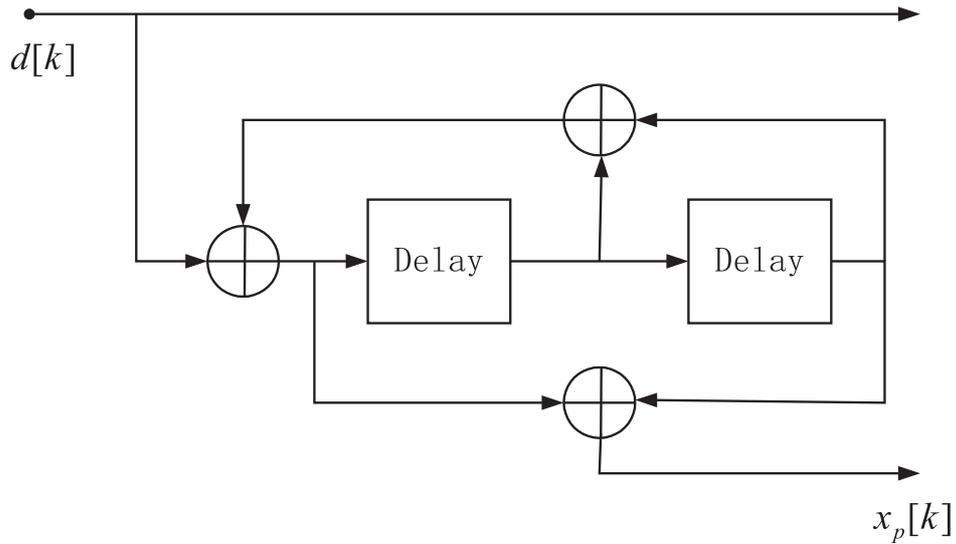
Figure 2.7: Example of a Recursive Systematic Convolutional (RSC) encoder.

of such an interleaver is generated randomly, with the constraint that any two input bits at a distance smaller than s bits will be separated by at least s bits after interleaving.

2.3.2 Turbo decoder

Fig. 2.9 illustrates the turbo decoder corresponding to the encoder in Fig. 2.6. It is seen that two RSC decoders are linked by an deinterleaver/interleaver, which is similar to that used in the encoder.

The turbo decoder works iteratively and in each iteration the two RSC decoders exchange the decoded information to help each other. Before decoding iterations, the received signals $y[k] = (y_d[k], y_p^1[k], y_p^2[k])$ from the demodulator are demultiplexed to sequences $y_d[k]$, $y_p^1[k]$ and $y_p^2[k]$, respectively, where $y_d[k]$ corresponds to the received systematic codes, $y_p^1[k]$ corresponds to the received 1st parity bits, and $y_p^2[k]$ corresponds to the received 2nd parity bits. The first RSC decoder applies $y_d[k]$ and $y_p^1[k]$ as input sequences and the second RSC decoder applies $y_d[k]$ and $y_p^2[k]$. When the parity bits of a given RSC encoder are punctured before transmission, the corresponding decoder's inputs are set to zeros at the punctured positions. In the initial iteration, the first RSC decoder takes only $y_d[k]$ and $y_p^1[k]$ to generate soft information of the data bits, $L_{E,1}(\bar{d}[k])$. Then the second RSC decoder can perform decoding with the soft information of $L_{E,1}(\bar{d}[k])$

$$x_p^1[k]$$

Delay   Delay

$$d[k]$$

$$x_p^2[k]$$
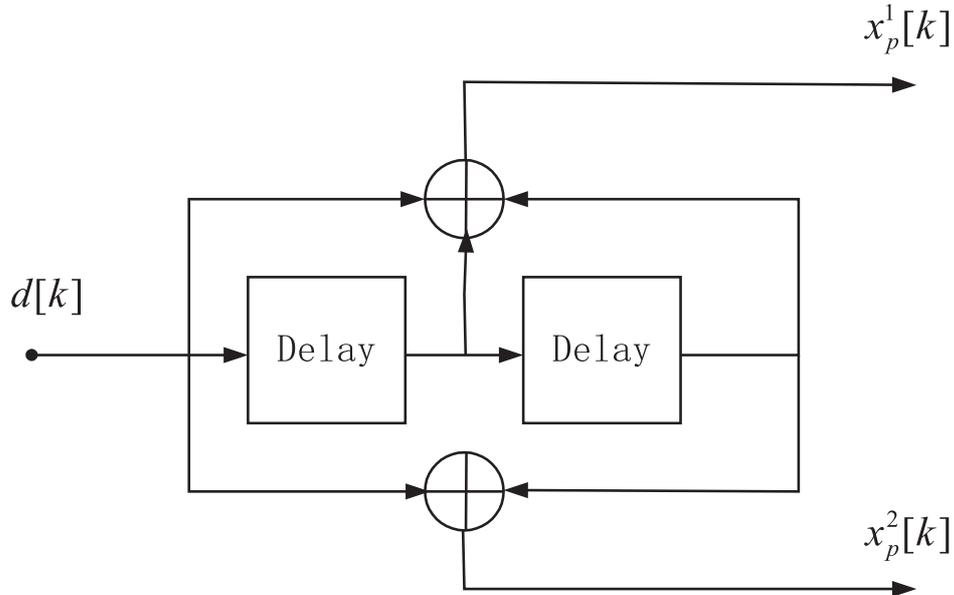
Figure 2.8: Example of a Non-Systematic Convolutional (NSC) encoder.

and $L_{ap,1}(d[k])$ from the first RSC decoder, in addition to the received $y_d[k]$ and $y_p^2[k]$. The output of the second decoder is another soft decoding information $L_{E,2}(\bar{d}[k])$, which will be deinterleaved to generate $L_{ap,2}(d[k])$ and fed back to the first RSC decoder. In the subsequent iterations, the first RSC decoder takes $L_{ap,2}(d[k])$ from the second RSC decoder in the previous iteration as additional information to $y_d[k]$ and $y_p^1[k]$, to generate $L_{E,1}(\bar{d}[k])$. The performance of the turbo decoder improves as the number of iterations increases. However, the improvement from iteration to iteration decreases as the number of iterations increases. This process is repeated iteratively until two RSC decoders's estimates of the original data bits converge. Eight iterations are commonly used, as a compromise between the performance and complexity [81]. Finally, the output *a posteriori* information $L(\bar{d}[k])$ of a data bit $d[k]$ delivered from the second RSC decoder is deinterleaved and used for the final hard decision.

In general, each RSC decoder performs decoding by using its input received signals ($y_d[k]$ and $y_p^i[k], i = 1, 2$) and the *a priori* information ($L_{ap,i}(d[k]), i = 1, 2$) from the other RSC decoder, and provides the extrinsic information $L_{E,i}(\bar{d}[k])$ for the other decoder. Note that the extrinsic information is only exchanged between decoders as intermediate information during the decoding process. The soft information exchanging between two RSC decoders is the reason why the turbo decoder is called soft-input soft-output decoder,

$$L_{ap,2}(d[k]) \qquad\qquad L_{E,2}(\bar{d}[k])$$

Deinterleaver

$$L_{E,1}(\bar{d}[k]) \qquad\qquad L_{ap,1}(d[k])$$

Interleaver

$y_d[k]$

Decoder 1

Demultiplexer

$y_p^1[k]$

Interleaver

$y_p[k]$

Decoder 2 $L(\bar{d}[k])$

Decision

$y_p^2[k]$
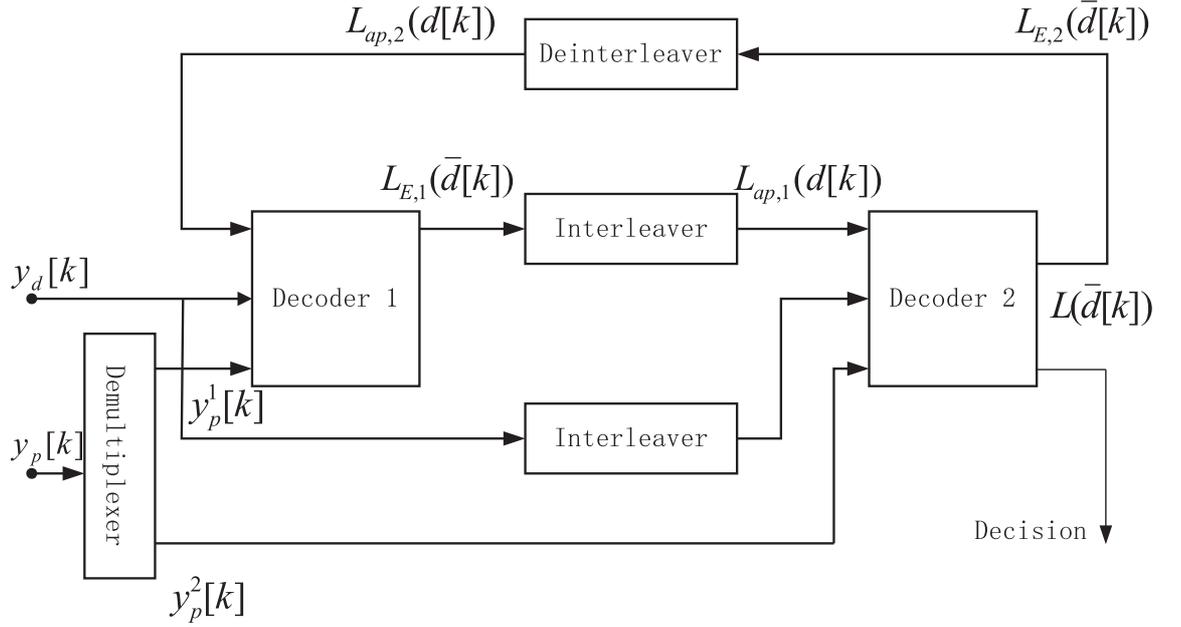
Figure 2.9: Structure of a Turbo decoder.

which accepts soft *a priori* information $L_{ap,i}(d[k])$ at one of its inputs from the previous decoding process and generates soft information $L_{E,i}(\bar{d}[k])$ as its output. Soft information means that besides decoded bits, the associated probability that each bit has been decoded correctly is also provided, usually in the form of log-likelihood ratio (LLR). This indicates that the decoder yields not only the coded bits but also how reliable they are. As its name implies, the LLR is the logarithm of the ratio of two probabilities in the case of binary transmission, e.g., the output *a posteriori* information ($L(\bar{d}[k])$) is generally given by

$$L(\bar{d}[k]) = \log \frac{P(d[k] = +1|\mathbf{y})}{P(d[k] = -1|\mathbf{y})}, \tag{2.47}$$

where the numerator and denominator are probabilities of the transmitted bit $d[k] = +1$ and $d[k] = -1$ conditioned on the received sequence $\mathbf{y}$. Based on (2.47), the more positive the value of $L(\bar{d}[k])$ is, the more reliably the transmitted bit was '1', or the more negative the value of $L(\bar{d}[k])$ is, the more likely '0' was transmitted.

In the turbo decoder shown in Fig. 2.9, the output $L(\bar{d}[k])$, *a posteriori* information of an information bit $d[k]$, is a sum of

$$L(\bar{d}[k]) = L(\tilde{d}[k]) + L_{ap}(d[k]) + L_E(\bar{d}[k]) \quad , \tag{2.48}$$

where $L(\tilde{d}[k])$ is the channel information, $L_{ap}(d[k])$ is the *a priori* information and $L_E(\bar{d}[k])$ is the extrinsic information.

The channel information $L(\tilde{d}[k])$ can be extracted directly from $y[k]$ which are the received signals of $d[k]$. Supposing that $d[k]$ are transmitted with E_s transmitted energy per symbol, over an AWGN channel, the received signal is

$$y[k] = a \cdot d[k] + n[k] \quad , \tag{2.49}$$

where n_k denotes an AWGN with a variance of σ_n^2, and a is the fading amplitude in a fading channel or a constant value in an AWGN channel. For such a model, the channel information is given by [81]

$$
\begin{aligned}
L(\tilde{d}[k]) &= \log \frac{\exp\left(-\frac{E_s}{2\sigma_n^2}(y[k]-a)^2\right)}{\exp\left(-\frac{E_s}{2\sigma_n^2}(y[k]+a)^2\right)} \\
&= \log\left(\exp\left(\frac{2aE_s}{\sigma_n^2}y[k]\right)\right) \\
&= L_c \cdot y[k],
\end{aligned}
\tag{2.50}
$$

where $L_c = 2aE_s/\sigma_n^2$ is the channel reliability factor, which reflects the reliability of estimating the transmitted signal from the received signal. For example, L_c will be large if SNR in the channel is high, and we can estimate the transmitted signal from the received signal correctly with a high probability. In such a case, the received signal will impact heavily on the final output *a posteriori* LLR.

The *a priori* information $L_{ap}(d[k])$ used here is the deinterleaved extrinsic information from the other RSC decoder.

The extrinsic information $L_E(\bar{d}[k])$ is the information that decoder exploits from the whole received sequence and *a priori* information, but excludes these of the bits which are currently being decoded in this iteration. It is only the extrinsic information that the decoders exchange between each other since the same information should not be used more than once at each decoding step.

Taking all these three types of information above into account, the turbo decoder delivers the *a posteriori* information of data bits. The final decision of the decoding is based on the *a posteriori* information of data bits.

There are 3 typical decoding algorithms applied widely: MAP algorithm, MAX-Log-MAP algorithm and Log-MAP algorithm [80, 92, 99–102]. All of these algorithm will be introduced briefly in follows.

The Maximum *a Posteriori* algorithm was firstly proposed by Bahl, Cocke, Jelinek and Raviv in [92] and modified by Berrou, Glavieuv and Thitmajshima in [80]. Compared with the conventional maximum likelihood sequence estimation (MLSE) algorithm which can be efficiently implemented by the Viterbi algorithm [103], the MAP algorithm is a symbol-by-symbol detection algorithm based on maximum *a posteriori* information. It is optimal in the sense of minimizing the probability of a symbol error by taking *a priori* information of the coded bits into account and providing soft information about estimated bits. The performance of the MAP and MLSE algorithms would be the same when there is no *a priori* information to be exploited. However, when *a priori* information is available, for example, in the soft-input soft-output turbo decoder exchanging the extrinsic information between two RSC decoders, the MAP algorithm will outperform the conventional MLSE one [80].

Although the MAP algorithm is the optimal decoding scheme, it is too complicated to be realized for implementation since the exact representation of probabilities used in the MAP algorithm requires a high dynamic range [81]. Moreover, there are many non-linear functions and numerous multiplications proposed in the scheme [81]. Working in the logarithmic domain instead of the linear domain for the probability used in the MAP algorithm and invoking the approximation

$$\ln(e^{x_1} + \ldots + e^{x_n}) \approx \max_{i \in 1,2\ldots,n} x_i, \tag{2.51}$$

the MAX-Log-MAP algorithm reduces the complexity significantly. However, it is obvious that the MAX-Log-MAP algorithm is suboptimal since only a part of information is exploited due to the approximation.

This approximation can be avoided by applying the Jacobian logarithm to calculate $x = \ln(e^{x_1} + \ldots + e^{x_n})$. The Jacobian logarithm [104, 105] is given by

$$
\begin{aligned}
\ln(e^{x_1} + e^{x_2}) &= \max(x_1, x_2) + \ln(1 + e^{-|x_1 - x_2|}) \\
&= \max(x_1, x_2) + f_c(|x_1 - x_2|) \\
&= g_c(x_1, x_2),
\end{aligned}
\tag{2.52}
$$

where $f_c(|x_1 - x_2|)$ can be regarded as a correction term. Robertson, Hoeher and Ville-brun in [100] proposed a method to show how to use the Jacobian logarithm to calculate $\ln(e^{x_1} + \ldots + e^{x_n})$ accurately. They supposed $x = \ln(e^{x_1} + \ldots + e^{x_{n-1}})$ is known. Then, they obtained

$$\begin{aligned} \ln(e^{x_1} + \ldots + e^{x_n}) \quad &= \ln(e^x + e^{x_n}) \\ &= \max(x, x_n) + f_c(|x - x_n|). \end{aligned} \quad (2.53)$$

This method is referred to as the Log-MAP algorithm [100]. It was also shown that the correction term $f_c(|x_1 - x_2|)$ can be implemented efficiently by an one dimension look-up table to avoid real time computation, moreover, only a few values are needed for the table. By applying the Jacobian logarithm, the Log-MAP algorithm retains the optimality of the original MAP algorithm, while preserves the computational simplicity of the MAX-Log-MAP algorithm. For interested readers, more details about the MAP, MAX-Log-MAP and Log-MAP algorithms can be obtained from [99, 101, 102].
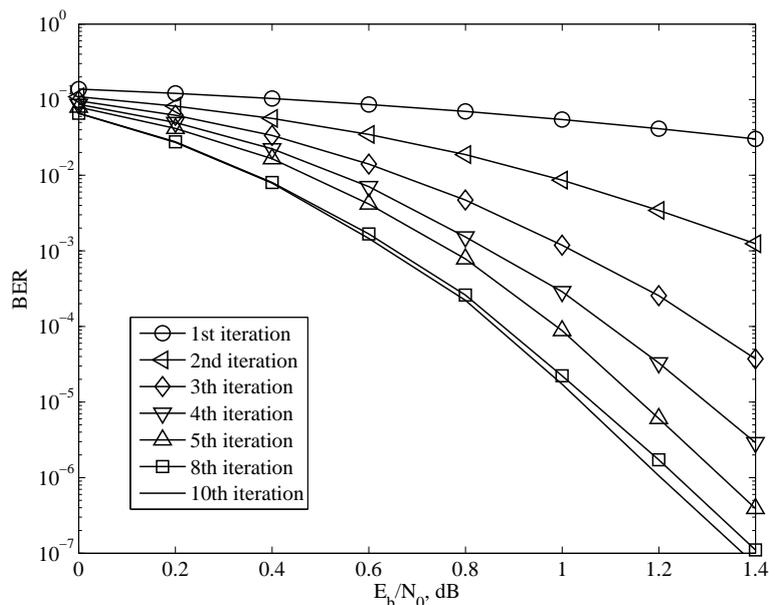
Figure 2.10: BER performance of turbo codes with rate 1/3, 8 states, 1024 bits, Log-MAP, over AWGN channels.

Fig. 2.10 shows the BER performance of the turbo codes over the AWGN channels. A turbo code with rate $1/3$, 8 states is employed in the simulation. The generator polynomials of the RSC encoders are $[13, 15]$ in octal for their feedback and output connections respectively. The length of frame is 1024, giving a sufficient s-parameter of 20 for the s-random interleaver. The turbo decoder applies the Log-MAP algorithm and the number

of decoding iterations are 1, 2, 3, 4, 5, 8 and 10. Simulation results show that the BER performance of turbo codes improves but the improvement between two consecutive iterations decreases as the number of iterations increases. When the number of iterations is larger than 8, the improvement almost decreases to zero. Thus, in this thesis, we will apply 8 turbo decoding iterations in simulations.

2.4 Conclusions

In this chapter, we have introduced fundamental techniques, such as the simulator of time-variant channels, BEMs and Turbo codes, which will be used throughout this thesis. After briefly introducing the reference Clarke's model and the deficiencies of Jakes' model and Pop-Beaulieu simulator, we have adopted the model introduced by Zheng and Xiao [51, 57] to generate multiple fading channels. We have also provided simulation results for second-order statistics of this model, which have shown that this model can accurately match to the desired Clarke's model. Therefore, all time-variant channels considered in this thesis are generated by using this improved simulation model.

We have also introduced BEMs used to approximate the fading channels. Advantages and disadvantages of the most widely used BMEs, such as KL, DPS, GCE and BS BEMs, have been briefly discussed. In the next chapter, we will compare these four BEMs in application to time-variant fading channels and will show that the BS BEM is the best practical choice providing good performance and low complexity.

Finally, we have introduced the turbo encoder and decoder with different decoding algorithms, such as MAP algorithm, MAX-Log-MAP algorithm and Log-MAP algorithm. After describing the advantages and disadvantages of each decoding algorithm, we have chosen Log-MAP decoding algorithm and will be using it in the subsequent chapters to realize the soft-input soft-output turbo decoding schemes.

Chapter 3

Channel Estimation of Time-Varying Channels Based on Basis Expansion Models

Contents

3.1 Introduction

In mobile communications, the Doppler effect causes the time variant fading. The parameter used to measure this Doppler effect is the Doppler spread. Usually, this fading is

well described by Jakes' model [50], or more generally, by Clarke's model [49, 106]. For Jakes' model, the channel gain is a stationary random process with a correlation function described as the zero-order Bessel function of the first kind parameterized by the Doppler spread.

In order to approximate the time-variant fading channels, basis expansion models (BEMs) are widely used [19, 20, 25–32, 64, 65]. The most often used BEMs are the Karhunen-Loeve (KL) functions [31, 32], discrete prolate spheroidal (DPS) functions [20, 29, 30, 64], generalized complex exponential (GCE) functions [19, 25, 64], and B-splines (BS) [26–28, 65]. With a BEM, estimation of a realization of the random process describing the time-variant channel is transformed into estimation of a few time-invariant expansion coefficients [20].

The MSE performance of BEM-based channel estimators is considered in [20, 32, 64, 69, 107–109]. Analytical results in [32, 107, 108] are based on the assumption that the Doppler spread is perfectly known, whereas in practice, the Doppler spread is estimated with some errors [110]. In [20], instead of assuming that the Doppler spread is perfectly known, the maximum Doppler spread is used to generate the DPS basis functions. This method is widely used in the works investigating the DPS BEM [64, 109]. Although the use of the maximum Doppler spread is more practical and may reduce the complexity of BEM-based estimators, it may also lead to a significant degradation in the MSE performance.

The Doppler spread can be estimated based on correlation and variation of channel estimates. For example, in [111], the Doppler estimation scheme based on the autocorrelation of complex channel estimates is described. Instead of using channel estimates, the received signal can also be used directly in estimating Doppler spread information. In [112], the mobile speed is estimated as a function of the deviation of the averaged signal envelope in flat fading channels. In [110], an efficient Doppler estimation algorithms for wireless mobile radio systems is introduced by using an ML approach relying on a periodic channel estimation. In [69], the performance of BEM-based MMSE estimators using the mismatched Doppler spread has been investigated. However, the derivation of MSE in [69] is limited to the MMSE estimator using BEMs with orthogonal basis functions; as a result, it cannot be applied to the BS and GCE BEMs.

In this chapter, we derive the MSE of a generic linear channel estimator using linearly independent basis functions, and specify the error for MMSE and ML estimators. The MSE performance of the Wiener solution is considered and used as a lower MSE bound. We then investigate and compare the MSE performance and complexity for three approaches to estimating time-variant channels with perfect or inaccurate knowledge of the Doppler spread:

1) channel estimation using perfect knowledge of the Doppler spread;

2) channel estimation using the maximum Doppler spread as suggested in [20];

3) channel estimation using an estimate of the Doppler spread.

The first approach provides the best MSE performance; however, it might be very complicated depending on the BEM used and the assumption of a perfectly known Doppler spread is impractical. Both the second and third approaches are based on inaccurate knowledge of the Doppler spread; we will investigate which one can provide a better MSE performance and lower complexity. The complexity of each approach depends on the BEM used. Therefore, we will compare the MSE performance and complexity of estimators using different BEMs for each approach, and give a practical choice of the approach and BEM providing the best performance and the lowest complexity.

The remainder of this chapter is organized as follows. In Section 3.2, the transmission model and different BEMs are introduced. Section 3.3 describes the MSE derivation for a generic BEM-based linear channel estimator. The performance and complexity of the first approach using perfect knowledge of the Doppler spread are described in Section 3.4. Then, the MSE performance and complexity of the second and third approaches both based on inaccurate knowledge of the Doppler spread are investigated and compared in Section 3.5 and Section 3.6, respectively, followed by conclusions in Section 3.7.

3.2   Transmission model and BEMs

We consider a PSAM system and assume that a block of N pilot symbols is transmitted and there are N_d data symbols transmitted between two neighboring pilot symbols, e.g. as shown in Fig. 3.1. The received pilot signal can be written in the time-domain as

$$z(i) = s(i)h(i) + n(i), \quad i = 0, \cdots, N-1, \tag{3.1}$$
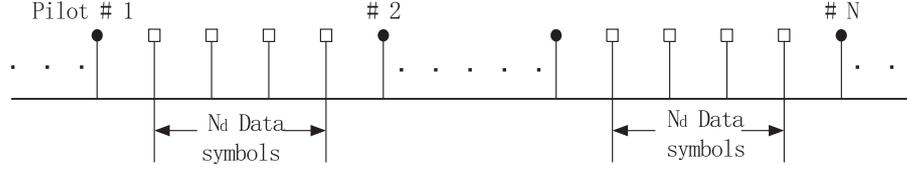
Figure 3.1: Structure of transmitted block.

where at the ith observed instance, $s(i)$ is the transmitted pilot symbol of a power $|s(i)|^2 = \sigma_s^2$, $h(i)$ is the (time variant) channel coefficient and $n(i)$ is the complex additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2. The matrix form of (3.1) is given by

$$\mathbf{z} = \mathbf{Sh} + \mathbf{n}, \tag{3.2}$$

where $\mathbf{z}$, $\mathbf{h}$ and $\mathbf{n}$ are $N \times 1$ vectors with elements $z(i)$, $h(i)$ and $n(i)$, respectively, and $\mathbf{S}$ is an $N \times N$ diagonal matrix with diagonal elements $s(i)$.

We consider time-variant Rayleigh fading channels following Jakes' model [50, 51]. The covariance matrix of such a channel is an $N \times N$ matrix $\mathbf{\Upsilon}$ with elements

$$[\mathbf{\Upsilon}]_{t_1, t_2} = \rho(t_1 - t_2), \tag{3.3}$$

where $t_1, t_2 = 1, \ldots, N$, and $\rho(\tau)$ is the autocorrelation function [50]

$$\rho(\tau) = \sigma_h^2 J_0(2\pi\nu\tau), \tag{3.4}$$

σ_h^2 is the variance of channel coefficients, $J_0(\cdot)$ is the zero-order Bessel function of the first kind and ν is the Doppler spread.

The time-variant fading channel can be represented by a BEM. We define an $N \times M$ matrix $\mathbf{B}$ containing samples of basis functions corresponding to N transmitted symbols, where M is the number of basis functions. The time-variant channel can be modeled as

$$\bar{\mathbf{h}} = \mathbf{Ba}, \tag{3.5}$$

where the vector $\mathbf{a} = [a_1, \ldots, a_M]^T$ contains the expansion coefficients and $[\cdot]^T$ denotes matrix transpose. The difference between $\bar{\mathbf{h}}$ and $\mathbf{h}$ is due to a modeling error. By applying

a BEM, the task of estimating N time varying channel coefficients is transformed into estimating only M expansion coefficients with usually $M << N$. As mentioned above, we will consider the following BEMs: KL, DPS, GCE, and BS basis functions.

KL BEM

The KL BEM is introduced in [32,107] as a BEM with perfect knowledge of the statistical information of fading channels. The KL basis functions $v_m(n)$ are the eigenvectors of the covariance matrix of the fading $\boldsymbol{\Upsilon}$, which is obtained in (3.3). We order the eigenvalues λ_m of $\boldsymbol{\Upsilon}$ as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$. Then, the matrix $\mathbf{B}$ of the KL BEM can be represented as

$$[\mathbf{B}]_{n,m} = v_m(n), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{3.6}$$

DPS BEM

The DPS functions are a set of orthogonal functions bandlimited to the range $[-\nu, \nu]$. We consider M basis functions $u_m(n)$ of length N. Such sequences are defined as the real-value solution of the following equation [20]

$$\sum_{q=0}^{N-1} \frac{\sin(2\pi\nu(q-n))}{\pi(q-n)} u_m(q) = \lambda_m u_m(n), \quad n = 0, \ldots, N-1, \tag{3.7}$$

where λ_m and u_m are the eigenvalues and eigenvectors of the matrix

$$[\mathbf{C}]_{i,j} = \frac{\sin(2\pi\nu(i-j))}{\pi(i-j)}, \quad i, j \in \{1, \ldots, N\}. \tag{3.8}$$

λ_m indicates the fraction of energy of $u_m(n)$ contained in the range $[-\nu, \nu]$ [67]. The eigenvalues are ordered in the descending order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$. Then, the matrix $\mathbf{B}$ of the DPS BEM can be represented as

$$[\mathbf{B}]_{n,m} = u_m(n), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{3.9}$$

GCE BEM

For the GCE BEM, elements of the matrix $\mathbf{B}$ are given by [25, 70]

$$[\mathbf{B}]_{n,m} = e^{\frac{j2\pi}{\kappa N}(n-1)((m-1)-(M-1)/2)}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N, \qquad (3.10)$$

where κ is a real number larger than 1. For $\kappa = 1$, the GCE BEM becomes the complex exponential (CE) BEM. However, the CE BEM results in a large modeling error [25, 70]. The GCE BEM introduces a lower modeling error; usually, $\kappa = 2$ is used [25].

BS BEM

The B-splines have previously been investigated in application to estimating the Clarke's model [26–28, 65]. To build the basis functions, we use the B-spline of order q [76]

$$B_q(t) = \frac{1}{q!} \sum_{i=0}^{q+1} (-1)^i \binom{q+1}{i} \left(\frac{t}{T} + \frac{q+1}{2} - i \right)_+^q, \qquad (3.11)$$

where $T = (N-1)/(M-q)$ is the sampling interval separating two adjacent B-spline functions, and $(x)_+ = \max\{0, x\}$. In this case, elements of the basis function matrix are given by

$$[\mathbf{B}]_{n,m} = B_q \left((n-1) - \left(m - \frac{q+1}{2} \right) T \right), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \qquad (3.12)$$

The accuracy and complexity of B-spline approximation depend on q, and normally, the B-spline of order $q = 3$ is widely used since it provides the trade-off between complexity and accuracy [27, 76]. Here, we use the cubic BS BEM and compare its performance and complexity with those of the other BEMs. For this case, from (3.11) we have

$$B_3(t) = \begin{cases} \frac{2}{3} - \frac{t^2}{T^2} + \frac{|t|^3}{2T^3}, & \text{if } |t| < T, \\ \frac{1}{6}(2 - \frac{|t|}{T})^3, & \text{if } T \leq |t| < 2T, \\ 0, & \text{otherwise}, \end{cases} \qquad (3.13)$$

and (3.12) becomes

$$[\mathbf{B}]_{n,m} = B_3 \left((n-1) - (m-2)T \right), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \qquad (3.14)$$

It is clear that for cubic B-splines, the $N \times M$ matrix $\mathbf{B}$ is a sparse matrix, and there are
only 4 nonzero elements in each row and K nonzero elements in each column, where

$$K = \lceil 4T \rceil = \lceil 4(N-1)/(M-3) \rceil, \tag{3.15}$$

and $\lceil x \rceil$ denotes the smallest integer value larger than or equal to x.

3.3 MSE of a generic linear channel estimator

3.3.1 BEM-based estimator

A BEM-based channel estimate is given by

$$\hat{\mathbf{h}} = \mathbf{B}\hat{\mathbf{a}}, \tag{3.16}$$

where $\hat{\mathbf{a}}$ is a vector of estimates of expansion coefficients, the estimation of $\hat{\mathbf{h}}$ is transfer-
ring to the estimation of $\mathbf{a}$. In this chapter, we consider linear channel estimators based
on BEMs, i.e., ML channel estimator and MMSE channel estimator.

The ML channel estimator has the asymptotic properties of being unbiased and have
a Gaussian PDF [113]. We consider the case with white Gaussian noise with PDF
$\mathcal{N}_C(0, \sigma_n^2 \mathbf{I}_N)$. Under these condition the PDF of received signal $\mathbf{z}$ is [113]

$$p(\mathbf{z}; \mathbf{a}) = \frac{1}{(2\pi)^{N/2} \sigma_n N} \exp\left[-\frac{1}{2} (\mathbf{z} - \mathbf{SBa})^H (\mathbf{z} - \mathbf{SBa}) \right], \tag{3.17}$$

therefore, the ML channel estimation of $\mathbf{a}$ is found by minimizing

$$J(\mathbf{a}) = (\mathbf{z} - \mathbf{SBa})^H (\mathbf{z} - \mathbf{SBa}). \tag{3.18}$$

Since this is a quadratic function of the elements of $\mathbf{a}$ and $\sigma_n^2 \mathbf{I}_N$ is a positive definite
matrix, differentiation will produce the global minimum [113]. Thus, we have

$$\frac{\partial \ln p(\mathbf{z}; \mathbf{a})}{\partial \mathbf{a}} = \frac{\partial J(\mathbf{a})}{\partial \mathbf{a}} = \sigma_n^{-2} \frac{\partial (\mathbf{SBa})^H}{\partial \mathbf{a}} (\mathbf{z} - \mathbf{SBa}). \tag{3.19}$$

By setting (3.19) equal zero we have

$$\mathbf{SB}^H (\mathbf{z} - \mathbf{SB}\hat{\mathbf{a}}) = \mathbf{0}. \tag{3.20}$$

Therefore, the ML estimate of $\mathbf{a}$ is given by [113]

$$\hat{\mathbf{a}}_{\text{ML}} = (\mathbf{B}^H \mathbf{S}^H \mathbf{S} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{S}^H \mathbf{z}. \tag{3.21}$$

The MMSE channel estimator solving the estimation problem by minimizing

$$J(\mathbf{a}) = E\left\{ (\mathbf{z} - \mathbf{S}\mathbf{B}\mathbf{a})^H (\mathbf{z} - \mathbf{S}\mathbf{B}\mathbf{a}) \right\}. \tag{3.22}$$

In (3.22), since the matrices $\mathbf{S}$ and $\mathbf{B}$ and vector $\mathbf{z}$ are perfectly known, the only random unknown variable is the vector $\mathbf{a}$. Therefore, the expectation in (3.22) is over $\mathbf{a}$. We can solve the estimation problem by set

$$\frac{\partial J(\mathbf{a})}{\partial \mathbf{a}} = 0. \tag{3.23}$$

After some algebra, we obtain the MMSE estimates of $\mathbf{a}$ as [113]

$$\hat{\mathbf{a}}_{\text{MMSE}} = E\{\mathbf{a}\} + \left(\mathbf{B}^H \mathbf{S}^H \mathbf{S} \mathbf{B} + \sigma_n^2 \left(E\{\mathbf{a}\mathbf{a}^H\} \right)^{-1} \right)^{-1} \mathbf{B}^H \mathbf{S}^H (\mathbf{z} - \mathbf{S}\mathbf{B}E\{\mathbf{a}\}). \tag{3.24}$$

Here we consider the Rayleigh fading channels and therefore, we have $E\{\mathbf{a}\} = \mathbf{0}$ and we define that $E\{\mathbf{a}\mathbf{a}^H\} = \mathbf{R}_a$ is the covariance matrix of expansion coefficients, and represented as [65]

$$\mathbf{R}_a = (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{\Upsilon} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1}. \tag{3.25}$$

Then, (3.24) is simplified as [113]

$$\hat{\mathbf{a}}_{\text{MMSE}} = \left(\mathbf{B}^H \mathbf{S}^H \mathbf{S} \mathbf{B} + \sigma_n^2 \mathbf{R}_a^{-1} \right)^{-1} \mathbf{B}^H \mathbf{S}^H \mathbf{z}. \tag{3.26}$$

Considering a general expression of a BEM-based linear channel estimator, the vector $\hat{\mathbf{a}}$ can be represented as

$$\hat{\mathbf{a}} = \mathbf{A}\mathbf{z}, \tag{3.27}$$

where $\mathbf{A}$ is a matrix defined by the channel estimation scheme. For MMSE estimation, we have

$$\mathbf{A}_{\text{MMSE}} = (\mathbf{B}^H \mathbf{S}^H \mathbf{S} \mathbf{B} + \sigma_n^2 \mathbf{R}_a^{-1})^{-1} \mathbf{B}^H \mathbf{S}^H. \tag{3.28}$$

For ML channel estimation, we have

$$\mathbf{A}_{\text{ML}} = (\mathbf{B}^H \mathbf{S}^H \mathbf{S} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{S}^H. \tag{3.29}$$

Here we derive the MSE of a generic BEM-based linear estimator and specify it for
different estimation schemes. The MSE is represented as

$$\text{MSE} = \frac{1}{\text{tr}\left\{E\left[\mathbf{hh}^H\right]\right\}}\text{tr}\left\{E\left[\left(\mathbf{h} - \hat{\mathbf{h}}\right)\left(\mathbf{h} - \hat{\mathbf{h}}\right)^H\right]\right\}, \qquad (3.30)$$

where $\text{tr}\{\cdot\}$ denotes the trace operator, $(\cdot)^H$ denotes the Hermitian transpose and $E[\cdot]$
denotes the expectation. By substituting (3.16) and (3.27) into (3.30), we obtain

$$\begin{aligned}
\text{MSE} &= \frac{1}{\text{tr}\left\{\boldsymbol{\Upsilon}\right\}}\text{tr}\left\{E\left[\left(\mathbf{h} - \mathbf{BA}(\mathbf{Sh} + \mathbf{n})\right)\left(\mathbf{h} - \mathbf{BA}(\mathbf{Sh} + \mathbf{n})\right)^H\right]\right\} \\
&= \frac{1}{\sigma_h^2 N}\text{tr}\left\{E\left[\mathbf{hh}^H\right] - E\left[\mathbf{h}\left(\mathbf{Sh} + \mathbf{n}\right)^H\mathbf{A}^H\mathbf{B}^H\right] - E\left[\mathbf{BA}\left(\mathbf{Sh} + \mathbf{n}\right)\mathbf{h}\right]\right. \\
&\quad \left. + E\left[\mathbf{BA}(\mathbf{Sh} + \mathbf{n})(\mathbf{Sh} + \mathbf{n})^H\mathbf{A}^H\mathbf{B}^H\right]\right\} \\
&= \frac{1}{\sigma_h^2 N}\text{tr}\left\{\boldsymbol{\Upsilon} - 2\Re\left\{E\left[\mathbf{hh}^H\right]\mathbf{S}^H\mathbf{A}^H\mathbf{B}^H\right\} + \mathbf{BAS}\left\{E\left[\mathbf{hh}^H\right]\right\}\mathbf{S}^H\mathbf{A}^H\mathbf{B}^H\right. \\
&\quad \left. + \mathbf{BA}\left\{E\left[\mathbf{nn}^H\right]\right\}\mathbf{A}^H\mathbf{B}^H\right\} \\
&= \frac{1}{\sigma_h^2 N}\text{tr}\left\{\boldsymbol{\Upsilon} + \mathbf{BAS}\boldsymbol{\Upsilon}\mathbf{S}^H\mathbf{A}^H\mathbf{B}^H - 2\Re\left\{\mathbf{BAS}\boldsymbol{\Upsilon}\right\} + \sigma_n^2\mathbf{BAA}^H\mathbf{B}^H\right\}, \quad (3.31)
\end{aligned}$$

where $\Re\{\cdot\}$ denotes the real part. The MSE in (3.31) describes the overall error of a BEM-
based channel estimator and can be easily specified for any particular channel estimator,
i.e., the MMSE estimator and the ML estimator with perfect or inaccurate knowledge of
the Doppler spread.

3.3.2 Wiener solution

To evaluate the MSE performance of the BEM-based estimators, the MSE performance
of the Wiener solution is considered and used as a lower bound indicating the optimal
performance that can be achieved. The channel coefficients $\mathbf{h}$ can be estimated by using
the Wiener solution [113]:

$$\hat{\mathbf{h}}_W = \left(\mathbf{S}^H\mathbf{S} + \sigma_n^2\boldsymbol{\Upsilon}^{-1}\right)^{-1}\mathbf{S}^H\mathbf{z}. \qquad (3.32)$$

The MSE of the Wiener solution is minimum and given by [113]

$$\text{MSE}_W = \frac{\sigma_n^2}{\sigma_h^2 N}\text{tr}\left\{\left(\mathbf{S}^H\mathbf{S} + \sigma_n^2\boldsymbol{\Upsilon}^{-1}\right)^{-1}\right\}. \qquad (3.33)$$

We will consider MSE_W as a lower bound. The Wiener solution relies on an N-parameter
model, while a BEM-based estimator exploits an M-parameter model with $M << N$.

The computational load of the Wiener solution is very high. The matrix inversion in (3.32) requires $\mathcal{O}(N^3)$ complex multiplications. This matrix inversion has to be calculated once for the channel coefficient at each data position, and therefore, there are $\mathcal{O}(N_d N^4)$ complex multiplications required to estimate the channel coefficients over the whole transmission block with NN_d data symbols. The use of BEMs can significantly reduce the complexity but may lead to performance loss. In the next section, we will investigate this performance loss and compare the complexity of the BEM-based estimators using perfect knowledge of the Doppler spread. Normally, constant envelope signals are used as pilot symbols and we will assume that PSK pilot signals are transmitted, and therefore $\mathbf{S}^H \mathbf{S} = \sigma_s^2 \mathbf{I}_N$, where $\mathbf{I}_N$ denotes an $N \times N$ identity matrix.

3.4 Approach 1: Channel estimation using perfect knowledge of the Doppler spread

In this section, we will specify the MSE in (3.31) to MMSE and ML estimators using perfect knowledge of the Doppler spread. Then, we will compare the MSE performance and complexity of estimators applying different BEMs.

3.4.1 BEM-based MMSE estimator

The BEM-based MMSE channel estimator in the first approach is given by

$$\hat{\mathbf{h}}_{\text{MMSE,1}} = \mathbf{B} \mathbf{A}_{\text{MMSE,1}} \mathbf{z}, \tag{3.34}$$

where $\mathbf{A}_{\text{MMSE,1}} = (\sigma_s^2 \mathbf{B}^H \mathbf{B} + \sigma_n^2 \mathbf{R}_a^{-1})^{-1} \mathbf{B}^H \mathbf{S}^H$. By substituting $\mathbf{A}_{\text{MMSE,1}}$ into (3.31), the MSE of a BEM-based MMSE estimator is represented as

$$\text{MSE}_{\text{MMSE,1}} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \frac{\sigma_n^2}{\sigma_s^2} \left[\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \left(\mathbf{B}^H \mathbf{B} \mathbf{R}_a \right)^{-1} \right]^{-2} \right.$$
$$\left. + \left[\mathbf{I} - \mathbf{B} \left(\mathbf{B}^H \mathbf{B} + \frac{\sigma_n^2}{\sigma_s^2} \mathbf{R}_a^{-1} \right)^{-1} \mathbf{B}^H \right]^2 \mathbf{\Upsilon} \right\}. \tag{3.35}$$
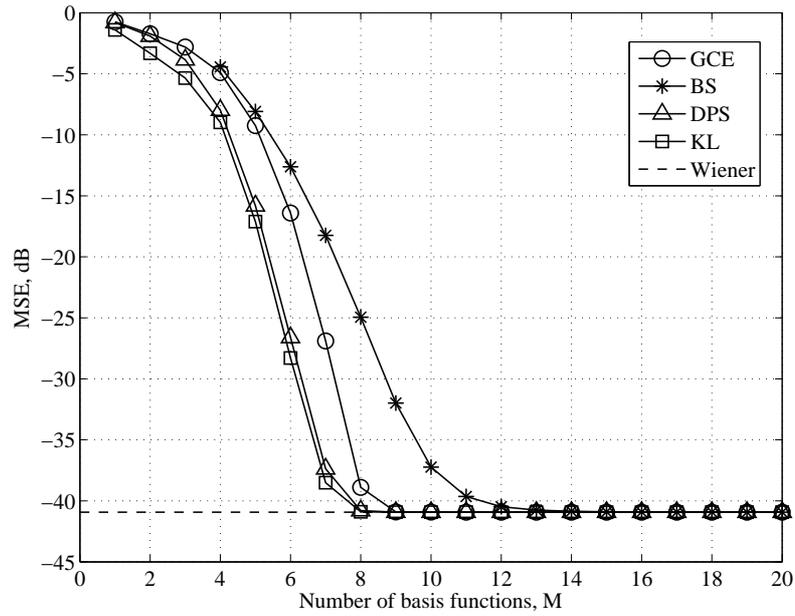
Figure 3.2: MSE performance of the BEM-based MMSE channel estimators versus the number of basis functions, M, with perfect knowledge of the Doppler spread, $N = 100$, SNR = 30 dB, $\nu T_s = 0.02$.
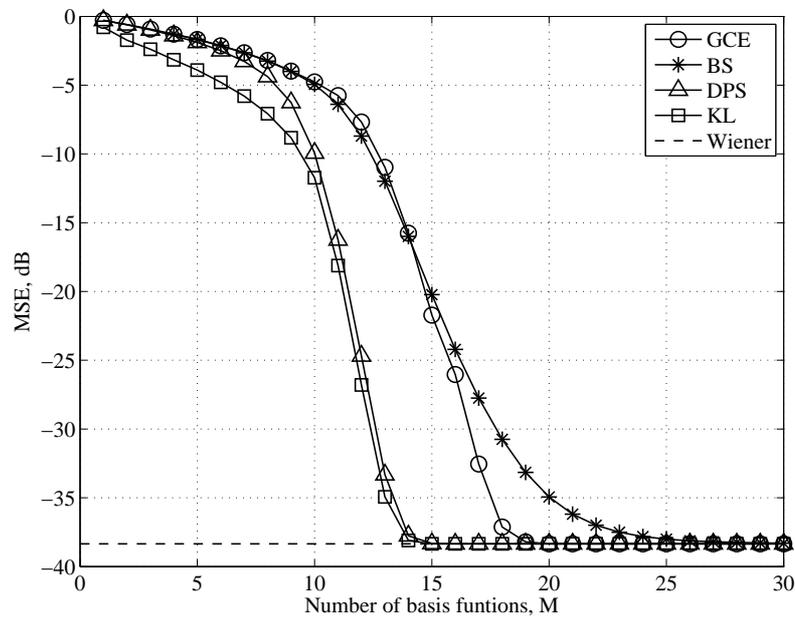
Figure 3.3: MSE performance of the BEM-based MMSE channel estimators versus the number of basis functions, M, with perfect knowledge of the Doppler spread, $N = 100$, SNR = 30 dB, $\nu T_s = 0.05$.

Fig. 3.2 and Fig. 3.3 show the MSE performance of the BEM-based MMSE channel
estimators for time-variant channels with $\nu T_s = 0.02$ and $\nu T_s = 0.05$, respectively. We
set $N = 100$ and SNR $= 30$dB. It can be seen that the MSE of the BEM-based MMSE
estimators decrease when more basis functions are applied, and for different BEMs, the
number of basis functions required to achieve the lower bound is different. The MMSE
estimator using BS BEM requires a larger number of basis functions than the others to
provide the same performance. This indicates that the modeling error of the BS BEM is
larger than that of the others. However, the MMSE estimators using all BEMs can achieve
the identical lower bound and when the number of basis functions approaches N, there is
no performance loss. Note that these numerical results are based on (3.35), which match
to the results obtained by simulations in which the MSE of estimated are calculated and
averaged over time variant Rayleigh fading channels generated by using the improved
simulator (2.22) introduced in Chapter 3.

The complexity of the BEM-based estimators depends on the BEM used. The gener-
ation of KL and DPS basis functions requires the knowledge of the Doppler spread and
uses the singular value decomposition (SVD) to calculate the eigenvectors of the $N \times N$
matrix $\mathbf{\Upsilon}$. Therefore, for any ν, the generation of KL and DPS basis functions requires
$\mathcal{O}(N^3)$ complex multiplications [114]. Note that the matrix $\mathbf{B}^H \mathbf{B}$ for KL or DPS BEM
is an $M \times M$ identity matrix, since the basis functions are orthogonal.

For BS and GCE BEMs, which do not require the knowledge of the Doppler spread,
the basis functions can be precalculated and the matrices $\mathbf{B}$, $\mathbf{B}^H \mathbf{B}$, $\left(\mathbf{B}^H \mathbf{B}\right)^{-1} \mathbf{B}^H$ can be
saved in a memory, and therefore, the complexity of generating basis functions is negli-
gible. Moreover, since the matrix $\mathbf{B}$ of the BS BEM is a sparse matrix, the complexity of
the MMSE estimator using BS BEM is lower than that of the estimator using GCE BEM
as shown in Table. 3.1.

Table 3.1 shows the number of complex multiplications required by MMSE estimators
using different BEMs in the first approach, where q is the order of the B-splines and
$K = \lceil \frac{4(N-1)}{M-q} \rceil$ is the number of non-zero elements in each column of the matrix $\mathbf{B}$ of
the BS BEM. It can be seen that the complexity of the MMSE estimators using KL and
DPS BEMs is of the same order. The number of complex multiplications required by the
MMSE estimator using GCE BEM is $\mathcal{O}(N^3)$ less than that required by estimators using

Calculation	KL	DPS	BS	GCE
$\mathbf{B}$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	$-$	$-$
$\mathbf{R}_a$	$MN^2 + M^2N$	$MN^2 + M^2N$	$MN^2 + M^2N$	$MN^2 + M^2N$
$\mathbf{R}_a^{-1}$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$
$\mathbf{V} = \left(\sigma_s^2\mathbf{B}^H\mathbf{B} + \sigma_n^2\mathbf{R}_a^{-1}\right)^{-1}$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$
$\mathbf{A}_{\text{MMSE,1}} = \mathbf{V}\mathbf{B}^H\mathbf{S}^H$	$(M^2+M)N$	$(M^2+M)N$	$(q+1)N + (q+1)KM$	$(M^2+M)N$
$\hat{\mathbf{a}}_{\text{MMSE,1}} = \mathbf{A}_{\text{MMSE,1}}\mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	MN_dN	MN_dN	$(q+1)N_dN$	MN_dN

Table 3.1: The number of complex multiplications required by MMSE estimators using different BEMs in the first approach using perfect knowledge of the Doppler spread.

KL and DPS BEMs. When $M << N$, the term $\mathcal{O}(N^3)$ dominates in the complexity of the estimators using KL and DPS BEMs. For this case, the complexity of the MMSE estimator using the GCE BEM is much lower than that of the estimators using KL and DPS BEMs. When $M < 5$, the complexity of the MMSE estimators using cubic BS ($q = 3$) and GCE BEMs is close. However, when M increases, the complexity of the MMSE estimator using the GCE BEM is much higher than that of the estimator using the BS BEM.

3.4.2 BEM-based ML estimator

The BEM-based ML channel estimator is given by

$$\hat{\mathbf{h}}_{\mathbf{ML,1}} = \mathbf{B}\mathbf{A}_{\text{ML,1}}\mathbf{z}, \tag{3.36}$$

where $\mathbf{A}_{\text{ML,1}} = (\mathbf{B}^H\mathbf{S}^H\mathbf{S}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{S}^H$. By substituting $\mathbf{A}_{\text{ML,1}}$ into (3.31) and taking $\mathbf{S}^H\mathbf{S} = \sigma_s^2\mathbf{I}_N$ into account, the MSE of the BEM-based ML estimator is represented as

$$
\begin{aligned}
\text{MSE}_{\text{ML,1}} &= \frac{1}{\sigma_h^2 N}\text{tr}\left\{\left(\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\right)\boldsymbol{\Upsilon} + \frac{\sigma_n^2}{\sigma_s^2}\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\right\} \\
&= \frac{1}{\sigma_h^2 N}\text{tr}\left\{\left(\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\right)\boldsymbol{\Upsilon}\right\} + \frac{\sigma_n^2 M}{\sigma_s^2\sigma_h^2 N} \tag{3.37} \\
&= \delta_{\text{m,ML}}^2 + \delta_{\text{s,ML}}^2, \tag{3.38}
\end{aligned}
$$

where $\delta_{\text{m,ML}}^2 = \frac{1}{\sigma_h^2 N}\text{tr}\left\{\left(\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\right)\boldsymbol{\Upsilon}\right\}$ and $\delta_{\text{s,ML}}^2 = \frac{\sigma_n^2 M}{\sigma_s^2\sigma_h^2 N} = \frac{M}{\beta N}$, and $\beta = \frac{\sigma_h^2\sigma_s^2}{\sigma_n^2}$ as the signal-to-noise ratio (SNR). When M increases, $\delta_{\text{m,ML}}^2$ reduces but $\delta_{\text{s,ML}}^2$ rises up.
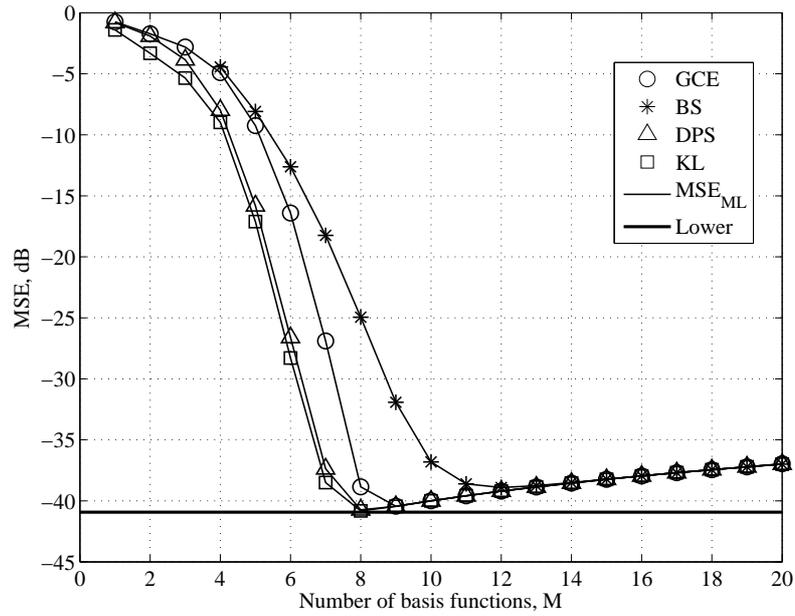
Figure 3.4: MSE performance of the BEM-based ML channel estimators versus the number of basis functions, M, with perfect knowledge of the Doppler spread, $N = 100$, SNR $= 30$ dB, $\nu T_s = 0.02$.
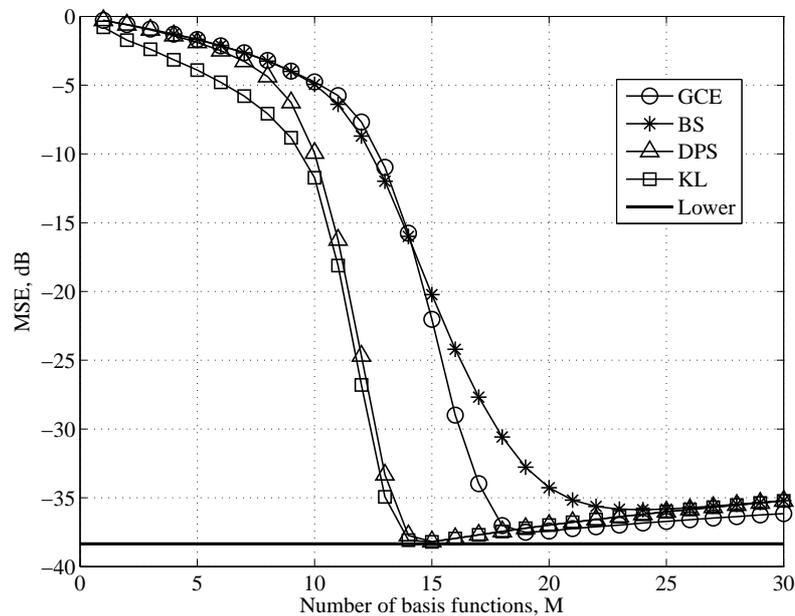
Figure 3.5: MSE performance of the BEM-based ML channel estimators versus the number of basis functions, M, with perfect knowledge of the Doppler spread, $N = 100$, SNR $= 30$ dB, $\nu T_s = 0.05$.

For the case with fixed N, νT_s and SNR, there is a value of M, below which, the term $\delta^2_{m,\mathrm{ML}}$ dominates in $\mathrm{MSE}_{\mathrm{ML},1}$ and above which $\delta^2_{s,\mathrm{ML}}$ dominates in $\mathrm{MSE}_{\mathrm{ML},1}$. We call this the optimal number of basis functions, because, for this M the ML estimator provides the best performance. The optimal values of M for different BEMs can be identified, as shown in Fig. 3.4 and Fig. 3.5 for the cases with $\nu T_s = 0.02$ and $\nu T_s = 0.05$, respectively. We set $N = 100$ and SNR to 30dB. Note that these numerical results also match to the simulation results. By comparing Fig. 3.4 and Fig. 3.5 with Fig. 3.2 and Fig. 3.3, we find that unlike the MMSE estimators, the MSE performance of the BEM-based ML estimators degrades when the number of basis functions is larger than the optimal value of M.

Calculation	KL	DPS	BS	GCE
$\mathbf{B}$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	–	–
$\mathbf{A}_{\mathrm{ML},1} = \left(\sigma_s^2 \mathbf{B}^H \mathbf{B}\right)^{-1} \mathbf{B}^H \mathbf{S}^H$	$MN + M^2$	$MN + M^2$	$MN + M^2$	$MN + M^2$
$\hat{\mathbf{a}}_{\mathrm{ML},1} = \mathbf{A}_{\mathrm{ML},1} \mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	MN_dN	MN_dN	$(q+1)N_dN$	MN_dN

Table 3.2: The number of complex multiplications required by ML estimators using different BEMs in the first approach using perfect knowledge of the Doppler spread.

We also investigate the complexity of the BEM-based ML estimators. Table. 3.2 shows the number of complex multiplications required by ML estimators with different BEMs in the first approach using perfect knowledge of the Doppler spread. By comparing Table. 3.2 with Table. 3.1, we find that for BS and GCE BEMs, the complexity of the ML estimator is significantly lower than that of the MMSE estimator, since the ML estimators do not require the matrix $\mathbf{R}_a$. However, the complexity of the ML estimators using KL and DPS BEMs is still high due to the computational load required for the generation of basis functions.

The results show that in the first approach, the MMSE estimator outperforms the ML estimator but requires more complex multiplications. When M increases, the MMSE estimators for all BEMs can provide a good performance close to that of the Wiener solution. The complexity of the estimators using KL and DPS BEMs is much higher than that of the estimators using BS and GCE BEMs. Among all BEM-based estimators, the complexity of the one using BS BEM is the lowest.

Although the performance of the first approach can be very close to the Wiener solution, the estimators using KL and DPS are too complicated to be implemented in practice. Moreover, the assumption of perfectly known Doppler spread is also impractical. In the next section, we will investigate the performance and complexity of the second approach using the maximum Doppler spread and compare them with those of the first approach.

3.5 Approach 2: Channel estimation using the maximum Doppler spread

In practical scenarios, the Doppler spread is not always available. Therefore, in [20], the maximum Doppler spread, $\nu_{\max}$, is suggested to generate the DPS basis functions. This method can also be applied for the KL BEM. For this approach, the generation of KL and DPS basis functions depends on $\nu_{\max}$. Therefore, like the BS and GCE BEMs, the KL and DPS basis functions can be precalculated and saved in memory. This method can significantly reduce the complexity of the estimators using KL and DPS BEMs, but may also lead to degradation of the performance due to the mismatch between the maximum Doppler spread and the real Doppler spread. In this section, we will investigate the decrease in the complexity and degradation in the MSE performance of the second approach compared to the first approach where the Doppler spread is perfectly known.

In the second approach, we can precalculate the mismatched covariance matrix $\bar{\boldsymbol{\Upsilon}}$ for the Doppler spread $\nu_{\max}$ and save it in memory. Elements of $\bar{\boldsymbol{\Upsilon}}$ are calculated as

$$[\bar{\boldsymbol{\Upsilon}}]_{t_1,t_2} = \bar{\rho}(t_1 - t_2), \tag{3.39}$$

where

$$\bar{\rho}(\tau) = \sigma_h^2 J_0(2\pi\nu_{\max}\tau). \tag{3.40}$$

The corresponding mismatched covariance matrix of the expansion coefficients is calculated by

$$\bar{\mathbf{R}}_a = (\bar{\mathbf{B}}^H\bar{\mathbf{B}})^{-1}\bar{\mathbf{B}}^H\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{B}}(\bar{\mathbf{B}}^H\bar{\mathbf{B}})^{-1}. \tag{3.41}$$

For KL and DPS BEMs, $\bar{\mathbf{B}}$ is a matrix containing samples of the basis functions generated by using $\bar{\nu}_{\max}$. For BS and GCE BEMs, $\bar{\mathbf{B}} = \mathbf{B}$.

By applying matrices of $\bar{\mathbf{B}}$ and $\bar{\mathbf{R}}_a$, the expressions of Wiener solution, MMSE estimator and ML estimators are modified as below.

Mismatched Wiener solution

In the second approach, the mismatched Wiener solution is given by

$$\bar{\mathbf{h}}_{W,2} = \left(\sigma_s^2 \mathbf{I}_N + \sigma_n^2 \bar{\boldsymbol{\Upsilon}}^{-1}\right)^{-1} \mathbf{S}^H \mathbf{z}. \tag{3.42}$$

The MSE of the mismatched Wiener solution is represented as

$$\text{MSE}_{W,2} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \boldsymbol{\Upsilon} - \left[2\boldsymbol{\Upsilon} - \boldsymbol{\Upsilon} \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \boldsymbol{\Upsilon}^{-1} \right) \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \bar{\boldsymbol{\Upsilon}}^{-1} \right)^{-1} \right] \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \bar{\boldsymbol{\Upsilon}}^{-1} \right)^{-1} \right\}. \tag{3.43}$$

BEM-based MMSE estimator

The MMSE channel estimator in the second approach is given by

$$\bar{\mathbf{h}}_{\text{MMSE},2} = \mathbf{A}_{\text{MMSE},2} \mathbf{z}, \tag{3.44}$$

where $\mathbf{A}_{\text{MMSE},2} = \left(\sigma_s^2 \bar{\mathbf{B}}^H \bar{\mathbf{B}} + \sigma_n^2 \bar{\mathbf{R}}_a\right)^{-1} \bar{\mathbf{B}}^H \mathbf{S}^H$, and the MSE is represented as:

$$\text{MSE}_{\text{MMSE},2} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \frac{\sigma_n^2}{\sigma_s^2} \left[\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \left(\bar{\mathbf{B}}^H \bar{\mathbf{B}} \bar{\mathbf{R}}_a\right)^{-1} \right]^{-2} \right.$$
$$\left. + \left[\mathbf{I} - \bar{\mathbf{B}} \left(\bar{\mathbf{B}}^H \bar{\mathbf{B}} + \frac{\sigma_n^2}{\sigma_s^2} \bar{\mathbf{R}}_a^{-1} \right)^{-1} \bar{\mathbf{B}}^H \right]^2 \boldsymbol{\Upsilon} \right\}. \tag{3.45}$$

BEM-based ML estimator

The ML estimator using in the second approach is given by

$$\bar{\mathbf{h}}_{\text{ML},2} = \mathbf{A}_{\text{ML},2} \mathbf{z}, \tag{3.46}$$

where $\mathbf{A}_{\text{ML},2} = \left(\sigma_s^2 \bar{\mathbf{B}}^H \bar{\mathbf{B}}\right)^{-1} \bar{\mathbf{B}}^H \mathbf{S}^H$, and the MSE in represented as

$$\text{MSE}_{\text{ML},2} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \left(\mathbf{I} - \bar{\mathbf{B}}(\bar{\mathbf{B}}^H \bar{\mathbf{B}})^{-1} \bar{\mathbf{B}}^H\right) \boldsymbol{\Upsilon} \right\} + \frac{\sigma_n^2 M}{\sigma_s^2 \sigma_h^2 N}. \tag{3.47}$$

3.5.1 MSE performance
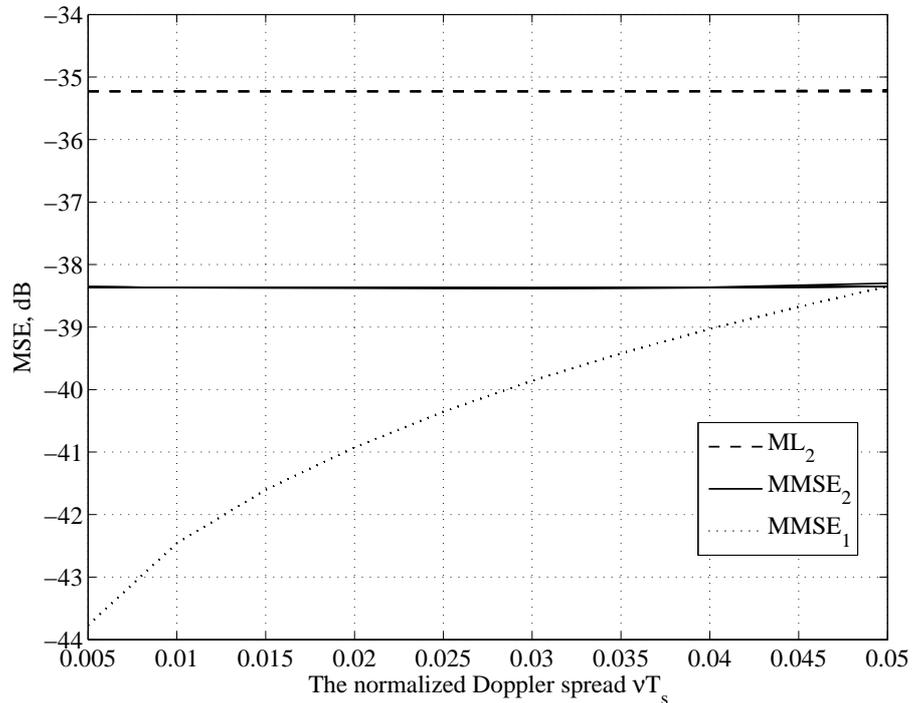
Figure 3.6: MSE performance of estimators with all BEMs using the maximum Doppler spread, $\bar{\nu}_{\max}T_s = 0.05$, $N = 100$, $M = 26$ and SNR = 30dB.

Fig. 3.6 shows the degradation in the MSE performance of the second approach compared with the first approach. We set $N = 100$, SNR = 30dB and $\nu_{\max}T_s = 0.05$. Based on the results shown in Fig. 3.3, we use $M = 26$ basis functions to guarantee that all BEMs provide the best performance for $\nu_{\max}T_s = 0.05$. In Fig. 3.6, ML_2 and $MMSE_2$ indicate the ML and MMSE estimators in the second approach, respectively, and $MMSE_1$ denotes the MMSE estimators in the first approach. It can be seen that, in the second approach, the MMSE estimators for all BEMs provide a similar performance as that of the mismatched Wiener solution using $\nu_{\max}$. In the second approach, the ML estimators using all BEMs provide a similar performance, which is 3.2dB inferior to that of the MMSE estimators.

Compared with the performance of MMSE estimators in the first approach, the performance of the MMSE estimator in the second approach degrades when ν significantly mismatches with $\nu_{\max}$. For example, when $\nu = \nu_{\max}T_s = 0.05$, which indicates that there is no mismatch, the performance of the MMSE estimator in the second approach is the

same as that of the MMSE estimators in the first approach. However, when $\nu T_s = 0.01$, the performance of the MMSE estimators using $\nu_{\max} T_s = 0.05$ is 4.2dB inferior to the performance of the MMSE estimators in the first approach.

3.5.2 Complexity

Calculation	KL	DPS	BS	GCE
$\bar{\mathbf{B}}$	–	–	–	–
$\bar{\mathbf{R}}_a$	–	–	–	–
$\bar{\mathbf{R}}_a^{-1}$	–	–	–	–
$\mathbf{V} = \left(\sigma_s^2 \bar{\mathbf{B}}^H \bar{\mathbf{B}} + \sigma_n^2 \bar{\mathbf{R}}_a^{-1} \right)^{-1}$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$
$\mathbf{A}_{\text{MMSE,2}} = \mathbf{V}\bar{\mathbf{B}}^H \mathbf{S}^H$	$(M^2 + M)N$	$(M^2 + M)N$	$(M^2 + M)N$	$(M^2 + M)N$
$\hat{\mathbf{a}}_{\text{MMSE,2}} = \mathbf{A}_{\text{MMSE,2}}\mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	MN_dN	MN_dN	$(q+1)N_dN$	MN_dN

Table 3.3: The number of complex multiplications required by MMSE estimators using different BEMs in the second approach.

Calculation	KL	DPS	BS	GCE
$\bar{\mathbf{B}}$	–	–	–	–
$\mathbf{A}_{\text{ML,2}} = \left(\sigma_s^2 \bar{\mathbf{B}}^H \bar{\mathbf{B}} \right)^{-1} \mathbf{B}^H \mathbf{S}^H$	$MN + M^2$	$MN + M^2$	$MM + M^2$	$MN + M^2$
$\hat{\mathbf{a}}_{\text{ML,2}} = \mathbf{A}_{\text{ML,2}}\mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	MN_dN	MN_dN	$(q+1)N_dN$	MN_dN

Table 3.4: The number of complex multiplications required by ML estimators using different BEMs in the second approach.

Now, we will investigate the complexity of BEM-based estimators in the second approach. Table. 3.3 shows the number of complex multiplications required by MMSE estimators using different BEMs. The matrices $\bar{\mathbf{B}}$, $\bar{\boldsymbol{\Upsilon}}$ and $\bar{\mathbf{R}}_a^{-1}$ can be precalculated and saved in memory. Therefore, the estimators with KL and DPS BEMs are significantly simplified. However, for the estimators using BS and GCE BEMs, the simplification is not significant. We also find that the MMSE estimators using all BEMs require a similar number of multiplications when $M < 5$. For the case with $M > 5$, the complexity of the estimator using the BS BEM is lower than that of the others.

Table. 3.4 shows the number of complex multiplications required by ML estimators using different BEMs in the second approach. By comparing with Table. 3.2, it can be seen that the complexity of the ML estimators using KL and DPS BEMs is significantly

reduced. However, the complexity of the ML estimators using BS and GCE BEMs is the
same as that of the estimators in the first approach.

The results shown in this section indicate that the use of the maximum Doppler spread
can significantly simplify the estimation when using KL and DPS BEMs. However, for
the estimation using BS and GCE BEMs, the simplification is not significant. Note that
in the second approach, the complexity of the estimator using BS BEM is still the lowest.
The MSE performance of the estimators in the second approach is worse than that of the
estimators in the first approach, especially when $\nu_{\max} >> \nu$. In the next section, we will
investigate the third approach using an estimate of the Doppler spread, and compare its
performance and complexity with those of the second approach.

3.6 Approach 3: Channel estimation using an estimate of the Doppler spread

In this approach, we use $\hat{\nu}$, an estimate of the Doppler spread, to compute $\mathbf{R}_a$ and generate
the KL and DPS basis functions. The Doppler spread can be estimated based on corre-
lation and variation of channel estimates. For example, in [111], the Doppler estimation
scheme based on the autocorrelation of complex channel estimates is described. Instead
of using channel estimates, the received signal can also be used directly in estimating
Doppler spread information. In [112], the mobile speed is estimated as a function of the
deviation of the averaged signal envelope in flat fading channels. By using the estimate
of the Doppler spread, elements of the estimated covariance matrix, $\hat{\mathbf{\Upsilon}}$, can be calculated
by

$$[\hat{\mathbf{\Upsilon}}]_{t_1,t_2} = \hat{\rho}(t_1 - t_2), \tag{3.48}$$

where

$$\hat{\rho}(\tau) = \sigma_h^2 J_0(2\pi\hat{\nu}\tau). \tag{3.49}$$

The estimated covariance matrix of the expansion coefficients is given by

$$\hat{\mathbf{R}}_a = (\hat{\mathbf{B}}^H\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}^H\hat{\mathbf{\Upsilon}}\hat{\mathbf{B}}(\hat{\mathbf{B}}^H\hat{\mathbf{B}})^{-1}. \tag{3.50}$$

For KL and DPS BEMs, $\hat{\mathbf{B}}$ is a matrix containing samples of the basis functions generated by using $\hat{\nu}$. For BS and GCE BEMs, $\hat{\mathbf{B}} = \mathbf{B}$.

By applying matrices $\hat{\mathbf{B}}$ and $\hat{\mathbf{R}}_a$, we obtain the MSE for the mismatched Wiener solution, as well as the MMSE and ML estimators.

Mismatched Wiener solution

The mismatched Wiener solution in the third approach is given by

$$\hat{\mathbf{h}}_{W,3} = \left(\sigma_s^2 \mathbf{I}_N + \sigma_n^2 \hat{\boldsymbol{\Upsilon}}^{-1} \right)^{-1} \mathbf{S}^H \mathbf{z}. \tag{3.51}$$

The MSE of the mismatched Wiener solution is represented as

$$\text{MSE}_{W,3} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \boldsymbol{\Upsilon} - \left[2\boldsymbol{\Upsilon} - \boldsymbol{\Upsilon} \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \boldsymbol{\Upsilon}^{-1} \right) \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \hat{\boldsymbol{\Upsilon}}^{-1} \right)^{-1} \right] \left(\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \hat{\boldsymbol{\Upsilon}}^{-1} \right)^{-1} \right\}. \tag{3.52}$$

BEM-based MMSE estimators

The MMSE channel estimator in the third approach is given by

$$\hat{\mathbf{h}}_{\text{MMSE},3} = \mathbf{A}_{\text{MMSE},3} \mathbf{z}, \tag{3.53}$$

where $\mathbf{A}_{\text{MMSE},3} = \left(\sigma_s^2 \hat{\mathbf{B}}^H \hat{\mathbf{B}} + \sigma_n^2 \hat{\mathbf{R}}_a \right)^{-1} \hat{\mathbf{B}}^H \mathbf{S}^H$, and the MSE is represented as:

$$\text{MSE}_{\text{MMSE},3} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \frac{\sigma_n^2}{\sigma_s^2} \left[\mathbf{I} + \frac{\sigma_n^2}{\sigma_s^2} \left(\hat{\mathbf{B}}^H \hat{\mathbf{B}} \hat{\mathbf{R}}_a \right)^{-1} \right]^{-2} \right.$$
$$\left. + \left[\mathbf{I} - \hat{\mathbf{B}} \left(\hat{\mathbf{B}}^H \hat{\mathbf{B}} + \frac{\sigma_n^2}{\sigma_s^2} \hat{\mathbf{R}}_a^{-1} \right)^{-1} \hat{\mathbf{B}}^H \right]^2 \boldsymbol{\Upsilon} \right\}. \tag{3.54}$$

BEM-based ML estimator

The ML estimator in the third approach is given by

$$\hat{\mathbf{h}}_{\text{ML},3} = \mathbf{A}_{\text{ML},3} \mathbf{z}, \tag{3.55}$$

where $\mathbf{A}_{\text{ML},3} = \left(\sigma_s^2 \hat{\mathbf{B}}^H \hat{\mathbf{B}} \right)^{-1} \hat{\mathbf{B}}^H \mathbf{S}^H$, and the MSE is represented as

$$\text{MSE}_{\text{ML},3} = \frac{1}{\sigma_h^2 N} \text{tr} \left\{ \left(\mathbf{I} - \hat{\mathbf{B}}(\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^H \right) \mathbf{\Upsilon} \right\} + \frac{\sigma_n^2 M}{\sigma_s^2 \sigma_h^2 N}. \tag{3.56}$$

3.6.1 MSE performance

Generally, an estimate of ν can be represented as

$$\hat{\nu} = \nu + \Delta\nu, \tag{3.57}$$

where $\Delta\nu$ corresponds to overestimation ($\Delta\nu > 0$) or underestimation ($\Delta\nu < 0$) of the Doppler spread. The sensitivity of the estimators to overestimation and underestimation is different.

Fig. 3.7 shows the MSE performance of the BEM-based MMSE estimators in the third appraoch versus $\Delta\nu$ for the case with $N = 100$, SNR $= 30$dB, and $\nu T_s = 0.02$. We set $M = 13$. The mismatched Wiener solution and MMSE estimators using all BEMs are sensitive to underestimation of the Doppler spread ($\Delta\nu < 0$). In the case of overestimation ($\Delta\nu > 0$), the performance of the MMSE estimators using BS and GCE BEMs is similar and close to that of the mismatched Wiener solution. However, the performance of the MMSE estimators using KL and DPS BEMs degrades when the Doppler spread is significantly overestimated.

Fig. 3.8 shows the MSE performance of the BEM-based ML estimators in the third approach versus $\Delta\nu$ for the case with $N = 100$, SNR $= 30$dB, and $\nu T_s = 0.02$. We also set $M = 13$. It is seen that the performance of ML estimators using BS and GCE BEMs is not that significantly affected by $\Delta\nu$. However, the performance of the ML estimators using KL and DPS BEMs degrades significantly when ν is underestimated or significantly overestimated. Therefore, the estimators using BS and GCE BEMs are more robust than the estimators using KL and DPS BEMs.

In Fig. 3.9, we plot the MSE performance of the MMSE and ML estimation in the third approach together and find that the MMSE estimators are more sensitive to underestimation of the Doppler spread, but provide better performance than the ML estimators
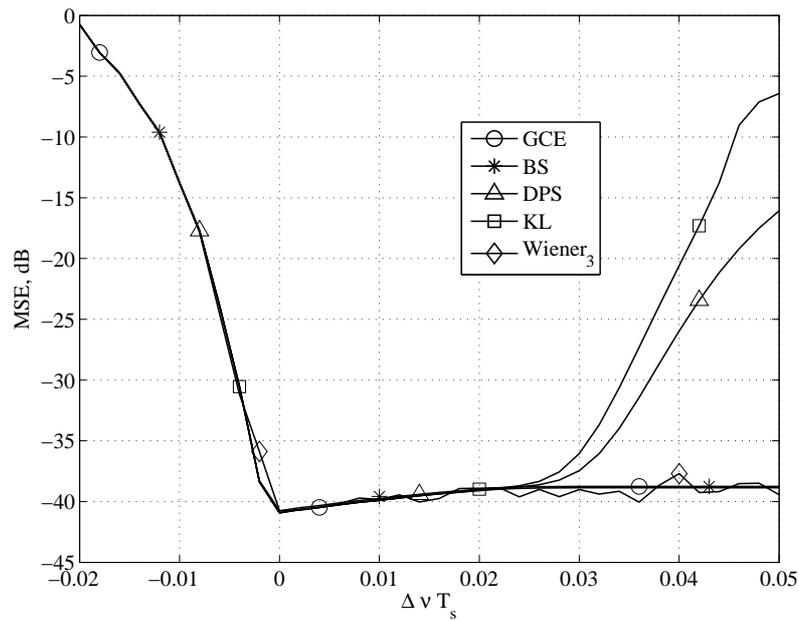
Figure 3.7: MSE performance of MMSE estimators in the third approach versus the change of the Doppler spread, $\nu T_s = 0.02$, $N = 100$, $M = 13$ and SNR = 30dB.
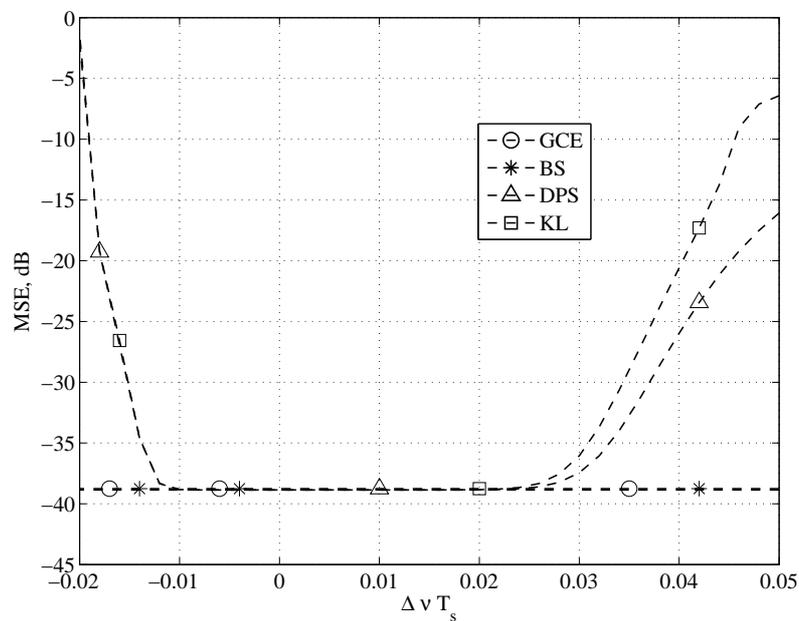
Figure 3.8: MSE performance of ML estimators in the third approach versus the change of the Doppler spread, $\nu T_s = 0.02$, $N = 100$, $M = 13$ and SNR = 30dB.
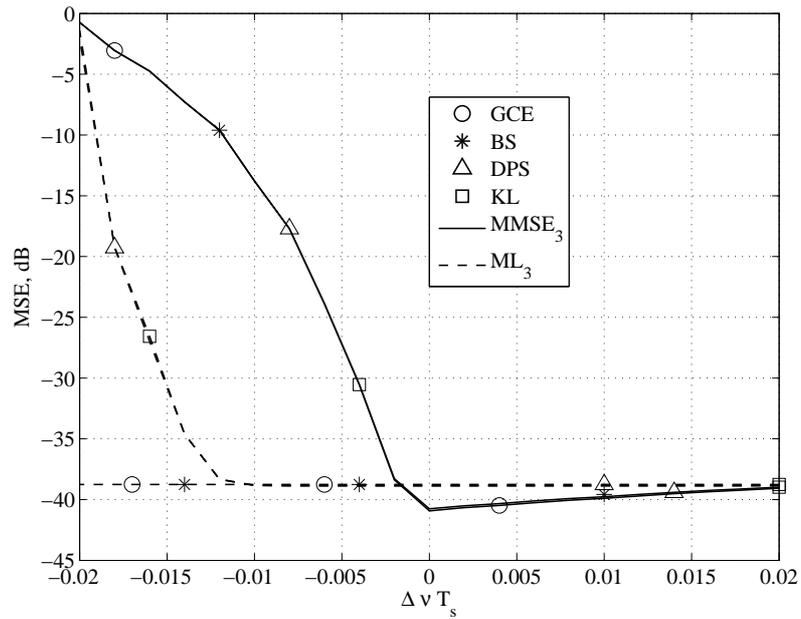
Figure 3.9: MSE performance of MMSE and ML estimators in the third approach versus
the change of the Doppler spread, $\nu T_s = 0.02$, $N = 100$, $M = 13$ and SNR = 30dB.
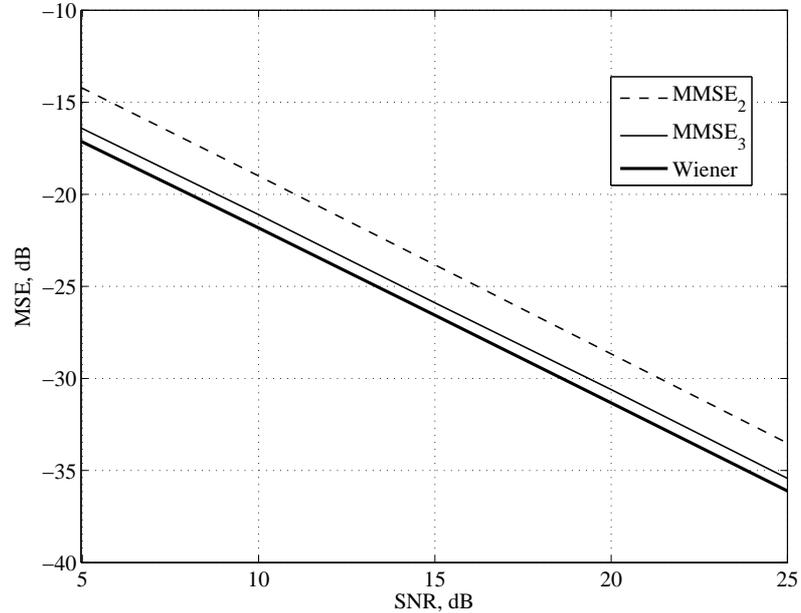
Figure 3.10: MSE performance of the BEM-based MMSE estimators using all BEMs
with different estimated Doppler spreads, $N = 100$, $M = 26$, $\nu_{\max} T_s = 0.05$, $\rho = 0.1$ and
$20\% \hat{\nu}$ overestimation.

over the range $0 < \Delta\nu T_s < 0.02$. Therefore, in the third approach, the Doppler spread is suggested to be slightly overestimated for the BEM-based MMSE estimation. Normally, a Doppler spread estimator provides an unbiased estimation of the Doppler spread [110], which can often be represented as a normal random variable with a mean $\nu_m = \nu$ and variance ρ^2, where $\rho = \alpha\nu$ and $\alpha \in [0, 1]$ characterizes the accuracy of estimating the Doppler spread. In order to reduce the chance of underestimation, we suggest to add a small positive value, $20\%\hat{\nu}$ to the unbiased estimate $\hat{\nu}$.

Now we compare the MSE performance of the MMSE estimators in the third approach using the suggested overestimate of the Doppler spread with that of the MMSE estimators in the second approach using $\nu_{\max}$. We consider a case with $\nu T_s = 0.02$, $N = 100$. For the second approach, we set $\nu_{\max}T_s = 0.05$ and $M = 26$. For the third approach, we set $M = 26$, $\alpha = 10\%$ and $20\%\hat{\nu}$ overestimation. The performance of the Wiener solution in the first approach using perfect knowledge of ν is considered as a lower MSE bound. The results are shown in Fig. 3.10; they have been obtained by averaging over 1000 simulation trials. It can be seen that the performance of the BEM-based MMSE estimators in the third approach using a slight ($20\%\hat{\nu}$) overestimate of the Doppler spread is only 0.5dB inferior to the lower bound, and it outperforms that of the BEM-based MMSE estimators in the second approach using $\bar{\nu}_{\max}T_s = 0.05$ by 2.3dB for all SNRs.

3.6.2 Complexity

Calculation	KL	DPS	BS	GCE
$\hat{\mathbf{B}}$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	$-$	$-$
$\hat{\mathbf{R}}_a$	$MN^2 + M^2N$	$MN^2 + M^2N$	$MN^2 + M^2N$	$MN^2 + M^2N$
$\hat{\mathbf{R}}_a^{-1}$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$
$\mathbf{V} = \left(\sigma_s^2\hat{\mathbf{B}}^H\hat{\mathbf{B}} + \sigma_n^2\hat{\mathbf{R}}_a^{-1}\right)^{-1}$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$	$\mathcal{O}(M^3)$
$\mathbf{A}_{\text{MMSE,3}} = \mathbf{V}\hat{\mathbf{B}}^H\mathbf{S}^H$	$(M^2 + M)N$	$(M^2 + M)N$	$(q+1)N + (q+1)KM$	$(M^2 + M)N$
$\hat{\mathbf{a}}_{\text{MMSE,3}} = \mathbf{A}_{\text{MMSE}}\mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	MN_dN	MN_dN	$(q+1)N_dN$	MN_dN

Table 3.5: The number of complex multiplications required by MMSE estimators using different BEMs in the third approach using an estimate of the Doppler spread.

The number of complex multiplications required by the MMSE and ML estimators using different BEMs in the third approach is shown in Table. 3.5 and Table. 3.6,

Calculation	KL	DPS	BS	GCE
$\hat{\mathbf{B}}$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	–	–
$\mathbf{A}_{\mathrm{ML},3} = \left(\sigma_s^2 \hat{\mathbf{B}}^H \hat{\mathbf{B}}\right)^{-1} \hat{\mathbf{B}}^H \mathbf{S}^H$	$MN + M^2$	$MN + M^2$	$MN + M^2$	$MN + M^2$
$\hat{\mathbf{a}}_{\mathrm{ML},3} = \mathbf{A}_{\mathrm{ML},3} \mathbf{z}$	MN	MN	MN	MN
To estimate NN_d channel coefficients	$MN_d N$	$MN_d N$	$(q+1)N_d N$	$MN_d N$

Table 3.6: The number of complex multiplications required by ML estimators using different BEMs in the third approach using an estimate of the Doppler spread.

respectively. It is clear that the complexity of the third approach is the same as that of the first approach. The complexity of the estimators using KL and DPS BEMs in the third approach is much higher than that of the estimators in the second approach. For the estimation using BS and GCE BEMs in the third approach, the complexity is close to that of the estimation in the second approach.

By comparing the third approach using a slight $(20\%\hat{\nu})$ overestimate of the Doppler spread with the second approach using the maximum Doppler spread, we find that although the complexity of the second approach is lower, its performance is worse than that of the third approach. Moreover, the increase of the complexity in the third approach is only significant for the estimation using KL and DPS BEMs. For the estimation using BS and GCE BEMs, the complexity of the third approach is close to that of the second approach. It is also worthy to emphasize that the complexity of the estimators using BS BEM is the lowest in all approaches. Therefore, the BS BEM is the most practical choice due to its highest robustness and lowest complexity.

3.7   Conclusions

In this chapter, BEM-based estimators for time-variant fading channels have been investigated. The MSE of a generic linear channel estimator with linearly independent BEMs has been derived. We have investigated and compared the MSE performance and complexity for three approaches using different BEMs, such as KL, DPS, GCE and BS BEMs for estimation with perfect or inaccurate knowledge of the Doppler spread:
1) channel estimation using perfect knowledge of the Doppler spread;

2) channel estimation using the maximum Doppler spread as suggested in [20];

3) channel estimation with an estimate of the Doppler spread.

In the first approach based on the impractical assumption of perfect knowledge of the Doppler spread, when the number of basis functions increases, all BEM-based MMSE estimators can provide a similar performance close to that of the Wiener solution. However, the performance of the ML estimator degrades after an optimal number of basis functions. Although the B-splines require slightly more basis functions than the other BEMs to achieve the same estimation performance, the complexity of the estimator using B-splines is still lower than that of the others.

In the second approach, the maximum Doppler spread is used to calculate the fading statistics and generate the KL and DPS basis functions. Although the estimation is simplified with respect to the first approach, the significant mismatch between the maximum Doppler spread and the real Doppler spread leads to degradation in the MSE performance. Moreover, the decrease of the complexity is only significant for the estimation using KL and DPS BEMs, but not significant for the estimation using BS and GCE BEMs.

The third approach uses an estimate of the Doppler spread to compute the fading statistics and generate the KL and DPS basis functions. In this approach, all estimators are sensitive to underestimation of the Doppler spread but may have little sensitivity to overestimation depending on the BEM used. The estimators using BS and GCE BEMs are more robust than the estimators using KL and DPS BEMs. The estimation in this approach using a slight (20%) overestimate of the Doppler spread outperforms the estimation in the second approach using the maximum Doppler spread. The MSE performance of the MMSE estimation in the third approach is very close to the performance of the Wiener solution with perfect knowledge of the Doppler spread. In this approach, the estimation with KL and DPS BEMs is more complicated than those in the second approach. However, the increase of complexity for the estimation using BS and GCE BEMs in this approach is close to that in the second approach.

It is also worthy to emphasize that the complexity of the estimation using BS BEM is the lowest in each approach. Therefore, the BS-based estimation using a slight (e.g., 20%) overestimate of the Doppler spread is a good practical choice providing a good

performance, high robustness and low complexity.

In the following chapters, we will investigate the best performance that can be obtained by using the optimal detection with imperfect channel estimation. Therefore, we will assume that the Doppler spread is perfectly known. Moreover, the cubic BS BEM will be used to approximate the time-variant channels.

Chapter 4

Optimal and Mismatched Detection in SISO Frequency-Flat Fading Channels with Imperfect Channel Estimation

Contents

4.1   Introduction

In communication systems transmitting data through unknown channels, traditional detection techniques are based on channel estimation (e.g., by using pilot signals), and then

treating the estimates as perfect in a minimum distance detector; we call such detectors
mismatched (after [35]). A better detection performance can be obtained in an *optimal*
detector that does not estimate the channel explicitly but jointly processes the received
pilot and data symbols to recover the data [35]. The optimal detector in [35] was obtained
for communication scenarios with space-time coding in a channel with uncorrelated fad-
ing and additive white noise. In this chapter, we consider a more general scenario that is
applicable to channels with correlated fading. We derive the generic optimal detector and
specify it for frequency-flat fading channels. We then compare its detection performance
with that of mismatched detectors using different channel estimation techniques for both
cases of time-invariant and time-variant fading.

In time-variant fading channels, the channel estimation is difficult, especially in sys-
tems with powerful channel codes, such as turbo codes, generally operating at low SNR
where pilot-based channel estimates are often of low accuracy. Iterative channel esti-
mation and decoding over flat fading channels has been proposed to improve the esti-
mates [63, 115]. In this chapter, we will consider representation of channel gain time
variations by using basis functions. As shown in Chapter 3, compared with the other
BEMs, the B-splines provide high accuracy of approximation and require lower complex-
ity. Therefore, we consider approximation of the channel time variations by B-splines,
and, for this case, investigate the detection performance of the optimal detector. We then
investigate iterative receivers that exchange channel and data estimates in a system trans-
mitting turbo-encoded data. Four channel estimation schemes are considered: the ML
estimator, MMSE estimator, regularized ML estimator, and an estimator providing statis-
tics for the optimal detector.

This chapter is organized as follows. In Section 4.2, the transmission model is in-
troduced. The generic optimal detection and mismatched detection are derived and de-
scribed in Section 4.3 and Section 4.4, respectively. In Section 4.5, the optimal detector is
specified for time-invariant channels. Section 4.6 specifies the optimal detector for time-
variant fading channels approximated by B-splines and describes iterative receivers with
soft-input hard-output and/or soft-input soft-output turbo decoding scheme. Numerical
results are given in Section 4.7. Section 4.8 concludes the chapter.

4.2 Transmission model

We assume that the data transmission is split into two parts, pilot and data transmission. For the pilot transmission, the received signal is modeled as

$$\mathbf{z}_p = \mathbf{\Psi}_p \mathbf{a} + \mathbf{n}_p \tag{4.1}$$

where $\mathbf{z}_p$ is an $N_p \times 1$ received signal vector and $\mathbf{\Psi}_p$ is an $N_p \times M$ matrix formed from pilot symbols (for time-invariant channels) or formed from the multiplication of pilot symbols and basis functions (for time-variant channels). The $N_p \times 1$ complex-valued noise vector $\mathbf{n}_p$ has a zero mean Gaussian PDF $\mathcal{N}_{\mathcal{C}}(\mathbf{0}, \mathbf{R}_p)$ with covariance matrix $\mathbf{R}_p = E\{\mathbf{n}_p \mathbf{n}_p^H\}$. The vector $\mathbf{a} = [a_1, \dots, a_M]^T$ is an $M \times 1$ vector of complex-valued channel gains (for time-invariant channels) or basis coefficients (for time-invariant channels) with the Gaussian PDF

$$f(\mathbf{a}) = \mathcal{N}_{\mathcal{C}}(\mathbf{0}, \mathbf{R}_a) \tag{4.2}$$

where $\mathbf{R}_a = E\{\mathbf{a}\mathbf{a}^H\}$ is an $M \times M$ covariance matrix. The function $f(\mathbf{a})$ defines a Rayleigh fading channel. The definition of the vector $\mathbf{a}$ depends on the considered channel. In time-invariant channels, $\mathbf{a}$ contains the channel gains, i.e., $\mathbf{a} = h$, and in time-variant channels, $\mathbf{a}$ contains the expansion coefficients, i.e., $\mathbf{a} = [a_1, \dots, a_M]^H$. Correspondingly, the structure of $\mathbf{\Psi}_p$ and is modified. In time-invariant channels, $\mathbf{\Psi}_p = \mathbf{s}_p = [s_p(1), \dots, s_p(i), \dots, s_p(N_p)]^H$ where $s_p(i)$ is an transmitted pilot symbols. In time-variant channels, $\mathbf{\Psi}_p$ and $\mathbf{\Psi}_d$ become matrices as

$$\mathbf{\Psi}_p = \mathbf{D}_p \mathbf{B}_p, \tag{4.3}$$

where $\mathbf{D}_p$ being a diagonal matrix defined as

$$\mathbf{D}_p = \mathrm{diag}\{s_p(t_1), \dots, s_p(t_{N_p})\}, \tag{4.4}$$

The matrices $\mathbf{B}_p$ contain samples of the basis functions at the pilot instants. The details will be introduced in section 4.5 and section 4.6.1 corresponding to time-invariant channels and time-variant channels, respectively.

The PDF of the received signal vector $\mathbf{z}_p$ for a given vector $\mathbf{a}$ is

$$p(\mathbf{z}_p|\mathbf{a}) = \mathcal{N}_{\mathcal{C}}(\mathbf{\Psi}_p \mathbf{a}, \mathbf{R}_p). \tag{4.5}$$

For the data transmission, the received signal is modeled as

$$\mathbf{z}_d = \boldsymbol{\Psi}_d \mathbf{a} + \mathbf{n}_d \tag{4.6}$$

where $\mathbf{z}_d$ is the received $N_d \times 1$ data vector and $\boldsymbol{\Psi}_d$ is an $N_d \times M$ matrix which depends on a vector $\mathbf{d}$ of transmitted data symbols. In time-variant channels, $\boldsymbol{\Psi}_d = [d(1), \ldots, d(i), \ldots, d(N_d)]^H$ where $d(i)$ is an transmitted data symbol, and $\boldsymbol{\Psi}_d = \mathbf{D}_d \mathbf{B}_d$, where $\mathbf{D}_d = \text{diag}\{d_1, \ldots, d_{N_d}\}$ and $\mathbf{B}_d$ contains samples of the basis functions at the data instants. The $N_d \times 1$ noise vector $\mathbf{n}_d$ has the Gaussian PDF $\mathcal{N}_\mathcal{C}(\mathbf{0}, \mathbf{R}_d)$ with covariance matrix $\mathbf{R}_d = E\{\mathbf{n}_d \mathbf{n}_d^H\}$. The PDF of the vector $\mathbf{z}_d$ for given vectors $\mathbf{d}$ and $\mathbf{a}$ is also Gaussian:

$$p(\mathbf{z}_d | \mathbf{d}, \mathbf{a}) = \mathcal{N}_\mathcal{C}(\boldsymbol{\Psi}_d \mathbf{a}, \mathbf{R}_d). \tag{4.7}$$

The transmission model presented in (4.1) and (4.6) is quite general. It can be used to describe single-input single-output or multi-input multi-output systems in both frequency-flat and frequency-selective fading channels or in time invariant or variant channels. For different channels, the structure of matrices or vectors defined above, i.e., $\boldsymbol{\Psi}_d$, $\boldsymbol{\Psi}_p$, $\mathbf{z}_p$, $\mathbf{z}_d$ and $\mathbf{a}$ should be modified. The modifications of these matrices and vectors corresponding are declared in following chapters corresponding to different channels considered.

4.3   Generic optimal detection

For the described transmission model, the optimal detector is derived by maximizing the PDF $p(\mathbf{z}_d | \mathbf{d}, \mathbf{z}_p)$ of the signal $\mathbf{z}_d$ received at the data stage, conditioned on the transmitted symbols $\mathbf{d}$ and the signal $\mathbf{z}_p$ received at the pilot stage, over the data set (alphabet) $\mathcal{A}$:

$$
\begin{aligned}
\hat{\mathbf{d}}_{\text{opt}} &= \arg \max_{\mathbf{d} \in \mathcal{A}} \{ p(\mathbf{z}_d | \mathbf{d}, \mathbf{z}_p) \} \\
&= \arg \max_{\mathbf{d} \in \mathcal{A}} \{ \ln [p(\mathbf{z}_d | \mathbf{d}, \mathbf{z}_p)] \} .
\end{aligned} \tag{4.8}
$$

In the case of QAM transmission, the alphabet $\mathcal{A}$ includes all symbols corresponding to the QAM constellation points. The PDF $p(\mathbf{z}_d | \mathbf{d}, \mathbf{z}_p)$ is obtained from the PDF $p(\mathbf{z}_d | \mathbf{d}, \mathbf{a})$ in (4.7) by integrating out the channel parameters $\mathbf{a}$ which are now treated as *nuisance parameters*:

$$p(\mathbf{z}_d | \mathbf{d}, \mathbf{z}_p) = \int p(\mathbf{z}_d | \mathbf{d}, \mathbf{a}) f(\mathbf{a} | \mathbf{z}_p) d\mathbf{a} \tag{4.9}$$

where the *posterior* PDF $f(\mathbf{a}|\mathbf{z}_p)$ of channel parameters is conditioned on the received pilot signal $\mathbf{z}_p$. Since (4.1) is the *Bayesian general linear model*, the PDF $f(\mathbf{a}|\mathbf{z}_p)$ is also Gaussian (see [113], pp.326),

$$f(\mathbf{a}|\mathbf{z}_p) = \mathcal{N}_{\mathcal{C}}(\mathbf{m_a}, \mathbf{S_a}), \tag{4.10}$$

with mean $\mathbf{m_a}$ and covariance $\mathbf{S_a}$ given by

$$\mathbf{m_a} = (\mathbf{\Gamma}_p + \mathbf{R}_a^{-1})^{-1}\mathbf{L}_p, \tag{4.11}$$

$$\mathbf{S_a} = (\mathbf{\Gamma}_p + \mathbf{R}_a^{-1})^{-1}, \tag{4.12}$$

where

$$\mathbf{L}_p = \mathbf{\Psi}_p^H \mathbf{R}_p^{-1} \mathbf{z}_p, \tag{4.13}$$

$$\mathbf{\Gamma}_p = \mathbf{\Psi}_p^H \mathbf{R}_p^{-1} \mathbf{\Psi}_p. \tag{4.14}$$

If $\mathbf{R}_p = \sigma_n^2 \mathbf{I}_{N_p}$ (i.e., the noise is white with variance σ_n^2), then we have

$$\mathbf{L}_p = \sigma_n^{-2}\mathbf{\Psi}_p^H \mathbf{z}_p, \tag{4.15}$$

$$\mathbf{\Gamma}_p = \sigma_n^{-2}\mathbf{\Psi}_p^H \mathbf{\Psi}_p, \tag{4.16}$$

where $\mathbf{I}_{N_p}$ denotes an $N_p \times N_p$ identity matrix.

By substituting (4.7) in (4.9), we obtain

$$p(\mathbf{z}_d|\mathbf{d}, \mathbf{z}_p) = c_3 \int e^{2\Re(\mathbf{a}^H \mathbf{L}_d) - \mathbf{a}^H \mathbf{\Gamma}_d \mathbf{a}} f(\mathbf{a}|\mathbf{z}_p) d\Re(\mathbf{a}) d\Im(\mathbf{a}) \tag{4.17}$$

where $\Re(\cdot)$ denotes the real part and $\Im(\cdot)$ denotes the imaginary part,

$$c_3 = \frac{1}{\pi^{N_d}|\mathbf{R}_d|} e^{-\mathbf{z}_d^H \mathbf{R}_d^{-1} \mathbf{z}_d} > 0$$

is a constant, and $|\mathbf{R}_d|$ denotes the determinant of the matrix $\mathbf{R}_d$. By substituting (4.10) in (4.17) and taking into account (7.3) in Appendix A, we obtain

$$p(\mathbf{z}_d|\mathbf{d}, \mathbf{z}_p) = \frac{c_3}{|\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M|} \exp\left\{(\mathbf{L}_d + \mathbf{L}_p)^H (\mathbf{\Gamma}_d + \mathbf{S_a}^{-1})^{-1}(\mathbf{L}_d + \mathbf{L}_p)\right\} \tag{4.18}$$

$$= \frac{c_3}{|\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M|} \exp\left\{2\Re[\mathbf{L}_d^H(\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{m_a}] + \mathbf{L}_d^H(\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{S_a}\mathbf{L}_d \right.$$
$$\left. -\mathbf{m_a}^H\mathbf{\Gamma}_d(\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{m_a}\right\} \tag{4.19}$$

where

$$\mathbf{L}_d = \mathbf{\Psi}_d^H \mathbf{R}_d^{-1} \mathbf{z}_d, \tag{4.20}$$

$$\mathbf{\Gamma}_d = \mathbf{\Psi}_d^H \mathbf{R}_d^{-1} \mathbf{\Psi}_d. \tag{4.21}$$

For the white noise case, we have $\mathbf{R}_d = \sigma_n^2 \mathbf{I}_{N_d}$ and

$$
\mathbf{L}_d = \sigma_n^{-2} \boldsymbol{\Psi}_d^H \mathbf{z}_d, \tag{4.22}
$$

$$
\boldsymbol{\Gamma}_d = \sigma_n^{-2} \boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d. \tag{4.23}
$$

Finally, the optimal detector (4.8) is given by

$$
\hat{\mathbf{d}}_{\text{opt}} = \arg \min_{\mathbf{d} \in \mathcal{A}} \{\lambda(\mathbf{d})\}. \tag{4.24}
$$

where the metric $\lambda(\mathbf{d})$ to be minimized is given by

$$
\begin{aligned}
\lambda(\mathbf{d}) &= -(\mathbf{L}_d + \mathbf{L}_p)^H (\boldsymbol{\Gamma}_d + \boldsymbol{\Gamma}_p + \mathbf{R}_a^{-1})^{-1} (\mathbf{L}_d + \mathbf{L}_p) + \ln|\boldsymbol{\Gamma}_d + \boldsymbol{\Gamma}_p + \mathbf{R}_a^{-1}| \quad (4.25) \\
&= \ln|\mathbf{S}_{\mathbf{a}}\boldsymbol{\Gamma}_d + \mathbf{I}_M| - 2\Re[\mathbf{L}_d^H(\mathbf{S}_{\mathbf{a}}\boldsymbol{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{m}_{\mathbf{a}}] - \mathbf{L}_d^H(\mathbf{S}_{\mathbf{a}}\boldsymbol{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{S}_{\mathbf{a}}\mathbf{L}_d \\
&\quad + \mathbf{m}_{\mathbf{a}}^H \boldsymbol{\Gamma}_d (\mathbf{S}_{\mathbf{a}}\boldsymbol{\Gamma}_d + \mathbf{I}_M)^{-1}\mathbf{m}_{\mathbf{a}}. \quad (4.26)
\end{aligned}
$$

The first presentation (4.25) of the optimal metric $\lambda(\mathbf{d})$ shows how this metric is expressed in terms of the channel statistic $\mathbf{L}_p$, which is a vector of outputs of filters matched to the pilot signals, and the correlation matrix $\boldsymbol{\Gamma}_p$ of the pilot signals. The second presentation (4.26) shows how the optimal metric is expressed in terms of the mean $\mathbf{m}_{\mathbf{a}}$ and covariance $\mathbf{S}_{\mathbf{a}}$ of the posterior PDF $f(\mathbf{a}|\mathbf{z}_p)$.

If the perfect channel information (PCI) is available, we can write $\mathbf{m}_{\mathbf{a}} = \mathbf{a}$ and $\mathbf{S}_{\mathbf{a}} = \mathbf{0}_M$, where $\mathbf{0}_M$ is an $M \times M$ zero matrix. In this case, the metric (4.26) takes the form

$$
\lambda(\mathbf{d}) = -2\Re(\mathbf{L}_d^H \mathbf{a}) + \mathbf{a}^H \boldsymbol{\Gamma}_d \mathbf{a}. \tag{4.27}
$$

The detector minimizing the metric (4.27) is equivalent to the classical minimum distance detector

$$
\hat{\mathbf{d}}_{\text{PCI}} = \arg \min_{\mathbf{d} \in \mathcal{A}} \left\{ \|\mathbf{z}_d - \boldsymbol{\Psi}_d \mathbf{a}\|_{\mathbf{R}_d^{-1}}^2 \right\}. \tag{4.28}
$$

In what follows, we will only consider scenarios with additive white noise.

4.4 Generic mismatched detection

The vector $\mathbf{m}_{\mathbf{a}}$ in (4.11) is known to be the MMSE estimate of the channel parameters $\mathbf{a}$, i.e.,

$$
\hat{\mathbf{a}}_{\text{MMSE}} = \mathbf{m}_{\mathbf{a}}, \tag{4.29}
$$

and this estimate is unbiased and has the covariance matrix $\mathbf{S_a}$ (4.12) [113]. At high SNR,
i.e., $\sigma_n^2 \to 0$, we obtain

$$f(\mathbf{a}|\mathbf{z}_p) \to \delta(\mathbf{a} - \mathbf{m_a}), \tag{4.30}$$

where $\delta(\mathbf{a} - \mathbf{m_a})$ is an analog of the Dirac delta function for a vector argument. Then the
integration in (4.9) results in

$$p(\mathbf{z}_d|\mathbf{d}, \mathbf{z}_p) \to p(\mathbf{z}_d|\mathbf{d}, \mathbf{a} = \hat{\mathbf{a}}_{\text{MMSE}}), \tag{4.31}$$

and the optimal detector (4.8) becomes a mismatched detector where the MMSE channel
estimates

$$\hat{\mathbf{a}}_{\text{MMSE}} = (\mathbf{\Gamma}_p + \mathbf{R}_a^{-1}\mathbf{L}_p \tag{4.32}$$

are treated as perfect when minimizing the Euclidean distance:

$$\hat{\mathbf{d}}_{\text{MMSE}} = \arg\min_{\mathbf{d}\in\mathcal{A}} \left\{ (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_{\text{MMSE}})^H (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_{\text{MMSE}}) \right\}. \tag{4.33}$$

This motivates us to compare the optimal detector with the MMSE-mismatched detec-
tor. The mismatched detector with MMSE channel estimates exploits the same *a priori*
information as the optimal detector, but, in a different way. It minimizes the error of
channel estimates, while the optimal detector minimizes the probability of detection er-
rors. Therefore, when comparing the detection performance of the two detectors, we
expect the optimal detector to outperform the mismatched detector with MMSE channel
estimates.

The optimal detector and mismatched detector with MMSE channel estimates require
the knowledge of the fading statistics that are not always available. Therefore, it is of
interest to consider the mismatched detector with ML channel estimates

$$\hat{\mathbf{d}}_{\text{ML}} = \arg\max_{\mathbf{d}\in\mathcal{A}} \left\{ (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_{\text{ML}})^H (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_{\text{ML}}) \right\} \tag{4.34}$$

where the ML channel estimates are given by

$$\hat{\mathbf{a}}_{\text{ML}} = (\mathbf{\Psi}_p^H \mathbf{\Psi}_p)^{-1} \mathbf{\Psi}_p^H \mathbf{z}_p = \mathbf{\Gamma}_p^{-1}\mathbf{L}_p. \tag{4.35}$$

When comparing (4.32) and (4.35), it is seen that the inverse of the fading covariance ma-
trix $\mathbf{R}_a$ in (4.32) plays a regularization role. This, however, requires the fading covariance
to be known. Therefore, we will also consider a mismatched detector

$$\hat{\mathbf{d}}_\epsilon = \arg\max_{\mathbf{d}\in\mathcal{A}} \left\{ (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_\epsilon)^H (\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}_\epsilon) \right\} \tag{4.36}$$

with channel estimates using regularization based on the diagonal loading

$$\hat{\mathbf{a}}_\epsilon = (\mathbf{\Gamma}_p + \epsilon \mathbf{I}_M)^{-1} \mathbf{L}_p \tag{4.37}$$

where $\epsilon \geq 0$ is a regularization parameter. Such regularization does not require the fading statistics to be available. Note that for $\epsilon = 0$, we have $\hat{\mathbf{a}}_\epsilon = \hat{\mathbf{a}}_{\mathrm{ML}}$.

Thus, we are going to investigate the detection performance of the following detectors:

1) optimal detector defined by (4.24) and (4.25);

2) mismatched detector with MMSE channel estimates given by (4.33);

3) mismatched detector with ML channel estimates given by (4.34); and

4) mismatched detector with ϵ-ML channel estimates given by (4.36).

The relationship (4.24) describes the optimal detector applicable to many communications scenarios. However, in this chapter we are only interested in investigating single-user systems in SISO frequency-flat time-invariant and time-variant Rayleigh fading channels.

4.5 Optimal and mismatched detection in time invariant SISO channels

In this section, we specify the optimal detector for time-invariant fading channels with QAM transmission and show that for signals with constant envelope the mismatched detectors are optimal.

Consider the transmission in a SISO time invariant channel, described as

$$\mathbf{z}_p = h\mathbf{s}_p + \mathbf{n}_p \tag{4.38}$$

$$\mathbf{z}_d = h\mathbf{s}_d + \mathbf{n}_d \tag{4.39}$$

i.e., $M = 1$, $\mathbf{a} = h$ is a complex-valued scalar, $\mathbf{R}_a = \sigma_h^2 = E\{|h|^2\}$, and $\mathbf{\Psi}_p = \mathbf{s}_p$ is a $N_p \times 1$ vector of pilot symbols. We consider symbol-by-symbol detection, i.e., $N_d = 1$

and $\boldsymbol{\Psi}_d = d$. We only consider the white noise case, i.e., $\mathbf{R}_d = \sigma_n^2$ and $\mathbf{R}_p = \sigma_n^2 \mathbf{I}_{N_p}$.
Denote $\eta = \sigma_a^2/\sigma_n^2$, $\gamma_p = \boldsymbol{s}_p^H \mathbf{z}_p$, and $\gamma_d = d^* z_d$. Then, from (4.12) and (4.11), we obtain
$\mathbf{S_a} = \sigma_a^2 (E_p \eta + 1)^{-1}$ and $\mathbf{m_a} = \eta \gamma_p (E_p \eta + 1)^{-1}$, where $E_p = \boldsymbol{s}_p^H \boldsymbol{s}_p$ is the energy of the
pilot signal. We also obtain $\mathbf{L}_p = \gamma_p \sigma_n^{-2}$, $\boldsymbol{\Gamma}_p = E_p \sigma_n^{-2}$, $\mathbf{L}_d = \gamma_d \sigma_n^{-2}$, and $\boldsymbol{\Gamma}_d = |d|^2 \sigma_n^{-2}$.
With these notations, from (4.25) we arrive at the optimal detector

$$\hat{d}_{\text{opt}} \quad = \quad \arg \max_{d \in \mathcal{A}} \left\{ \frac{|d^* z_d + \gamma_p|^2}{\sigma_n^2 (|d|^2 + E_p + 1/\eta)} - \ln \left(|d|^2 + E_p + \frac{1}{\eta} \right) \right\}. \quad (4.40)$$

For signals with constant envelope $|d| = $ const, such as PSK signals, from (4.40) we
obtain

$$\hat{d}_{\text{opt}} = \arg \max_{d \in \mathcal{A}} \left\{ \Re(\gamma_p^* \gamma_d) \right\} = \hat{d}_{\text{ML}} = \hat{d}_{\text{MMSE}} = \hat{d}_\epsilon. \quad (4.41)$$

The data symbol estimate (4.41) is equivalent to the mismatched detectors with the fol-
lowing estimates of the channel gain:

$$\hat{a}_{\text{ML}} \quad = \quad \gamma_p / E_p, \quad (4.42)$$

$$\hat{a}_{\text{MMSE}} \quad = \quad \gamma_p / (E_p + 1/\eta), \quad (4.43)$$

$$\hat{a}_\epsilon \quad = \quad \gamma_p / (E_p + \sigma_n^2), \quad (4.44)$$

correcting the received signal according to these estimates ($z_0 = \hat{a}_{\text{ML}}^* z_d$ or $z_0 = \hat{a}_{\text{MMSE}}^* z_d$
or $z_0 = \hat{a}_\epsilon^* z_d$), and, finally, deciding on the transmitted symbol d by mapping z_0 to the PSK
constellation $\mathcal{A}$. Thus, for constant envelope signals (such as PSK signals) all the three
mismatched detectors are optimal. In other cases of QAM signals, the optimal detector is
given by (4.40).

4.6 Optimal and mismatched detection in SISO time variant channels

4.6.1 Transmission model

Now we consider single-user transmission in time variant channels. We assume that a
data block of N symbols is transmitted, N_p of which are pilot symbols and the other
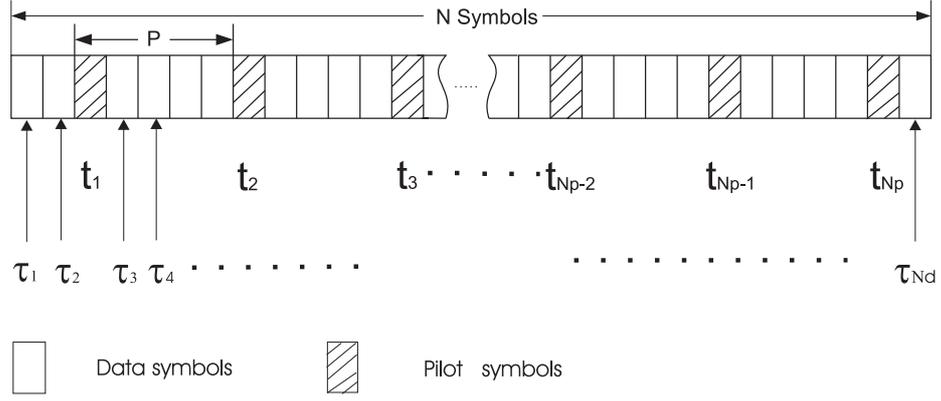
Figure 4.1: Structure of the transmitted data block.

$N_d = N - N_p$ are data symbols as shown in Fig. 4.1. The received signal corresponding to the pilot and data parts of the data block are modeled, respectively, as

$$z_p(t_k) = s_p(t_k)h(t_k) + n(t_k), \quad k = 1, \ldots, N_p, \tag{4.45}$$

$$z_d(\tau_k) = s_d(\tau_k)h(\tau_k) + n(\tau_k), \quad k = 1, \ldots, N_d, \tag{4.46}$$

where $s_d(\tau_k) = d_k$ is a data symbol transmitted at time τ_k, $n(t)$ is the noise, and $h(t)$ states a Rayleigh fading channel following Jakes' model [50, 51]. The covariance matrix of such fading channels is an $N \times N$ matrix with elements

$$[\mathbf{\Upsilon}]_{t_1,t_2} = \rho(t_1 - t_2), \tag{4.47}$$

where $t_1, t_2 = 1, \ldots, N$, and $\rho(\tau)$ is the autocorrelation function of Jakes' fading process [50] as shown by (3.4) in Chapter 3

$$\rho(\tau) = \sigma_h^2 J_0(2\pi\nu\tau), \tag{4.48}$$

σ_h^2 is the variance of the channel coefficients, $J_0(\cdot)$ is the zero-order Bessel function of the first kind, ν is the Doppler frequency.

As shown in Chapter 3, BEMs can be used to approximate the time-variant fading channels following Jakes' model. With a BEM, the task of estimating N time variant channel gains transforms to estimating only M time invariant expansion coefficients, where $M << N$, and the time-varying channel is represented as a series

$$h(t) = \sum_{m=1}^{M} a_m \varphi_m(t), \quad t = 0, \ldots, N-1, \tag{4.49}$$

where $\{\varphi_m(t)\}_{m=1}^M$ are basis functions. In the matrix form, (4.49) can be represented by

$$\mathbf{h} = \mathbf{B}\mathbf{a}, \tag{4.50}$$

where $\mathbf{B}$ is an $N \times M$ matrix with elements $[\mathbf{B}]_{t,m} = \varphi_m(t)$, $t = 0, \ldots, N - 1$, and $\mathbf{a} = [a_1, \ldots, a_m, \ldots, a_M]^T$ is an $M \times 1$ vector corresponding to the expansion coefficients.

We can represent the received data and pilot signals in the matrix form (4.1) and (4.6), respectively, with

$$\mathbf{\Psi}_p = \mathbf{D}_p\mathbf{B}_p, \quad \mathbf{\Psi}_d = \mathbf{D}_d\mathbf{B}_d, \tag{4.51}$$

and $\mathbf{D}_p$ and $\mathbf{D}_d$ being diagonal matrices defined as

$$\mathbf{D}_p = \mathrm{diag}\{s_p(t_1), \ldots, s_p(t_{N_p})\}, \tag{4.52}$$

$$\mathbf{D}_d = \mathrm{diag}\{d_1, \ldots, d_{N_d}\}. \tag{4.53}$$

The matrices $\mathbf{B}_p$ and $\mathbf{B}_d$ contain samples of the basis functions at the pilot and data symbol instants, respectively:

$$[\mathbf{B}_p]_{k,m} = \varphi_m(t_k), \quad [\mathbf{B}_d]_{k,m} = \varphi_m(\tau_k). \tag{4.54}$$

Using these notations, and denoting

$$\boldsymbol{\beta}_d = \mathbf{D}_d^H\mathbf{z}_d, \quad \boldsymbol{\beta}_p = \mathbf{D}_p^H\mathbf{z}_p \tag{4.55}$$

and

$$\mathbf{F}_d = \mathbf{D}_d^H\mathbf{D}_d, \quad \mathbf{F}_p = \mathbf{D}_p^H\mathbf{D}_p, \tag{4.56}$$

we obtain:

$$\mathbf{L}_d = \sigma_n^{-2}\mathbf{B}_d^H\boldsymbol{\beta}_d, \qquad \mathbf{L}_p = \sigma_n^{-2}\mathbf{B}_p^H\boldsymbol{\beta}_p, \tag{4.57}$$

and

$$\mathbf{\Gamma}_d = \sigma_n^{-2}\mathbf{B}_d^H\mathbf{F}_d\mathbf{B}_d, \quad \mathbf{\Gamma}_p = \sigma_n^{-2}\mathbf{B}_p^H\mathbf{F}_p\mathbf{B}_p. \tag{4.58}$$

4.6.2 Optimal detection

The optimal detector becomes very complicated for high N_d. We want to consider the simplest case of symbol-by-symbol detection of data symbols in a data block. In this

case, expressions above are simplified: $\mathbf{D}_d = d$, $\mathbf{F}_d = |d|^2$ and $\boldsymbol{\beta}_d = d^* z_d$ are now
scalars; $\mathbf{B}_d$ is a $(1 \times M)$ vector whose elements are values of the basis functions at the data
symbol instant; $\mathbf{L}_d = \sigma^{-2} d^* z_d \mathbf{B}_d^H$; and $\boldsymbol{\Gamma}_d = \sigma^{-2} |d|^2 \mathbf{B}_d^H \mathbf{B}_d$. The optimal detector (4.24)
minimizes the metric $\lambda(\mathbf{d})$ which is now given by

$$
\begin{aligned}
\lambda(d) \;=\; & -\frac{1}{\sigma_n^2} \left(\mathbf{B}_d^H \boldsymbol{\beta}_d + \mathbf{B}_p^H \boldsymbol{\beta}_p \right)^H \left(|d|^2 \mathbf{B}_d^H \mathbf{B}_d + \mathbf{B}_p^H \mathbf{F}_p \mathbf{B}_p + \sigma_n^2 \mathbf{R_a}^{-1} \right)^{-1} \\
& \times \left(\mathbf{B}_d^H \boldsymbol{\beta}_d + \mathbf{B}_p^H \boldsymbol{\beta}_p \right) + \ln \left| |d|^2 \mathbf{B}_d^H \mathbf{B}_d + \mathbf{B}_p^H \mathbf{F}_p \mathbf{B}_p + \sigma_n^2 \mathbf{R_a}^{-1} \right| . \quad (4.59)
\end{aligned}
$$

The optimal detector based on modeling time-variant fading using a BEM requires an
explicit expression for the fading covariance $\mathbf{R_a}$. To obtain $\mathbf{R_a}$, we can use the transform
(3.25)

$$
\mathbf{R_a} = (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \boldsymbol{\Upsilon} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1}. \quad (4.60)
$$

If the perfect channel information is available, for the PDF $f(\mathbf{a}|\mathbf{z}_p) = \mathcal{N}_\mathcal{C}(\mathbf{m_a}, \mathbf{S_a})$ we
can write $\mathbf{m_a} = \mathbf{a}$ and $\mathbf{S_a} = \mathbf{0}_M$. In this case, the optimal metric (4.59) takes the form
$\lambda(d) = -2\Re(\mathbf{L}_d^H \mathbf{a}) + \mathbf{a}^H \boldsymbol{\Gamma}_d \mathbf{a}$. The detector minimizing this metric is equivalent to the
classical minimum distance detector

$$
\hat{d}_{\mathrm{PCI}} = \arg \min_{d \in \mathcal{A}} \left\{ (\mathbf{z}_d - \boldsymbol{\Psi}_d \mathbf{a})^H (\mathbf{z}_d - \boldsymbol{\Psi}_d \mathbf{a}) \right\} . \quad (4.61)
$$

4.6.3 Mismatched detection

As introduced above, we will consider the mismatched detectors using MMSE, ML and
ϵ-ML channel estimates and compare their performance with that of the optimal detector.
Since we have used a BEM to approximate time-variant channels here, and therefore, the
task of the channel estimator becomes to estimate M expansion coefficients. The MMSE
channel estimates are given by

$$
\hat{\mathbf{a}}_{\mathrm{MMSE}} = (\mathbf{B}_p^H \mathbf{F}_p \mathbf{B}_p + \sigma_n^2 \mathbf{R_a}^{-1})^{-1} \mathbf{B}_p^H \mathbf{D}_p^H \mathbf{z}_p. \quad (4.62)
$$

Correspondingly, a mismatched detector that treats $\hat{\mathbf{a}}_{\mathrm{MMSE}}$ as perfect and minimizes the
Euclidean distance is represented as

$$
\hat{d}_{\mathrm{MMSE}} = \arg \min_{d \in \mathcal{A}} \left\{ (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\mathrm{MMSE}})^H (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\mathrm{MMSE}}) \right\} . \quad (4.63)
$$

The ML channel estimates are given by

$$\hat{\mathbf{a}}_{\mathrm{ML}} = (\mathbf{B}_p^H \mathbf{F}_p \mathbf{B}_p)^{-1} \mathbf{B}_p^H \mathbf{D}_p^H \mathbf{z}_p. \tag{4.64}$$

and a mismatched detector applying the ML channel estimates is represented as

$$\hat{d}_{\mathrm{ML}} = \arg \min_{d \in \mathcal{A}} \left\{ (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\mathrm{ML}})^H (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\mathrm{ML}}) \right\} \tag{4.65}$$

The ϵ-ML channel estimates are given by

$$\hat{\mathbf{a}}_\epsilon = (\mathbf{B}_p^H \mathbf{F}_p \mathbf{B}_p + \epsilon \sigma_n^2 \mathbf{I}_M)^{-1} \mathbf{B}_p^H \mathbf{D}_p^H \mathbf{z}_p, \tag{4.66}$$

where $\epsilon \geq 0$ is a regularization parameter and $\mathbf{I}_M$ denotes an $M \times M$ identity matrix. Such regularization does not require the fading statistics to be available. A mismatched detector using the ϵ-ML estimates is represented as

$$\hat{d}_\epsilon = \arg \min_{d \in \mathcal{A}} \left\{ (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_\epsilon)^H (\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_\epsilon) \right\}. \tag{4.67}$$

Note that for $\epsilon = 0$, we have $\hat{\mathbf{a}}_{\mathrm{ML}} = \hat{\mathbf{a}}_\epsilon$. It can be shown that the ϵ-ML channel estimate provides the minimum MSE if $\epsilon = \sigma_a^{-2}$; this value of ϵ is used in our simulations.

4.6.4 Iterative receivers
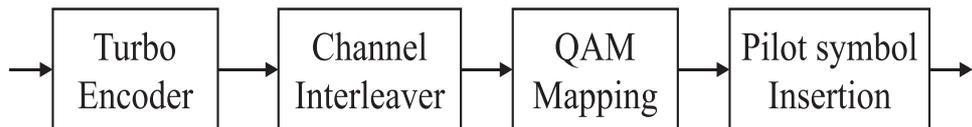
Figure 4.2: Transmitter.

The transmission system with QAM modulation is shown in Fig .4.2. In the transmitter, information bits are firstly encoded by a turbo encoder of 1/3 rate with generator polynomials [013,015] in octal notation. The output bits of the turbo encoder are channel-interleaved and grouped into QAM symbols. Then, pilot symbols are inserted periodically every $(P - 1)$ data symbols as shown in Fig. 4.1.

Fig. 4.3 or Fig. 4.4 show the structure of soft-input hard-output (SIHO) or soft-input soft-output (SISO) turbo decoders, respectively.

The receiver applying SIHO turbo decoder (Fig. 4.3) recovers the transmitted coded bits by applying a turbo encoder same as the one used in the transmitter to encoding the hard output of the decoder. The soft metrics output from the detector are de-interleaved and then passed to a turbo decoder. The SIHO turbo decoder in Fig. 4.3 outputs the decoded bits which are then turbo encoded. This receiver is too complicated for practical scenarios, and moreover, its performance is much worse than that of the iterative receiver with SISO Turbo decoder as shown in following simulations.

The receiver applying SISO turbo decoder (Fig. 4.4) performs several iterations, in which channel estimation and decoding are refined. The receiver uses an SISO turbo decoder. For every bit $c_k = \pm 1$ of a received symbol, $k = 1, \ldots, K$, the *a posteriori* log-likelihood ratio (LLR) is computed as [116, 117]

$$\lambda_{c_k} = \ln \left[\frac{\sum_{d \in \mathcal{A}_k^+} e^{-\lambda(d)} \prod_{i \neq k} P(c_i)}{\sum_{d \in \mathcal{A}_k^-} e^{-\lambda(d)} \prod_{i \neq k} P(c_i)} \right], \tag{4.68}$$

where the *a priori* probability $P(c_i)$ of a symbol bit is expressed in terms of its *a priori* LLR $L(c_i)$ [118]:

$$P(c_i) = \frac{1}{2} \left[1 + c_i \tanh \left(\frac{1}{2} L(c_i) \right) \right],$$

$\mathcal{A}_k^\pm = \{ d \in \mathcal{A} | c_k = \pm 1 \}$, and the metric $\lambda(d)$ depends on the detector used. For the first iteration, we have

$$\lambda_{c_k} = \ln \sum_{d \in \mathcal{A}_k^+} e^{-\lambda(d)} - \ln \sum_{d \in \mathcal{A}_k^-} e^{-\lambda(d)}. \tag{4.69}$$

The LLRs λ_{c_k} are de-interleaved and passed to a turbo-decoder that outputs both a sequence of the symbol bit LLRs and decoded bits; the LLRs are then transformed to recover coded bits by hard decision. After interleaving, QAM mapping, and adding the pilot symbols, the whole recovered sequence of the QAM symbols is used for channel estimation in the next iteration.

Actually, this iterative receiver with SISO turbo decoder is not optimal since the hard decision after the output of turbo decoder leads to a loss of soft information of coded bits and causes extra error in the QAM Mapper. A better performance can be obtained by using the soft mapping scheme to generate the data symbols based on the soft output of the turbo decoder directly [119]. However, the complexity of this soft mapping is higher. In this thesis, we focus on the hard mapping and the performance of the iterative receiver with soft mapping scheme will be investigated in further works.
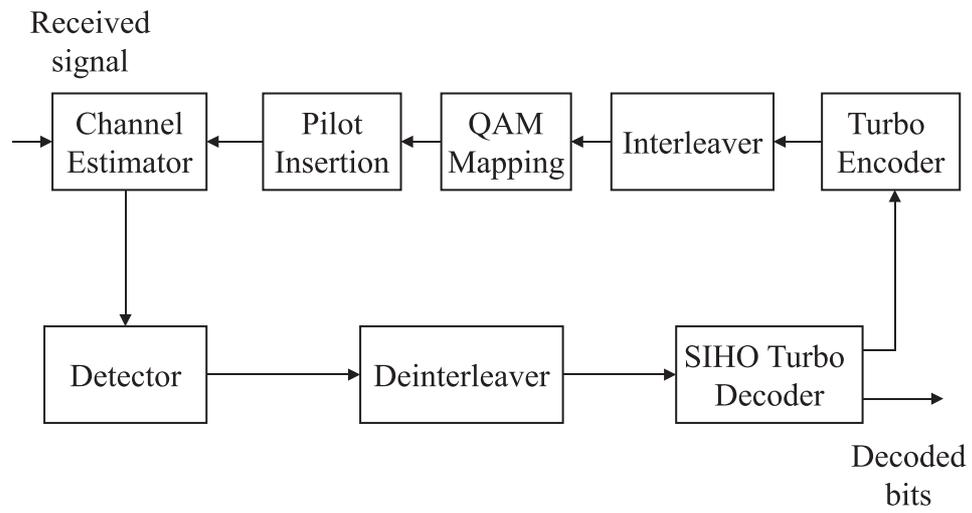
Figure 4.3: Receiver with soft-input hard-output (SIHO) turbo-decoder.
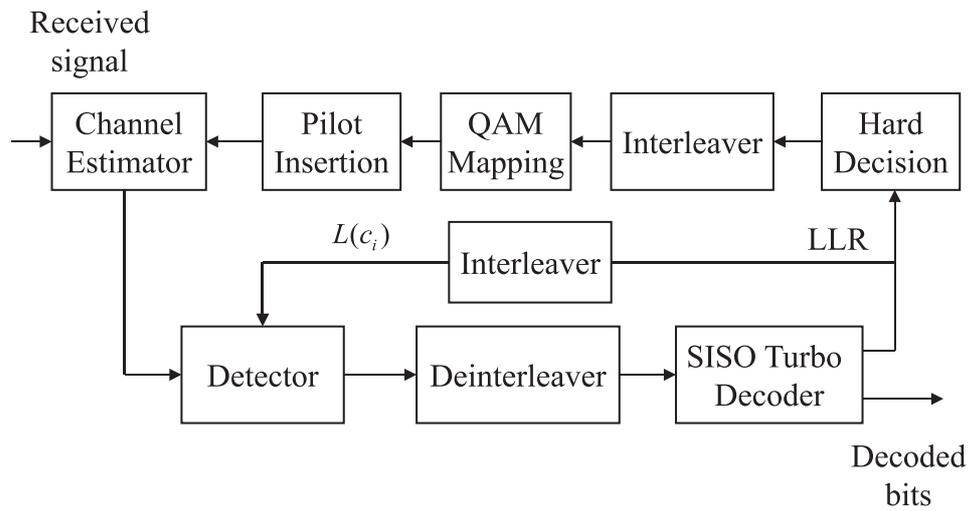
Figure 4.4: Receiver with soft-input soft-output (SISO) turbo-decoder.

Functions of the channel estimator and detector are also varying depending on the detector used and whether it is the first or a subsequent iteration:

1) *ML-ML* receiver: The ML channel estimation is used in all iterations. At the first iteration, the channel estimator provides the ML channel estimate $\hat{\mathbf{a}}_{\mathrm{ML}}$ according to (4.64). In the following iterations, it provides ML estimates with re-defined matrices $\mathbf{D}_p$ and $\mathbf{F}_p$ to include all (pilot and data) symbols; the matrix $\mathbf{B}_p$ is replaced by the matrix $\mathbf{B}$. The metric $\lambda(d)$ is calculated as

$$\lambda(d) = \sigma_n^{-2}|z_d - \hat{a}d|^2, \tag{4.70}$$

where $\hat{a}$ is a channel estimate given by $\hat{a} = \mathbf{B}_d\hat{\mathbf{a}}_{\mathrm{ML}}$.

2) *ϵ-ML-ϵ-ML* receiver: The receiver is similar to the *ML-ML* receiver with the channel estimates given by $\hat{a} = \mathbf{B}_d\hat{\mathbf{a}}_\epsilon$.

3) *MMSE-MMSE* receiver: The receiver is similar to the *ML-ML* receiver with the channel estimates given by $\hat{a} = \mathbf{B}_d\hat{\mathbf{a}}_{\mathrm{MMSE}}$.

4) *Opt-ϵ-ML* receiver: At the first iteration, the channel estimator provides the vector $\mathbf{L}_p = \sigma_n^{-2}\mathbf{B}_p^H\mathbf{D}_p^H\mathbf{z}_p$ required for the optimal detector. In subsequent iterations, it provides ϵ-ML estimates with re-defined matrices $\mathbf{D}_p$ and $\mathbf{F}_p$ to include all symbols; the matrix $\mathbf{B}_p$ is replaced by the matrix $\mathbf{B}$. At the first iteration, the LLR (4.69) with $\lambda(d)$ given by (4.59) is calculated. At other iterations, the LLR (4.68) is calculated with $\lambda(d)$ from (4.70) and $\hat{a} = \mathbf{B}_d\hat{\mathbf{a}}_\epsilon$.

5) *Opt-MMSE* receiver: The receiver is similar to the *Opt-ϵ-ML* receiver with channel estimates given by $\hat{a} = \mathbf{B}_d\hat{\mathbf{a}}_{\mathrm{MMSE}}$.

4.7 Simulation results

We first consider the time-invariant fading channel. Fig. 4.5 shows the BER performance of the optimal detector in Rayleigh fading channel with 16QAM modulation. In this scenario, only one pilot symbol is transmitted ($N_p = 1$). In the simulation trials, the pilot symbol is chosen randomly from the alphabet $\mathcal{A}$. It is seen a small gain (less than 0.1
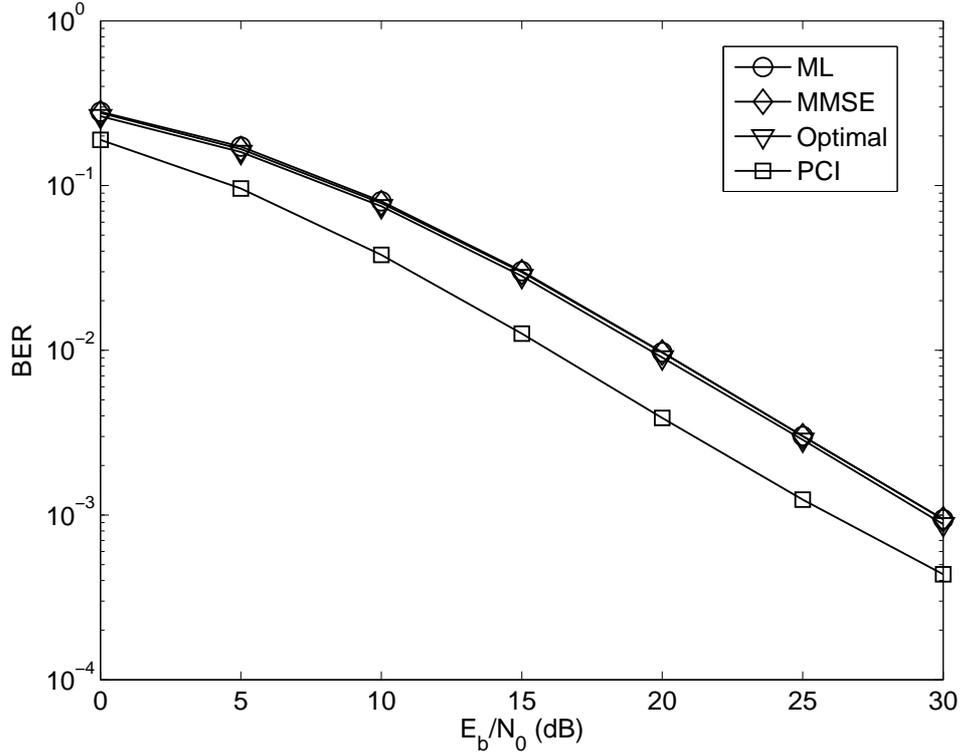
Figure 4.5: BER performance of the optimal detector in time-invariant frequency-flat
Rayleigh fading channel with 16QAM modulation; $N_p = 1$.

dB) due to the use of the optimal detector with respect to the ML and MMSE mismatched
detectors. Simulation for other modulation schemes has shown even a smaller gain. Thus,
in time-invariant fading channels, the optimal detector provides little improvement in the
detection performance compared to that of the mismatched detectors.

Then, we consider a time-variant fading channel with the Doppler spread factor $\nu T_s = 0.01$, where T_s is the duration of a symbol. The time variant channel coefficients are
modeled by cubic B-splines with basis functions calculated by [76, 120]

$$\varphi(t) = B_3(t) = \begin{cases} \frac{2}{3} - \frac{t^2}{T^2} + \frac{|t|^3}{2T^3}, & \text{if } |t| < T, \\ \frac{1}{6}\left(2 - \frac{|t|}{T}\right)^3, & \text{if } T \leq |t| < 2T, \\ 0, & \text{otherwise} \end{cases} \tag{4.71}$$

where T is a sampling interval. For approximation of $h(t)$ on an interval $t \in [0, N-1]$,
we set $T = (N-1)/(M-3)$; then the basis functions $\varphi_m(t)$ are given by

$$\varphi_m(t) = \varphi(t - mT + 2T), \quad m = 1, \ldots, M. \tag{4.72}$$

Fig. 4.6 shows the mean squared error (MSE) of approximation of the fading process
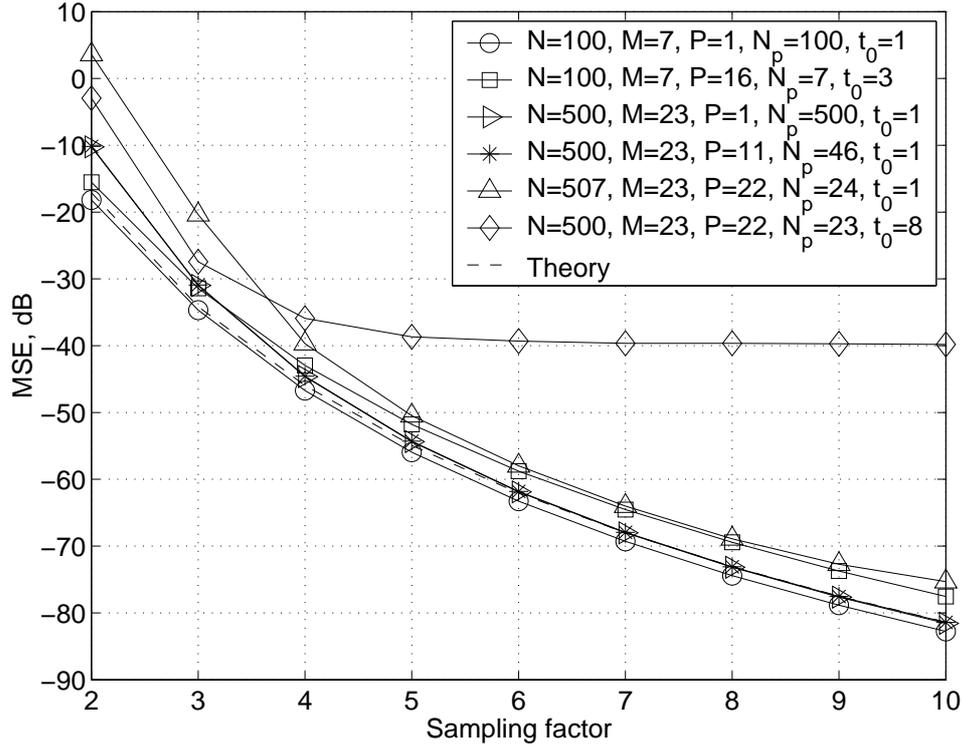
Figure 4.6: MSE performance of approximation of the fading Jake's model by cubic B-splines; no noise; M the number of basis functions; $P - 1$ is the number of data symbols between 2 neighboring pilot symbols; N_p is the number of pilot symbols in the block and t_0 is the position of the first pilot symbol.

$h(t)$ with Jakes' autocorrelation by cubic BS BEM with spline coefficients (4.64) as a function of the sampling factor $\gamma = 1/(\nu T)$. These results are obtained by simulation for the case of no additive noise. The MSE depends on the length N of the data block, the number N_p and positions $t_n = t_0 + (n - 1)P$ of pilot symbols within the block. If all the symbols are pilots ($N_p = N$), then the MSE is very close to the 'theoretical' MSE of approximation of Jakes' model by using the optimal splines of an arbitrary order q [27]

$$\varepsilon_q^2 \approx \frac{\pi^{2q+2}B_{2q+2}}{[(q+1)!]^2\gamma^{2q+2}} + \frac{\pi^{2q+4}(q+1)(2q+3)B_{2q+4}}{[(q+2)!]^2\gamma^{2q+4}}, \tag{4.73}$$

where B_m are Bernoulli numbers [79]. For the cubic B-splines ($q = 3$), from (4.73) we have

$$\varepsilon_3^2 \approx \frac{0.549}{\gamma^8} + \frac{17.736}{\gamma^{10}}. \tag{4.74}$$

The MSE is still close to the theoretical calculation, if there are at least 1.5 pilot symbols per sampling interval T. If the number of pilot symbols is close to one symbol per sampling interval, the MSE performance becomes sensitive to positions of the pilot symbols.

When comparing two scenarios, $[N = 500, N_p = 23]$ and $[N = 507, N_p = 24]$, it is seen that there is a significant difference between the MSE performance of these two cases. However, even for the worst-case scenario, the MSE is better than -20 dB for $\gamma \geq 3$ and better than -36 dB for $\gamma \geq 4$. To avoid degradation in the detection performance with respect to the case of perfect channel information, "the estimation error should be negligible compared to the reciprocal of the signal-to-noise ratio" [121]. Thus, in the scenarios considered, the MSE is low enough for many modulation techniques (operating at SNR lower than 30 dB), i.e., for these cases, in our derivations, we can neglect the modeling error of cubic B-spline approximation of Jakes' model of time-variant fading channels. Therefore, we set $\gamma = 4$ in all simulations following.

Fig. 4.7 shows simulation results for a scenario with 16QAM modulation in a system without coding. It is seen that for BER= 10^{-2}, the optimal detector outperforms the mismatched detector with ML channel estimates by about 5.6 dB and is inferior to the receiver with perfect channel knowledge by about 2.1 dB. The mismatched detector with ϵ-ML channel estimates is inferior to the optimal detector by 1.5 dB. However, the mismatched detector with MMSE channel estimates provides nearly the same performance as the optimal one.

Now, we will consider the iterative receiver with turbo decoder.

Firstly, we will consider the upper and lower bound of the MSE performance of channel estimators in the iterative receiver. Fig. 4.8 shows the dependence of the MSE on E_b/N_0 for 16QAM modulation and $\nu T_s = 0.01$. At the first iteration, the channel estimators only deal with pilot symbols. When the number of pilot symbols is small with respect to the number of data symbols ($P = 22$), the ML estimator provides the worst MSE, and the MMSE estimator provides the best performance, while the ϵ-ML estimator with the optimal $\epsilon = \sigma_a^{-2}$ has an intermediate MSE performance. These curves for $P = 22$ show the upper MSE bounds for channel estimation performance in the iterative receivers, or, in other words, they are equivalent to the performance of the receiver without iterations. In following iterations, after correcting by the FEC decoder, recovered data symbols are also treated as pilot and involved in channel estimation. The MSE performance will be improved and approach the lower bound given by the case $P = 1$, i.e., when all symbols are pilot symbols. It is seen that, in this case, the ML and ϵ-ML estimators have similar
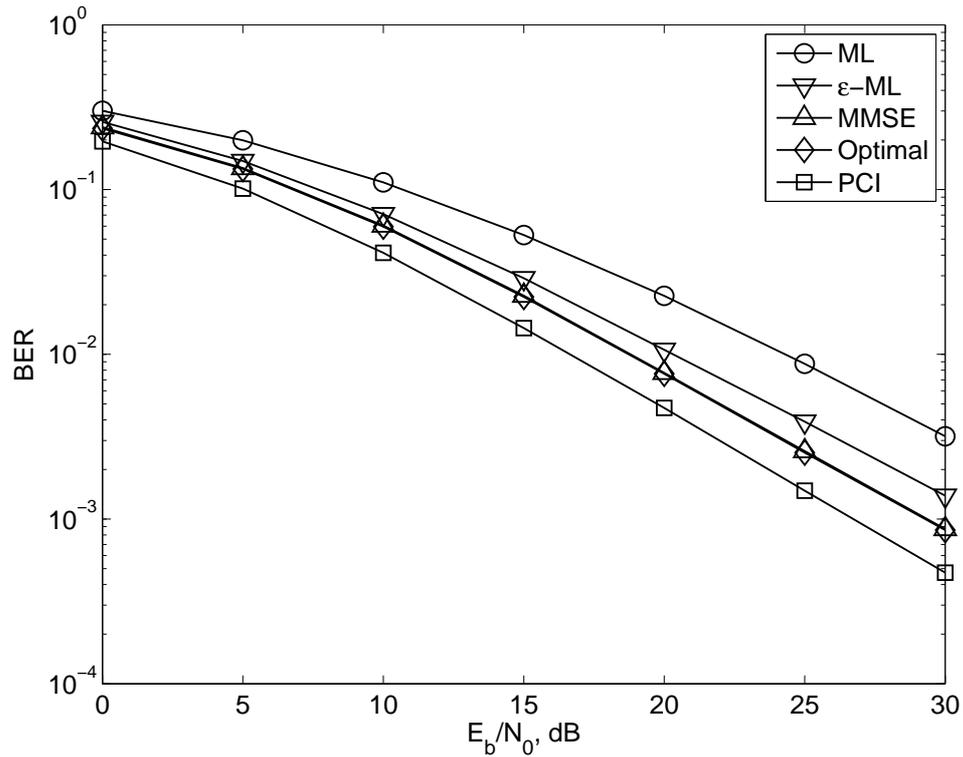
Figure 4.7: BER performance of the optimal and mismatched detectors in time-variant frequency-flat Rayleigh fading channel with 16QAM modulation; $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $P = 22$, $t_1 = 1$.

MSE performance, whereas, at low SNRs, the MMSE estimator outperforms the others. The curves for $P = 1$ show the lower bounds for channel estimation performance in the iterative performance. We can expect that the MSE performance of the channel estimation at the initial iteration is the same as the upper bounds and improves iteration by iteration.

Fig. 4.9 shows the MSE performance of the *MMSE-MMSE* iterative receiver versus the number of iterations increases. The SISO turbo decoder is implemented here. It can be seen that the MSE performance of channel estimation in the *MMSE-MMSE* iterative receiver does improve and the improvement between two consecutive iterations reduces when the number of iterations increases. However, the improvement of MSE performance can't be distinguished after 4 iterations. Therefore, in this thesis, the iterative receiver will perform 4 iterations at which it can provide acceptable performance and require affordable complexity.

Two iterative receivers with SIHO and SISO turbo decoders are considered here. Fig.
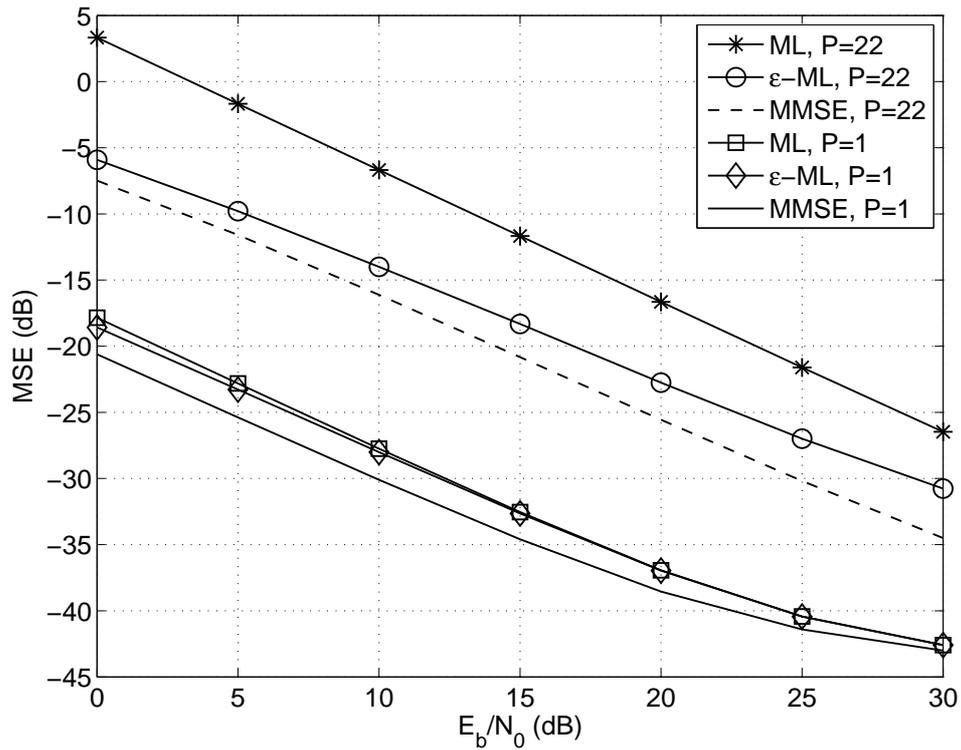
Figure 4.8: MSE performance of the ML, ϵ-ML, and MMSE estimators of Jake's fading model; $N = 507$, $M = 23$, $\nu T_s = 0.01$, $t_1 = 1$.

4.10 shows the BER performance of the iterative receivers with a SIHO turbo decoder after 4th iteration in a scenario with 16QAM modulation and $1/3$ rate turbo code. The *Optimal-MMSE* and *Optimal-ϵ-ML* iterative receivers, i.e., receivers using the optimal detector at the first iteration, significantly outperform the iterative receivers with ML channel estimation. At BER= 10^{-2}, the improvement in the detection performance is about 5.4 dB against the *ML-ML* iterative receiver and about 1.9 dB against the ϵ-ML-ϵ-ML iterative receiver. However, the *MMSE-MMSE* iterative receiver is only 0.3 dB inferior to the iterative receivers with optimal detection. Fig. 4.11 shows the MSE performance of channel estimation for this scenario. When comparing Fig. 4.11 and Fig. 4.8, it is seen that, at high SNRs, MSEs of the estimators approach corresponding lower MSE bounds shown in Fig. 4.8. However, it can be seen from Fig. 4.10 that the gap between the detection performance of receivers with the optimal detector and a receiver with perfect channel information is still significant (about 2.6 dB). This gap can be reduced if a SISO turbo decoder is used.

Fig. 4.12 shows the BER performance of the iterative receivers with a SISO turbo
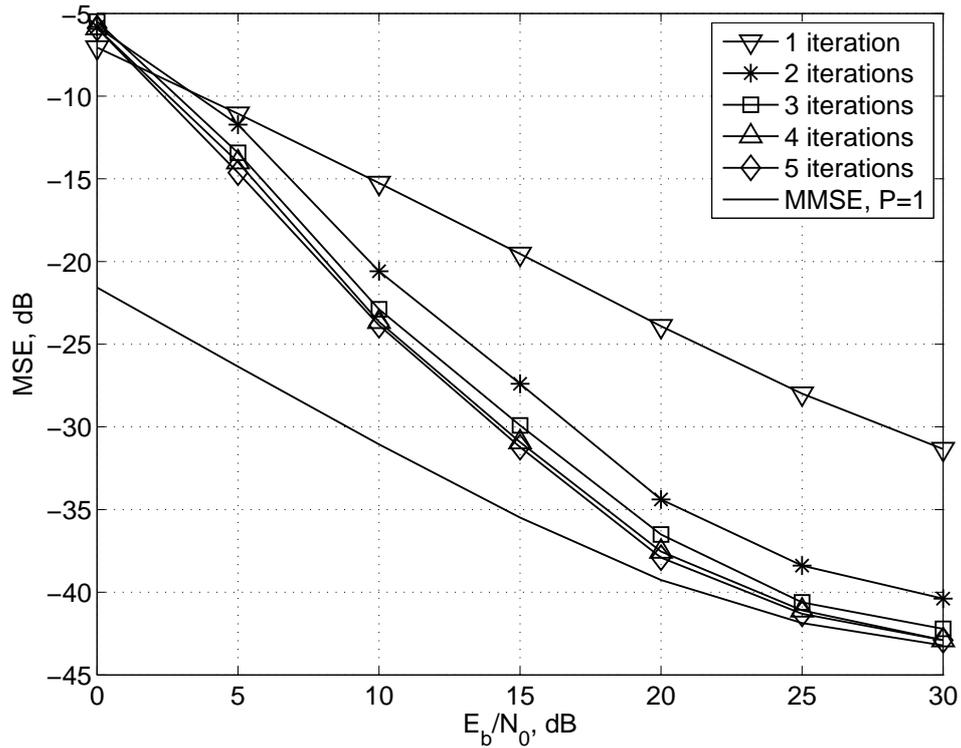
Figure 4.9: MSE performance of the *MMSE-MMSE* iterative receiver with a soft-input soft-output turbo decoder versus E_b/N_0 with respect to the number of iterations; code rate-$1/3$, $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $t_1 = 1$.

decoder in the same scenario. The gap between the detection performance of the *Optimal-MMSE* receiver and a receiver with PCI is only 0.3 dB at BER=10^{-3}. The gap between the *Optimal-MMSE* and ϵ-*ML*-ϵ-*ML* receivers is only 1.5 dB. The performance of the *MMSE-MMSE* receiver is only 0.16 dB to that of the *Optimal-MMSE* receiver.

Fig. 4.13 shows the MSE performance of channel estimation for this scenario. When comparing Fig. 4.13 with Fig. 4.11, it is seen a significant improvement in the MSE performance at low SNRs when using SISO decoder.

From the simulation results, we can conclude that the MMSE channel estimation allows the detection performance of the mismatched detector to approach that of the optimal detector in both uncoded and coded systems; the difference in the performance is at most 0.16 dB. However, both the optimal detector and the mismatched detector with MMSE channel estimates require the channel covariance matrix to be known. The mismatched detector with regularized ML channel estimates does not need the knowledge of the fading
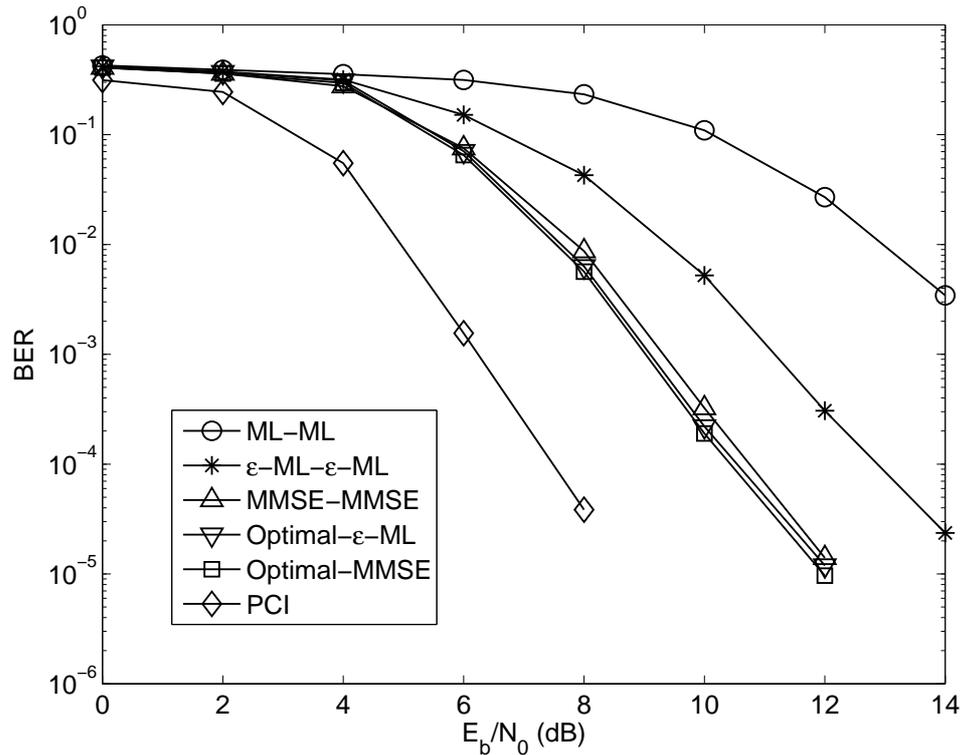
Figure 4.10: BER performance of the iterative receivers with a soft-input hard-output turbo decoder after 4th iteration in a time-variant frequency-flat Rayleigh fading channel with 16QAM modulation; code rate-$1/3$, $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $t_1 = 1$.

statistical characteristics. The payment for this *a priori* uncertainty is a worse detection performance. However, the performance degradation is not significant.

4.8 Conclusions

We have derived an optimal detector for pilot-assisted transmission in Rayleigh fading channels with unknown parameters. The results obtained are applicable to a wide range of communications scenarios, including single-input single-output and multi-input multi-output systems, single-user and multiuser systems in frequency-flat and frequency-selective time-invariant and time-variant fading channels. They can be used for correlated fading channels and non-white additive noise.
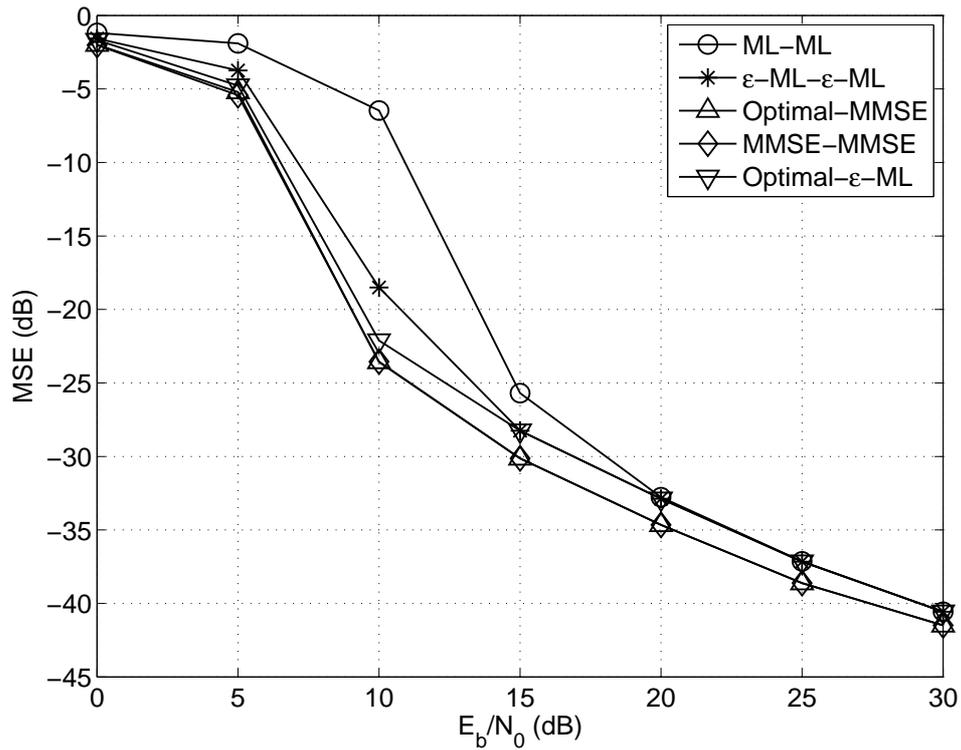
Figure 4.11: MSE performance of the iterative receivers with a soft-input hard-output turbo decoder after 4th iteration in a time-variant frequency-flat Rayleigh fading channel with 16QAM modulation; code rate-$1/3$, $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $t_1 = 1$.

In this chapter, we were only interested in investigation of single-input single-ouput systems in frequency-flat fading channels. For slow fading channels, it has been shown that, in the case of constant-envelope (e.g., PSK) modulation, the mismatched detectors are equivalent to the optimal detector, while, in a general case of QAM modulation, the optimal detector outperforms the mismatched detectors. In time-variant fading channels, we have considered B-spline approximation of the channel gain time variations. Simulation results for uncoded data transmission have shown that, in such channels, the optimal detector can significantly improve the detection performance compared to that of the mismatched detectors exploiting ML channel estimates. However, the MMSE-mismatched detector provides nearly optimal detection performance. We have also investigated the detection performance of iterative receivers that exchange information between a channel estimator and decoder. It is shown by simulation that the iterative receiver with the optimal detector at the first iteration outperforms the receiver using ML or regularized ML channel estimates. However, the use of MMSE channel estimates makes the detection
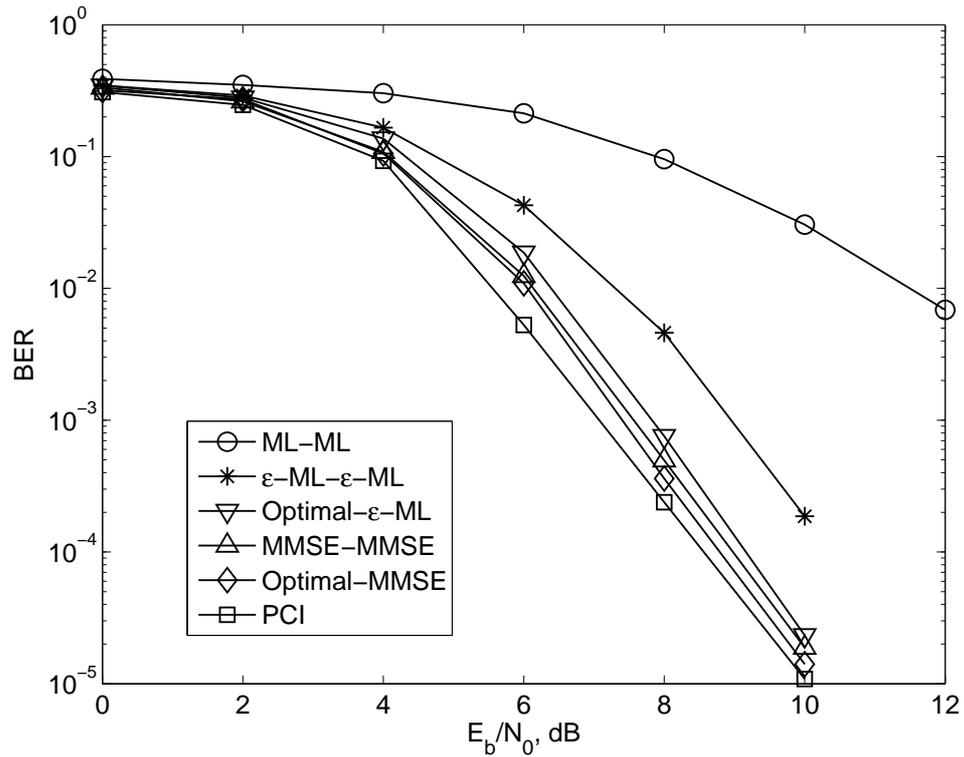
Figure 4.12: BER performance of the iterative receivers with a soft-input soft-output
after 4th iteration in a time-variant frequency-flat Rayleigh fading channel with 16QAM
modulation; code rate-$1/3$, $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $P = 22$, $t_1 = 1$.

performance close to that of the receiver with the optimal detector at the first iteration.

   In the next chapter, we will apply this optimal detector to MIMO Rayleigh flat fad-
ing channels, and compare its performance with those of the mismatched detectors. We
expect that the improvement of the performance by using this optimal detector will be-
come more obvious. We will focus on the comparison between the optimal detector and
the mismatched detector with MMSE channel estimation and investigate the conditions
under which these two detectors are equivalent.
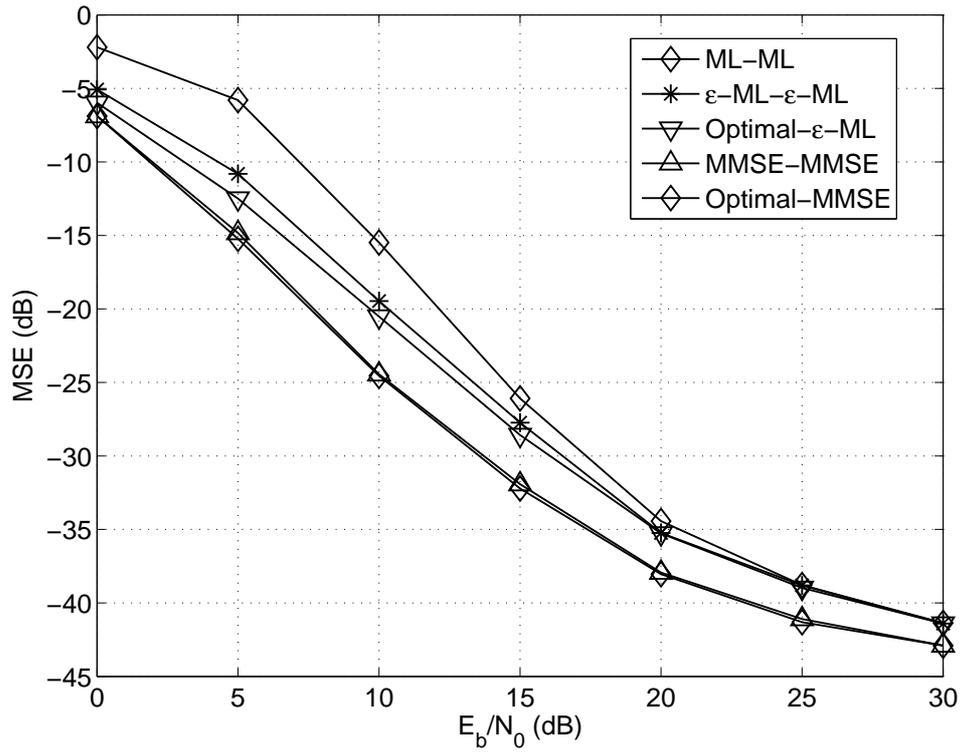
Figure 4.13: MSE performance of the iterative receivers with a soft-input soft-output turbo decoder after 4th iteration in a time-variant frequency-flat Rayleigh fading channel with 16QAM modulation; code rate-$1/3$, $\nu T_s = 0.01$, $N = 507$, $M = 23$, $N_p = 24$, $t_1 = 1$.

Chapter 5

Optimal and Mismatched Detection in MIMO Frequency-Flat Fading Channels with Imperfect Channel Estimation

Contents

5.1 Introduction

In the previous chapter, we have investigated a pilot assisted optimal detector which out-
performs the mismatched detectors in SISO Rayleigh fading channel. The optimal detec-
tor does significantly outperform mismatched detectors with ML or ϵ-ML channel esti-
mates, but slightly outperforms the one with MMSE channel estimates even for the case
with turbo code and iterative receiver. Similar conclusions are also presented in [36].
Simulation results in [36] show that the BER performance of the optimal detection is
close to that of the mismatched detector with MMSE channel estimates.

We can expect that the use of the optimal detection will bring a more significant ben-
efit in channels with a large number of unknown parameters. In this chapter, we con-
sider a more general channel, the MIMO channel. After deriving the optimal detector for
spatially correlated MIMO Rayleigh time-invariant fading channels, we then extend the
optimal detector to MIMO Rayleigh time-variant fading channels, in particular with the
fading correlation following Jakes' model [50], and specify it for MIMO time-variant fad-
ing channels. The time-variant fading channel is modeled by using BEMs; specifically,
cubic B-spline functions are used [76].

In [35, 36], the analysis and simulation results show that in SISO channels with PSK
modulation, the symbol-by-symbol optimal detection is equivalent to the mismatched de-
tection with MMSE channel estimates. However, this equivalence is only true for SISO
channels. In this chapter, we consider more general cases and prove that in spatially
uncorrelated SIMO channels the optimal symbol-by-symbol detection of PSK signals is
equivalent to the mismatched detection with MMSE channel estimation. However, this is
not the case for signals with non-constant envelope and/or multi-antenna transmission.

The rest of this chapter is structured as follows. In Section 5.2, the optimal detector is
specified for spatially correlated MIMO time-invariant Rayleigh fading channels. Section
5.3 describes the proposed optimal and mismatched detectors for MIMO time-variant
Rayleigh fading channels. The conditions of equivalence between the optimal detector
and the mismatched detector with MMSE channel estimates are discussed in Section 5.4.
Simulation results are given in Section 5.5, followed by conclusions in Section 5.6.

5.2 Optimal and mismatched detection in MIMO time invariant channels

In this section, we consider a system with N_t transmit antennas and N_r receive antennas over MIMO time-invariant Rayleigh fading channels. We set $\mathbf{\Psi}_d = \mathbf{I}_{N_r} \otimes \mathbf{s}_d$, where $\mathbf{s}_d = [s_1, \ldots, s_k, \ldots, s_{N_t}]$ is an $1 \times N_t$ vector, and s_k is the data symbol transmitted from the kth transmit antenna, and $\mathbf{I}_{N_r}$ is a $N_r \times N_r$ identity matrix. Correspondingly, we also define $\mathbf{\Psi}_p = \mathbf{I}_{N_r} \otimes \mathbf{S}_p$, where $\mathbf{S}_p$ is an $N_p \times N_t$ matrix with element $[\mathbf{S}_p]_{k,i} = p_k(i)$, and N_p is the number of pilot symbols transmitted from each antenna and $p_k(i)$ is the pilot symbol transmitted from kth transmit antenna at the ith instance, and $i = 1, \ldots, N_p$. $\otimes$ denotes the Kronecker product. The received pilot and data signals are given by

$$\mathbf{z}_p = \mathbf{\Psi}_p \mathbf{h} + \mathbf{n}_p, \tag{5.1}$$

$$\mathbf{z}_d = \mathbf{\Psi}_d \mathbf{h} + \mathbf{n}_d, \tag{5.2}$$

where $\mathbf{h}$, an $N_t N_r \times 1$ vector of channel coefficients, is given by

$$\mathbf{h} = [\mathbf{h}_1, \ldots, \mathbf{h}_r, \ldots, \mathbf{h}_{N_r}]^T,$$

$$\mathbf{h}_r = [h_{r,1}, \ldots, h_{r,k}, \ldots, h_{r,N_t}],$$

$h_{r,k}$ is the channel coefficient between the kth transmit antenna and rth receive antenna, and $\mathbf{n}_p$ and $\mathbf{n}_d$ are the noise observed at the data symbol and pilot symbol positions, respectively. We consider the scenario where the noise samples at different receive antennas are uncorrelated and assume that the noise temporal covariance matrix, $\mathbf{R}_{n0}$, which characterizes time-correlation of noise samples for a single receive antenna, is the same for all receive antennas. We also assume that the variance of the path between any pair of transmit and receive antennas is normalized to 1 ($\sigma^2_{h_{r,k}} = 1$). We then define the spatial correlation matrix of transmit antennas as an $N_t \times N_t$ symmetric matrix $\mathbf{R}_t$ with elements $[\mathbf{R}_t]_{i,i} = 1$ and $[\mathbf{R}_t]_{i,j} = \rho, \quad i \neq j$, where $i, j = 1, \ldots, N_t$, while the spatial correlation matrix of receive antennas is an $N_r \times N_r$ symmetric matrix $\mathbf{R}_r$ with elements $[\mathbf{R}_r]_{i,i} = 1$ and $[\mathbf{R}_r]_{i,j} = \rho, \quad i \neq j$, where $i, j = 1, \ldots, N_r$. The joint spatial covariance matrix of the MIMO channel is given by [122, 123]

$$\mathbf{\Upsilon} = E\{\mathbf{h}\mathbf{h}^H\} = \mathbf{R}_r \otimes \mathbf{R}_t. \tag{5.3}$$

5.2.1 Optimal detector

By substituting these notations into the general expression (4.24), the optimal detector for spatial multiplexing signals in MIMO time-invariant fading channels is given by

$$
\begin{aligned}
\hat{\mathbf{s}}_{d,\text{opt}} &= \arg \max_{\mathbf{s}_d \in \mathcal{A}^{N_t}} \left\{ \ln \left[\int p(\mathbf{z}_d | \mathbf{s}_d, \mathbf{a}) f(\mathbf{a} | \mathbf{z}_p) d\mathbf{a} \right] \right\} \\
&= \arg \max_{\mathbf{s}_d \in \mathcal{A}^{N_t}} \left\{ \sigma_n^{-2} \left(\boldsymbol{\Psi}_d^H \mathbf{z}_d + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right)^H \left(\boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \boldsymbol{\Upsilon}^{-1} \right)^{-1} \right. \\
&\qquad \left. \times \left(\boldsymbol{\Psi}_d^H \mathbf{z}_d + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right) - \ln \left| \boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \boldsymbol{\Upsilon}^{-1} \right| \right\}. \quad (5.4)
\end{aligned}
$$

5.2.2 Mismatched detectors

Correspondingly to the notations above, the channel estimators described in the last chapter are also modified. In MIMO time-invariant fading channels, the ML channel estimates are given by

$$
\hat{\mathbf{h}}_{\text{ML}} = \left(\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p \right)^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p. \tag{5.5}
$$

The ϵ-ML channel estimates become:

$$
\hat{\mathbf{h}}_\epsilon = \left(\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \epsilon \sigma_n^2 \mathbf{I}_{N_t N_r} \right)^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p, \tag{5.6}
$$

where the regularization parameter $\epsilon = 1$.

The MMSE channel estimates that take the joint spatial covariance matrix of the MIMO channel fading into account are given by

$$
\hat{\mathbf{h}}_{\text{MMSE}} = \left(\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \boldsymbol{\Upsilon}^{-1} \right)^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p. \tag{5.7}
$$

A mismatched detector uses the minimum distance detector that treats the channel estimates as perfect channel information and decides on the transmitted data symbols by minimizing the Euclidean distance

$$
\hat{\mathbf{s}}_{d,\text{mis}} = \arg \min_{\mathbf{s}_d \in \mathcal{A}^{N_t}} \left\{ \frac{\left\| \mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{h}} \right\|^2}{\sigma_n^2} \right\}, \tag{5.8}
$$

where $\hat{\mathbf{h}} = \hat{\mathbf{h}}_{\text{ML}}$ for ML channel estimates (5.5), or $\hat{\mathbf{a}} = \hat{\mathbf{h}}_\epsilon$ for ϵ-ML channel estimates (5.6) or $\hat{\mathbf{a}} = \hat{\mathbf{h}}_{\text{MMSE}}$ for MMSE channel estimates (5.7).
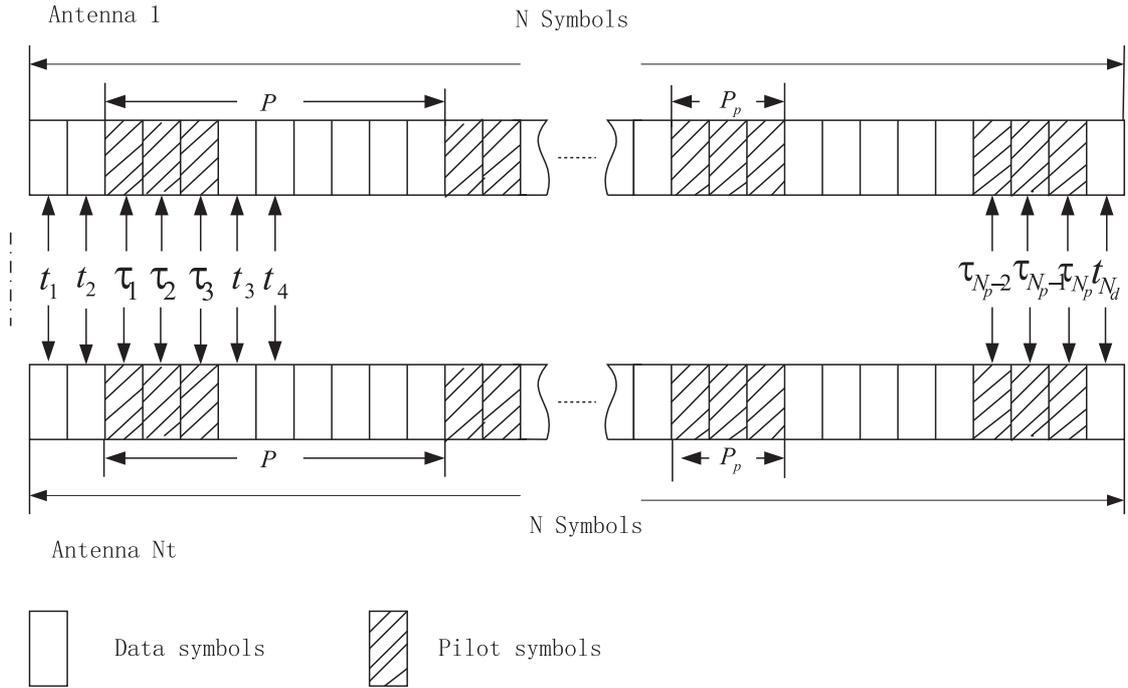
Figure 5.1: Structure of transmitted data blocks transmitted from all antennas.

5.3   Optimal and mismatched detection in MIMO time variant channels

5.3.1   Transmission Model

We now consider the transmission over $N_t \times N_r$ MIMO time-variant Rayleigh fading channels. We assume that a data matrix of $N \times N_t$ symbols is transmitted, $N_p \times N_t$ of which are pilot symbols and the others $N_d \times N_t$ are data symbols, where $N_d = N - N_p$. In this chapter, we consider the case that the pilot symbols are transmitted in groups of P_p symbols with the group period of P symbols, which is similar to the pattern scheme introduced in [24]. We assume that $P_p \geq N_t$. The structure of data blocks transmitted from transmit antennas is shown in Fig. 5.1. Note that from different antennas, random pilot symbols are transmitted at the same time instances. There are some other designs of pilot patterns for estimating MIMO channels, i.e., the optimal placement of pilot symbols [124] and the non-overlapping pilot structure [125]. The performance of the optimal detection with these pilot patterns will be investigated in further work. The received signal at the

rth receive antenna at time i can be written as:

$$z_r(i) = \sum_{k=1}^{N_t} h_{rk}(i)s_k(i) + n_r(i), \quad i = 0, \cdots, N-1, \quad r = 1, \cdots, N_r, \tag{5.9}$$

where $h_{rk}(i)$ is the channel coefficient linking the kth transmit antenna to the rth receive antenna; $s_k(i)$ is a symbol transmitted from the kth transmit antenna and $n_r(i)$ is the additive white Gaussian noise observed at the rth receive antenna. We denote a vector of the received signal as $\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_r^T, \ldots, \mathbf{z}_{N_r}^T]^T$, where $\mathbf{z}_r = [z_r(0), \ldots, z_r(N-1)]^T$; a noise vector $\mathbf{n} = [\mathbf{n}_1^T, \ldots, \mathbf{n}_r^T, \ldots, \mathbf{n}_{N_r}^T]^T$, where $\mathbf{n}_r = [n_r(0), \ldots, n_r(N-1)]^T$; a vector of the MIMO channel coefficients $\mathbf{h} = [\mathbf{h}_1, \ldots, \mathbf{h}_r, \ldots, \mathbf{h}_{N_r}]^T$, $\mathbf{h}_r = [\mathbf{h}_r^{(0)}, \ldots, \mathbf{h}_r^{(i)}, \ldots, \mathbf{h}_r^{(N-1)}]$ and $\mathbf{h}_r^{(i)} = [h_{r1}(i), \ldots, h_{rk}(i), \ldots, h_{rN_t}(i)]$; $(\cdot)^T$ denotes matrix transpose. We also assume that the noise samples at different receive antennas are uncorrelated and assume that $\mathbf{R}_{n0}$ is the same for all receive antenna. The $N \times N$ matrix $\mathbf{R}_{n0}$ characterizes time-correlation of noise samples for a single receive antenna. Here we assume that the noise samples are uncorrelated in time and $\mathbf{R}_{n0} = \sigma_n^2 \mathbf{I}_N$, where σ_n^2 is the noise variance. Then the total noise covariance matrix is given by $\mathbf{R}_n = \mathbf{I}_{N_r} \otimes \mathbf{R}_{n0}$.

We consider MIMO time variant channels, which are temporally correlated following Jakes' model [50, 51]. The temporal covariance matrix of the time-variant fading channel is a $N \times N$ matrix with elements $[\mathbf{\Upsilon}_0]_{t_1,t_2} = \rho(t_1 - t_2)$, $t_1, t_2 = 1, \ldots, N$, where $\rho(\tau)$ is the autocorrelation function of Jakes' fading process

$$\rho(\tau) = \sigma_{h_{r,k}}^2 J_0(2\pi\nu\tau), \tag{5.10}$$

$\sigma_{h_{r,k}}^2$ is the variance of the channel coefficients which is the same for all paths, $J_0(\cdot)$ is the zero-order Bessel function of the first kind, and ν is the Doppler frequency. The spatial correlation matrix of transmit antennas is $\mathbf{R}_t$, and the spatial correlation matrix of receive antennas is $\mathbf{R}_r$ as defined in Section 5.2. The joint spatial and temporal covariance matrix of the MIMO channel is given by [122]

$$\mathbf{\Upsilon} = E\{\mathbf{h}\mathbf{h}^H\} = \mathbf{R}_r \otimes \mathbf{\Upsilon}_0 \otimes \mathbf{R}_t. \tag{5.11}$$

Time variations in the MIMO time-variant fading channels are represented by the cubic B-splines which is given by

$$\varphi(t) = \begin{cases} \frac{2}{3} - \frac{t^2}{T^2} + \frac{|t|^3}{2T^3}, & \text{if } |t| < T, \\ \frac{1}{6}\left(2 - \frac{|t|}{T}\right)^3, & \text{if } T \le |t| < 2T, \\ 0, & \text{otherwise,} \end{cases} \tag{5.12}$$

where $T = (N-1)/(M-3)$ and M is the number of basis functions $\varphi_m(t) = \varphi[t - (m-2)T]$, $\quad m = 1, \cdots, M$. These basis functions are used to model the time-variant fading channel linking the kth transmit antenna and the rth receive antenna:

$$\bar{h}_{rk}(i) = \sum_{m=1}^{M} a_m^{(rk)} \varphi_m(i), \tag{5.13}$$

where $a_m^{(rk)}$ are expansion spline coefficients. An approximation (model) error between $h_{rk}(i)$ and $\bar{h}_{rk}(i)$ can be neglected if the number of basis functions M is large enough [27].

In the matrix form, the series (5.13) is given by

$$\bar{\mathbf{h}}_{rk} = \mathbf{B}\mathbf{a}^{(rk)}, \tag{5.14}$$

where $\mathbf{a}^{(rk)} = [a_1^{(rk)}, \ldots, a_M^{(rk)}]^T$ and $\mathbf{B}$ is an $N \times M$ matrix with elements

$$[\mathbf{B}]_{i,m} = \varphi_m(i), \quad i = 0, \ldots, N-1, \quad m = 1, \ldots, M, \tag{5.15}$$

which are samples of the basis functions at the symbol positions. The matrix $\mathbf{B}$ can be split into two parts as follows. The $N_p \times M$ matrix $\mathbf{B}_p$ contains samples of basis functions at the pilot symbol instants: $[\mathbf{B}_p]_{i,m} = \phi_m(\tau_i)$. The $N_d \times M$ matrix $\mathbf{B}_d$ contains samples of basis functions at the data symbol instants: $[\mathbf{B}_d]_{i,m} = \phi_m(t_i)$. According to these notations, the received signal can be represented as

$$\mathbf{z} = \mathbf{\Psi}\mathbf{a} + \mathbf{n}, \tag{5.16}$$

where

$$\mathbf{a} = [\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(r)}, \ldots, \mathbf{a}^{(N_r)}]^T,$$
$$\mathbf{a}^{(r)} = [\mathbf{a}^{(r1)}, \ldots, \mathbf{a}^{(rk)}, \ldots, \mathbf{a}^{(rN_t)}],$$
$$\mathbf{a}^{(rk)} = [a_1^{(rk)}, \ldots, a_m^{(rk)}, \ldots, a_M^{(rk)}],$$

$\mathbf{\Psi} = \mathbf{I}_{N_r} \otimes \mathbf{\Psi}_{N_t}$, $\mathbf{\Psi}_{N_t} = [\mathbf{\Psi}^{(1)}, \ldots, \mathbf{\Psi}^{(k)}, \ldots, \mathbf{\Psi}^{(N_t)}]$, $\mathbf{\Psi}^{(k)} = \mathbf{S}_k\mathbf{B}$, and $\mathbf{S}_k = \text{diag}[s_k(0), \ldots, s_k(i), \ldots, s_k(N-1)]$. The received signal corresponding to data and pilot parts of the transmitted data block are modeled, respectively, as

$$\mathbf{z}_d = \mathbf{\Psi}_d\mathbf{a} + \mathbf{n}_d, \quad \mathbf{z}_p = \mathbf{\Psi}_p\mathbf{a} + \mathbf{n}_p. \tag{5.17}$$

According to (5.17), the vector $\mathbf{z}$ can be split into a vector of received data symbols: $\mathbf{z}_d = [\mathbf{v}_1^T, \ldots, \mathbf{v}_r^T, \ldots, \mathbf{v}_{N_r}^T]^T$, where $[\mathbf{v}_r]_i = z_r(t_i)$, and a vector of received pilot symbols $\mathbf{z}_p =$

$[\mathbf{u}_1^T, \ldots, \mathbf{u}_r^T, \ldots, \mathbf{u}_{N_r}^T]^T$, where $[\mathbf{u}_r]_i = z_r(\tau_i)$. The noise vector $\mathbf{n}$ is also split into $\mathbf{n}_d =$
$[\mathbf{x}_1^T, \cdots, \mathbf{x}_r^T, \cdots, \mathbf{x}_{N_r}^T]^T$ where $[\mathbf{x}_r^T]_i = n_r(t_i)$ and $\mathbf{n}_p = [\mathbf{w}_1^T, \cdots, \mathbf{w}_r^T, \cdots, \mathbf{w}_{N_r}^T]^T$ where
$[\mathbf{w}_r^T]_i = n_r(\tau_i)$. Correspondingly, the matrix $\mathbf{\Psi}$ can be split into a matrix of transmitted
data symbols $\mathbf{\Psi}_d = \mathbf{I}_{N_r} \otimes \mathbf{\Psi}_d^{N_t}$, where

$$\mathbf{\Psi}_d^{N_t} = [\mathbf{\Psi}_d^{(1)}, \ldots, \mathbf{\Psi}_d^{(k)}, \ldots, \mathbf{\Psi}_d^{(N_t)}], \quad \mathbf{\Psi}_d^{(k)} = \mathbf{S}_d^{(k)}\mathbf{B}_d,$$

$$\mathbf{S}_d^{(k)} = \mathrm{diag}[s_k(0), \ldots, s_k(t_i), \ldots, s_k(N_d - 1)],$$

and a matrix of transmitted pilot symbols $\mathbf{\Psi}_p = \mathbf{I}_{N_r} \otimes \mathbf{\Psi}_p^{N_t}$, where

$$\mathbf{\Psi}_p^{N_t} = [\mathbf{\Psi}_p^{(1)}, \ldots, \mathbf{\Psi}_p^{(k)}, \ldots, \mathbf{\Psi}_p^{(N_t)}], \quad \mathbf{\Psi}_p^{(k)} = \mathbf{S}_p^{(k)}\mathbf{B}_p,$$

$$\mathbf{S}_p^{(k)} = \mathrm{diag}[s_k(0), \ldots, s_k(\tau_i), \ldots, s_k(N_p - 1)].$$

5.3.2 Optimal detection

The task of the generic optimal detector in (4.8) now becomes to find a data matrix $\mathbf{S}_d$ by
maximizing the PDF $p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)$ of the received signal $\mathbf{z}_d$, conditioned on the transmitted
data symbols $\mathbf{S}_d$ and the received pilot signal $\mathbf{z}_p$:

$$\hat{\mathbf{S}}_{d,\mathrm{opt}} = \arg \max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}} \{\ln[p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)]\}$$

$$= \arg \max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}} \left\{ \ln \left[\int p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a}) f(\mathbf{a}|\mathbf{z}_p) d\mathbf{a} \right] \right\}.$$

By taking above notations and after some algebra, we arrive at

$$\hat{\mathbf{S}}_{d,\mathrm{opt}} = \arg \max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}} \left\{ \sigma_n^{-2} \left(\mathbf{\Psi}_d^H \mathbf{z}_d + \mathbf{\Psi}_p^H \mathbf{z}_p \right)^H \left(\mathbf{\Psi}_d^H \mathbf{\Psi}_d + \mathbf{\Psi}_p^H \mathbf{\Psi}_p + \sigma_n^2 \mathbf{R}_\mathbf{a}^{-1} \right)^{-1} \right.$$

$$\left. \times \left(\mathbf{\Psi}_d^H \mathbf{z}_d + \mathbf{\Psi}_p^H \mathbf{z}_p \right) - \ln \left| \mathbf{\Psi}_d^H \mathbf{\Psi}_d + \mathbf{\Psi}_p^H \mathbf{\Psi}_p + \sigma_n^2 \mathbf{R}_\mathbf{a}^{-1} \right| \right\}. \quad (5.18)$$

The optimal detection requires the joint spatial-temporal correlation matrix $\mathbf{R}_a$ of the
expansion coefficients $\mathbf{a}$. The matrix $\mathbf{R}_a$ is given by

$$\mathbf{R}_a = E\{\mathbf{a}\mathbf{a}^H\} = \mathbf{R}_r \otimes \mathbf{R}_t \otimes \mathbf{\Lambda}_0, \quad (5.19)$$

where $\mathbf{\Lambda}_0 = E\{\mathbf{a}^{(rk)}(\mathbf{a}^{(rk)})^H\}$ is the $M \times M$ correlation matrix of the expansion coeffi-
cients $\mathbf{a}^{(rk)}$ that can be obtained from the fading covariance matrix $\mathbf{\Upsilon}_0$ defined by (5.10)
by requiring

$$E\{\bar{\mathbf{h}}_{rk}\bar{\mathbf{h}}_{rk}^H\} = E\{\mathbf{h}_{rk}\mathbf{h}_{rk}^H\} = \mathbf{\Upsilon}_0. \quad (5.20)$$

This requirement means that the correlation matrix of spline coefficients $\mathbf{\Lambda}_0$ results in time correlation of the time-variant fading that is equivalent to $\mathbf{\Upsilon}_0$, the time correlation of Jakes' model. By substituting (5.14) in (5.20), we obtain

$$\mathbf{B}\mathbf{\Lambda}_0\mathbf{B}^H = \mathbf{\Upsilon}_0. \tag{5.21}$$

As shown in Chapter 3, by multiplying both sides of (5.21) by $\mathbf{\Omega} = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ from the left and by $\mathbf{\Omega}^H$ from the right, we arrive at

$$\mathbf{\Lambda}_0 = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\mathbf{\Upsilon}_0\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}. \tag{5.22}$$

The optimal detector for MIMO time variant channels is given by (5.18). However, it is not feasible to solve this optimization problem for a high number $N_d N_t$ of data symbols due to extremely high complexity. If QAM modulated symbols with K constellation points are transmitted, we have to calculate this metric $K^{N_t N_d}$ times. In order to reduce the complexity, we only detect N_t symbols at once. In this case, the expressions above are simplified: $\mathbf{z}_d = [z_1(i), \ldots, z_r(i), \ldots, z_{N_r}(i)]^T$, $\mathbf{\Psi}_d^{(k)} = s_k(i)\mathbf{B}_d$, and $\mathbf{B}_d$ becomes a $1 \times M$ vector corresponding to the BS samples at ith instants. Now, we only calculate the optimal metric $N_d K^{N_t}$ times to recover all $N_d N_t$ data symbols.

5.3.3 Mismatched detection

A mismatched detector uses the minimum distance detector that treats the channel estimates as perfect channel information and decides on the transmitted data symbols by minimizing the Euclidean distance

$$\hat{\mathbf{S}}_{d,\text{mis}} = \arg\min_{\mathbf{S}_d \in \mathcal{A}^{N_t}} \left\{ \frac{\|\mathbf{z}_d - \mathbf{\Psi}_d\hat{\mathbf{a}}\|^2}{\sigma_n^2} \right\}. \tag{5.23}$$

where $\hat{\mathbf{a}}$ is the estimate of expansion coefficients, and it is depends on the applied estimation schemes. Here we also consider the ML, ϵ-ML and MMSE channel estimation.

For the ML channel estimation, the vector $\hat{\mathbf{a}} = \hat{\mathbf{a}}_{\text{ML}}$ and is given by

$$\hat{\mathbf{a}}_{\text{ML}} = (\mathbf{\Psi}_p^H\mathbf{\Psi}_p)^{-1}\mathbf{\Psi}_p^H\mathbf{z}_p. \tag{5.24}$$

For the ϵ-ML channel estimation, the vector $\hat{\mathbf{a}} = \hat{\mathbf{a}}_\epsilon$ and is given by

$$\hat{\mathbf{a}}_\epsilon = (\mathbf{\Psi}_p^H \mathbf{\Psi}_p + \epsilon \sigma_n^2 \mathbf{I}_{MN_tN_r})^{-1} \mathbf{\Psi}_p^H \mathbf{z}_p. \tag{5.25}$$

For the MMSE channel estimation, the vector $\hat{\mathbf{a}} = \hat{\mathbf{a}}_{\text{MMSE}}$ and is given by

$$\hat{\mathbf{a}}_{\text{MMSE}} = (\mathbf{\Psi}_p^H \mathbf{\Psi}_p + \sigma_n^2 \mathbf{R}_a^{-1})^{-1} \mathbf{\Psi}_p^H \mathbf{z}_p. \tag{5.26}$$

5.4 The equivalence between the optimal detector and the mismatched detector with MMSE channel estimates in SIMO channels with PSK modulation

In [35,36], the analysis and simulation results show that in SISO channels with PSK modulation and white Gaussian noise, the symbol-by-symbol optimal detection is equivalent to the mismatched detection with MMSE channel estimation. We find that the equivalence between the optimal detector and mismatched detector with MMSE channel estimates can be extended to the SIMO spatially uncorrelated Rayleigh fading channels with white Gaussian noise. Note that for the case with non-white Gaussian noise, this equivalence does not exist. The proof is given below.

Now, we consider a PSAM system with one transmit ($N_t = 1$) and N_r receive antennas, and the received signal in (5.9) becomes:

$$z_r(i) = h_r(i)s(i) + n_r(i), \tag{5.27}$$

where $h_r(i)$ is the channel coefficient between the transmit antenna and the rth receive antenna; $s(i)$ is a transmitted symbol and $n_r(i)$ is the additive white Gaussian noise observed at the rth receive antenna. Thus, the vectors and matrices defined in Section 5.3.1 are transformed to

$$\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_r^T, \ldots, \mathbf{z}_{N_r}^T]^T, \tag{5.28}$$

where $\mathbf{z}_r = [z_r(0), \ldots, z_r(N-1)]^T$; a noise vector

$$\mathbf{n} = [\mathbf{n}_1^T, \ldots, \mathbf{n}_r^T, \ldots, \mathbf{n}_{N_r}^T]^T, \tag{5.29}$$

where $\mathbf{n}_r = [n_r(0), \ldots, n_r(N-1)]^T$; a vector of SIMO channel coefficients as

$$\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_r^T, \ldots, \mathbf{h}_{N_r}^T]^T, \tag{5.30}$$

where $\mathbf{h}_r = [h_r(0), \ldots, h_r(i), \ldots, h_r(N-1)]^T$. The matrix of transmitted data symbols becomes

$$\boldsymbol{\Psi}_d = \mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d, \tag{5.31}$$

where

$$\bar{\boldsymbol{\Psi}}_d = \mathbf{S}_d \mathbf{B}_d, \quad \mathbf{S}_d = \text{diag}[s(\tau_1), \ldots, s(\tau_i), \ldots, s(\tau_{N_d})],$$

and the matrix of transmitted pilot symbols becomes

$$\boldsymbol{\Psi}_p = \mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_{N_p}, \tag{5.32}$$

where

$$\bar{\boldsymbol{\Psi}}_p = \mathbf{S}_p \mathbf{B}_p, \quad \mathbf{S}_p = \text{diag}[s(t_1), \ldots, s(t_i), \ldots, s(t_{N_p})].$$

Note that $\bar{\boldsymbol{\Psi}}_d$ and $\bar{\boldsymbol{\Psi}}_p$ are equivalent to $\boldsymbol{\Psi}_d^{(1)}$ and $\boldsymbol{\Psi}_p^{(1)}$ defined in Section 5.3.1.

In SIMO time-variant fading channels, the spatial correlation matrix of transmit antenna is $\mathbf{R}_t = 1$, while the spatial correlation matrix of receive antennas is an $N_r \times N_r$ matrix. We consider channels with no space correlation, i.e., $\mathbf{R}_r = \mathbf{I}_{N_r}$. The joint spatial and temporal covariance matrix of the SIMO channel is given by

$$\boldsymbol{\Upsilon} = \mathbf{R}_r \otimes \boldsymbol{\Upsilon}_0. \tag{5.33}$$

We intend to prove that the optimal symbol-by-symbol detector and the mismatched detector with MMSE channel estimation are equivalent, both maximizing the metric

$$\lambda(d) = \Re \left\{ \mathbf{z}_d^H \left[\mathbf{I}_{N_r} \otimes \left(\bar{\boldsymbol{\Psi}}_d (\bar{\boldsymbol{\Psi}}_p^H \bar{\boldsymbol{\Psi}}_p + \sigma_n^2 \boldsymbol{\Lambda}_0^{-1})^{-1} \bar{\boldsymbol{\Psi}}_p^H) \right] \mathbf{z}_p \right\}. \tag{5.34}$$

5.4.1 Mismatched detector with MMSE channel estimates

Firstly, we consider the mismatched detector treating MMSE channel estimates as perfect. Now, the task of the mismatched detector shown in (5.23) is to detect one data symbol

$$
\begin{aligned}
\hat{d}_{\text{mis}} \;&= \arg\min_{d\in\mathcal{A}} \left\{ \frac{\left\| \mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\text{MMSE}} \right\|^2}{\sigma_n^2} \right\} \\
&= \arg\max_{d\in\mathcal{A}} \left\{ 2\Re\{\mathbf{z}_d^H \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\text{MMSE}}\} - \hat{\mathbf{a}}_{\text{MMSE}}^H \boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\text{MMSE}} \right\},
\end{aligned}
$$

$$(5.35)$$

where the alphabet $\mathcal{A}$ includes all symbols corresponding to the PSK constellation points. Note that in (5.35), since $|d|$ is constant, the term

$$
\begin{aligned}
\hat{\mathbf{a}}_{\text{MMSE}}^H \boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\text{MMSE}} \;&= \hat{\mathbf{a}}_{\text{MMSE}}^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d)^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d) \hat{\mathbf{a}}_{\text{MMSE}} \\
&= \hat{\mathbf{a}}_{\text{MMSE}}^H [\mathbf{I}_{N_r} \otimes (|d|^2 \mathbf{B}_d^H \mathbf{B}_d)] \hat{\mathbf{a}}_{\text{MMSE}}
\end{aligned}
$$

$$(5.36)$$

does not depend on d, so it can be removed from (5.35) without affecting the decision result. Finally, after some algebra, we arrive at

$$
\begin{aligned}
\hat{d}_{\text{mis}} \;&= \arg\max_{d\in\mathcal{A}} \left\{ \Re\{\mathbf{z}_d^H \boldsymbol{\Psi}_d \hat{\mathbf{a}}_{\text{MMSE}}\} \right\} \\
&= \arg\max_{d\in\mathcal{A}} \left\{ \Re\{\mathbf{z}_d^H \boldsymbol{\Psi}_d (\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_a^{-1})^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p\} \right\} \\
&= \arg\max_{d\in\mathcal{A}} \left\{ \Re\left\{ \mathbf{z}_d^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d) \left[\mathbf{I}_{N_r} \otimes (\bar{\boldsymbol{\Psi}}_p^H \bar{\boldsymbol{\Psi}}_p + \sigma_n^2 \boldsymbol{\Lambda}_0)^{-1} \right]^{-1} (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_p^H) \mathbf{z}_p \right\} \right\} \\
&= \arg\max_{d\in\mathcal{A}} \{\lambda(d)\}.
\end{aligned}
$$

$$(5.37)$$

5.4.2 Optimal detector

We now derive the optimal symbol-by-symbol detector of PSK signals in SIMO time-variant fading channels. After some algebra, (5.18) becomes

$$
\hat{d}_{\text{opt}} = \arg\max_{d\in\mathcal{A}} \left\{ \frac{1}{\sigma_n^2} \left(\boldsymbol{\Psi}_d^H \mathbf{z}_d + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right)^H \mathbf{Y} \left(\boldsymbol{\Psi}_d^H \mathbf{z}_d + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right) - \ln \left| \mathbf{Y}^{-1} \right| \right\}, \quad (5.38)
$$

where $\mathbf{Y} = \left(\boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_\mathbf{a}^{-1} \right)^{-1}$. Similar to (5.36), since $|d|$ is constant, we find that $\boldsymbol{\Psi}_d^H \boldsymbol{\Psi}_d$ does not depend on the transmitted PSK symbol d. As a result, the term $\ln \left| \mathbf{Y}^{-1} \right|$ can be removed from (5.38) and we obtain

$$
\hat{d}_{\text{opt}} = \arg\max_{d\in\mathcal{A}} \left\{ \mathbf{z}_p^H \boldsymbol{\Psi}_p \mathbf{Y} \boldsymbol{\Psi}_p^H \mathbf{z}_p + \mathbf{z}_d^H \boldsymbol{\Psi}_d \mathbf{Y} \boldsymbol{\Psi}_d^H \mathbf{z}_d + 2\Re \left\{ \mathbf{z}_d^H \boldsymbol{\Psi}_d \mathbf{Y} \boldsymbol{\Psi}_p^H \mathbf{z}_d \right\} \right\}. \quad (5.39)
$$

The first term in (5.39) does not depend on d. The second term can be transformed into

$$
\begin{aligned}
\mathbf{z}_d^H \boldsymbol{\Psi}_d \mathbf{Y} \boldsymbol{\Psi}_d^H \mathbf{z}_d \; &= \mathbf{z}_d^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d) \mathbf{Y} (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d)^H \mathbf{z}_d \\
&= |d|^2 \mathbf{z}_d^H (\mathbf{I}_{N_r} \otimes \bar{\mathbf{B}}_d) \mathbf{Y} (\mathbf{I}_{N_r} \otimes \bar{\mathbf{B}}_d)^H \mathbf{z}_d;
\end{aligned} \tag{5.40}
$$

it also does not depend on d since $|d|^2$ is constant. We can now simplify (5.39) as

$$
\begin{aligned}
\hat{d}_{\text{opt}} \; &= \arg \max_{d \in \mathcal{A}} \left\{ \Re \{ \mathbf{z}_d^H \boldsymbol{\Psi}_d \mathbf{Y} \boldsymbol{\Psi}_p^H \mathbf{z}_p \} \right\} \\
&= \arg \max_{d \in \mathcal{A}} \left\{ \Re \left\{ \mathbf{z}_d^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d) \left[(\mathbf{I}_{N_r} \otimes \boldsymbol{\Psi}_d)^H \right. \right. \right. \\
&\qquad \left. \times (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_d) + (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_p)^H (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_p) \right. \\
&\qquad \left. \left. + \sigma_n^2 \mathbf{I}_{N_r} \otimes \boldsymbol{\Lambda}_0^{-1} \right]^{-1} (\mathbf{I}_{N_r} \otimes \bar{\boldsymbol{\Psi}}_p)^H \mathbf{z}_p \right\} \right\} \\
&= \arg \max_{d \in \mathcal{A}} \left\{ \Re \left\{ \mathbf{z}_d^H \left[\mathbf{I}_{N_r} \otimes \left(\bar{\boldsymbol{\Psi}}_d (\bar{\boldsymbol{\Psi}}_d^H \bar{\boldsymbol{\Psi}}_d + \mathbf{X}_p)^{-1} \bar{\boldsymbol{\Psi}}_p^H \right) \right] \mathbf{z}_p \right\} \right\}, \tag{5.41}
\end{aligned}
$$

where $\mathbf{X}_p = \bar{\boldsymbol{\Psi}}_p^H \bar{\boldsymbol{\Psi}}_p + \sigma_n^2 \boldsymbol{\Lambda}_0^{-1}$. By using the matrix inversion lemma [126], we obtain

$$
\begin{aligned}
\hat{d}_{\text{opt}} \; &= \arg \max_{d \in \mathcal{A}} \left\{ \Re \left\{ \mathbf{z}_d^H \left[\mathbf{I}_{N_r} \otimes \left(\bar{\boldsymbol{\Psi}}_d (\mathbf{X}_p^{-1} - \frac{\mathbf{X}_p^{-1} \bar{\boldsymbol{\Psi}}_d^H \bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1}}{1 + \bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1} \bar{\boldsymbol{\Psi}}_d^H}) \bar{\boldsymbol{\Psi}}_p^H \right) \right] \mathbf{z}_p \right\} \right\} \\
&= \arg \max_{d \in \mathcal{A}} \left\{ \Re \left\{ \mathbf{z}_d^H \left[\mathbf{I}_{N_r} \otimes \left(\bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p - \frac{\bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1} \bar{\boldsymbol{\Psi}}_d^H}{1 + \bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1} \bar{\boldsymbol{\Psi}}_d^H} \bar{\boldsymbol{\Psi}}_d \mathbf{X}_p^{-1} \bar{\boldsymbol{\Psi}}_p^H \right) \right] \mathbf{z}_p \right\} \right\} \\
&= \arg \max_{d \in \mathcal{A}} \left\{ \frac{\lambda(d)}{1 + |d|^2 \mathbf{B}_d \mathbf{X}_p^{-1} \mathbf{B}_d^H} \right\}. \tag{5.42}
\end{aligned}
$$

As $|d|^2 \mathbf{B}_d \mathbf{X}_p^{-1} \mathbf{B}_d^H$ is a constant which does not depend on d, (5.42) transforms into (5.37). Thus, we proved that the mismatched detector with MMSE channel estimates is equivalent to the optimal symbol-by-symbol detector for PSK signals in spatially uncorrelated SIMO Rayleigh fading channels. Note that in spatially correlated channels or with non-white Gaussian noise, these two detectors are not equivalent.

5.5   Simulation results

In this section, numerical results obtained by simulation are presented to compare the performance of the optimal detector and the mismatched detectors for BPSK and 16QAM uncoded signals. We assume that the average energy of each pilot and data symbol is equal to E_s.

The average SNR is defined as

$$\zeta = \frac{E\{(\mathbf{\Psi a})^H \mathbf{\Psi a}\}}{\text{tr}\{\mathbf{R}_n\}} = \frac{\text{tr}\{\mathbf{\Gamma R}_a\}}{\text{tr}\{\mathbf{R}_n\}}, \tag{5.43}$$

where $\mathbf{\Gamma} = E\{\mathbf{\Psi}^H \mathbf{\Psi}\}$. For the additive noise uncorrelated at different receive antennas we have $\text{tr}\{\mathbf{R}_n\} = N_r \text{tr}\{\mathbf{R}_{n0}\}$ and (5.43) becomes

$$\zeta = \frac{\text{tr}[\mathbf{\Gamma R}_a]}{N_r \text{tr}[\mathbf{R}_{n0}]}. \tag{5.44}$$

The matrix $\mathbf{R}_a = \mathbf{I}_{N_r} \otimes \mathbf{I}_{N_t} \otimes \mathbf{\Lambda}_0$ is block-diagonal and the noise is white, i.e., $\mathbf{R}_{n0} = \sigma_n^2 \mathbf{I}_N$; then taking $\mathbf{\Gamma} = \mathbf{I}_{N_r} \otimes \mathbf{\Gamma}_0$ into account, we obtain

$$\zeta = \frac{\text{tr}[\mathbf{\Gamma}_0 (\mathbf{I}_{N_t} \otimes \mathbf{\Lambda}_0)]}{N \sigma_n^2}, \tag{5.45}$$

where $\mathbf{\Gamma}_0 = E\{\mathbf{\Psi}_{N_t}^H \mathbf{\Psi}_{N_t}\} = N E_s \mathbf{I}_{N_t} \otimes (\mathbf{B}^H \mathbf{B})$. Finally, the average SNR is given by

$$\zeta = \frac{N_t E_s \sigma_{h_{r,k}}^2}{\sigma_n^2}. \tag{5.46}$$

The average bit energy to noise ratio is defined as $E_b/N_0 = \zeta/(N_t \log_2 K)$. The simulation results here represent the BER versus E_b/N_0.

5.5.1   MIMO time invariant channels

Firstly we will consider the performance of the optimal and mismatched detectors in MIMO time-invariant fading channels. We set $N_d = 1$, $N_p = N_t + 1$.

Fig. 5.2 shows the performance of the optimal and mismatched detectors in 2×2 MIMO system with BPSK signals. It is seen that all detectors provide similar performance in both spatially uncorrelated (Fig. 5.2.a) and high spatially correlated (Fig. 5.2.b) MIMO channels.

Fig. 5.3 and Fig. 5.4 show the performance of the optimal and mismatched detectors for BPSK signals in 2×4 and 4×4 systems, respectively. It is clear that the improvement in the performance caused by using the optimal detector increases when the number of antennas in the system increases. However, the spatial correlation between antennas can not affect this improvement.
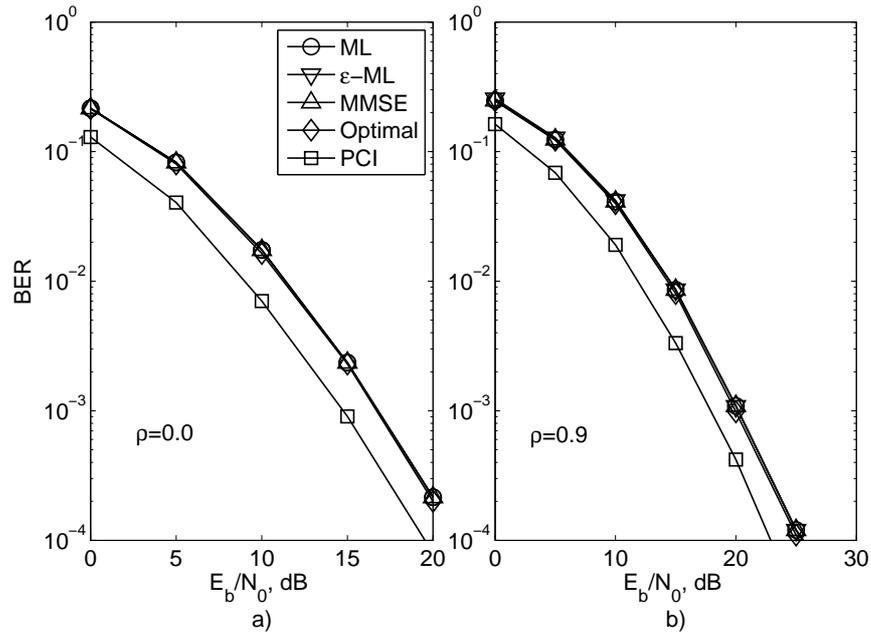
Figure 5.2: BER performance of the optimal and mismatched detectors for BPSK signals in 2×2 MIMO time-invariant fading channels, $N_t = 2$, $N_r = 2$, $N_p = 3$; a) $\rho = 0$ and b) $\rho = 0.9$.
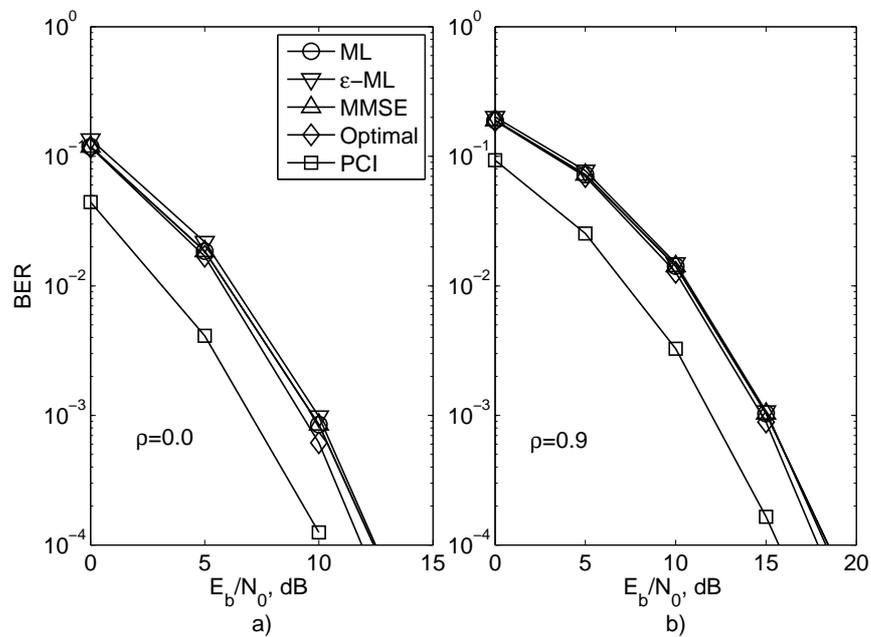
Figure 5.3: BER performance of the optimal and mismatched detectors for BPSK signals in 2×4 MIMO time-invariant fading channels, $N_t = 2$, $N_r = 4$, $N_p = 3$; a) $\rho = 0$ and b) $\rho = 0.9$.
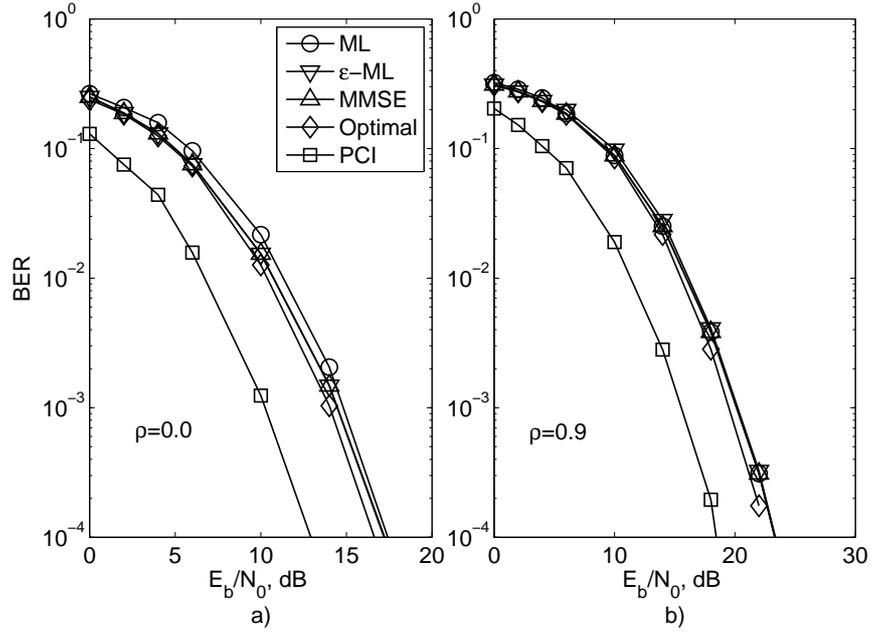
Figure 5.4: BER performance of the optimal and mismatched detectors for BPSK signals in 4×4 MIMO time-invariant fading channels, $N_t = 4$, $N_r = 4$, $N_p = 5$; a) $\rho = 0$ and b) $\rho = 0.9$.

Fig. 5.5 and Fig. 5.6 for the transmission with 16QAM signals result in a similar conclusion. Therefore, we will consider the channels with no spatial correlation in the following simulation, i.e., $\mathbf{R}_t = \mathbf{I}_{N_t}, \mathbf{R}_r = \mathbf{I}_{N_r}$.

5.5.2   MIMO time variant channels

We now investigate the performance of the optimal and mismatched detectors in spatially uncorrelated MIMO time-variant fading channels. The number of pilot symbols inserted into each transmitted block of one transmit antenna is $N_p = (M + 1)P_p$ and there are $N_t P_p (M + 1)$ pilot symbols in total, where P_p is the length of a group of pilot symbols as shown in Fig. 5.1. In the simulation we set $P_p = N_t$, $N = 507$, $P = 22$ and $M = 23$. This corresponds to as little as 4.3% overhead due to the use of pilot symbols.

Fig. 5.7 and Fig. 5.8 show the BER performance of the detectors for 16QAM signals in MIMO time-variant Rayleigh fading channels with $\nu T_s = 0.01$, where T_s is a symbol duration. Fig. 5.7 shows the BER performance in a SISO channel $(N_r = N_t = 1)$ and in
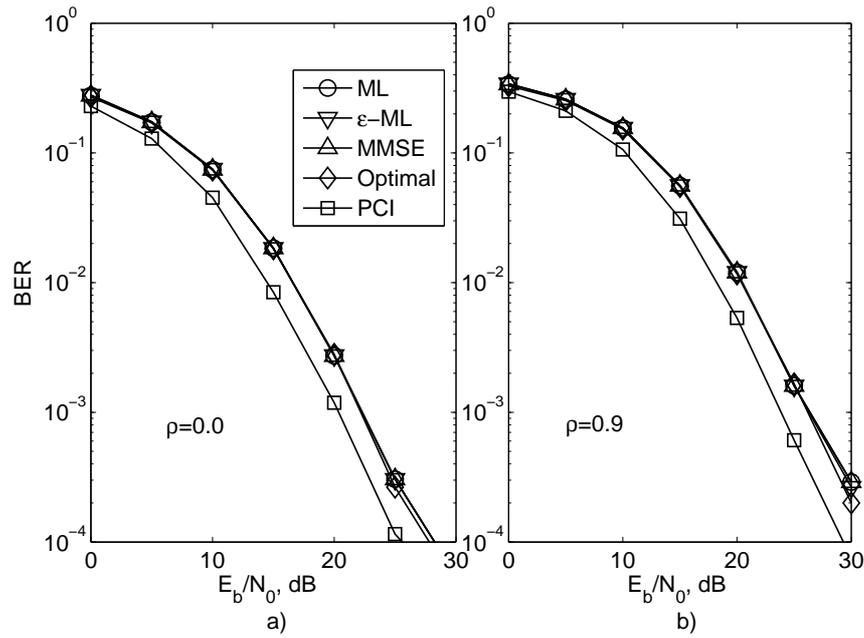
Figure 5.5: BER performance of the optimal and mismatched detectors for 16QAM signals in 2×2 MIMO time-invariant fading channels, $N_t = 2$, $N_r = 2$, $N_p = 3$; a) $\rho = 0$ and b) $\rho = 0.9$.
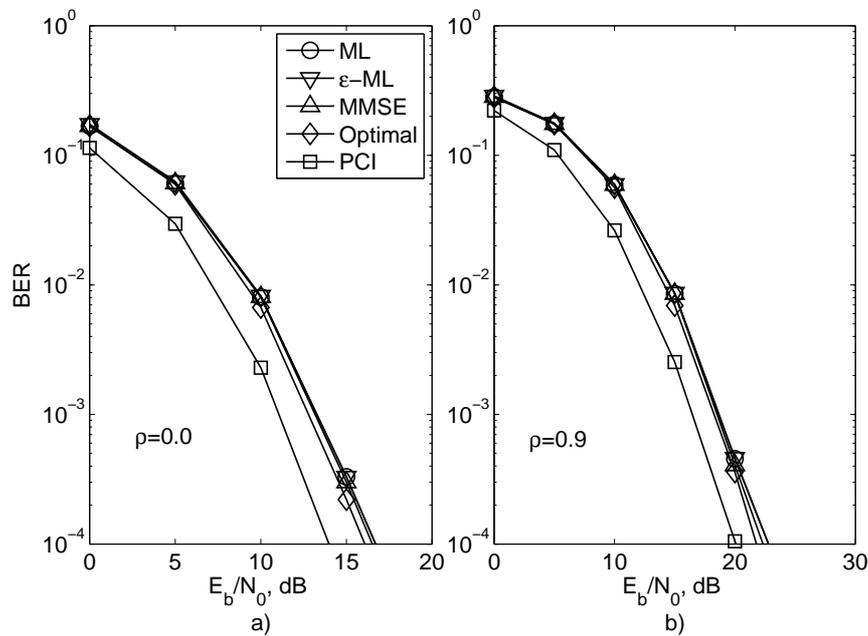
Figure 5.6: BER performance of the optimal and mismatched detectors for 16QAM signals in 2×4 MIMO time-invariant fading channels, $N_t = 2$, $N_r = 4$, $N_p = 3$; a) $\rho = 0$ and b) $\rho = 0.9$.
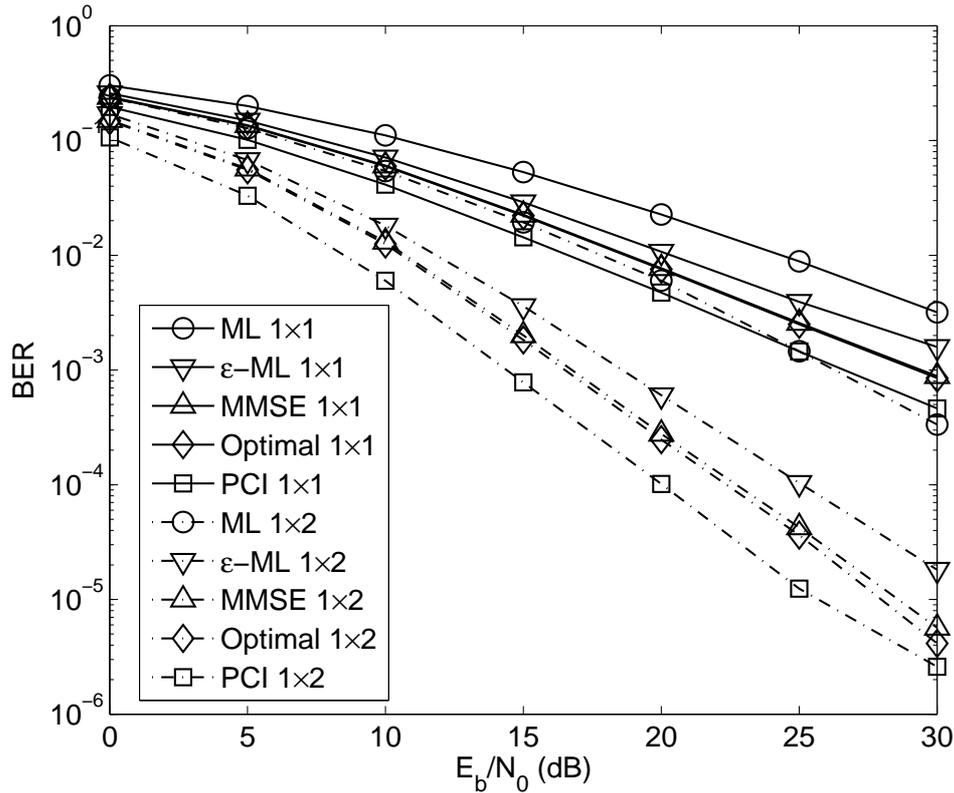
Figure 5.7: BER performance of the optimal and mismatched detectors for 16QAM signals in 1×1 and 1×2 channels; $N = 507$, $P = 22$, $P_p = 1$, $M = 23$.

a 1×2 SIMO channel $(N_t = 1, N_r = 2)$. We set $P_p = 1$ and $N_p = 24$ in both these cases. In the SISO channel, at BER=10^{-2}, the optimal detector outperforms the mismatched detector with ML channel estimates by 6.1 dB and with ϵ-ML channel estimates by 1.2 dB. However, the BER performance of the optimal detector and the mismatched detector with MMSE channel estimates are similar. These detectors are inferior to the minimum distance detector with perfect channel information (PCI) by 2.4 dB. In the 1×2 SIMO channel, when BER=10^{-3}, the improvement due to the use of the optimal detector increases up to 10 dB compared with the mismatched detector with ML channel estimates and up to 2.1 dB compared with ϵ-ML channel estimates. The BER performance of the optimal detector and that of the mismatched detector with MMSE channel estimates are close; the difference in the performance is 0.2 dB. The gap between the BER curve of the optimal detector and that of the minimum distance detector with PCI is 2 dB when BER=10^{-2}.

Fig. 5.8 shows simulation results for a 2×2 MIMO channel $(N_t = N_r = 2)$; here, we
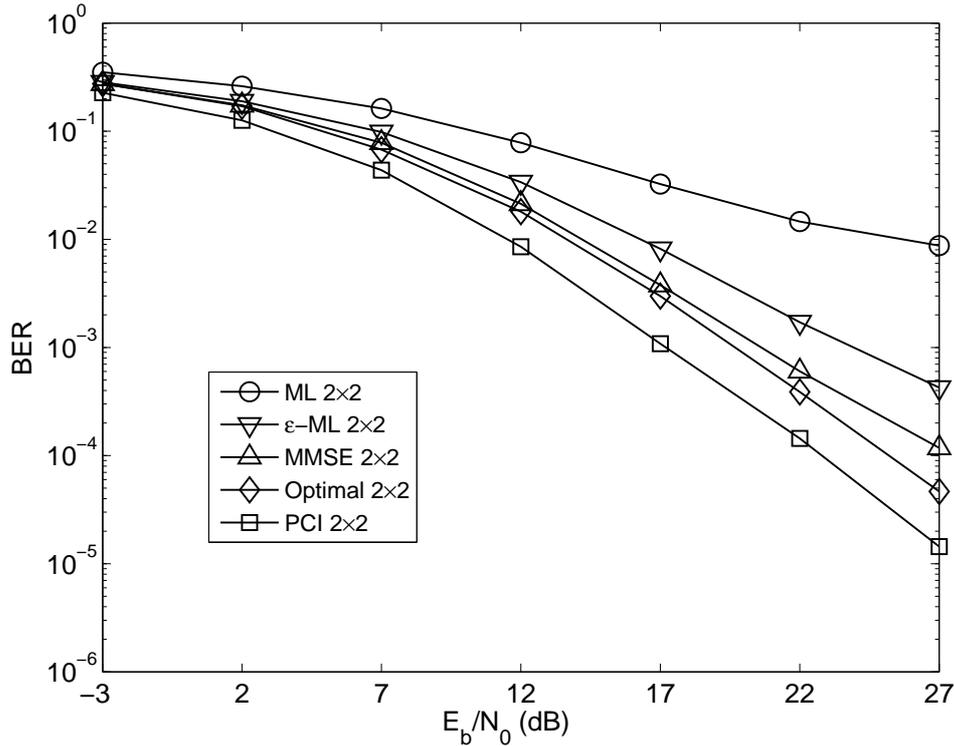
Figure 5.8: BER performance of the optimal and mismatched detectors for 16QAM signals in a 2×2 channel ; $N = 507$, $P = 22$, $P_p = 2$, $M = 23$.

set $P_p = 2$ and $N_p = 47$. The optimal detector outperforms the mismatched detector with ML channel estimates by 13 dB at BER=10^{-2}; it outperforms the mismatched detector with ϵ-ML channel estimates by 4.8 dB at BER=10^{-3}; it also outperforms the mismatched detector with MMSE channel estimates by 2.2 dB when BER=10^{-4}. From this figure, we find that the optimal detector provides better BER performance than mismatched detectors which treat channel estimates as perfect for 16QAM signals, and the improvement in the BER performance is increased when the number of antennas increases.

Fig. 5.9 and Fig. 5.10 show the BER performance of the optimal detector and mismatched detectors for BPSK signals. As shown in Fig. 5.9, in a SISO channel, when BER=10^{-2}, the optimal detector outperforms the mismatched detector with ML channel estimates by 5.5 dB and the one with ϵ-ML channel estimates by 0.5 dB. Similar to the case of 16QAM signals in Fig. 5.7, the BER curves for the optimal detector and the mismatched detector with MMSE channel estimates are close. In a 1×2 SIMO channel, when BER=10^{-4}, the benefit due to the use of the optimal detector is 7.4 dB compared with the mismatched detector with ML channel estimates. Compared with the mismatched detec-
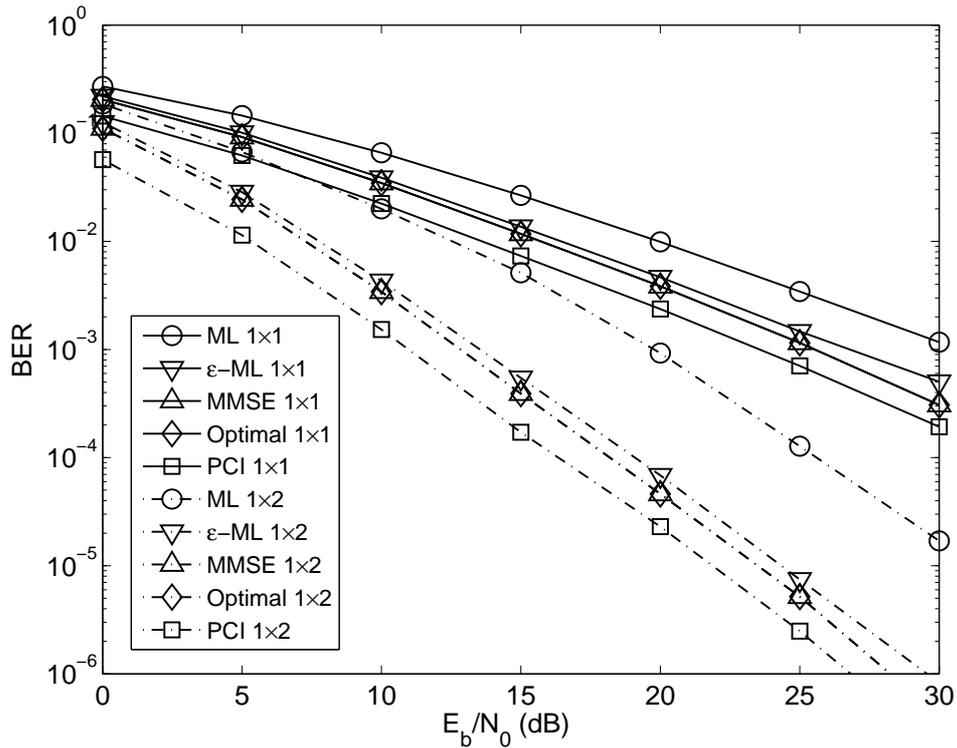
Figure 5.9: BER performance of the optimal and mismatched detectors for BPSK signals
in 1×1 and 1×2 channels; $N = 507$, $P = 22$, $P_p = 1$, $M = 23$.

tor with ϵ-ML channel estimates, this benefit is 1.1 dB. It can be seen that the performance
of the optimal detector is similar with that of the mismatched detector with MMSE chan-
nel estimates, and it is inferior to the performance of the minimum distance detector with
PCI by 1.6 dB when BER$= 10^{-4}$.

Fig. 5.10 shows simulation results in a 2×2 and 2×4 MIMO channels. In the 2×2
MIMO channel, the mismatched detector with ML estimates provides significantly worse
performance compared with others. The optimal detector provides significantly better per-
formance than the mismatched detector with ϵ-ML channel estimates and it outperforms
the mismatched detector with MMSE channel estimates by 3 dB when BER$=10^{-5}$. In
the 2×4 MIMO channels, the optimal detector significantly outperforms the mismatched
detectors with ϵ-ML channel estimates. It also outperforms the mismatched detector with
MMSE channel estimates by 5.7 dB when BER$=10^{-6}$.
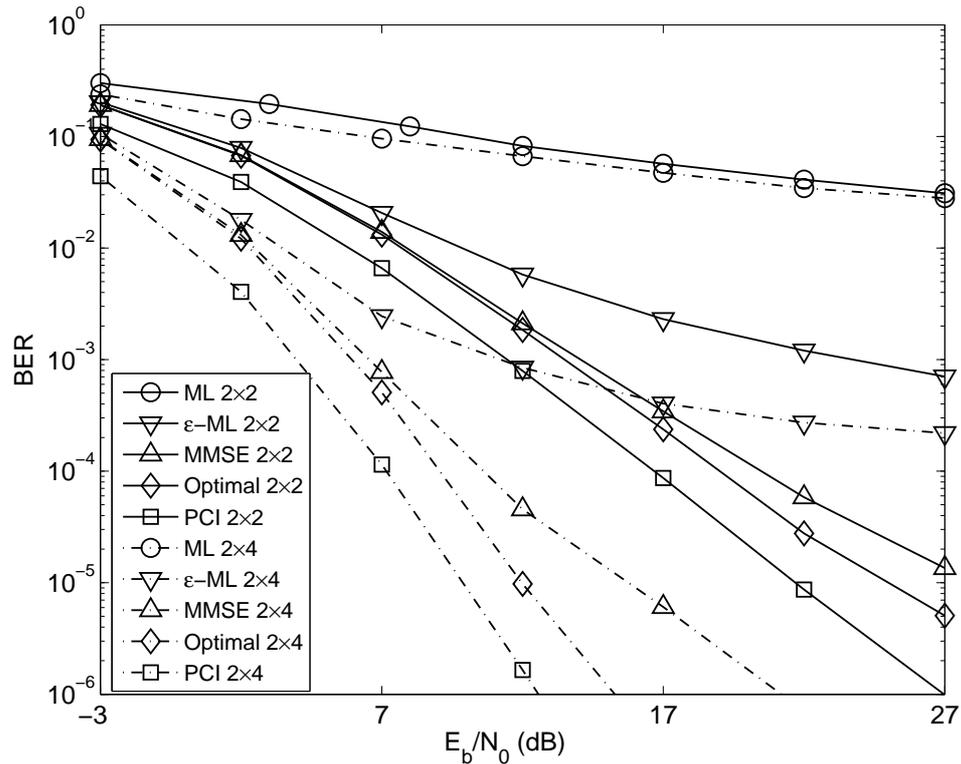
Figure 5.10: BER performance of the optimal and mismatched detectors for BPSK signals in 2×2 and 2×4 channels; $N = 507$, $P = 22$, $P_p = 2$, $M = 23$.

5.6 Conclusions

We have proposed and investigated an optimal detector for PSAM systems in MIMO Rayleigh fading channels. In MIMO time-invariant Rayleigh fading channels, comparing with mismatched detectors, the benefit on performance caused by using the optimal detector becomes significant when the number of antennas increases. However, the simulation results show that the spatial correlation between antennas does not influence upon the difference between the performance of the optimal detector and that of mismatched detectors. We have also extended the optimal detector to spatially uncorrelated MIMO time-variant fading channels, the time variation of which is modeled by BS basis functions. We have investigated the optimal detector and compared its performance with that of traditional mismatched detectors with ML, regularized ML or MMSE channel estimates and compared their performance in MIMO time-variant fading channels with 16QAM and BPSK modulation. Among these mismatched detectors, the one exploiting MMSE channel estimates provides the best performance and its performance is close to that of the

optimal detector in SISO time-variant fading channels when QAM signals are transmitted. However, the optimal detector significantly outperforms the mismatched detectors in spatially uncorrelated MIMO time-variant fading channels when the number of antennas increases. In this chapter, we have also shown that if the SIMO channel is spatially uncorrelated, the optimal symbol-by-symbol detector of PSK signals is equivalent to the mismatched detector with MMSE channel estimates. Note that in the general case of QAM symbols, the term $\mathbf{Y}$ will depend on d and these two detectors are not equivalent.

Chapter 6

Optimal Detection of OFDM Signals in Frequency-Selective Fading Channels with Imperfect Channel Estimation

Contents

6.1 Introduction

In previous chapters, we have investigated the optimal detector and compared it with mismatched detectors in frequency-flat Rayleigh fading channels. The optimal detector can also be used to detect OFDM signals in frequency-selective fading channels. In this chapter, we derive and investigate the optimal detection of OFDM signals.

In OFDM systems, channel estimation is usually performed by employing pilot tones [17, 21, 24, 127–138]. Then, the channel estimates are treated as perfect in the traditional minimum distance detector. Motivated by the benefit caused by using the optimal detector as shown in previous chapters, we derive an optimal detector for OFDM signals and specify it for spatially uncorrelated MIMO frequency-selective fading channels. We compare the BER performance of this detector with that of mismatched detectors with ML, regularized ML (ϵ-ML) or MMSE channel estimates for uncoded transmission. We also investigate the performance of iterative receivers incorporating the optimal detector. Specifically, four iterative receivers are considered: receivers with mismatched detectors using ML, $\epsilon-$ML or MMSE channel estimates, and a receiver with optimal detector at the first iteration and the mismatched detector based on MMSE channel estimates in subsequent iterations.

In order to approximate the channel frequency response at data positions by using channel estimates at positions of pilot symbols, many channel estimation schemes for OFDM systems have been proposed in the literature [21, 70, 77, 129–131]. In [129], a low rank approximation to the frequency domain linear MMSE channel estimator was proposed by using singular value decomposition. In [130], the Wiener filter has been investigated, and a robust MMSE channel estimator exploiting correlation in both time and frequency domains was proposed in [131]. In this chapter, we also use channel estimation based on BEMs, such as CE model [19, 21–24], GCE model [25], B-spline functions [26–28], Slepian sequences [20, 29, 30] or KL basis functions [31, 32] to approximate correlated fading channels. After comparing the MSE performance of MMSE channel estimators corresponding to these BEMs in Rayleigh frequency-selective fading channels, we use the cubic B-splines to represent the channel frequency response.

The rest of this chapter is structured as follows. In Section 6.2, the transmission model and communication scenarios are introduced. Different BEMs used to represent the channel frequency response are specified for frequency domain approximation in Section 6.3. Section 6.4 describes the proposed optimal detector and mismatched detectors with different channel estimation schemes, and Section 6.5 describes the iterative receivers. Simulation results are given in Section 6.6, followed by conclusions in Section 6.7.

6.2 Transmission model

We consider an MIMO OFDM system with N subcarriers, N_t transmit and N_r receive antennas. We assume that a data matrix of $N \times N_t$ symbols is transmitted, $N_p \times N_t$ of which are pilot symbols and the others $N_d \times N_t$ are data symbols, where $N_d = N - N_p$. The duration of an OFDM symbol without a cyclic prefix (CP) is $T_s = 1/\Delta f$, where Δf is the space between two neighboring subcarriers. In frequency domain, the pilot symbols are inserted in groups of P_p symbols with the group period of P symbols to construct an OFDM symbol transmitted from one antenna as shown in Fig. 6.1. Here we follow the the design of group pilot insertion in [24] and guarantee $P_p \geq N_t$. This OFDM symbol is inverse Fourier transformed and a CP is added before the transmission.
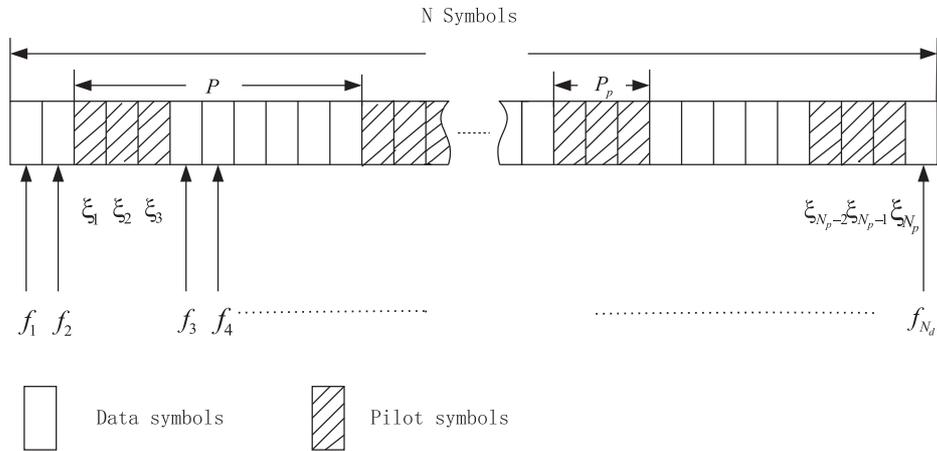
Figure 6.1: Structure of an OFDM symbol transmitted from one transmit antenna.

We consider transmission over MIMO time invariant frequency-selective fading channels with L path components, and assume that the inter-symbol interference (ISI) between consecutive OFDM symbols is eliminated by using a CP of length $L_{\max}T$ chosen to be longer than the maximum channel delay, where $T = T_s/N$. The channel from the kth transmit antenna to the rth receive antenna can be represented by the channel impulse response

$$g_{rk}(\tau) = \sum_{l=0}^{L-1} \alpha_{rk}(l)\delta(\tau - \tau_{rk}(l)), \qquad (6.1)$$

where $\delta(\tau)$ is the Dirac delta function, $\tau_{rk}(l)$ and $\alpha_{rk}(l)$ are, respectively, the delay and complex amplitude of the lth path in the channel. The channel frequency response at the

ith subcarrier from the kth transmit antenna to the rth receive antenna is given by

$$h_{rk}(i) = \sum_{l=0}^{L-1} \alpha_{rk}(l) e^{-j2\pi i \Delta f \tau_{rk}(l)}. \tag{6.2}$$

We denote an $N \times 1$ vector $\mathbf{h}_{rk} = [h_{rk}(0), \ldots, h_{rk}(i), \ldots, h_{rk}(N-1)]^T$ and an $N_t N_r M \times 1$ vector of the MIMO channel frequency response $\mathbf{h} = [\mathbf{h}_1, \ldots, \mathbf{h}_r, \ldots, \mathbf{h}_{N_r}]^T$, where $\mathbf{h}_r = [\mathbf{h}_r^{(0)}, \ldots, \mathbf{h}_r^{(i)}, \ldots, \mathbf{h}_r^{(N-1)}]$ is a $1 \times N$ row vector, $\mathbf{h}_r^{(i)} = [h_{r1}(i), \ldots, h_{rk}(i), \ldots, h_{rN_t}(i)]$ is a $1 \times N_t$ row vector.

The path amplitudes $\alpha_{rk}(l)$ are independent zero-mean complex Gaussian random variables with exponential power delay profile given by [129]

$$\vartheta(\tau) = e^{-\tau/\tau_{\text{rms}}}, \tag{6.3}$$

where τ_{rms} is the root-mean square width of $\vartheta(\tau)$. The probability density function (PDF) of a random delay τ_{rk} is uniform and given by

$$f_\tau(\tau) = \begin{cases} 1/(L_{\max}T) & \text{if } \tau \in [0, L_{\max}T], \\ 0, & \text{otherwise.} \end{cases} \tag{6.4}$$

With these definitions, elements of the covariance matrix $\mathbf{\Upsilon}_0 = E\{\mathbf{h}_{rk}\mathbf{h}_{rk}^H\}$ of the fading in the frequency domain can be represented as [129]

$$[\mathbf{\Upsilon}_0]_{m,n} = \frac{1 - e^{-L_{\max}T[(1/\tau_{\text{rms}})+j2\pi\Delta f(m-n)]}}{(1 - e^{-L_{\max}T/\tau_{\text{rms}}})(1 + j2\pi\Delta f(m-n)\tau_{\text{rms}})}, \tag{6.5}$$

where n and m denote two subcarriers of the OFDM symbol. Therefore, the dimension of $\mathbf{\Upsilon}_0$ is $N \times N$.

The spatial correlation matrix of transmit antennas is an $N_t \times N_t$ symmetric matrix $\mathbf{R}_t$, while the spatial correlation matrix of receive antennas is an $N_r \times N_r$ symmetric matrix $\mathbf{R}_r$. We consider channels with no spatial correlation, i.e., $\mathbf{R}_t = \mathbf{I}_{N_t}, \mathbf{R}_r = \mathbf{I}_{N_r}$, where $\mathbf{I}_S$ is an $S \times S$ identity matrix. The joint spatial and frequency covariance matrix of the MIMO channel is given by

$$\mathbf{\Upsilon} = E\{\mathbf{h}\mathbf{h}^H\} = \mathbf{R}_r \otimes \mathbf{\Upsilon}_0 \otimes \mathbf{R}_t. \tag{6.6}$$

Note that the fading channel described above is only an example used in our simulation. The results obtained below for the optimal detection can be applied to the case of an arbitrary fading covariance matrix $\mathbf{\Upsilon}_0$.

At the receiver side, the CP is removed and the received signal is Fourier transformed. In the frequency domain, the received OFDM symbol at the rth receive antenna can be written as:

$$z_r(i) = \sum_{k=1}^{N_t} h_{rk}(i)s_k(i) + n_r(i), \quad r = 1, \cdots, N_r, \tag{6.7}$$

where $s_k(i)$ is a symbol transmitted at the ith subcarrier from the kth transmit antenna and $n_r(i)$ is the additive white Gaussian noise observed at the rth receive antenna. We denote an $NN_r \times 1$ vector of the received signal as $\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_r^T, \ldots, \mathbf{z}_{N_r}^T]^T$, where $\mathbf{z}_r = [z_r(0), \ldots, z_r(N-1)]^T$ is an $N \times 1$ vector; an $NN_r \times 1$ noise vector is given by $\mathbf{n} = [\mathbf{n}_1^T, \ldots, \mathbf{n}_r^T, \ldots, \mathbf{n}_{N_r}^T]^T$, where $\mathbf{n}_r = [n_r(0), \ldots, n_r(N-1)]^T$ is an $N \times 1$ vector. We consider scenarios where the noise samples at different receive antennas are uncorrelated and assume that in the frequency domain, the noise covariance matrix $\mathbf{R}_{n0} = E\{\mathbf{n}_r\mathbf{n}_r^H\}$ is the same for all receive antennas. The $N \times N$ matrix $\mathbf{R}_{n0}$ characterizes correlation of noise samples for a single receive antenna. Here we assume that the noise samples are uncorrelated in frequency domain and $\mathbf{R}_{n0} = \sigma_n^2\mathbf{I}_N$, where σ_n^2 is the noise variance. Then the total noise covariance matrix in frequency domain is given by $\mathbf{R}_n = \mathbf{I}_{N_r} \otimes \mathbf{R}_{n0}$.

6.3   BEM of channel frequency response

A channel frequency response $h(f)$ can be represented by a BEM as

$$h(f) \approx \bar{h}(f) = \sum_{m=1}^{M} a_m B(f, m), \tag{6.8}$$

where $B(f, m)$ are basis functions, a_m are expansion coefficients, and M is the number of basis functions. The BEM allows transforming the nonparametric estimation problem to a parametric one: we need to estimate M unknown expansion coefficients instead of estimating $h(f)$ as a function of f. The difference between $h(f)$ and $\bar{h}(f)$ represents a modeling error, which can be made negligible by choosing M large enough [20,27]; then, we can assume that $h(f) = \bar{h}(f)$. However, depending on an estimation technique, a large M may also result in a high noise error (as opposed to the modeling error) [26]. Moreover, a large M will lead to high complexity of the receiver. Therefore, it is important to choose an appropriate M to guarantee that the receiver provides a high estimation performance and requires a low computational load.

We denote $\mathbf{B}$ as an $N \times M$ matrix $\mathbf{B}$ containing samples of basis functions at subcarrier frequencies. The channel frequency response between the kth transmit antenna and the rth receive antenna is then modeled as

$$\bar{\mathbf{h}}_{rk} = \mathbf{B}\mathbf{a}^{(rk)}, \tag{6.9}$$

where the $M \times 1$ vector $\mathbf{a}^{(rk)} = [a_1^{(rk)}, \ldots, a_M^{(rk)}]^T$ represents the BEM coefficients between the kth transmit antenna and the rth receive antenna, and these coefficients are constant over an OFDM symbol. The matrix $\mathbf{B}$ can be split into two parts as follows. The $N_p \times M$ matrix $\mathbf{B}_p$ contains samples of basis functions at subcarriers occupied by pilot symbols: $[\mathbf{B}_p]_{i,m} = [\mathbf{B}]_{\xi_i,m}$. The $N_d \times M$ matrix $\mathbf{B}_d$ contains samples of basis functions at subcarriers occupied by data symbols: $[\mathbf{B}_d]_{i,m} = [\mathbf{B}]_{f_i,m}$. With these notations, the received signal can be represented as

$$\mathbf{z} = \mathbf{\Psi}\mathbf{a} + \mathbf{n}, \tag{6.10}$$

where

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}^{(1)} \\ \vdots \\ \mathbf{a}^{(r)} \\ \vdots \\ \mathbf{a}^{(N_r)} \end{bmatrix}, \quad \mathbf{a}^{(r)} = \begin{bmatrix} \mathbf{a}^{(r1)} \\ \vdots \\ \mathbf{a}^{(rk)} \\ \vdots \\ \mathbf{a}^{(rN_t)} \end{bmatrix},$$

and $\mathbf{\Psi} = \mathbf{I}_{N_r} \otimes \mathbf{\Psi}_{N_t}$ is an $NN_r \times MN_tN_r$ matrix, $\mathbf{\Psi}_{N_t} = [\mathbf{\Psi}^{(1)}, \ldots, \mathbf{\Psi}^{(k)}, \ldots, \mathbf{\Psi}^{(N_t)}]$ is an $N \times MN_t$ matrix, $\mathbf{\Psi}^{(k)} = \mathbf{S}_k\mathbf{B}$ where the $N \times N$ matrix $\mathbf{S}_k$ is given by

$$\mathbf{S}_k = \mathrm{diag}[s_k(0), \ldots, s_k(i), \ldots, s_k(N-1)]. \tag{6.11}$$

The received signal corresponding to subcarriers occupied by data and pilot symbols are modeled, respectively, as

$$\mathbf{z}_d = \mathbf{\Psi}_d\mathbf{a} + \mathbf{n}_d, \quad \mathbf{z}_p = \mathbf{\Psi}_p\mathbf{a} + \mathbf{n}_p. \tag{6.12}$$

According to (6.12), the $NN_r \times 1$ vector $\mathbf{z}$ is split into a vector of received data symbols: $\mathbf{z}_d = [\mathbf{v}_1^T, \ldots, \mathbf{v}_r^T, \ldots, \mathbf{v}_{N_r}^T]^T$ is an $N_dN_r \times 1$ vector, where $[\mathbf{v}_r]_i = z_r(f_i)$, and an $N_pN_r \times 1$ vector of received pilot symbols $\mathbf{z}_p = [\mathbf{u}_1^T, \ldots, \mathbf{u}_r^T, \ldots, \mathbf{u}_{N_r}^T]^T$, where $[\mathbf{u}_r]_i = z_r(\xi_i)$. Similarly, the noise vector $\mathbf{n}$ is split into an $N_dN_r \times 1$ vector $\mathbf{n}_d = [\mathbf{x}_1^T, \cdots, \mathbf{x}_r^T, \cdots, \mathbf{x}_{N_r}^T]^T$, where $[\mathbf{x}_r^T]_i = n_r(f_i)$ and an $N_pN_r \times 1$ vector $\mathbf{n}_p = [\mathbf{w}_1^T, \cdots, \mathbf{w}_r^T, \cdots, \mathbf{w}_{N_r}^T]^T$, where

$[\mathbf{w}_r^T]_i = n_r(\xi_i)$. Correspondingly, the matrix $\boldsymbol{\Psi}$ is split into a matrix of transmitted data symbols $\boldsymbol{\Psi}_d = \mathbf{I}_{N_r} \otimes \boldsymbol{\Psi}_d^{N_t}$, where

$$\boldsymbol{\Psi}_d^{N_t} = [\boldsymbol{\Psi}_d^{(1)}, \ldots, \boldsymbol{\Psi}_d^{(k)}, \ldots, \boldsymbol{\Psi}_d^{(N_t)}], \quad \boldsymbol{\Psi}_d^{(k)} = \mathbf{S}_d^{(k)} \mathbf{B}_d,$$

$$\mathbf{S}_d^{(k)} = \mathrm{diag}[s_k(0), \ldots, s_k(f_i), \ldots, s_k(N_d - 1)],$$

and a matrix of transmitted pilot symbols $\boldsymbol{\Psi}_p = \mathbf{I}_{N_r} \otimes \boldsymbol{\Psi}_p^{N_t}$, where

$$\boldsymbol{\Psi}_p^{N_t} = [\boldsymbol{\Psi}_p^{(1)}, \ldots, \boldsymbol{\Psi}_p^{(k)}, \ldots, \boldsymbol{\Psi}_p^{(N_t)}], \quad \boldsymbol{\Psi}_p^{(k)} = \mathbf{S}_p^{(k)} \mathbf{B}_p,$$

$$\mathbf{S}_p^{(k)} = \mathrm{diag}[s_k(0), \ldots, s_k(\xi_i), \ldots, s_k(N_p - 1)].$$

For different BEMs, the calculation of the matrix $\mathbf{B}$ and the vector $\mathbf{a}$ are different. In this chapter, we consider the following BEMs: CE BEM, GCE BEM, cubic B-splines, Slepian sequences and KL basis functions.

6.3.1 CE basis functions

The CE BEM is widely used [19, 21–23], but it can result in large modeling errors. For the CE model, elements of the matrix $\mathbf{B}$ are given by [21]

$$[\mathbf{B}]_{n,m} = e^{j2\pi\Delta f n(m - \frac{M}{2})}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{6.13}$$

6.3.2 GCE basis functions

An improved modeling performance is obtained by using the GCE BEM applying a set of complex exponentials with the period longer than the window length related to the CE BEM [25, 70]. For the GCE model, elements of the matrix $\mathbf{B}$ are given by [25, 70]

$$[\mathbf{B}]_{n,m} = e^{j2\pi\frac{\Delta f}{\kappa} n(m - \frac{M}{2})}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N, \tag{6.14}$$

where κ is a real number which is larger than 1; usually, $\kappa = 2$ is used [25].

Note that there is another ways to build the generalized complex exponential functions, i.e.,

$$[\mathbf{B}]_{n,m} = e^{-j2\pi\Delta f(n-1)(m-1)\frac{L_{\max}T}{M-1}}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{6.15}$$

By using (6.15), we assume that the considered channel (6.7) is equivalent to a multipath channel

$$\bar{g}_{rk}(\tau) = \sum_{m=0}^{M-1} \bar{\alpha}_{rk}(m)\delta\left(\tau - \frac{m}{M}L_{\max}T\right),\tag{6.16}$$

and the task of the estimator becomes to estimate M corresponding $\bar{\alpha}_{rk}$. It indicates that the number of basis functions does not depend on N anymore. In this chapter, we just consider the GCE defined in (6.14), and the details of the performance of the modified GCE as (6.15) can be found in our previous publications [41].

6.3.3   Cubic B-splines functions

Cubic B-splines have previously been used for estimating the channel frequency response in the underwater acoustic channel [139]. To build basis functions, we use the cubic B-splines [76]

$$\varphi(f) = \begin{cases} \frac{2}{3} - \frac{f^2}{F^2} + \frac{|f|^3}{2F^3}, & \text{if } |f| < F, \\ \frac{1}{6}\left(2 - \frac{|f|}{F}\right)^3, & \text{if } F \le |f| < 2F, \\ 0, & \text{otherwise,} \end{cases}\tag{6.17}$$

where $F = (N-1)\Delta f/(M-3)$ and the basis functions are given by $\varphi_m(f) = \varphi[f - (m-2)\Delta f]$, where $m = 1,\ldots,M$. Elements of the matrix $\mathbf{B}$ are given by

$$[\mathbf{B}]_{n,m} = \varphi_m(n-1), \quad m = 1,\ldots,M, \quad n = 1,\ldots,N.\tag{6.18}$$

6.3.4   Slepian sequences

Slepian sequences are a set of orthogonal functions which are widely used for channel estimation both in time and frequency domains [20, 30, 67]. Let consider M Slepian sequences $u_m(n)$ with length N bandlimited to the frequency range $[-\frac{1}{2}\tau_{\max}\Delta f, \frac{1}{2}\tau_{\max}\Delta f]$. Such sequences are the eigenvectors of the following equation

$$\sum_{q=1}^{N} \frac{\sin(\pi\tau_{\max}\Delta f(q-n))}{\pi(q-n)}u_m(q) = \lambda_m u_m(n),\tag{6.19}$$

where λ_m is an eigenvalue indicating the fraction of energy contained in the frequency range $[-\frac{1}{2}\tau_{\max}\Delta f, \frac{1}{2}\tau_{\max}\Delta f]$ of the corresponding eigenvector [67]. The eigenvalues are

ordered starting with the maximum one: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$. Therefore, $u_m(n)$ is the mth most concentrated Slepian sequence. According to [29], M should be chosen to provide λ_m close to 1 when $m \leq M$ and close to 0 when $m \geq M$. We intend to use the Slepian sequences over the frequency range $[0, \tau_{\max}\Delta f]$. Then, the basis functions can be represented as

$$[\mathbf{B}]_{n,m} = u_m(n)e^{-j\pi n\tau_{\max}\Delta f}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{6.20}$$

6.3.5 KL BEM

The KL BEM is optimal in terms of the mean square error (MSE) [31, 32], which is a reduced-rank decomposition of channels whose statistical information is known at the receiver side. The KL basis functions $v_m(n)$ are eigenvectors of the fading covariance matrix. Specifically, the covariance matrix of the fading in the frequency domain is $\boldsymbol{\Upsilon}_0$ which is obtained in (6.5). We also order the eigenvalues λ_m of $\boldsymbol{\Upsilon}_0$ as: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$, and assume that when m is larger than a fixed value $M << N$, λ_m decreases rapidly and can be neglected [32]. Then, the KL basis functions can be represented as

$$[\mathbf{B}]_{n,m} = v_m(n), \quad m = 1, \ldots, M, \quad n = 1, \ldots, N. \tag{6.21}$$

6.4 Optimal and mismatched detectors

6.4.1 Optimal detection

We now derive an optimal detector by maximizing the PDF $p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)$ of the received signal $\mathbf{z}_d$ and the received pilot signal $\mathbf{z}_p$, conditioned on the transmitted data symbols $\mathbf{S}_d$:

$$\begin{aligned}
\hat{\mathbf{S}}_{d,\mathrm{opt}} &= \arg \max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}} \{p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)\} \\
&= \arg \max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}} \{\lambda_{\mathrm{opt}}(\mathbf{S}_d)\},
\end{aligned} \tag{6.22}$$

where the metric $\lambda_{\mathrm{opt}}(\mathbf{S}_d)$ is given by

$$\lambda_{\mathrm{opt}}(\mathbf{S}_d) = \ln\left[p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)\right], \tag{6.23}$$

and the alphabet $\mathcal{A}$ includes symbols corresponding to all constellation points. The PDF $p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p)$ can be obtained from the PDF $p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a})$ of the received signal vector $\mathbf{z}_d$ conditioned on the transmitted data $\mathbf{S}_d$ and channel parameters (expansion coefficients) $\mathbf{a}$ by integrating out the channel parameters $\mathbf{a}$ which are treated as *nuisance parameters*:

$$p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{z}_p) = \int p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a}) f(\mathbf{a}|\mathbf{z}_p) d\mathbf{a}, \tag{6.24}$$

where $p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a})$ is given by

$$p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a}) = \frac{1}{\pi^{N_d}\sigma_n^2}\exp\left\{-\frac{\|\mathbf{z}_d - \boldsymbol{\Psi}_d\mathbf{a}\|^2}{\sigma_n^2}\right\}. \tag{6.25}$$

The *posterior* PDF $f(\mathbf{a}|\mathbf{z}_p)$ of the expansion coefficients $\mathbf{a}$ subject to the received pilot signals $\mathbf{z}_p$ is Gaussian with the mean vector $(\boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \sigma_n^2\mathbf{R}_a^{-1})^{-1}\boldsymbol{\Psi}_p^H\mathbf{z}_p$ and covariance matrix $(\boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \sigma_n^2\mathbf{R}_a^{-1})$ [113], where the covariance matrix $\mathbf{R}_a$ is given by $\mathbf{R}_a = E\{\mathbf{a}\mathbf{a}^H\} = \mathbf{R}_r \otimes \mathbf{R}_t \otimes \boldsymbol{\Lambda}_0$, and $M \times M$ matrix $\boldsymbol{\Lambda}_0 = E\left\{(\mathbf{a}^{(rk)})^H\mathbf{a}^{(rk)}\right\}$ is the covariance matrix of the expansion coefficients $\mathbf{a}^{(rk)}$ for the frequency response between the kth transmit antenna and the rth receive antenna. The matrix $\boldsymbol{\Lambda}_0$ can be obtained from the fading covariance matrix $\boldsymbol{\Upsilon}_0$ by requiring that

$$E\{\bar{\mathbf{h}}_{rk}\bar{\mathbf{h}}_{rk}^H\} = E\{\mathbf{h}_{rk}\mathbf{h}_{rk}^H\} = \boldsymbol{\Upsilon}_0. \tag{6.26}$$

This requirement means that the correlation matrix $\boldsymbol{\Lambda}_0$ results in fading correlation $\boldsymbol{\Upsilon}_0$ in the frequency domain. By substituting (6.9) in (6.26), we obtain

$$\mathbf{B}\boldsymbol{\Lambda}_0\mathbf{B}^H = \boldsymbol{\Upsilon}_0. \tag{6.27}$$

Multiplying both sides of (6.27) by $\boldsymbol{\Omega} = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ from the left and by $\boldsymbol{\Omega}^H$ from the right, we arrive at

$$\boldsymbol{\Lambda}_0 = \boldsymbol{\Omega}\boldsymbol{\Upsilon}_0\boldsymbol{\Omega}^H. \tag{6.28}$$

The optimal detector finds a data matrix $\mathbf{S}_d$ that maximizes the metric (6.23) which is obtained by substituting the Gaussian PDF $f(\mathbf{a}|\mathbf{z}_p)$ in (6.24) and then the result of integration - in (6.23). After some algebra, we arrive at

$$
\begin{aligned}
\hat{\mathbf{S}}_{d,\text{opt}} &= \arg\max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}}\left\{\ln\left[\int p(\mathbf{z}_d|\mathbf{S}_d, \mathbf{a}) f(\mathbf{a}|\mathbf{z}_p)d\mathbf{a}\right]\right\}\\
&= \arg\max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}}\left\{\ln\left[\int e^{\sigma_n^{-2}\left[2\Re(\mathbf{a}^H\boldsymbol{\Psi}_d^H\mathbf{z}_d) - \mathbf{a}^H\boldsymbol{\Psi}_d^H\boldsymbol{\Psi}_d\mathbf{a}\right]} f(\mathbf{a}|\mathbf{z}_p)d\mathbf{a}\right]\right\}\\
&= \arg\max_{\mathbf{S}_d \in \mathcal{A}^{N_d N_t}}\left\{\sigma_n^{-2}\left(\boldsymbol{\Psi}_d^H\mathbf{z}_d + \boldsymbol{\Psi}_p^H\mathbf{z}_p\right)^H\left(\boldsymbol{\Psi}_d^H\boldsymbol{\Psi}_d + \boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \sigma_n^2\mathbf{R}_{\mathbf{a}}^{-1}\right)^{-1}\right.\\
&\quad\times\left.\left(\boldsymbol{\Psi}_d^H\mathbf{z}_d + \boldsymbol{\Psi}_p^H\mathbf{z}_p\right) - \ln\left|\boldsymbol{\Psi}_d^H\boldsymbol{\Psi}_d + \boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \sigma_n^2\mathbf{R}_{\mathbf{a}}^{-1}\right|\right\}, \tag{6.29}
\end{aligned}
$$

However, it is unfeasible to solve this optimization problem for a high N_d due to extremely high computational complexity. If QAM symbols with 2^K constellation points are transmitted, we have to calculate the metric $2^{KN_dN_t}$ times. In order to reduce the computational load, we separate the data into groups of symbols, each containing G data symbols. We repeat the detection N_d/G times to recover all the symbols. In this case, the optimal metric should be calculated $\frac{N_d}{G}2^{KN_tG}$ times, which can be significantly smaller compared with $2^{KN_dN_t}$.

As the simplest case, we consider $G = 1$, which indicates that the optimal symbol-by-symbol detection scheme is applied, and we only detect N_t data symbols at once. In this case, the expressions above are simplified: $\mathbf{z}_d = [z_1(f_i), \ldots, z_r(f_i), \ldots, z_{N_r}(f_i)]^T$, $\boldsymbol{\Psi}_d^{(k)} = s_k(f_i)\mathbf{B}_d$, and $\mathbf{B}_d$ becomes a $1 \times M$ vector corresponding to samples of basis functions at the ith subcarrier frequency. Although the computational load is reduced, the detection performance will be degraded. Therefore, we also consider the optimal detection with $1 < G \ll N_d$ to trade off the complexity and detection performance. In this case, expressions above are modified: $\mathbf{z}_d = [\mathbf{v}_1^T, \ldots, \mathbf{v}_r^T, \ldots, \mathbf{v}_{N_r}^T]^T$ and $\mathbf{v}_r = [z_r(f_i), \ldots, z_r(f_{i+G-1})]^T$; $\mathbf{B}_d$ becomes an $G \times M$ matrix corresponding to samples of basis functions at the f_i-th to $(f_i + G - 1)$-th subcarriers. Investigating the improvement of BER performance due to using the optimal detector compared with traditional mismatched detectors is the main target of this chapter, and methods to analyze the complexity and reduce the computational load of this optimal detector will be discussed in further works.

6.4.2   Mismatched detection

Now, we consider mismatched detectors applying three different channel estimators: ML, ϵ-ML or MMSE estimators. The ML channel estimate is given by

$$\hat{\mathbf{a}}_{\mathrm{ML}} = \left(\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p\right)^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p. \tag{6.30}$$

The performance of the ML channel estimator is significantly degraded in noisy scenarios. A better performance is obtained when using regularized ML channel estimation based on the diagonal loading:

$$\hat{\mathbf{a}}_{\epsilon} = \left(\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \epsilon\sigma_n^2 \mathbf{I}_{MN_tN_r}\right)^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p, \tag{6.31}$$

where ϵ is a positive regularization parameter and in our simulation we use $\epsilon = 1$. We also consider the MMSE channel estimation that takes the statistical information of the channel fading into account,

$$\hat{\mathbf{a}}_{\text{MMSE}} = (\boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_a^{-1})^{-1} \boldsymbol{\Psi}_p^H \mathbf{z}_p. \tag{6.32}$$

A mismatched detector uses the minimum distance detector that treats the channel estimates as perfect channel information and decides on the transmitted data symbols by minimizing the Euclidean distance

$$
\begin{aligned}
\hat{\mathbf{S}}_{d,\text{mis}} &= \arg \min_{\mathbf{S}_d \in \mathcal{A}^{N_t}} \left\{ \lambda_{\text{mis}}(\mathbf{S}_d) \right\}, \\
&= \arg \min_{\mathbf{S}_d \in \mathcal{A}^{N_t}} \left\{ \frac{\|\mathbf{z}_d - \boldsymbol{\Psi}_d \hat{\mathbf{a}}\|^2}{\sigma_n^2} \right\},
\end{aligned}
\tag{6.33}
$$

where $\hat{\mathbf{a}} = \hat{\mathbf{a}}_{\text{ML}}$ for ML channel estimates (6.30), or $\hat{\mathbf{a}} = \hat{\mathbf{a}}_\epsilon$ for ϵ-ML channel estimates (6.31) or $\hat{\mathbf{a}} = \hat{\mathbf{a}}_{\text{MMSE}}$ for MMSE channel estimates (6.32).

6.4.3 Complexity analysis

Now, we analyze the complexity of the optimal detector with symbol-by-symbol detection scheme ($G = 1$) for SISO systems. In this case, quantities in (6.29) are significantly simplified: $\mathbf{S}_d = d$, where d is a data symbol; $\mathbf{z}_d = z_1(f_i)$; $\boldsymbol{\Psi}_d = d\mathbf{B}_d$, where $\mathbf{B}_d$ becomes a $1 \times M$ vector. Accordingly, we can simplify (6.29) as

$$
\begin{aligned}
\hat{d}_{\text{opt}} &= \arg \max_{d \in \mathcal{A}} \left\{ \sigma_n^{-2} \left(d^* z_d \mathbf{B}_d^H + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right)^H \left(|d|^2 \mathbf{B}_d^H \mathbf{B}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_{\mathbf{a}}^{-1} \right)^{-1} \right. \\
&\quad \left. \times \left(d^* z_d \mathbf{B}_d^H + \boldsymbol{\Psi}_p^H \mathbf{z}_p \right) - \ln \left| |d|^2 \mathbf{B}_d^H \mathbf{B}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_{\mathbf{a}}^{-1} \right| \right\} \\
&= \arg \max_{d \in \mathcal{A}} \left\{ \sigma_n^{-2} |d|^2 |z_d|^2 \mathbf{B}_d \mathbf{Y} \mathbf{B}_d^H + 2\sigma_n^{-2} \Re\{ d z_d^* \mathbf{B}_d \mathbf{Y} \boldsymbol{\Psi}_p \mathbf{z}_p \} \right. \\
&\quad \left. + \sigma_n^{-2} \mathbf{z}_p^H \boldsymbol{\Psi}_p \mathbf{Y} \boldsymbol{\Psi}_p^H \mathbf{z}_p + \ln |\mathbf{Y}| \right\},
\end{aligned}
\tag{6.34}
$$

where $\mathbf{Y} = \left(|d|^2 \mathbf{B}_d^H \mathbf{B}_d + \boldsymbol{\Psi}_p^H \boldsymbol{\Psi}_p + \sigma_n^2 \mathbf{R}_{\mathbf{a}}^{-1} \right)^{-1}$ is an $M \times M$ matrix and $(\cdot)^*$ denotes complex conjugate. If we assume that (i) the same pilot symbols are inserted into every OFDM symbol ($\boldsymbol{\Psi}_p$ is the same for all OFDM symbols); (ii) the covariance matrix of fading $\mathbf{R}_a$ and σ_n^2 are constant over N_c OFDM symbols and known at the receiver side, the matrix $\mathbf{Y}$ will depend only on the possible magnitudes of data symbols $|d|$, pilot

symbols and BEM samples $\mathbf{B}_d$ and $\mathbf{B}_p$. The pilot symbols and all possible values of $|d|^2$ (we denote this number as D) are known to the receiver. Moreover, $\mathbf{B}_d$ and $\mathbf{B}_p$ are also available, since positions of data and pilot symbols are fixed. Therefore, we can precompute scalars $\ln|\mathbf{Y}|$ for all possible $|d|$. This precomputation requires $\mathcal{O}(M^3 N_d D)$ complex multiplications. E.g., for 16QAM modulation, the number of possible $|d|$ is $D = 4$; thus the number of scalars $\ln|\mathbf{Y}|$ is $4N_d$. The scalar $\mathbf{B}_d \mathbf{Y} \mathbf{B}_d^H$ in the first term of (6.34), $1 \times N_p$ vector $\mathbf{B}_d \mathbf{Y} \mathbf{\Psi}_p$ in the second term of (6.34) and $N_p \times N_p$ matrix $\mathbf{\Psi}_p \mathbf{Y} \mathbf{\Psi}_p^H$ in the third term of (6.34) can also be precomputed for all D possible $|d|$. In total, there are $(N_p^2 + N_p + N_d)D$ complex numbers to be kept in memory. To precompute these numbers requires $\mathcal{O}(M^3 N_d D)$ complex multiplications. These complex numbers can be used for detecting all N_c OFDM symbols. Therefore, the average number of complex multiplications for all the precomputation is $\mathcal{O}\left(\frac{M^3 D + N_p M^2 + N_p^2 M}{N_c}\right)$ per one data symbol.

We now denote: $\rho = \mathbf{B}_d \mathbf{Y} \mathbf{B}_d^H$, $\mathbf{w} = \mathbf{B}_d \mathbf{Y} \mathbf{\Psi}_p$, and $\mathbf{\Xi} = \mathbf{\Psi}_p \mathbf{Y} \mathbf{\Psi}_p^H$. Then, the equation (6.34) is transformed to

$$\hat{d}_{\text{opt}} \quad = \quad \arg\max_{d \in \mathcal{A}} \left\{\sigma_n^{-2}|d|^2|z_d|^2\rho + 2\sigma_n^{-2}\Re\{dz_d^* \mathbf{w}\mathbf{z}_p\} + \sigma_n^{-2}\mathbf{z}_p^H \mathbf{\Xi} \mathbf{z}_p + \ln|\mathbf{Y}|\right\} \quad (6.35)$$

For each OFDM symbol, in (6.35), the scalars $\mathbf{w}\mathbf{z}_p$ and $\mathbf{z}_p \mathbf{\Xi} \mathbf{z}_p$ can be computed once and reused for detecting all data symbols. Thus, only $|d|^2|z_d|^2$ in the first term and dz_d^* in the second term require to be computed symbol-by-symbol. Therefore, with the precomputation as explained above, the average number of complex multiplications required to detect one data symbol is $\frac{N_p^2 + 2N_p + 4}{N_d} + 2^K + D$.

Now, we analyze the complexity of a mismatched detector treating the MMSE channel estimates as perfect. For a fixed σ_n^2, we can precompute the $M \times N_p$ matrix $(\mathbf{\Psi}_p^H \mathbf{\Psi}_p + \sigma_n^2 \mathbf{R}_a^{-1})^{-1}\mathbf{\Psi}_p^H$ in (6.32) and keep it in memory; the number of complex multiplications required for this precomputation is $\mathcal{O}(M^3 + N_p M^2)$. This matrix can also be calculated once and reused to detect all N_c OFDM symbols as well, for channels with fixed $\mathbf{R}_a$ and σ_n^2. Therefore, the number of complex multiplications for this precomputation is $\mathcal{O}\left(\frac{M^3 + N_p M^2}{N_c N_d}\right)$ per one data symbol. The number of complex multiplications required by a minimum distance detector (6.33) is $M2^K$. Table. 1 shows the complexity of the optimal and mismatched detectors.

It is seen that when N_c is large, which indicates that the channel statistics change slowly ($\mathbf{R}_a$ and σ_n^2 keep constant for numerous OFDM symbols), to detect one data sym-

Algorithms	Optimal detector	Mismatched detector
Number of complex numbers saved in memory	$(N_p^2 + N_p + N_d)D$	$N_p M$
Number of complex multiplications required for the precomputation per a data symbol	$\mathcal{O}\left(\frac{M^3 D + N_p M^2 + N_p^2 M}{N_c}\right)$	$\mathcal{O}\left(\frac{M^3 + N_p M^2}{N_c N_d}\right)$
Number of complex multiplications required for detecting a data symbol	$\frac{N_p^2 + 2N_p + 4}{N_d} + 2^K + D$	$M 2^K$

Table 6.1: Complexity of the optimal detector and the mismatched detector with MMSE channel estimates.

bol, the number of complex multiplications required by the optimal detector is comparable to that required by the mismatched detector. However, if the channel statistics change faster (N_c is small), all precomputed matrices, vectors and scalars have to be updated frequently and the complexity of the optimal detector significantly increases compared with the mismatched detector. For example, if the channel statistics change between two neighboring OFDM symbols ($N_c = 1$), for the optimal detector, the number of complex multiplications required for the precomputation becomes $\mathcal{O}(M^3 D)$, which is much larger than the number required by the mismatched detector, $\mathcal{O}\left(\frac{M^3 + N_p M^2}{N_d}\right)$, for a large N_d.

6.5 Iterative receiver

We are also interested in the performance of an iterative receiver incorporating the optimal detector for coded transmission in SISO frequency-selective channels. The structures of the transmitter and iterative receiver are shown in Fig. 6.2 and Fig. 6.3, respectively. If the optimal detection is used, the channel estimator and detector in Fig. 6.3 are replaced by the optimal detector. In the receiver, the CP is removed and the received signal is Fourier transformed before the first iteration. Channel estimators use the vectors $\mathbf{z}_p$ and $\mathbf{s}_p$ to estimate the channel frequency response in the first iteration and the vector $\mathbf{z}$ and recovered OFDM symbol $\hat{\mathbf{S}}_k$ in subsequential iterations. The channel estimates are used in the detector to calculate the soft metric λ_{c_k} for coded bits. For every bit $c_k = \pm 1$ of a
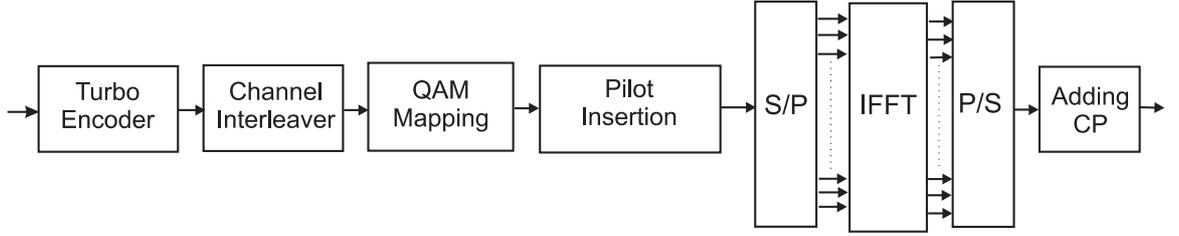
Figure 6.2: Block-diagram of the transmitter with turbo encoder and channel interleaver for SISO channels.
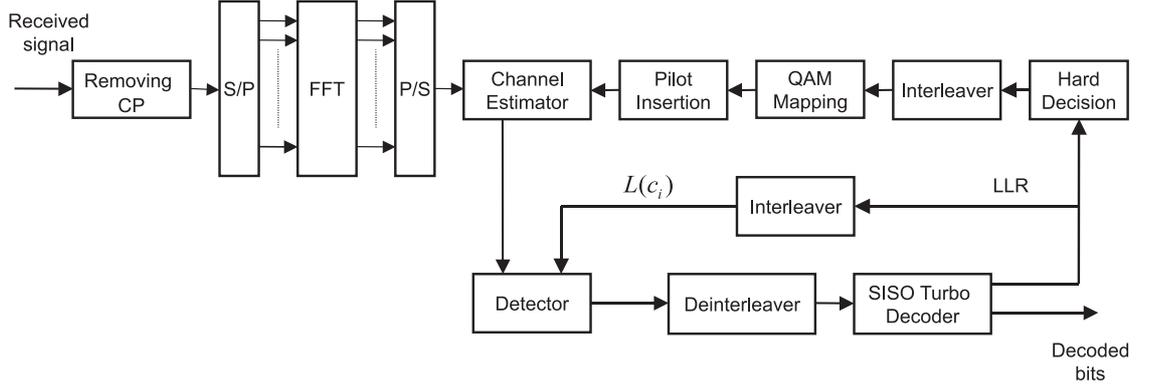
Figure 6.3: Block-diagram of the iterative receiver for SISO channels.

received symbol, $k = 1, \ldots, K$, the *a posteriori* log-likelihood ratio (LLR) is computed as

$$\lambda_{c_k} = \ln \left[\frac{\sum_{d \in \mathcal{A}_k^+} e^{-\lambda(d)} \prod_{i \neq k} P(c_i)}{\sum_{d \in \mathcal{A}_k^-} e^{-\lambda(d)} \prod_{i \neq k} P(c_i)} \right], \tag{6.36}$$

where the *a priori* probability $P(c_i)$ of a symbol bit is expressed in terms of its *a priori* LLR $L(c_i)$ [118]:

$$P(c_i) = \frac{1}{2} \left[1 + c_i \tanh \left(\frac{1}{2} L(c_i) \right) \right],$$

$\mathcal{A}_k^\pm = \{ d \in \mathcal{A} | c_k = \pm 1 \}$, and the metric $\lambda(d)$ depends on the detector used. For the first iteration, we have

$$\lambda_{c_k} = \ln \sum_{d \in \mathcal{A}_k^+} e^{-\lambda(d)} - \ln \sum_{d \in \mathcal{A}_k^-} e^{-\lambda(d)}. \tag{6.37}$$

After being de-interleaved, LLRs λ_{c_k} are decoded by a soft-input soft-output turbo decoder. The hard decisions of the LLRs of decoded bits are interleaved and mapped to the QAM constellation to rebuild the data symbols. The pilot symbols are inserted to recover the OFDM symbol in the frequency domain. The recovered OFDM symbol is feedback to the channel estimator or to the optimal detector. The channel estimates,

LLRs of coded bits and LLRs of decoded bits are refined once per iteration by treating
all recovered data symbols as pilot symbols. Since the channel estimator and detector
applied at the first iteration and those applied in the following iterations can be different,
the schemes used at the receiver are correspondingly modified. Depending on the detector
used and whether it is the first or a subsequent iteration, four different iterative receivers
are considered:

1. *ML-ML* receiver: The ML channel estimator is used in all iterations. At the first
iteration, the ML estimator (6.30) is used to estimate the channel frequency response
based on transmitted pilot symbols. In the following (three) iterations, the number of input
pilot symbols used to obtain the frequency response estimates $\hat{\mathbf{h}}_{\mathrm{ML}} = \mathbf{B}\hat{\mathbf{a}}_{\mathrm{ML}}$ is extended
from N_p to N, all recovered symbols are used as pilot symbols to refine the channel
estimation and signal detection, and consequently, $\mathbf{\Psi}_p$, $\mathbf{B}_p$, $\mathbf{z}_p$ in (6.30) are replaced by
$\mathbf{\Psi}$, $\mathbf{B}$ and $\mathbf{z}$, respectively.

2. ϵ-*ML*-ϵ-*ML* receiver: This receiver is similar to the *ML-ML* receiver with replace-
ment $\hat{\mathbf{a}}_{\mathrm{ML}}$ by $\hat{\mathbf{a}}_\epsilon$ according to (6.31).

3. *MMSE-MMSE* receiver: The receiver is similar to the *ML-ML* receiver with re-
placement $\hat{\mathbf{a}}_{\mathrm{ML}}$ by $\hat{\mathbf{a}}_{\mathrm{MMSE}}$ according to (6.32).

4. *Optimal-MMSE* receiver: The optimal detector is used at the first iteration; MMSE
channel estimation and minimum distance detector are used in the following three itera-
tions.

6.6 Simulation Results

In this section, numerical results obtained by simulation are presented. We consider
frequency-selective fading channels with $L = 6$ paths, and set $L_{\mathrm{max}} = 10$, $P_p = N_t$,
$N = 461$, $P = 20$ and $M = 23$. Firstly, we compare the performance of the MMSE
channel estimators corresponding to different BEMs for SISO multipath channels with
different τ_{rms}. Then, the performance of the optimal detector with group symbol detection
$(G > 1)$ in SISO channels is also considered. As the improvement of the BER perfor-

mance is not significant even if $G = 7$, in the following simulation scenarios, we only consider $G = 1$. We then compare the BER performance of four iterative receivers for turbo coded transmission in SISO multipath channels. Finally, we consider the transmission with 16QAM and BPSK modulation in MIMO channels and concentrate on the comparison between the BER performance of the optimal detector and that of the mismatched detector with MMSE channel estimates for uncoded transmission. The average SNR is given by

$$\zeta = \frac{N_t E_s \sigma_h^2}{\sigma_n^2}, \tag{6.38}$$

where E_s is the average energy of data symbols and σ_h^2 is the variance of channel frequency response between a pair of transmit and receive antennas. The average bit energy to noise ratio is defined as $E_b/N_0 = \zeta/(N_t K)$. Simulation results below represent the MSE or BER versus E_b/N_0.

Firstly, we compare the MSE performance of the mismatched detector with MMSE channel estimates based on different BEMs. The MSE in one simulation trial is calculated as

$$\text{MSE} = \frac{\sum_{r=1}^{N_r} \sum_{k=1}^{N_t} \sum_{i=0}^{N-1} |h_{r,k}(i) - \hat{h}_{r,k}(i)|^2}{\sum_{r=1}^{N_r} \sum_{k=1}^{N_t} \sum_{i=0}^{N-1} |h_{r,k}(i)|^2} . \tag{6.39}$$

Then these MSEs are averaged over all simulation trials. Fig. 6.4 shows the MSE performance of the MMSE channel estimators in SISO channels with $\tau_{\text{rms}} = 5T$. The CE BEM shows a poor performance compared to the other BEMs, while the KL BEM provides the best MSE performance. The other BEMs show the performance similar to that of the KL BEM. We have also considered a channel with $\tau_{\text{rms}} = T$; the simulation results show a similar performance. However, both the Slepian and KL BEMs require the knowledge of statistical characteristics of the fading, which is not practical for ML and ϵ-ML channel estimation. The cubic B-splines provide performance close to that of the KL BEM in most scenarios. However, as B-splines have a finite support that is significantly smaller than N, as seen from (6.17) and (6.18), the complexity of estimators using B-splines can be significantly lower than that of the other BEMs. Moreover, the complexity of the minimum distance detector applying B-splines channel estimates is lower than that of the minimum distance detector applying other BEMs channel estimates. Based on (6.33), in order to detect all N_d data symbols, the number of complex multiplications required by the minimum distance detector applying other BEMs channel estimates is $\mathcal{O}(MN_d 2^K)$, while the number of required complex multiplications for cubic B-splines is only $\mathcal{O}(4N_d 2^K)$, as
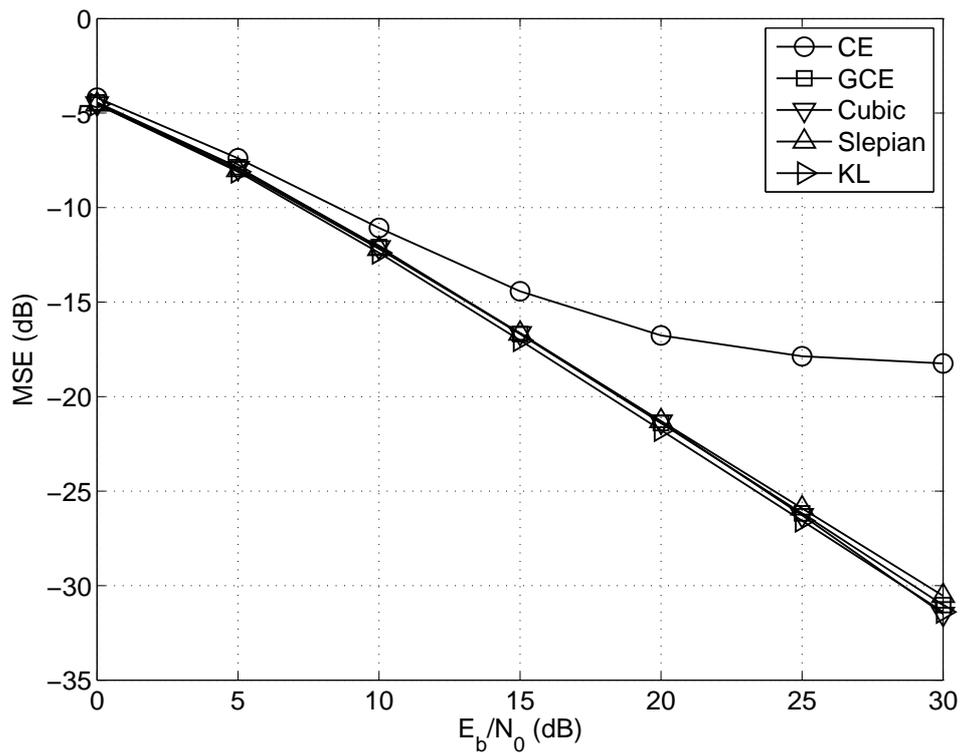
Figure 6.4: MSE performance of MMSE channel estimators with different BEMs for BPSK signals in SISO channels, $L = 6$, $L_{\max} = 10$, $N = 461$, $P = 20$, $P_p = 1$, $M = 23$, $\tau_{\mathrm{rms}} = 5T$.

seen from (6.17). For these reasons, we use B-splines as basis functions in the following simulations.

Simulation results for transmission of BPSK signals in SISO channels with different G are shown in Fig. 6.5. For BER$= 10^{-3}$, the optimal detector with $G = 1$ is inferior to the minimum distance detector with perfect channel information by 1.7 dB. For $G = 3$, this is reduced to 1.4 dB, and for $G = 7$ it is reduced to 1 dB. Thus, the performance of the optimal detector is slightly improved if the number of symbols in detection groups increases. However, compared with the symbol-by-symbol detection ($G = 1$), the complexity is also significantly increased. In the following simulation, we will only consider the case $G = 1$.

We investigate the performance of the optimal and mismatched detectors for coded transmission of 16QAM symbols ($K = 4$) in SISO channels. The iterative receivers and
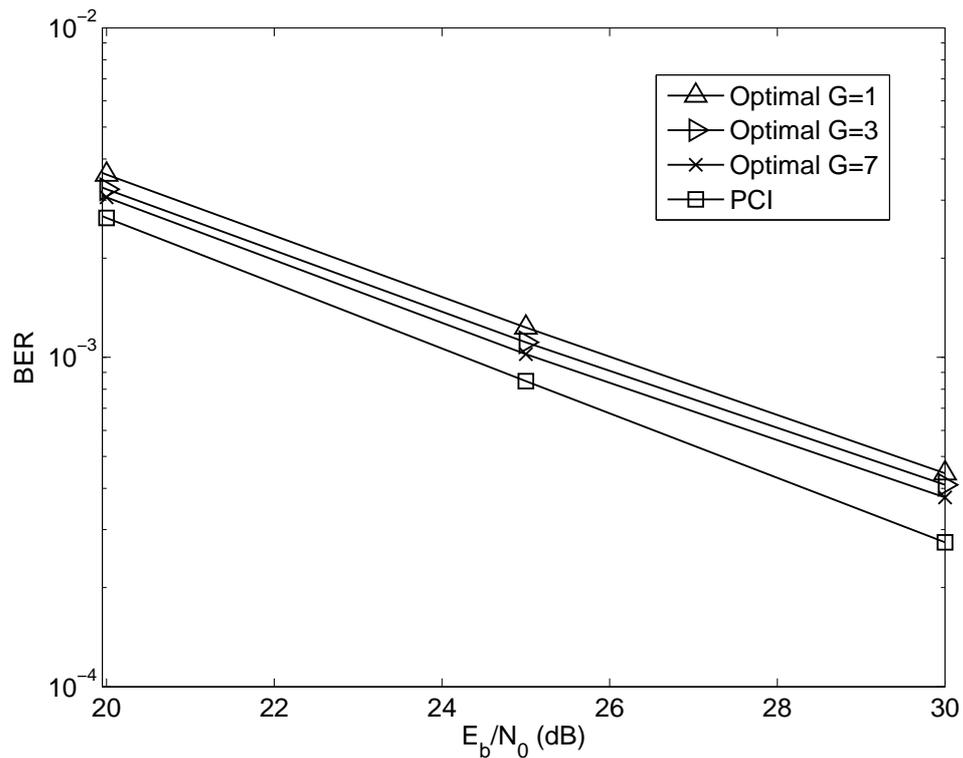
Figure 6.5: BER performance of the optimal detector against G for the transmission of
BPSK signals in SISO channels; $L = 6$, $L_{\max} = 10$, $\tau_{\mathrm{rms}} = 5T$, $N = 461$, $P = 20$,
$P_p = 1$, $M = 23$.

rate-$\frac{1}{3}$ turbo code with generating polynomial $[013, 015]$ are used to improve the BER and
MSE performance. Fig. 6.6 shows the BER performance of the iterative receivers after
4 iterations. The receiver using the optimal detection at the first iteration significantly
outperforms receivers using mismatched detectors with ML and ϵ-ML channel estimates.
At BER $= 10^{-3}$, the improvement in the detection performance is 6.3 dB against the
ML-ML receiver and 2.0 dB against the ϵ-*ML*-ϵ-*ML* receiver. Both the *MMSE-MMSE*
receiver and the *Optimal-MMSE* receiver provide the BER performance close to the case
of perfect channel information; the gap between the BER curves of the *Optimal-MMSE*
receiver and that of the *MMSE-MMSE* receiver is approximately 0.6 dB at BER=10^{-3}.

We now compare the performance of the optimal detector for uncoded symbols with
that of mismatched detectors in MIMO channels. Fig. 6.7 shows the BER performance of
the detectors for 16QAM signals in MIMO channels. Fig. 6.7(a) shows the BER perfor-
mance in SISO channels $(N_r = N_t = 1)$ and in 1×2 SIMO channels $(N_t = 1, N_r = 2)$.
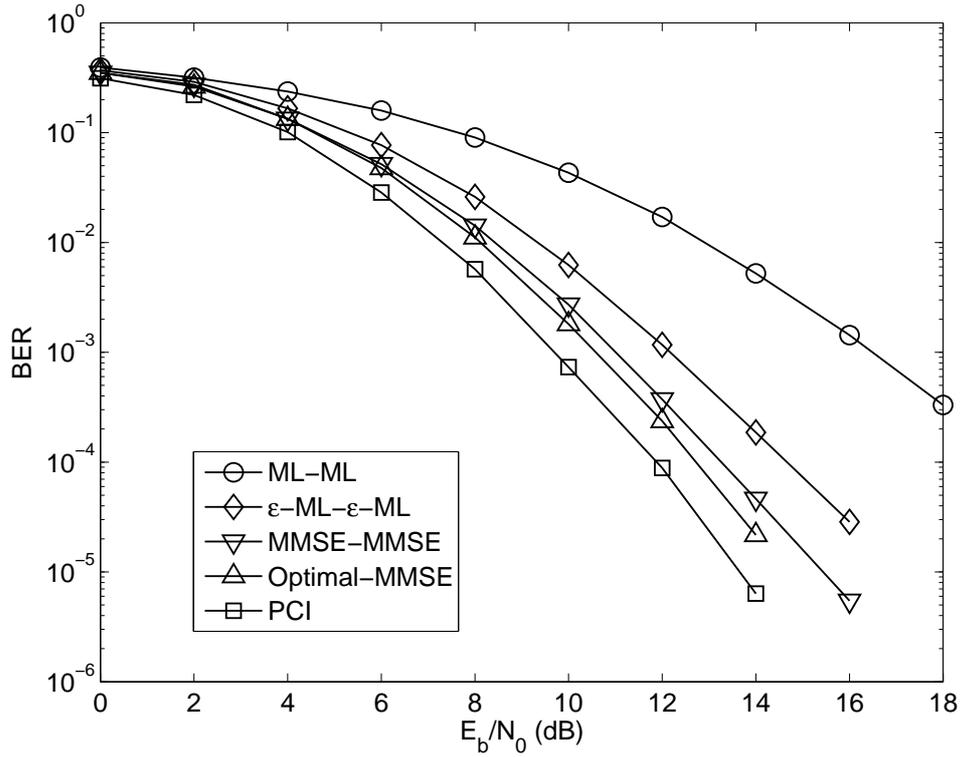
Figure 6.6: BER performance of iterative receivers applying optimal and/or mismatched detection for 16QAM signals in SISO channels, rate 1/3 turbo code, 4 iterations; $L = 6$, $L_{\max} = 10$, $\tau_{\mathrm{rms}} = 5T$, $N = 461$, $P = 20$, $P_p = 1$, $M = 23$.

We set $P_p = 1$ and $N_p = 24$ in both these cases. In 1×1 SISO channels, at BER=10^{-2}, the optimal detector outperforms the mismatched detector with ML channel estimates by 13 dB and with ϵ-ML channel estimates by 1.2 dB. However, the BER performance of the optimal detector and the mismatched detector with MMSE channel estimates are similar. These detectors are inferior to the minimum distance detector with perfect channel information (PCI) by 2 dB. In 1×2 SIMO channels, the mismatched detector with ML channel estimates is inferior to other detectors significantly. For BER=10^{-2}, the improvement due to the use of the optimal detector increases up to 18 dB compared with the mismatched detector with ML channel estimates and for BER=10^{-3}, the improvement is up to 2 dB compared with the mismatched detector using the ϵ-ML channel estimates. The BER performance of the optimal detector is close to that of the mismatched detector with MMSE channel estimates; the difference in the performance is approximately 0.2 dB. The gap between the BER curve of the optimal detector and that of the minimum distance detector with PCI is 2 dB at BER=10^{-3}.

Fig. 6.7(b) shows simulation results for 2×2 MIMO channels ($N_t = N_r = 2$); here, we set $P_p = 2$ and $N_p = 47$. It can be seen that using the mismatched detector with ML channel estimates does not allow achieving good performance. The optimal detector outperforms the mismatched detector with ϵ-ML channel estimates by 4.3 dB at BER=10^{-2}; it also outperforms the mismatched detector with MMSE channel estimates by 4 dB at BER=10^{-3}. From Fig. 6.7(a) and Fig. 6.7(b), it is seen that the improvement is increased when the number of antennas increases.

Fig. 6.8 shows the BER performance of the optimal and mismatched detectors for MIMO systems with BPSK signals. As seen in Fig. 6.8(a), in a SISO channel, for BER=10^{-2}, the optimal detector outperforms the mismatched detector with ML channel estimates by 13.5 dB and the one with ϵ-ML channel estimates by 0.7 dB. Similar to the case of 16QAM signals in Fig. 6.7(a), the BER curves for the optimal detector and the mismatched detector with MMSE channel estimates are close. In a 1×2 SIMO channel, for BER=10^{-2}, the benefit due to the use of the optimal detector is 15.6 dB compared with the mismatched detector with ML channel estimates. Compared with the mismatched detector with ϵ-ML channel estimates, this benefit is 1.1 dB at BER=10^{-4}. The performance of the optimal detector is the same as that of the mismatched detector with MMSE channel estimates, and it is inferior to the minimum distance detector with PCI by 1.6 dB when BER= 10^{-4}.

Fig. 6.8(b) shows simulation results for larger MIMO systems. In 2×2 MIMO channels, the optimal detector provides significantly better performance than the mismatched detectors with ML and ϵ-ML channel estimates and it outperforms the mismatched detector with MMSE channel estimates by 3 dB at BER=10^{-4}. In 2×4 MIMO channels, the optimal detector significantly outperforms all the mismatched detectors, e.g., it outperforms the mismatched detector with MMSE channel estimates by 6.5 dB at BER=2×10^{-6}. Thus, as the number of antennas in a MIMO system increases, the benefit of using the optimal detector becomes more significant.

6.7 Conclusions

We have proposed and investigated an optimal detector for OFDM signals with PSAM in spatially uncorrelated MIMO frequency-selective Rayleigh fading channels. The optimal detector does not estimate the channel explicitly but jointly processes the received data and pilot symbols to recover the data. After comparing the performance of the MMSE estimators of the channel frequency response with different BEMs for SISO channels and choosing B-splines, which provide good performance and require low computational load, we investigated the optimal detector and compared its performance with that of traditional mismatched detectors with ML, regularized ML or MMSE channel estimates for 16QAM and BPSK modulation for uncoded OFDM symbols. Among these mismatched detectors, the one exploiting MMSE channel estimates provides the best performance and in SISO channels its performance is close to that of the optimal detector. We have also investigated the detection performance of iterative receivers that exchange soft information between a turbo decoder and the optimal or mismatched detectors in SISO channels. The simulation results show that in SISO channels, although the iterative receiver using the optimal detector at the first iteration outperforms iterative receivers applying mismatched detectors in all iterations, the improvement of the detection performance is not significant compared with the iterative receiver using the mismatched detector with MMSE channel estimates in all iterations. However, for transmission with larger number of antennas, the optimal detector significantly outperforms the mismatched detectors in spatially uncorrelated MIMO frequency selective fading channels. We can expect that the improvement will become more significant for coded transmission with iterative receivers in MIMO channels, and we will consider this in our further work.
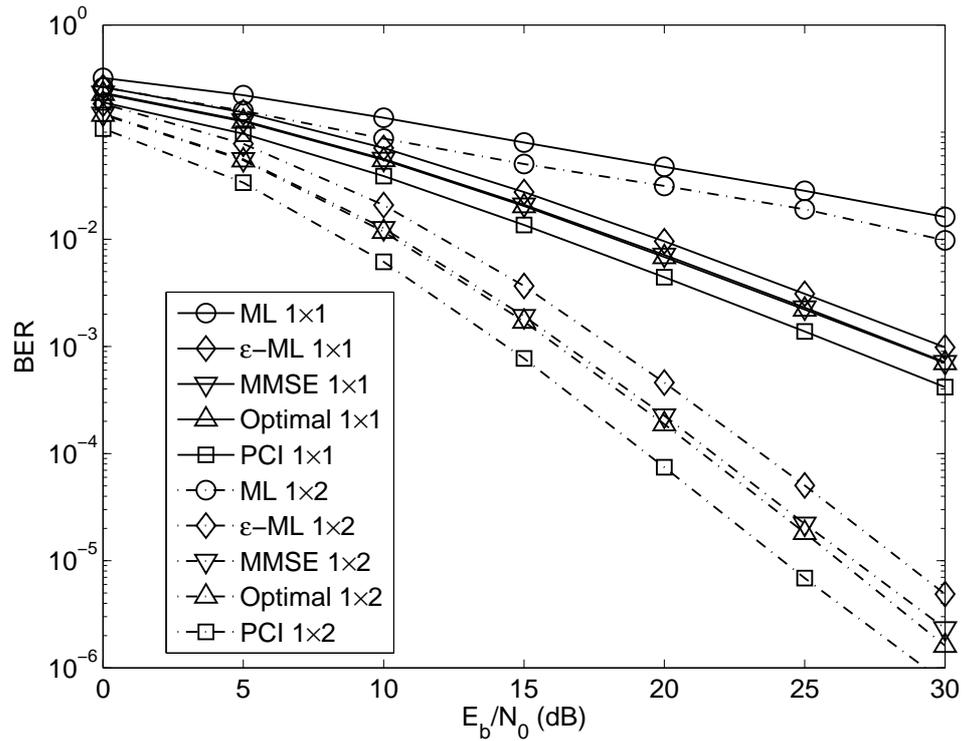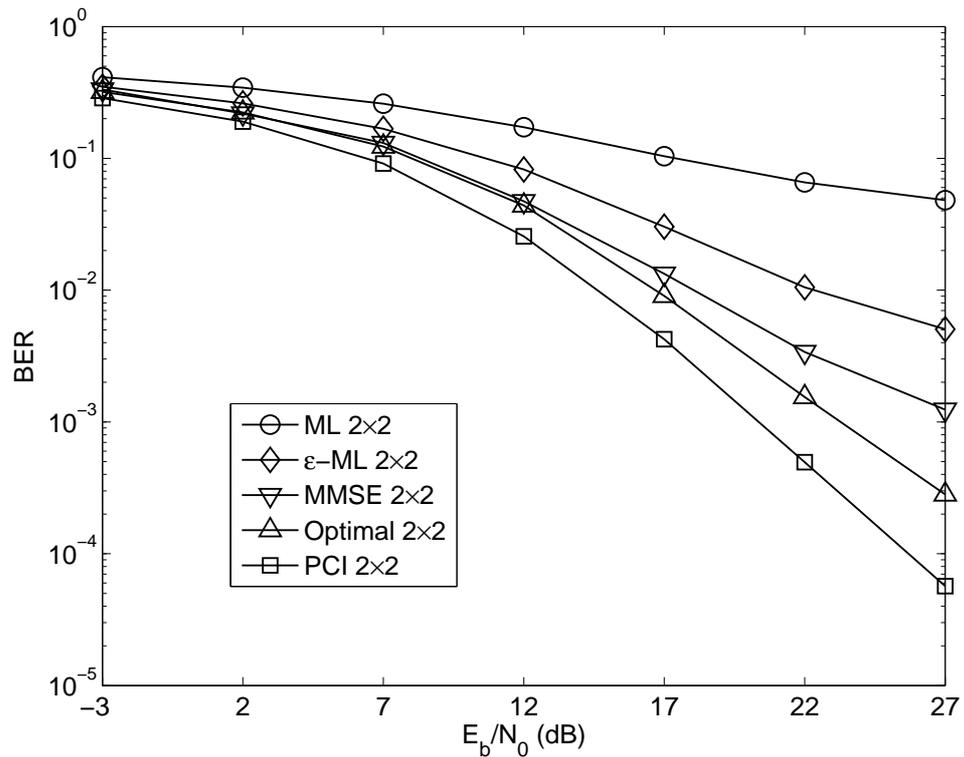
Figure 6.7: BER performance of the optimal and mismatched detectors for 16QAM signals in MIMO channels, $L = 6$, $L_{max} = 10$, $\tau_{rms} = 5T$, $N = 461$, $P = 20$, $P_p = N_t$, $M = 23$; (a) 1×1 and 1×2 MIMO channels, and (b) 2×2 MIMO channels.
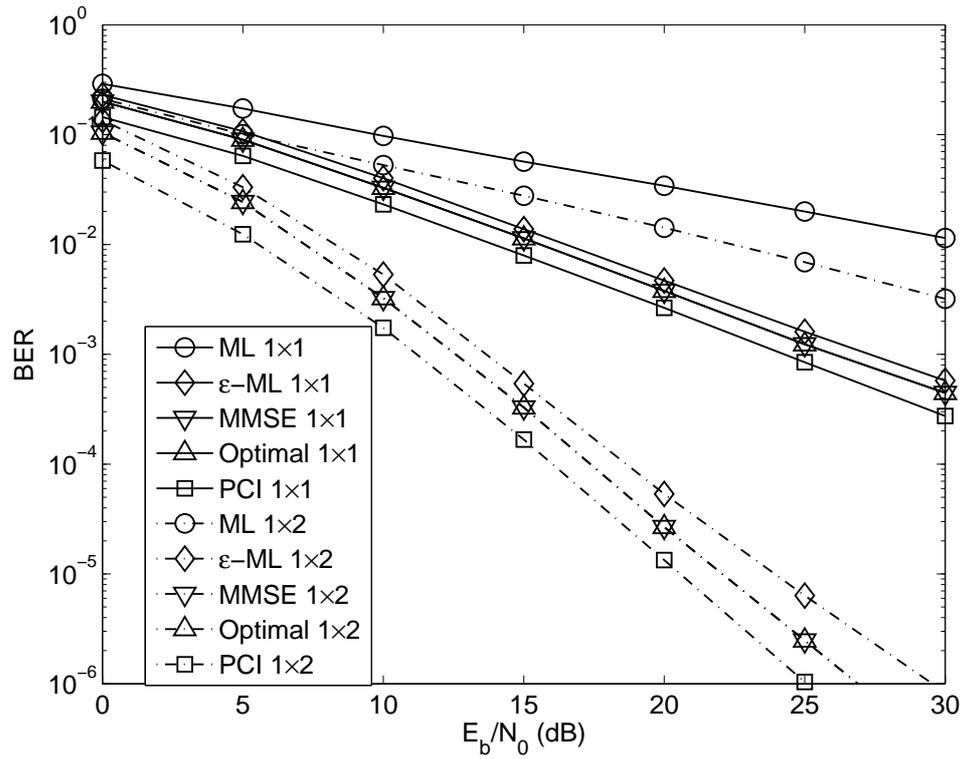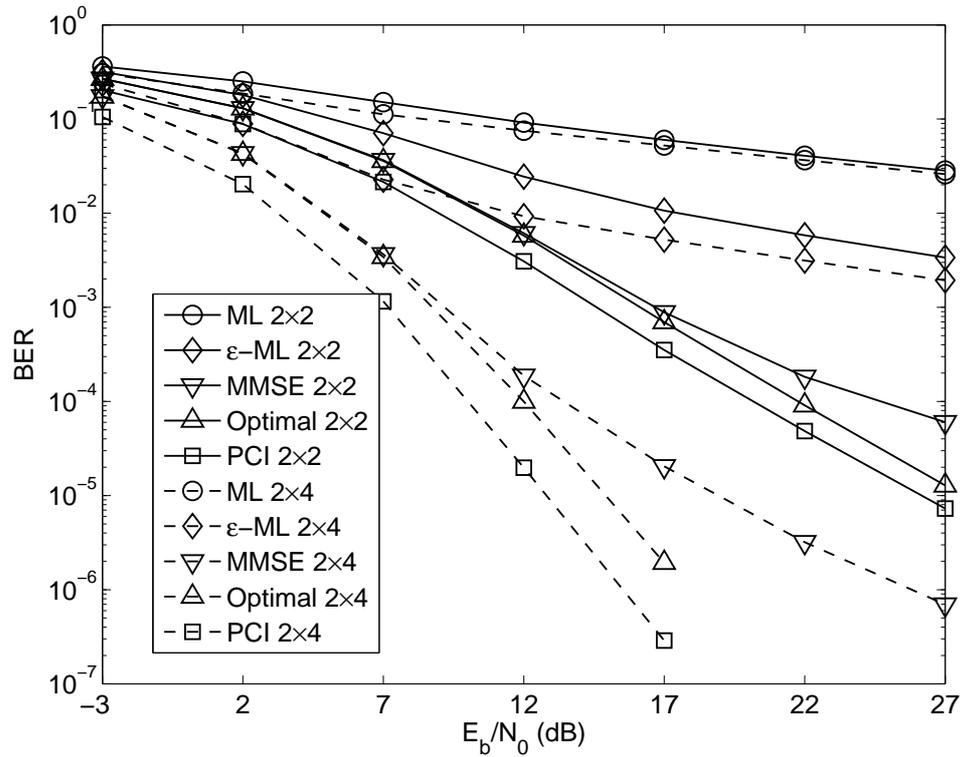
Figure 6.8: BER performance of the optimal and mismatched detectors for BPSK signals in MIMO channels, $L = 6$, $L_{\max} = 10$, $\tau_{\mathrm{rms}} = 5T$, $N = 461$, $P = 20$, $P_p = N_t$, $M = 23$; (a) 1×1 and 1×2 MIMO channels, and (b) 2×2 and 2×4 MIMO channels.

Chapter 7

Conclusions and Further Work

Contents

This thesis investigated the optimal detection in different scenarios of wireless communications with imperfect knowledge of the channel state information. We have investigated the MSE of BEM-based ML and MMSE channel estimators and sensitivity of the estimators to the mismatched Doppler spread (Chapter 3). We have derived an optimal detector which does not estimate the channel explicitly but jointly processes the received pilot and data symbols to recover the data and specify it to SISO time-invariant and time-variant channels (Chapter 4). We have extended this optimal detector to MIMO time-variant Rayleigh fading channels (Chapter 5) and investigated its detection performance. Finally, we have specified the optimal detector for scenarios with OFDM transmission and investigated the performance and complexity of this optimal detector in SISO and MIMO frequency selective channels (Chapter 6).

7.1 Conclusions

Chapter 1 has briefly introduced the whole work, and Chapter 2 has presented fundamental techniques including simulators of time-variant fading channels, BEMs and turbo

codes, which are used throughout the rest of the thesis.

In Chapter 3, we have derived the MSE of a generic BEM-based linear channel estimator with perfect or imperfect knowledge of the Doppler spread in time-variant channels. We have compared the performance and complexity of linear estimators based on various BEMs including Karhunen-Loeve (KL), discrete prolate spheroidal (DPS), generalized complex exponential (GCE) and B-spline BEMs, for the case with perfect or inaccurate knowledge of the Doppler spread. For the case with perfect knowledge of the Doppler spread, when the number of basis functions increases, all BEM-based MMSE estimators allow achievement of the optimal performance of the Wiener solution. We have also investigated the sensitivity of the BEM-based ML and MMSE estimators to the mismatched Doppler spread. The BEM-based estimators are very sensitive to underestimation of the Doppler spread but may have little sensitivity to overestimation. The estimators using BS and GCE BEMs are more robust to the Doppler spread mismatch than the estimators using KL and DPS BEMs. Although using the maximum Doppler spread to calculate the fading statistics and KL and DPS basis functions can simplify the estimation, it also leads to degradation in the MSE performance. The estimation using a slight overestimation of the Doppler spread outperforms the estimation using the maximum Doppler spread. It is worthy to emphasize that the complexity of the estimation using BS BEM is the lowest among all these 4 BEMs. Therefore, the estimation using B-splines with a slight (e.g., 20%) overestimate of the Doppler spread is a good practical choice providing a good performance, high robustness and low complexity.

In Chapter 4, we have proposed a pilot assisted optimal detector which does not require estimating the channel explicitly but jointly processes the received data and pilot symbols to recover the data with a minimum error. This optimal detector has been derived in the general case, i.e., time-variant channels, frequency-selective fading channels and MIMO channels. It outperforms the traditional detectors (named *mismatched*) treating channel estimates as perfect in systems applying QAM modulation schemes. However, we have found that the optimal detector cannot improve the performance in a system applying PSK modulation scheme in SISO channels. This optimal detector has been then specified to the SISO frequency-flat time-variant channels modeled by Jakes' model. We have compared performance of the optimal detector with that of the mismatched detectors applying ML, regularized ML or MMSE channel estimates. The simulation results have shown that

the optimal detector outperforms the mismatched detectors using ML and regularized ML estimates in the time-variant fading channel. However, the mismatched detector applying MMSE channel estimates provides nearly optimal detection performance. Due to the difficulty of estimating time-variant channels with high accuracy at low SNR, we have also investigated the detection performance of iterative receivers that exchange information between a channel estimator and turbo decoder. It has been shown by simulation that the iterative receiver with the optimal detector at the first iteration outperforms the receiver using the ML or regularized ML channel estimates. However, the use of MMSE channel estimates makes the detection performance close to that of the receiver with the optimal detector at the first iteration.

In Chapter 5, the optimal detector has been specified for MIMO Rayleigh frequency-flat fading channels. We have investigated the optimal detector and also compared its performance with that of traditional mismatched detectors with ML, regularized ML or MMSE channel estimation and compared their performance in MIMO time-invariant and time-variant flat fading channels with 16QAM and BPSK modulation. Simulation results for time-invariant fading channels have shown that the benefit on the performance caused by using the optimal detector becomes more significant when the number of antennas increases, but does not depend on the spatial correlation between antennas. We have then specified the optimal detector to spatially uncorrelated MIMO time-variant channels. The optimal detector significantly outperforms the mismatched detectors in spatially uncorrelated MIMO time-variant channels when the number of antennas increases. We have proved that if the SIMO channel is spatially uncorrelated, the optimal symbol-by-symbol detector of PSK signals is equivalent to the mismatched detector with MMSE channel estimation.

In Chapter 6, we have specified the optimal detector for OFDM signals in MIMO frequency-selective Rayleigh fading channels. We have compared the performance of the MMSE channel estimators with different BEMs for SISO channels and have chosen B-splines as basis functions to represent the channel frequency response. We have investigated the optimal detector and compared its performance with that of traditional mismatched detectors with ML, regularized ML or MMSE channel estimation for 16QAM and BPSK modulation for uncoded OFDM transmission. We have also investigated the detection performance of iterative receivers that exchange soft information between a

turbo decoder and the optimal or mismatched detectors in SISO channels. The simulation results have shown that in SISO channels, although the iterative receiver using the optimal detector at the first iteration outperforms iterative receivers applying mismatched detectors in all iterations, the improvement of the detection performance is not significant compared with the iterative receiver using the mismatched detector with MMSE channel estimates in all iterations. However, for transmission with a larger number of antennas, the optimal detector significantly outperforms the mismatched detectors in spatially uncorrelated MIMO frequency-selective fading channels.

7.2 Further Work

Some suggestions for further work based on this thesis are given below:

1. In this thesis, we have derived the optimal detector in the general case and compared its performance with that of the mismatched detectors in MIMO frequency-flat time-variant or frequency-selective time-invariant fading channels. Although this optimal detector outperforms the mismatched detectors for these cases, the improvement that can be obtained by using this optimal detector in doubly selective fading channels has not been investigated. We can expect that a more significant benefit can be obtained by using this optimal detector in MIMO doubly selective fading channels, which require a large number of unknown parameters to be estimated.

2. We have applied iterative receiver that exchanges information between a channel estimator and turbo decoder in SISO channels. We found that using the optimal detector only in the initial iteration can provide some improvement in the performance. A more significant benefit can be obtained by using iterative receivers with the optimal detector in MIMO channels. The challenge here is the significant increase in the complexity.

3. In this thesis, we have investigated the sensitivity of the MMSE channel estimator to the mismatched Doppler spread. However, we considered the optimal detector for the only case of perfect knowledge of statistical information of the fading. The sensitivity of the optimal detector to the mismatched estimation of the fading statistics is not clear. We can assume that the optimal detector is more robust to the mismatched fading statistics

than the mismatched detector with MMSE channel estimation, but this assumption needs to be proved by further research.

4. The optimal detector investigated in this thesis is based on the independence between received pilot and data symbols, and therefore, this detector cannot be extended directly to the transmission with superimposed training. We believe that an optimal detector for the transmission with superimposed training does exist, but the expression of this detector needs to be derived based on dependence of the received pilot and data symbols.

Appendix A

Let $\mathbf{x}$ and $\mathbf{v}$ be complex M-dimensional vectors and $\mathbf{C}$ be an $M \times M$ positive definite Hermitian matrix; then we have

$$\int e^{2\Re[\mathbf{x}^H\mathbf{v}]-\mathbf{x}^H\mathbf{C}^{-1}\mathbf{x}}d\Re[\mathbf{x}]d\Im[\mathbf{x}] = \pi^M|\mathbf{C}|e^{\mathbf{v}^H\mathbf{C}\mathbf{v}}. \tag{7.1}$$

To prove (7.1), consider the probability density $p(\mathbf{x})$ of a complex Gaussian vector $\mathbf{x}$: $p(\mathbf{x}) = \mathcal{N}_{\mathcal{C}}(\mathbf{u}, \mathbf{C})$ with mean $\mathbf{u}$ and covariance matrix $\mathbf{C}$. From $\int p(\mathbf{x})d\Re[\mathbf{x}]d\Im[\mathbf{x}] = 1$ and $(\mathbf{x}-\mathbf{u})^H\mathbf{C}^{-1}(\mathbf{x}-\mathbf{u}) = \mathbf{x}^H\mathbf{C}^{-1}\mathbf{x} - 2\Re(\mathbf{x}^H\mathbf{C}^{-1}\mathbf{u}) + \mathbf{u}^H\mathbf{C}^{-1}\mathbf{u}$, it follows

$$\int e^{2\Re[\mathbf{x}^H\mathbf{C}^{-1}\mathbf{u}]-\mathbf{x}^H\mathbf{C}^{-1}\mathbf{x}}d\Re[\mathbf{x}]d\Im[\mathbf{x}] = \pi^M|\mathbf{C}|e^{\mathbf{u}^H\mathbf{C}^{-1}\mathbf{u}}. \tag{7.2}$$

Using the substitution $\mathbf{v} = \mathbf{C}^{-1}\mathbf{u}$ and the equality $\mathbf{u}^H\mathbf{C}^{-1}\mathbf{u} = \mathbf{v}^H\mathbf{C}\mathbf{v}$, we finally obtain (7.1).

Let $f(\mathbf{a}|\mathbf{z}_p) = \mathcal{N}_{\mathcal{C}}(\mathbf{m_a}, \mathbf{S_a})$ be a PDF of a $M \times 1$ random vector $\mathbf{a}$ with mean $\mathbf{m_a}$ and covariance $\mathbf{S_a}$. Then, from (7.1) it follows that

$$\int e^{2\Re[\mathbf{a}^H\mathbf{L}_d]-\mathbf{a}^H\mathbf{\Gamma}_d\mathbf{a}}f(\mathbf{a}|\mathbf{z}_p)d\Re[\mathbf{a}]d\Im[\mathbf{a}]$$
$$= \frac{1}{|\mathbf{S_a}\mathbf{\Gamma}_d + \mathbf{I}_M|}e^{-\mathbf{m_a}^H\mathbf{S_a}^{-1}\mathbf{m_a}+W} \tag{7.3}$$

where

$$W = (\mathbf{L}_d + \mathbf{S_a}^{-1}\mathbf{m_a})^H(\mathbf{\Gamma}_d + \mathbf{S_a}^{-1})^{-1}(\mathbf{L}_d + \mathbf{S_a}^{-1}\mathbf{m_a}). \tag{7.4}$$

Bibliography

[1] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels", *IEEE Trans. on Veh. Technol.*, vol. 40, no. 4, pp. 686–693, 1991.

[2] J. K. Cavers, "Pilot symbol assisted modulation and differential detection infading and delay spread", *IEEE Trans. on Commun.*, vol. 43, no. 7, pp. 2206–2212, 1995.

[3] S. Sampei and T. Sunaga, "Rayleigh fading compensation method for 16QAM in digital land mobile radio channels", *Proc. IEEE 39th Veh. Technol. Conference, San Francisco, CA, USA*, pp. 640–646, 1989.

[4] M. K. Tsatsanis and Z. Xu, "Pilot symbol assisted modulation in frequency selective fading wireless channels", *IEEE Trans. on Signal Process.*, vol. 48, no. 8, pp. 2353–2365, 2000.

[5] P. Schramm and R. R. Muller, "Pilot symbol assisted BPSK on Rayleigh fading channels with diversity: Performance analysis and parameter optimization", *IEEE Trans. on Commun.*, vol. 46, no. 12, pp. 1560–1563, 1998.

[6] D. Subasinghe-Dias and K. Feher, "A coded 16 QAM scheme for fast fading mobile radio channels", *IEEE Trans. on Commun.*, vol. 43, no. 5, pp. 1906–1916, 1995.

[7] Q. Li, C. N. Georghiades, and X. Wang, "An iterative receiver for turbo-coded pilot-assisted modulation in fading channels", *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 145–147, 2001.

[8] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems", *Proc. IEEE 47th Veh. Technol. Conference, Costa Mesa, CA, USA*, vol. 3, pp. 1639–1643, 1997.

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

[10] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-symbol-aided channel estimation in time and frequency", *Proc. GLOBECOM'97, New York, USA*, vol. 4, pp. 90–96, 1997.

[11] C.L. Liu and K. Feher, "Pilot-symbol aided coherent M-ary PSK in frequency-selective fast Rayleigh fading channels", *IEEE Trans. on Commun.*, vol. 42, no. 1, pp. 54–62, 1994.

[12] D. Samardzija and N. Mandayam, "Pilot-assisted estimation of MIMO fading channel response and achievable data rates", *IEEE Trans. on Signal Process.*, vol. 51, no. 11, pp. 2882–2890, 2003.

[13] L. Cao and N. C. Beaulieu, "Exact error-rate analysis of diversity 16-QAM with channel estimation error", *IEEE Trans. on Commun.*, vol. 52, no. 6, pp. 1019–1029, 2004.

[14] L. Deneire, P. Vandenameele, L. Van der Perre, B. Gyselinckx, and M. Engels, "A low-complexity ML channel estimator for OFDM", *IEEE Trans. on Commun.*, vol. 51, no. 2, pp. 135–140, 2003.

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

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

[17] J. C. Lin, "Least-squares channel estimation for mobile OFDM communication on time-varying frequency-selective fading channels", *IEEE Trans. on Veh. Technol.*, vol. 57, no. 6, pp. 3538–3550, 2008.

[18] M.D.L.T. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels", *IEEE Trans. on Signal Process.*, vol. 52, no. 5, pp. 1403–1418, 2004.

[19] M. K. Tsatsanis and G. B. Giannakis, "Modelling and equalization of rapidly fading channels", *Int. J. Adaptive Contr. Signal Processing*, vol. 10, no. 2-3, pp. 159–176, 1996.

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

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

[22] T. A. Thomas and F. W. Vook, "Multi-user frequency-domain channel identification, interference suppression, and equalization for time-varying broadband wireless communications", *Proc. Sensor Array and Multichannel Signal Processing Workshop, 2000*, pp. 444–448, Mar. 2000.

[23] A. R. Kannu and P. Schniter, "MSE-optimal training for linear time-varying channels", *in Proc. IEEE ICASSP'05, Philadelphia, USA*, vol. 3, pp. 789–792, Mar. 2005.

[24] X. Dai, "Optimal training design for linearly time-varying MIMO/OFDM channels modelled by a complex exponential basis expansion", *IET Proc. Commun.*, vol. 1, no. 5, pp. 945–953, 2007.

[25] G. Leus, "On the estimation of rapidly time-varying channels", *Proc. EU-SIPCO'04, Vienna, Austria*, pp. 2227–2230, Sep. 2004.

[26] H. Mai, Y. V. Zakharov, and A. G. Burr, "Iterative channel estimation based on B-splines for fast flat fading channels", *IEEE Trans. on Wireless Commun.*, vol. 6, no. 4, pp. 1224–1229, Apr. 2007.

[27] Y. V. Zakharov, T. C. Tozer, and J. F. Adlard, "Polynomial spline-approximation of Clarke's model", *IEEE Trans. on Signal Process.*, vol. 52, no. 5, pp. 1198–1208, 2004.

[28] Y. V. Zakharov and T. C. Tozer, "Local spline approximation of time-varying channel model", *Electron. Lett.*, vol. 37, no. 23, pp. 1408–1409, 2001.

[29] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty, V: The discrete case", *Bell Syst. Tech. J.*, vol. 43, no. 6, pp. 3009–3058, 1964.

[30] P. Salvo Rossi and R. Muller, "Slepian-based two-dimensional estimation of time-frequency variant MIMO-OFDM channels", *IEEE Signal Process. Lett.*, vol. 15, 2008.

[31] G. W. Wornell, "A Karhunen-Loeve-like expansion for 1/f processes via wavelets", *IEEE Trans. on Inf. Theory*, vol. 36, no. 4, pp. 859–861, 1990.

[32] M. Visintin, "Karhunen-Loeve expansion of a fast Rayleigh fading process", *IEEE Electron . Lett.*, vol. 32, no. 18, pp. 1712–1713, 1996.

[33] X. Tang, M. S. Alouini, and A. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading", *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1856–1864, Dec. 1999.

[34] B. Xia and J. Wang, "Effect of channel-estimation error on QAM systems with antenna diversity", *IEEE Trans. on Commun.*, vol. 53, no. 3, pp. 481–488, 2005.

[35] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation", *IEEE Trans. on Wireless Commun.*, vol. 4, no. 4, pp. 1874–1888, 2005.

[36] Y. Chen and N. C. Beaulie, "Optimum pilot symbol assisted modulation", *IEEE Trans. on Commun.*, vol. 55, no. 8, pp. 1536–1546, Aug. 2007.

[37] J. Zhang, Y. Zakharov, and V. Baronkin, "Optimal detection in MIMO OFDM systems with imperfect channel estimation", *accepted by IET Communications*, 2009.

[38] Y. V. Zakharov, V. M. Baronkin, and J. Zhang, "Optimal detection of QAM signals in fast fading channels with imperfect channel estimation", *Proc. IEEE ICASSP'08, Las Vegas, USA*, pp. 3205–3208, Mar.-Apr. 2008.

[39] J. Zhang and Y. Zakharov, "Iterative B-spline estimator using superimposed training in doubly-selective fading channels", *Proc. IEEE 41st Asilomar Conf. Signals, Systems and Computers, ACSSC'07, Pacific Grove, CA, US*, Nov. 2007.

[40] Y. Zakharov, V. Baronkin, and J. Zhang, "Optimal detection of QAM signals in fast fading channels with imperfect channel estimation", *Proc. IEEE ICASSP'08, Las Vegas, USA*, Mar. 2008.

[41] J. Zhang, Y. Zakharov, and V. Baronkin, "Optimal detector of OFDM signals for imperfect channel estimation", *Proc. EUSIPCO'2008, Lausanne, Switzerland*, Aug. 2008.

[42] J. Zhang, Y. Zakharov, and V. Baronkin, "Optimal detection in MIMO Rayleigh fast fading channels with imperfect channel estimation", *Proc. IEEE 42st Asilomar Conf. Signals, Systems and Computers, ACSSC'08, Pacific Grove, CA, US*, Oct. 2008.

[43] R. N. Khal, Y. Zakharov, and J. Zhang, "Joint channel and frequency offset estimators for frequency-flat fast fading channels", *Proc. IEEE 42st Asilomar Conf. Signals, Systems and Computers, ACSSC'08, Pacific Grove, CA, US*, Oct. 2008.

[44] J. Zhang, Y. Zakharov, and R. N. Khal, "Optimal detection for STBC MIMO systems in spatially correlated Rayleigh fast fading channels with imperfect channel estimation", *Proc. IEEE 43st Asilomar Conf. Signals, Systems and Computers, ACSSC'09, Pacific Grove, CA, US*, Nov. 2009.

[45] R. N. Khal, J. Zhang, and Y. Zakharov, "Robustness of joint Bayesian frequency offset and channel estimation based on basis expansion models", *Proc. IEEE 43st Asilomar Conf. Signals, Systems and Computers, ACSSC'09, Pacific Grove, CA, US*, Nov. 2009.

[46] J. Zhang and Y. Zakharov, "Optimum detection in spatially uncorrelated SIMO Rayleigh fast fading channels with imperfect channel estimation", *under reveision by IEEE ICASSP'2010, Dallas, USA*.

[47] J. Zhang, R. N. Khal, and Y. Zakharov, "Sensitivity of MMSE channel estimator with B-splines to the mismatched Doppler frequency", *under reveision by IEEE ICASSP'10, Dallas, USA*.

[48] R. N. Khal, Y. Zakharov, and J. Zhang, "B-spline based joint channel and frequency offset estimation in doubly-selective fading channels", *under reveision by IEEE ICASSP'10, Dallas, USA*, 2010.

[49] R. H. Clarke, "A statistic theory of mobile-radio reception", *Bell Syst. Tech. J.*, pp. 957–1000, July 1968.

[50] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*, Wiley-IEEE Press, 1994.

[51] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels", *IEEE Trans. on Commun.*, vol. 51, no. 6, pp. 920–928, 2003.

[52] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited", *Electron. Lett.*, vol. 29, no. 13, pp. 1162–1163, 1993.

[53] M. Patzold, U. Killat, F. Laue, and Y. Li, "On the statistical properties of deterministic simulation models for mobile fading channels", *IEEE Trans. on Veh. Technol.*, vol. 47, no. 1, pp. 254–269, 1998.

[54] Y. Li and X. Huang, "The simulation of independent Rayleigh faders", *IEEE Trans. on Commun.*, vol. 50, no. 9, pp. 1503–1514, 2002.

[55] K. W. Yip and T. S. Ng, "A simulation model for Nakagami-m fading channels, m<1", *IEEE Trans. on Commun.*, vol. 48, no. 2, pp. 214–221, 2000.

[56] Y. R. Zheng and C. Xiao, "Improved models for the generation of multiple uncorrelated Rayleigh fading waveforms", *IEEE Commun. Letters*, vol. 6, no. 6, pp. 256–258, 2002.

[57] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels", *IEEE Trans. on Wireless Commun.*, vol. 5, no. 12, pp. 3667—3679, 2006.

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

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

[60] M. F. Pop and N. C. Beaulieu, "Limitations of sum-of-sinusoids fading channel simulators", *IEEE Trans. on Commun.*, vol. 49, no. 4, pp. 699–708, 2001.

[61] M. Patzold, R. García, and F. Laue, "Design of high-speed simulation models for mobile fading channels by using table look-up techniques", *IEEE Trans. on Veh. Technol.*, vol. 49, no. 4, pp. 1178–1190, 2000.

[62] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Second-order statistical properties of the WSS Jakes' fading channel simulator", *IEEE Trans. on Commun.*, vol. 50, no. 6, pp. 888–891, 2002.

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

[64] Z. Tang and G. Leus, "Time-multiplexed training for time-selective channels", *IEEE Signal Process. Lett.*, vol. 14, no. 9, pp. 585–588, 2007.

[65] Y. V. Zakharov, V. M. Baronkin, and J. Zhang, "Optimal and mismatched detection of QAM signals in fast fading channels with imperfect channel estimation", *IEEE Trans. on Wireless Commun.*, vol. 8, no. 2, pp. 617–621, 2009.

[66] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V- The discrete case", *Bell Syst. Tech. J.*, vol. 57, pp. 1371–1430, 1978.

[67] J. Kim, C. W. Wang, and W. E. Stark, "Frequency domain channel estimation for OFDM based on Slepian basis expansion", *Proc. IEEE ICC'07, Glasgow, Scotland*, pp. 3011–3015, 2007.

[68] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty III: the dimension of the space of essentially time- and band-limited signals", *Bell Syst. Tech. J.*, vol. 41, no. 4, pp. 1295–1336, Jul. 1962.

[69] T. Whitworth, M. Ghogho, and D. McLernon, "Optimized training and basis expansion model parameters for Doubly-Selective channel estimation", *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1490–1498, Mar. 2009.

[70] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems", *IEEE Trans. on Signal Process.*, vol. 55, no. 5, pp. 2226–2238, 2007.

[71] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of Doppler shifts", *IEEE Trans. on Signal Process.*, vol. 47, no. 6, pp. 1559–1569, 1999.

[72] M. Guillaud and D.T.M. Slock, "Channel modeling and associated inter-carrier interference equalization for OFDM systems with high Doppler spread", in *Proc. ICASSP'03, Hong Kong, China*, Apr. 2003, vol. 4, pp. 237–240.

[73] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blindidentification and equalization of time-varying channels", *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.

[74] G. Leus and M. Moonen, "Deterministic subspace based blind channel estimation for doubly-selective channels", in *4th IEEE Workshop on Signal Process. Adv. in Wireless Commun. (SPAWC'03), Rome, Italy*, Jun 2003, pp. 210–214.

[75] I. Barhumi, G. Leus, and M. Moonen, "Time-varying FIR equalization for doubly selective channels", *IEEE Trans. on Wireless Commun.*, vol. 4, no. 1, pp. 202–214, 2005.

[76] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I. Theory", *IEEE Trans. on Signal Process.*, vol. 41, no. 2, pp. 821–833, 1993.

[77] M. Unser, "Splines: a perfect fit for signal and image processing", *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, 1999.

[78] M. Unser, A. Aldroubi, and M. Eden, "Polynomial spline signal approximations: filter design and asymptotic equivalence with Shannon's sampling theorem", *IEEE Trans. on Inf. Theory*, vol. 38, no. 1, pp. 95–103, 1992.

[79] I. S. Gradshteyn and I. M. Ryzhik, "Table of Integrals, Series and Products (corrected and enlarged edition)", *Jeffrey Academic Press, New York, 1980*, 1980.

[80] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes.", *Proc. IEEE ICC'93, Geneva, Switzerland*, vol. 2, pp. 1064–1070, 1993.

[81] G. White, "Optimised turbo codes for wireless Channels", *Ph.d. Thesis, University of York*, 2001.

[82] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding And Decoding: Turbo-Codes", *IEEE Trans. on Commun.*, vol. 44, no. 10, pp. 1261, 1996.

[83] C. Berrou, "The ten-year-old turbo codes are entering into service", *IEEE Commun. Mag.*, vol. 41, no. 8, pp. 110–116, Aug. 2003.

[84] R. M. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes", *IEEE Trans. on Commun.*, vol. 46, no. 8, pp. 1003–1010, 1998.

[85] J. Garcia-Frias, "Compression of correlated binary sources using turbo codes", *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 417–419, 2001.

[86] S. Benedetto, R. Garello, and G. Montorsi, "A search for good convolutional codes to be used in the construction of turbo codes", *IEEE Trans. on Commun.*, vol. 46, no. 9, pp. 1101–1105, 1998.

[87] T. A. Summers and S. G. Wilson, "SNR mismatch and on–line estimation in turbo decoding", *IEEE Trans. on Commun.*, vol. 46, pp. 421–423, 1998.

[88] C. Berrou and M. Jezequel, "Non-binary convolutional codes for turbo coding", *Electron. Lett.*, vol. 35, no. 1, pp. 39–40, 1999.

[89] A. Burr, "Turbo-codes: the ultimate error control codes?", *Electronics & Communication Engineering Journal*, vol. 13, no. 4, pp. 155–165, 2001.

[90] D. Divsalar and F. Pollara, "Turbo codes for PCS applications", *Proc. IEEE ICC'95, Seattle, USA*, vol. 1, pp. 54–59, June 1995.

[91] B. Vucetic and J. Yuan, *Turbo codes: principles and applications*, Kluwer Academic Pub, 2000.

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

[93] M. M. Darmon and P. R. Sadot, "A new pseudo-random interleaving for antijamming applications", *Proc. IEEE MILCOM'89, Boston, MA, USA*, pp. 6–10, 1989.

[94] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations", *TDA Progress report*, vol. 42, no. 122, pp. 56–65, 1995.

[95] M. Ferrari, F. Scalise, S. Bellini, and D. E. e Inf, "Prunable s-random interleavers", *Proc. IEEE ICC'02, New York, USA*, vol. 3, pp. 1711–1715, 2002.

[96] L. Dinoi and S. Benedetto, "Design of fast-prunable S-random interleavers", *IEEE Trans. on Wireless Commun.*, vol. 4, no. 5, pp. 2540–2548, 2005.

[97] C. Heegard and S. B. Wicker, *Turbo coding*, Kluwer Academic Pub, 1999.

[98] P. Popovski, L. Kocarev, and A. Risteski, "Design of flexible-length S-random interleaver for turbo codes", *IEEE Commun. Lett.*, vol. 8, no. 7, pp. 461–463, 2004.

[99] S. Lin and D. J. Costello, *Error control coding*, Prentice-Hall Englewood Cliffs, NJ, 1983.

[100] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding", *Euro. Trans. on Telecom.*, vol. 8, no. 2, pp. 119–125, 1997.

[101] J. Vogt and A. Finger, "Improving the max-log-MAP turbo decoder", *Electron. lett.*, vol. 36, no. 23, pp. 1937–1939, 2000.

[102] A. Burr, *Modulation and coding: for wireless communications*, Prentice Hall, 2001.

[103] J. G. Proakis et al., *Digital Communication*, McGraw-Hill International 4th Editions, New York, 2001.

[104] J. Erfanian, S. Pasupathy, and G. Gulak, "Reduced complexity symbol detectors with parallel structure for ISI channels", *IEEE Trans. on Commun.*, vol. 42, no. 234 Part 3, pp. 1661–1671, 1994.

[105] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference", *Proc. IEEE Globecom'90, San Diego, CA, USA*, vol. 12, pp. 1679–1684, Nov. 1990.

[106] G. L. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publishers Norwell, MA, USA, 1996.

[107] K. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels", *Proc. IEEE GLOBECOM'05, St. Louis, Missouri, USA*, vol. 6, pp. 3503–3507, 2005.

[108] S. He and J. K. Tugnait, "On Doubly Selective Channel Estimation Using Super-imposed Training and Discrete Prolate Spheroidal Sequences", *IEEE Trans. on Signal Process.*, vol. 56, no. 7, pp. 3214–3228, 2008.

[109] P. S. Rossi and R. R. Muller, "Slepian-Based Two-Dimensional Estimation of Time-Frequency Variant MIMO-OFDM Channels", *IEEE Signal Process. Letters*, vol. 15, pp. 21, 2008.

[110] H. Arslan, L. Krasny, D. Koilpillai, and S. Chennakeshu, "Doppler spread estimation in mobile radio systems", *IEEE Commun. Lett.*, vol. 5, no. 5, pp. 197–199, 2001.

[111] M. Morelli, U. Mengali, and G. M. Vitetta, "Further results in carrier frequency estimation for transmissionsover flat fading channels", *IEEE commun. Letters*, vol. 2, no. 12, pp. 327–330, 1998.

[112] M. Wang, L. Silventoinen and Z. Honkasalo, "A new algorithm for estimating mobile speed at the TDMA-based cellular system", *Proc. IEEE Veh. Tech. Conf.96,*, p. 1145C1149, 1996.

[113] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall Inc., 1993.

[114] L. N. Trefethen and D. Bau, *Numerical linear algebra*, Society for Industrial Mathematics, 1997.

[115] H. J. Su and E. Geraniotis, "Low-complexity joint channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading", *IEEE Trans. on Commun.*, vol. 50, no. 2, pp. 249–261, 2002.

[116] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results", *IEEE Trans. on Commun.*, vol. 50, no. 5, pp. 754–767, 2002.

[117] R. Koetter, A. C. Singer, and M. Tuchler, "Turbo equalization", *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 67–80, Jan. 2004.

[118] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding forcoded CDMA", *IEEE Trans. on Commun.*, vol. 47, no. 7, pp. 1046–1061, 1999.

[119] R. Hoshyar and R. Tafazolli, "Soft decode and forward of MQAM modulations for cooperative relay channels", in *67-th IEEE Veh.Techn. Conf., 2008*, 2008, pp. 639–643.

[120] Y. V. Zakharov and V. P. Kodanev, "Doppler Scattering Adapted Reception in a Hydroacoustic Cornmunication Channel", *Acoustical Physics*, vol. 41, no. 2, pp. 219–223, Mar.–Apr. 1995.

[121] A. Lapidoth and S. Shamai, "Fading channels: how perfect need "perfect side information" be?", *IEEE Trans. on Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, 2002.

[122] C. N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading", *IEEE Trans. on Inf. Theory*, vol. 48, no. 3, pp. 637–650, 2002.

[123] A. Hedayat, H. Shah, and A. Nosratinia, "Analysis of space-time coding in correlated fading channels", *IEEE Trans. on Wireless Commun.*, vol. 4, no. 6, pp. 2882–2891, 2005.

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

[125] O. Sokoya and B. T. Maharaj, "Channel estimation for space-time trellis coded-OFDM systems based on nonoverlapping pilot structure", *Proc. SATNAC'08, Wild Coast, Eastern Cape*, pp. 7–10, 2008.

[126] R. Horn and C. Johnson, *Matrix analysis*, New York: Cambridge University Press, 1985.

[127] M. H. Hsieh and C. H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels", *IEEE Trans. on Consumer Electronics*, vol. 44, no. 1, pp. 217–225, 1998.

[128] S. Y. Park, Y. G. Kim, and C. G. Kang, "Iterative receiver for joint detection and channel estimation in OFDM systems under mobile radio channels", *IEEE Trans. on Veh. Technol.*, vol. 53, no. 2, pp. 450–460, 2004.

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

[130] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems", *IEEE Trans. on Wireless Commun.*, vol. 2, no. 5, pp. 876–882, 2003.

[131] Y. Li, L. J. Cimini Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels", *IEEE Trans. on Commun.*, vol. 46, no. 7, pp. 902–915, 1998.

[132] M. X. Chang and Y. T. Su, "Model-based channel estimation for OFDM signals in Rayleigh fading", *IEEE Trans. on Commun.*, vol. 50, no. 4, pp. 540–544, 2002.

[133] J. H. Yu and Y. T. Su, "Pilot-assisted maximum-likelihood frequency-offset estimation for OFDM systems", *IEEE Trans. on Commun.*, vol. 52, no. 11, pp. 1997–2008, 2004.

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

[135] A. J. Coulson, I. R. Ltd, L. Hutt, and N. Zealand, "Narrowband interference in pilot symbol assisted OFDM systems", *IEEE Trans. on Wireless Commun.*, vol. 3, no. 6, pp. 2277–2287, 2004.

[136] G. Auer, "Channel estimation in two dimensions for OFDM systems with multiple transmit antennas", *Proc. IEEE GLOBECOM'03, San Francisco, CA, USA*, vol. 1, pp. 322–326, 2003.

[137] X. Zhuang and F. W. Vook, "Iterative channel estimation and decoding for a turbo-coded OFDM system via the EM algorithm", *Proc. IEEE ICASSP'02, Orlando, FL, USA*, vol. 3, pp. 2337–2340, 2002.

[138] M. J. F. G. Garcia, S. Zazo, and J. M. Paez-Borrallo, "Pilot patterns for channel estimation in OFDM", *Electron. Lett.*, vol. 36, no. 12, pp. 1049–1050, 2000.

[139] Y. V. Zakharov and V. P. Kodanev, "Multipath-Doppler diversity of OFDM signals in an underwater acoustic channel", *Proc. IEEE ICASSP'00, Istanbul, Turkey*, vol. 5, pp. 2941–2944, June 2000.