國 立 交 通 大 學

電信工程研究所

碩 士 論 文

頻率選擇性衰減通道中對於外層渦輪碼正交頻分複用系統之
內層迴旋碼設計

Design of Non-coherent Inner Convolutional Codes
for Turbo-Coded OFDM System Over Frequency-Selective Fading Channels

研 究 生：呂志文

指導教授：陳伯寧　教授

中 華 民 國 一 〇 二 年 七 月

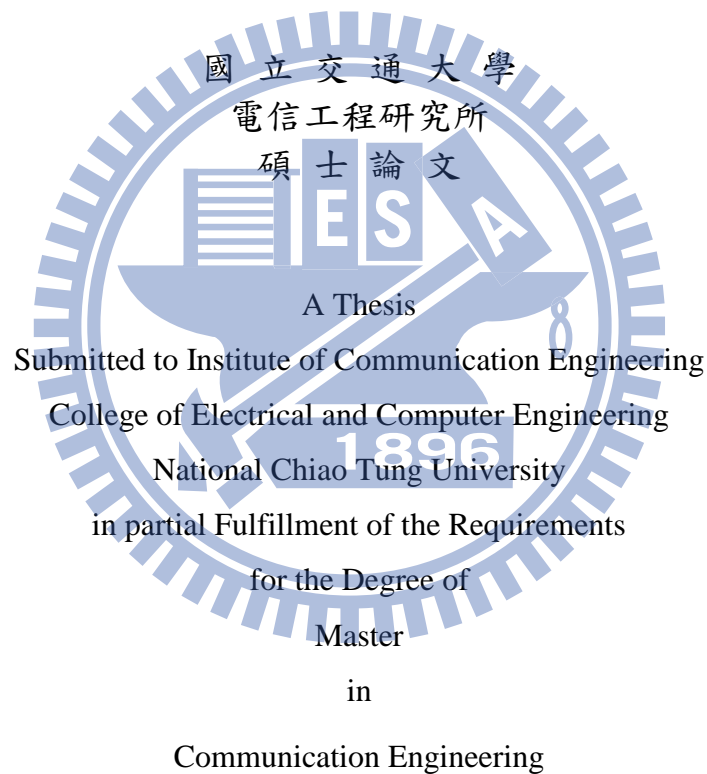頻率選擇性衰減通道中對於外層渦輪碼正交頻分複用系統之內層迴旋碼設計

Design of Non-coherent Inner Convolutional Codes
for Turbo-Coded OFDM System Over Frequency-Selective Fading Channels

研 究 生：呂志文　　　　　Student：Chih-Wen Lu

指導教授：陳伯寧　　　　　Advisor：Po-Ning Chen

國 立 交 通 大 學
電信工程研究所
碩 士 論 文

A Thesis

Submitted to Institute of Communication Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Communication Engineering

July 2013

Hsinchu, Taiwan, Republic of China

中華民國一〇二年七月

頻率選擇性衰減通道中對於外層渦輪碼正交頻分複用系統之內層迴旋 碼設計

學生: 呂志文　　　指導教授: 陳伯寧

國立交通大學電信工程研究所碩士班

摘要

在一個未知通道係數的頻率選擇性衰減通道中，通道估測以及通道等化基本上難以實現。因此在此篇論文，基於碼的累計距離函數(cumulative distance function)，我們提出一個簡單的法則來選擇內層迴旋碼，以與外層渦輪碼共同合作進行資料傳輸保護。在我們的設計中，內層迴旋碼其實相當於一個通道估測/等化器。由於我們所設計的內層迴旋碼的碼長極短，因此在接收端可以使用基於一般最大概度檢定測試準則(GLRT)的窮盡解碼法。模擬結果顯示，我們的方法在碼率為1/2時可以得到相當好的系統效能，但是當碼率提昇至2/3時，則所選出的迴旋碼有時仍非最佳選擇。我們最後經由模擬比較，確認使用我們所設計的內層迴旋碼確實可以達到比傳統最小平方(least square)通道估測更好的系統效能，故而我們的設計應該可以作為全盲通訊環境下，一個相當好的系統設計選項。

Design of Non-coherent Inner Convolutional Codes for Turbo-Coded OFDM System Over Frequency-Selective Fading Channels

Student: Chih-Wen Lu   Advisor: Po-Ning Chen
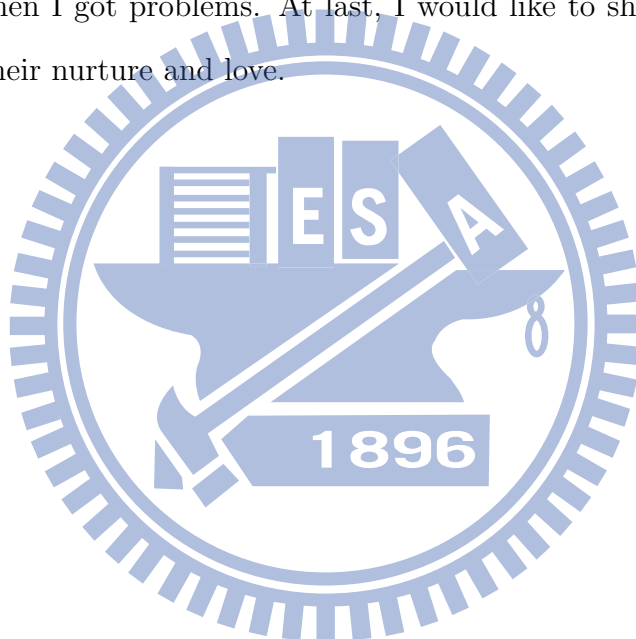
Institute of Communications Engineering

National Chiao Tung University

Abstract

In frequency-selective fading channels with unknown channel coefficients, channel estimation and equalization become infeasible. In this thesis, based on the cumulative distance functions of codes, we propose a simple method to select an inner convolutional code to co-work with an outer turbo code, which acts implicitly as a channel estimator and equalizer. Since the codeword length of the selected convolutional code is short, an exhaustive decoder based on the GLRT criterion can be implemented at the receiver end. Simulation results show that the concatenated coding system that use the convolutional codes of rate 1/2 performs well but may have rooms for improvement when the inner code rate increase to 2/3. In comparison with the tradition system that employs a least square (LS) estimator for the outer turbo code, our proposal can result in better performances and hence can be a good candidate design in a blind frequency-selective fading environment.
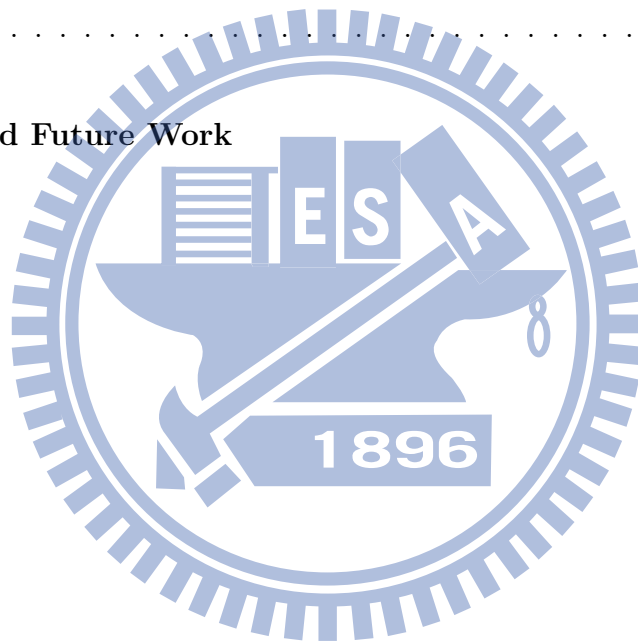
Acknowledgements

To begin with, I would like to sincerely appreciate my advisor, Prof. Po-Ning Chen, for his patient advising and guidance. I also would like to thank the members in the NTL-lab, who gave me support when I got problems. At last, I would like to show my deepest gratitude to my parents for their nurture and love.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

In this thesis, we research on the transceiving of orthogonal frequency-division multiplexing
(OFDM) modulated signals [1] over the unknown frequency-selective fading channels.

In the past decades, separate channel estimation and channel equalization have been
the major demodulation technologies for communication systems over fading channels [2,3].
However, such a system structure may not be well suited for a highly mobile environment,
which somehow motivates the use of noncoherent detection for communications among mov-
ing objects.

By [4], non-coherent detection jointly estimates the channel and data, and are often
regarded as a generalized likelihood ratio test (GLRT). It does not require a separate train-
ing sequence for the estimation and equalization of channel fading effects. Under a blind
environment, researches have shown that the GLRT demodulator can provide the best per-
formance [5,6].

At this background, we propose an error correcting coding scheme that concatenates
the turbo code [7] (as an outer code) and the convolutional code (as an inner code) in
sequence to alleviate the need of channel estimation and equalization. The convolutional

code is specifically designed for blind detection over fading channels. By analyzing the Euclidean distance among codewords [5] at the output of convolutional encoders, we provide a general method to determine which convolutional coding structure can yield a seemingly best performance in the sense of bit error rate (BER) under a blind environment. The proposed concatenated system is then compared with the conventional structure, where a turbo code is supported by a least square (LS) channel estimator [8] in stead of our convolutional code, provided that both systems have the same system transmission rate including the training sequence used for least square estimation.

The remaining of the thesis is organized as follows. Chapter 2 surveys some background technologies for this thesis, which includes non-coherent OFDM system, least square estimation and GLRT demodulation. Chapter 3 describes the proposed method on determination of the seemingly best convolutional code for a blind environment, as well as the entire system we construct for unknown frequency-selective fading channels. Chapter 4 presents the simulation results, while Chapter 5 concludes the thesis.

1.2 Notations

Symbol	Meaning
$\boldsymbol{x}$	a vector in time domain
$\boldsymbol{X}$	a vector in frequency domain
$\mathbb{Q}$	a matrix
$\|\boldsymbol{X}\|$	the norm of a vector $\boldsymbol{X}$
$\vec{\boldsymbol{S}}$	a combination of vectors
$\boldsymbol{X}^T$	transpose of a vector $\boldsymbol{X}$
$\boldsymbol{X}^\dagger$	Hermitian transpose of a vector $\boldsymbol{X}$
$\boldsymbol{X}^*$	complex conjugate of $\boldsymbol{X}$

Chapter 2

Preliminaries

In this chapter, the background knowledge on orthogonal frequency-division multiplexing (OFDM), the least square channel estimator, and generalized likelihood ratio test (GLRT) is introduced. Specifically, Section 2.1 presents the non-coherent OFDM system over fading channels, Section 2.2 provides the least square estimation for fading channels by following [9], and Section 2.3 gives the basic conception of GLRT-based signal detection.

2.1   A non-coherent OFDM system

In this thesis, we consider an OFDM system of N subchannels, for which binary phase-shift keying (BPSK) signals are non-coherently transmitted over a frequency-selective fading channel with fading coefficients $\boldsymbol{h} = \begin{bmatrix} h_\nu & \cdots & h_1 & h_0 \end{bmatrix}^{\mathbf{T}}$ unknown to both the transmitter and receiver.

Denote the information transmitted at time j as

$$\boldsymbol{S}_j = \begin{bmatrix} S_{1,j} \\ S_{2,j} \\ \vdots \\ S_{N,j} \end{bmatrix}_{N \times 1},$$

where each $S_{i,j} \in \{1, -1\}$. In terms of an N-point inverse discrete Fourier transform (iDFT),

the information is transformed to its time-domain counterpart as

$$\boldsymbol{s}_j = \mathbb{Q}^\dagger \boldsymbol{S}_j = \begin{bmatrix} s_{1,j} \\ s_{2,j} \\ \vdots \\ s_{N,j} \end{bmatrix}_{N \times 1}$$

where the superscript "†" indicates the matrix Hermitian transpose operation and the DFT matrix is given by

$$\mathbb{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} e^{-i\frac{2\pi}{N}(N-1)(N-1)} & e^{-i\frac{2\pi}{N}(N-2)(N-1)} & \cdots & e^{-i\frac{2\pi}{N}(N-1)} & 1 \\ e^{-i\frac{2\pi}{N}(N-1)(N-2)} & e^{-i\frac{2\pi}{N}(N-2)(N-2)} & \cdots & e^{-i\frac{2\pi}{N}(N-2)} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ e^{-i\frac{2\pi}{N}(N-1)} & e^{-i\frac{2\pi}{N}(N-2)} & \cdots & e^{-i\frac{2\pi}{N}} & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}_{N \times N} \tag{2.1}$$

Afterwards, a *cyclic prefix* [10] of length ν is added to the time-domain information $\boldsymbol{s}_j$ before its transmission over the fading channel, resulting a transmission signal of length $L = N + \nu$:

$$\boldsymbol{p}_j = \begin{bmatrix} p_{1,j} \\ p_{2,j} \\ \vdots \\ p_{L,j} \end{bmatrix} = \begin{bmatrix} s_{1,j} \\ s_{2,j} \\ \vdots \\ s_{N,j} \\ s_{1,j} \\ s_{2,j} \\ \vdots \\ s_{\nu,j} \end{bmatrix}_{L \times 1}$$

When transmitting $\boldsymbol{p}_j$ non-coherently over a fading channel, we should consider the impact of both the unknown channel coefficients and the background noise. Thus, the channel model is given by:

$$\boldsymbol{y}_j = \begin{bmatrix} y_{1,j} \\ y_{2,j} \\ \vdots \\ y_{L,j} \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\nu & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{\nu-1} & h_\nu & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_0 & h_1 & \cdots & h_\nu \\ 0 & 0 & \cdots & 0 & h_0 & \cdots & h_{\nu-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & h_0 \end{bmatrix} \begin{bmatrix} p_{1,j} \\ p_{2,j} \\ \vdots \\ p_{L,j} \end{bmatrix} + \begin{bmatrix} v_{1,j} \\ v_{2,j} \\ \vdots \\ v_{L,j} \end{bmatrix} \tag{2.2}$$

4

where

$$\boldsymbol{v}_j = \begin{bmatrix} v_{1,j} \\ v_{2,j} \\ \vdots \\ v_{L,j} \end{bmatrix}$$

represents a zero-mean additive Gaussian noise vector and $v_{1,j}, v_{2,j} \cdots, v_{L,j}$ are independent and identically distributed (*i.i.d.*) with covariance matrix $\sigma^2 \mathbb{I}_L$. In this thesis, $\mathbb{I}_L$ denotes an $L \times L$ identity matrix.

After the reception of $\boldsymbol{y}_j$, the cyclic prefix will be removed first, yielding a received vector of size N as:

$$\boldsymbol{x}_j = \begin{bmatrix} x_{1,j} \\ x_{2,j} \\ \vdots \\ x_{N,j} \end{bmatrix} = \begin{bmatrix} y_{1,j} \\ y_{2,j} \\ \vdots \\ y_{N,j} \end{bmatrix}.$$

By (2.2), we know that

$$\boldsymbol{x}_j = \begin{bmatrix} h_0 & h_1 & \cdots & h_{\nu-1} & h_\nu & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{\nu-2} & h_{\nu-1} & h_\nu & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_\nu \\ h_\nu & 0 & \cdots & 0 & 0 & h_0 & \cdots & h_{\nu-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_1 & h_2 & \cdots & h_\nu & 0 & 0 & \cdots & h_0 \end{bmatrix} \begin{bmatrix} s_{1,j} \\ s_{2,j} \\ \vdots \\ s_{N,j} \end{bmatrix} + \begin{bmatrix} v_{1,j} \\ v_{2,j} \\ \vdots \\ v_{N,j} \end{bmatrix}$$

$$= \mathbb{H}\boldsymbol{s_j} + \boldsymbol{w_j}$$

where $\boldsymbol{w}_j = \begin{bmatrix} w_{1,j} & w_{2,j} \cdots w_{N,j} \end{bmatrix}^T = \begin{bmatrix} v_{1,j} & v_{2,j} \cdots v_{N,j} \end{bmatrix}^T$. Since $\mathbb{H}$ is a circulant matrix, we can decompose it as:

$$\mathbb{H} = \mathbb{Q}^\dagger \Lambda \mathbb{Q} = \mathbb{Q}^\dagger \begin{bmatrix} \lambda_{N-1} & 0 & \cdots & 0 \\ 0 & \lambda_{N-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_0 \end{bmatrix} \mathbb{Q} \tag{2.3}$$

where the definition of $\mathbb{Q}$ can be found in (2.1), and Λ is a diagonal matrix with diagonals

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_{N-1} & \lambda_{N-2} & \cdots & \lambda_0 \end{bmatrix}^T \tag{2.4}$$

5

as indicated above. We can thus rewrite (2.3) as:

$$\boldsymbol{x}_j = \mathbb{Q}^\dagger \Lambda \mathbb{Q} \boldsymbol{s_j} + \boldsymbol{w}_j. \tag{2.5}$$

By $\mathbb{Q}\mathbb{Q}^\dagger = \mathbb{I}$, we can transform (2.5) into its frequency-domain counterpart as:

$$\boldsymbol{X}_j = \mathbb{Q}\boldsymbol{x}_j = \mathbb{Q}\mathbb{Q}^\dagger \Lambda \mathbb{Q} \boldsymbol{s}_j + \mathbb{Q}\boldsymbol{w}_j = \Lambda \boldsymbol{S}_j + \boldsymbol{W}_j$$

where $\boldsymbol{S}_j = \mathbb{Q}\boldsymbol{s}_j$ and $\boldsymbol{W_j} = \mathbb{Q}\boldsymbol{w}_j$ are the discrete Fourier transforms (DFTs) of signals $\boldsymbol{s}_j$ and $\boldsymbol{w}_j$, respectively. For derivation convenience, we rewrite the above system as:

$$\boldsymbol{X}_j = \mathbb{S}_j \boldsymbol{\lambda} + \boldsymbol{W}_j$$

where $\boldsymbol{\lambda}$ is defined in (2.4) and $\mathbb{S}_j$ is the diagonal matrix corresponding to $\boldsymbol{S_j}$, which is given by:

$$\mathbb{S}_j = \begin{bmatrix} S_{1,j} & 0 & \cdots & 0 \\ 0 & S_{2,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_{N,j} \end{bmatrix}_{N \times N}$$

We next introduce non-coherent OFDM detection. Under the assumption that both the transmitter and receiver know nothing about the channel coefficients, it can be derived that at least two OFDM symbols are needed to secure the detection error being better than that of a random guess. Thus, we assume that $T \geq 2T_s$, where T is the period, during which fading coefficients $\boldsymbol{h}$ remain constant, and T_s is the symbol duration of an OFDM symbol. For simplicity, we suppose two OFDM symbols are used for non-coherent detection at the receiver end, and its extension to more OFDM symbols can be similarly done. Let the two OFDM symbols be:

$$\boldsymbol{X}_1 = \mathbb{S}_1 \boldsymbol{\lambda} + \boldsymbol{W}_1 \tag{2.6}$$

and

$$\boldsymbol{X}_2 = \mathbb{S}_2 \boldsymbol{\lambda} + \boldsymbol{W}_2. \tag{2.7}$$

6

We then combine for convenience (2.6) and (2.7) into:

$$\vec{X} = \vec{S}\lambda + \vec{W} \tag{2.8}$$

where

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \vec{S} = \begin{bmatrix} \mathbb{S}_1 \\ \mathbb{S}_2 \end{bmatrix} \quad \text{and} \quad \vec{W} = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}.$$

Again, we can multiplex k OFDM symbols to make $\vec{S} = \begin{bmatrix} \mathbb{S}_1 & \mathbb{S}_2 \cdots \mathbb{S}_k \end{bmatrix}^{\mathrm{T}}$ if $T \geqslant kT_s$. However, the system complexity will significantly grow as k getting large. In addition, the requirement of $T \geq kT_s$ will restrict the applicability of our design to a less mobile environment for a moderately large k. For these reasons, we will focus only on the case of $k = 2$ in this thesis.

2.2   Least Square Estimator for Fading Channels

Consider the system model introduced in Section 2.1. Upon the reception of $\vec{X}$, the receiver intends to estimate λ such that $\|\vec{X} - \vec{S}\lambda\|^2$ is minimized when $\vec{S}$ is assumed known. This is therefore named the least square estimator (LSE). Denoting $\vec{Z} = \vec{S}\lambda$, we derive

$$
\begin{aligned}
J(\lambda) &= \|\vec{X} - \vec{Z}\|^2 \\
&= (\vec{X} - \vec{S}\lambda)^{\mathrm{T}}(\vec{X} - \vec{S}\lambda) \\
&= \vec{X}^{\mathrm{T}}\vec{X} - \vec{X}^{\mathrm{T}}\vec{S}\lambda - \lambda^{\mathrm{T}}\vec{S}^{\mathrm{T}}\vec{X} + \lambda^{\mathrm{T}}\vec{S}^{\mathrm{T}}\vec{S}\lambda.
\end{aligned}
$$

Note that $\vec{X}^{\mathrm{T}}\vec{X}$ is a constant scalar and hence we obtain the LSE of λ as:

$$\hat{\lambda} = (\vec{S}^T\vec{S})^{-1}\vec{S}^T\vec{X}. \tag{2.9}$$

With the availability of $\hat{\lambda}$, the estimated channel matrix $\hat{\mathbb{H}}$ can be computed through (2.3).

Together with the turbo coding, the LSE provides a conventional system design to combat the channel fading. This conventional system will be used as a benchmark to be compared with our proposed non-coherent scheme in this thesis.

2.3 GLRT Detection

As mentioned at the beginning of Section 2.1, $\boldsymbol{\lambda}$ that is used to determine $\mathbb{H}$ through (2.3) is unknown to the system. As no training sequence is presumed, its least square estimation cannot be practiced. However, by noting that what we really desire is the estimate of the transmitted signal $\vec{\boldsymbol{S}}$, a so-called general LRT (GLRT) detection has been proposed.

The GLRT criterion can be devised as follows.

$$
\begin{aligned}
\hat{\vec{\boldsymbol{S}}} &= \arg\min_{\vec{\boldsymbol{S}} \in \mathcal{S}} \left(\min_{\boldsymbol{\lambda} \in \mathcal{C}^N} \|\vec{\boldsymbol{X}} - \vec{\boldsymbol{S}}\boldsymbol{\lambda}\|^2 \right) \\
&= \arg\min_{\vec{\boldsymbol{S}} \in \mathcal{S}} \|\vec{\boldsymbol{X}} - \vec{\boldsymbol{S}}\hat{\boldsymbol{\lambda}}\|^2 \quad (2.10) \\
&= \arg\min_{\vec{\boldsymbol{S}} \in \mathcal{S}} \|\vec{\boldsymbol{X}} - \mathbb{P}_{\vec{\boldsymbol{S}}}\vec{\boldsymbol{X}}\|^2 \quad (2.11)
\end{aligned}
$$
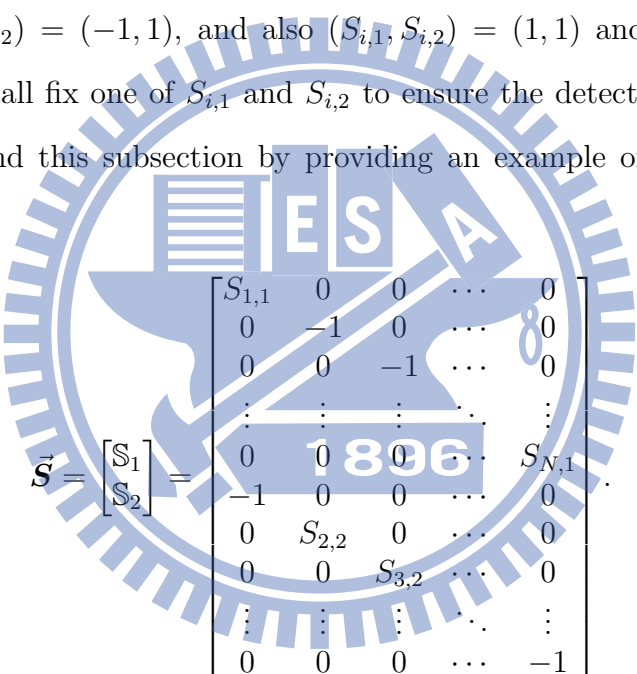
where $\mathcal{C}$ is the set of complex numbers, $\mathcal{S}$ is the set of all possible transmission signal $\vec{\boldsymbol{S}}$, $\hat{\boldsymbol{\lambda}}$ in (2.10) is given by (2.9), and

$$
\begin{aligned}
\mathbb{P}_{\vec{\boldsymbol{S}}} &= \vec{\boldsymbol{S}}(\vec{\boldsymbol{S}}^\dagger \vec{\boldsymbol{S}})^{-1} \vec{\boldsymbol{S}}^\dagger \\
&= \vec{\boldsymbol{S}} \begin{bmatrix} \frac{1}{|S_{1,1}|^2+|S_{1,2}|^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{|S_{2,1}|^2+|S_{2,2}|^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{|S_{N,1}|^2+|S_{N,2}|^2} \end{bmatrix} \vec{\boldsymbol{S}}^\dagger
\end{aligned}
$$

$$
= \begin{bmatrix}
\frac{|S_{1,1}|^2}{|S_{1,1}|^2+|S_{1,2}|^2} & 0 & \cdots & 0 & \frac{S_{1,1}S_{1,2}^*}{|S_{1,1}|^2+|S_{1,2}|^2} & 0 & \cdots & 0 \\
0 & \frac{|S_{2,1}|^2}{|S_{2,1}|^2+|S_{2,2}|^2} & \cdots & 0 & 0 & \frac{S_{2,1}S_{2,2}^*}{|S_{2,1}|^2+|S_{2,2}|^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{|S_{N,1}|^2}{|S_{N,1}|^2+|S_{N,2}|^2} & 0 & 0 & \cdots & \frac{S_{N,1}S_{N,2}^*}{|S_{N,1}|^2+|S_{N,2}|^2} \\
\frac{S_{1,1}^*S_{1,2}}{|S_{1,1}|^2+|S_{1,2}|^2} & 0 & \cdots & 0 & \frac{|S_{1,2}|^2}{|S_{1,1}|^2+|S_{1,2}|^2} & 0 & \cdots & 0 \\
0 & \frac{S_{2,1}^*S_{2,2}}{|S_{2,1}|^2+|S_{2,2}|^2} & \cdots & 0 & 0 & \frac{|S_{2,2}|^2}{|S_{2,1}|^2+|S_{2,2}|^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{S_{N,1}^*S_{N,2}}{|S_{N,1}|^2+|S_{N,2}|^2} & 0 & 0 & \cdots & \frac{|S_{N,2}|^2}{|S_{N,1}|^2+\|S_{N,2}\|^2}
\end{bmatrix}.
$$

$$(2.12)$$

Taking (2.12) into (2.11) yields:

$$
\begin{aligned}
\hat{\vec{S}} &= \arg\min_{\vec{S}\in\mathcal{S}} \|\vec{X} - \mathbb{P}_{\vec{S}}\vec{X}\|^2 \\
&= \arg\max_{\vec{S}\in\mathcal{S}} \vec{X}^\dagger \mathbb{P}_{\vec{S}}\vec{X} \\
&= \arg\max_{\vec{S}\in\mathcal{S}} \sum_{i=1}^{N} \frac{|X_{i,1}|^2|S_{i,1}|^2 + |X_{i,2}|^2|S_{i,2}|^2 + X_{i,1}S_{i,1}^*X_{i,2}^*S_{i,2} + X_{i,1}^*S_{i,1}X_{i,2}S_{i,2}^*}{|S_{i,1}|^2 + |S_{i,2}|^2} \\
&= \arg\max_{\vec{S}\in\mathcal{S}} \sum_{i=1}^{N} |S_{i,1}X_{i,1}^* + S_{i,2}X_{i,2}^*|^2
\end{aligned}
$$

The resulting formula for $\hat{\vec{S}}$ implies that the receiver cannot distinguish between $(S_{i,1}, S_{i,2}) = (1, -1)$ and $(S_{i,1}, S_{i,2}) = (-1, 1)$, and also $(S_{i,1}, S_{i,2}) = (1, 1)$ and $(S_{i,1}, S_{i,2}) = (-1, -1)$. Consequently, we shall fix one of $S_{i,1}$ and $S_{i,2}$ to ensure the detectability of information at the receiver. We end this subsection by providing an example of detectable information assignment:

$$
\vec{S} = \begin{bmatrix} \mathbb{S}_1 \\ \mathbb{S}_2 \end{bmatrix} = \begin{bmatrix}
S_{1,1} & 0 & 0 & \cdots & 0 \\
0 & -1 & 0 & \cdots & 0 \\
0 & 0 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & S_{N,1} \\
-1 & 0 & 0 & \cdots & 0 \\
0 & S_{2,2} & 0 & \cdots & 0 \\
0 & 0 & S_{3,2} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -1
\end{bmatrix}.
$$

Chapter 3

Design of Inner Convolutional Code for Turbo-Coded OFDM System

In this chapter, we will introduce how we determine a convolutional code that can provide a seemingly best performance, when it is co-worked with an outer turbo code, under channels with unknown fading coefficients. In addition, the entire inner and outer concatenated coding system will be defined.

3.1 Finding a Good Noncoherent Convolutional Code for Turbo-Coded OFDM System: The Approach

In order to construct a good convolutional coding structure without resorting to simulations, we adopt the "non-coherent" Euclidean distances (defined later in (3.1)) among codewords as our design criterion. This quantity may be computable directly and is seemingly related to the true error rate of the resulting code.

Consider the system that transmits a convolutional codeword $\boldsymbol{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_N \end{bmatrix}$ of length N over a frequency-selective fading channel as follows:

$$\boldsymbol{y} = \mathbb{B}\boldsymbol{h} + \boldsymbol{v}$$

where the codeword matrix is given by:

$$\mathbb{B} = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ b_2 & b_1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_N & b_{N-1} & \ddots & 0 \\ 0 & b_N & \ddots & b_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_N \end{bmatrix}_{L \times P}$$

with $L = N + \nu$, $\boldsymbol{h}$ is the channel coefficient vector of length $P = \nu + 1$, and $\boldsymbol{v}$ is a zero-mean additive white Gaussian noise vector. Note that in the above formula, ν is the memory order of the channel. Then, it can be derived that the "non-coherent" Euclidean distance between codewords i and j can be used to optimally detect the optimal codeword, which is defined as:

$$d_{i,j} = \|\text{vec}(\mathbb{P}_{B_i}) - \text{vec}(\mathbb{P}_{B_j})\|^2 \tag{3.1}$$

where

$$\mathbb{P}_B = \mathbb{B}(\mathbb{B}^T \mathbb{B})^{-1} \mathbb{B}^T.$$

In order to analyze the cumulative distribution function (CDF) of the "noncohrent" Euclidean distances of a convolutional codeword, we quantify $d_{i,j}$ into 1000 levels and compute the CDF functions for two convolutional codes to be compared. Denote by C_M and C_N the CDF functions of two convolutional codes with structure M and N, respectively. We then do the following computation in which we only take half of the level function values for this comparison:

$$a = \sum_{\texttt{level}=1}^{500} \texttt{index}(\texttt{level}) \tag{3.2}$$

where

$$\texttt{index}(\texttt{level}) = \begin{cases} 1, & \text{if} \quad C_M(\texttt{level}) < C_N(\texttt{level}) \\ -1, & \text{if} \quad C_M(\texttt{level}) > C_N(\texttt{level}) \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

11

From our observations, we note that if $a > 0$, then structure M mostly will yield a better performance than structure N. Hence, we propose to compare all the convolutional code design through this way, and determine the best one that with high possibility can yield a good (near optimal) error performance. We will later confirm our proposal via simulations.

Note that since the designed convolutional code will be co-worked with a turbo code, the convolutional code rate that we target in this research is 2/3 (instead of the usual 1/2). This convolutional code rate, when combined with an 1/3 turbo code, will give a system code rate of $(2/3)(1/3) = (2/9)$. Details will be given in the next section.

3.2 Concatenated Coding System Under an Unknown Fading Channel

The concatenated coding system we consider is depicted in Fig. 3.1.
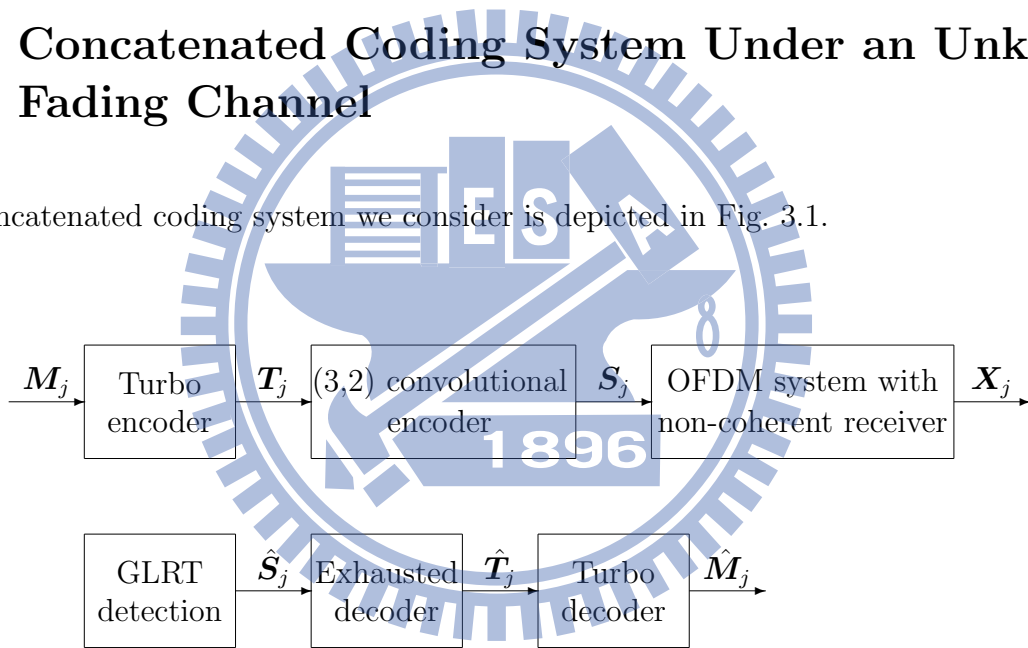
Figure 3.1: Concatenated encoding (above) and decoding (below) systems under an unknown frequency-selective fading channel. The non-coherent transmitter (resp., receiver) includes a turbo encoder (resp., decoder), a convolutional encoder (resp., decoder), and an OFDM modulater (resp., demodulator). The GLRT detecter is also employed at the receiver.

The idea behind our concatenated system, where we use the turbo code as an outer code and our non-coherent convolutional code as an inner code, is that the traditional turbo code design is based on the assumption that the system can perfectly recover the fading effect and

suffers only the additive white Gaussian noise. However, this may not be entirely valid in practical application. So instead of targeting a perfect channel estimator and equalizer, we use a non-coherent convolutional code, and are exempt from the need of designing channel estimator and equalizer.

Considering the code rate loss, we employ (non-coherent) convolutional codes of rate 2/3. Since our codeword length is small, an exhausted decoder can be implemented at the receiver. This is a replacement of the channel estimation and equalization on the channel fading effect with error protection enhancement. By this way, we can use the turbo outer code under a "blind" fading environment.

3.2.1 A Review of Turbo Coding

As shown in Fig. 3.2, a turbo encoder is structured with two recursive systematic convolutional (RSC) codes and one interleaver. In this thesis, three turbo codes with different RSC component codes and interleavers will be tested. They are respectively the (37,21) turbo code proposed by Berrou and Glavieux, one of the 3GPP/LTE specified turbo codes, and the turbo code with generator matrix $\left[1 \ (1+D+D^2+D^4)/(1+D^3+D^4)\right]$ plus a uniform interleaver as introduced in Shu Lin's book. To test how effective our internal non-coherent convolutional code is, we choose a practically short codeword length for the turbo code instead a usually long one. Thus, the codeword lengths for the three turbo codes are respectively 12288, 12288, and 3000, with information lengths respectively to be 4096, 4096 and 1000.

The decoding mechanism for turbo codes is the the maximum *a posteriori* probability (MAP) algorithm, which after a few iterations declares whether each information bit is 1 or
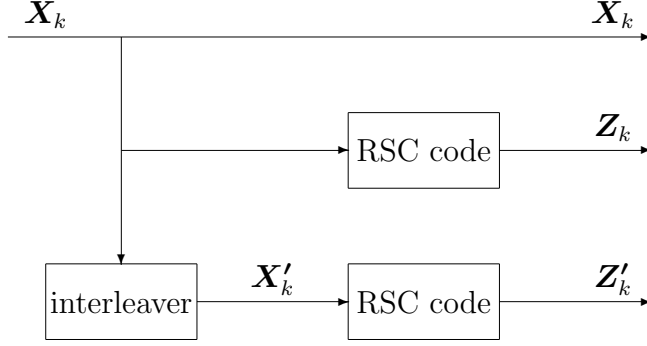
13

Figure 3.2: Architecture of the turbo encoder

0 according to the log-likelihood ratio (LLR):

$$\Lambda(i) = \log \frac{\Pr(c_i = 1 \mid \boldsymbol{r})}{\Pr(c_i = 0 \mid \boldsymbol{r})} = \log \frac{\displaystyle\sum_{(S_w^{i-1}, S_{\tilde{w}}^i) \in \mathcal{B}_i^{(1)}} \Pr(S_w^{i-1}, S_{\tilde{w}}^i, \boldsymbol{r})}{\displaystyle\sum_{(S_w^{i-1}, S_{\tilde{w}}^i) \in \mathcal{B}_i^{(0)}} \Pr(S_w^{i-1}, S_{\tilde{w}}^i, \boldsymbol{r})},$$

where $\boldsymbol{r}$ is the received vector, and $\mathcal{B}_i^{(k)}$ denotes the set of all transitions from nodes at level $i-1$ to nodes at level i corresponding to input $c_i = k \in \{0, 1\}$. Again, we remind the readers that S_w^i means the trellis state at level i. Apparently, the decision rule shall be $\hat{m}_i = 1$ if $\Lambda(i) > 0$, and $\hat{m}_i = 0$, otherwise.

3.2.2 Non-Coherent Convolutional Code and Exhausted Decoder

The non-coherent convolutional codes that we designed and tested in this thesis have three different codeword lengths: six, eight, and ten. For convenience, we will refer to them as type 3, type 2 and type 1 in the figures. Using the selection approach we propose, we found that the best type 1, type 2 and type 3 $(3, 2)$ non-coherent convolutional codes are [5 6 0 ; 0 7 5], [5 6 0 ; 0 7 5] and [7 6 0 ; 0 7 5] (in octal) under unknown fading channels. As an example, the structure of the type 1 non-coherent convolution code is depicted in Fig. 3.3.

Since the codeword length of our designed code is small, we propose to use the exhaustive

14

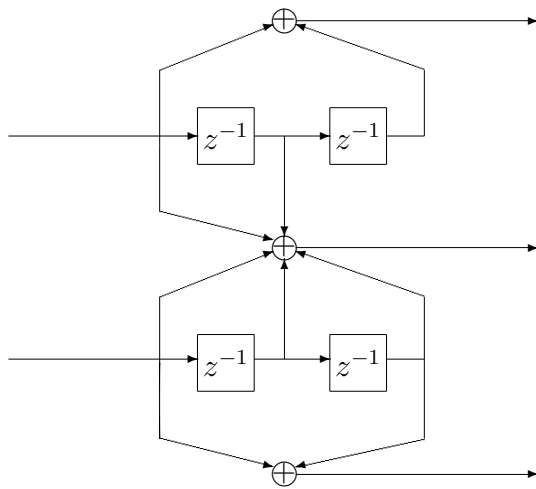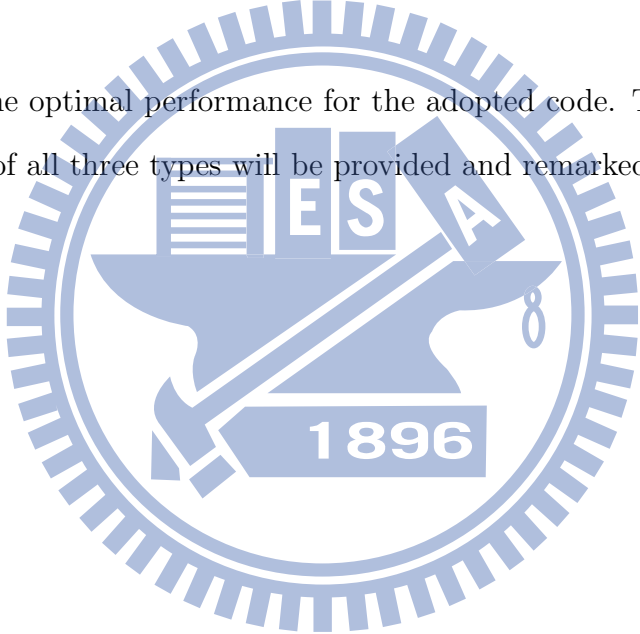Figure 3.3: The (3,2) non-coherent convolutional code which generating polynomial [5 6 0 ; 0 7 5].

decoder to secure the optimal performance for the adopted code. The simulation results of the designed codes of all three types will be provided and remarked in the next chapter.

Chapter 4

Simulation Results

In this chapter, we will examine how well our proposed method can locate a non-coherent convolutional code with acceptably good performance. Simulations on the found non-coherent convolutional codes concatenated with turbo codes will follow. For clarity, basic simulation results are provided in Section 4.1, and discussions on them as well as additional simulations to verify our interpretation will be placed in Section 4.2.

4.1   System Settings

In finding the optimal convolutional coding structure that gives the minimum decoding error, we first examine the relation between the convolutional code structures and corresponding minimum pair-wise non-coherent distances defined in (2.11).

In Figure 4.1, the legend mark "5-1;0.94" means that the convolutional code structure is defined via generator matrix $\begin{bmatrix} 1+D^2 & D^2 \end{bmatrix}$ and its resulting minimum pairwise non-coherent distance is 0.94. Similarly, the codes marked "5-2", "5-3", "5-4", "5-5", "5-6" and "5-7" are defined via the generator matrices $\begin{bmatrix} 1+D^2 & D \end{bmatrix}$, $\begin{bmatrix} 1+D^2 & D+D^2 \end{bmatrix}$, $\begin{bmatrix} 1+D^2 & 1 \end{bmatrix}$, $\begin{bmatrix} 1+D^2 & 1+D^2 \end{bmatrix}$, $\begin{bmatrix} 1+D^2 & 1+D \end{bmatrix}$ and $\begin{bmatrix} 1+D^2 & 1+D+D^2 \end{bmatrix}$, respectively. The simulation result shows that the minimum pairwise non-coherent distance may not be a good

decisive factor for performance. In particular, the one with the largest minimum pairwise non-coherent distance, i.e., "5-5", turns out to have the worst performance. Numbers of codeword pairs that result in the respective non-coherent distances should be taken into consideration.

For this reason, we examine in Figure 4.2 the so-called *cumulative distance function* (CDF) of the codes simulated in Figure 4.1. The y-axis in Figure 4.2 indicates the cumulative number of codeword pairs whose non-coherent distance is below the quantity indicated by the x-axis.

For a more well-rounded understanding of the relation between the CDFs of non-coherent convolutional codes and their performances, another set of convolutional codes is tested. The codes are structured as $\left[1 + D + D^2 \quad D^2\right], \ldots, \left[1 + D + D^2 \quad 1 + D + D^2\right]$, which are marked as "7-1", ... "7-7" in the legends of the respective figures. The results are summarized in Figures 4.3 and 4.4. Same notations, for example, "7-1;0.92", are used in these two figures. With the CDFs, we can locate good convolutional codes (termed type 1, type 2 and type 3 in Section 3.2.2) using the approach proposed in the previous chapter.

As previously mentioned, the convolutional code that we design is used as an inner code to combat the channel fading. An outer code such as the turbo code will be concatenated with it. Thus, our next focus is to examine the system performance of the proposed concatenated coding system. Three turbo codes are tested. They are the $(37, 21)$ turbo code [7], the 3GPP/LTE specified turbo code [12], and the turbo code with generator matrix $\left[1 \quad (1 + D + D^2 + D^4)/(1 + D^3 + D^4)\right]$ [11]. The results for these three turbo codes are respectively illustrated in Figures 4.5–4.7.

Finally, Figures 4.8, 4.9 and 4.10 show the performances of a traditional system structure that replaces our non-coherent inner convolutional code with a least square channel estimators. In these figures, "H'" means the estimated channel matrix by the LS channel

estimator.

Figure 4.1: Performances of non-coherent convolutional codes under blind fading channels. Here, the codeword length $N = 10$ and the channel memory order is $\nu = 2$. The code rate is $1/2$.

Figure 4.2: Cumulative distance function (CDF) of the convolutional codes under test. Here, we provide two plots that have different x-axis ranges to facilitate our interpretation of the results. The code rate is 1/2.
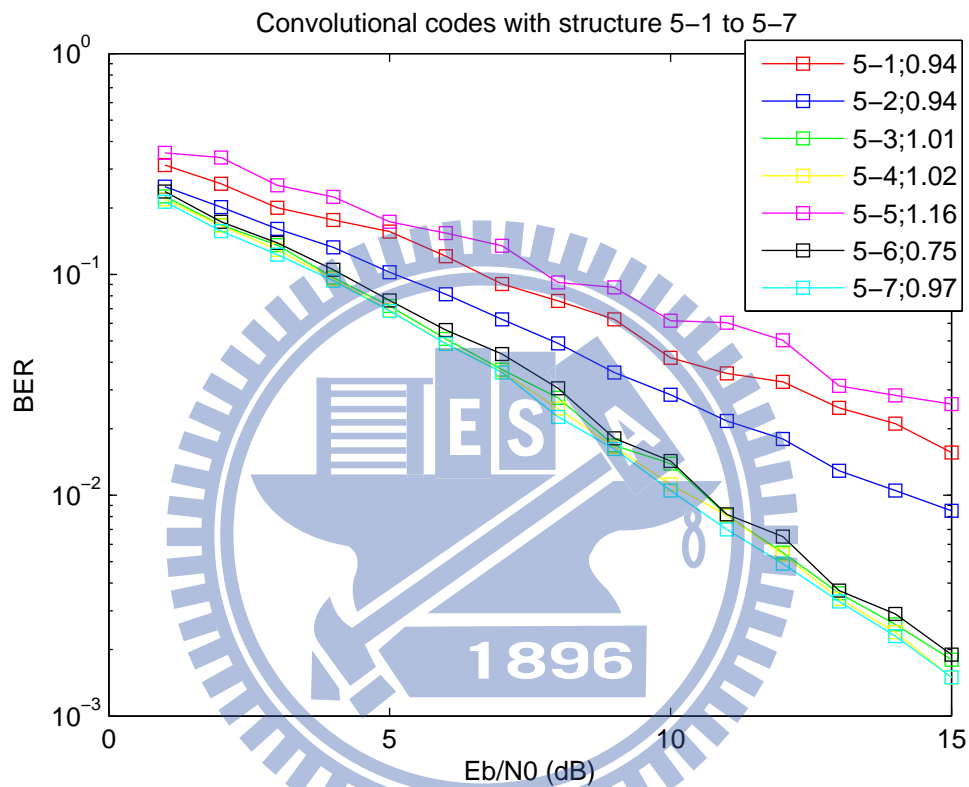
Figure 4.3: Performances of non-coherent convolutional codes under blind fading channels. Here, the codeword length $N = 10$ and the channel memory order is $\nu = 2$. The code rate is $1/2$.
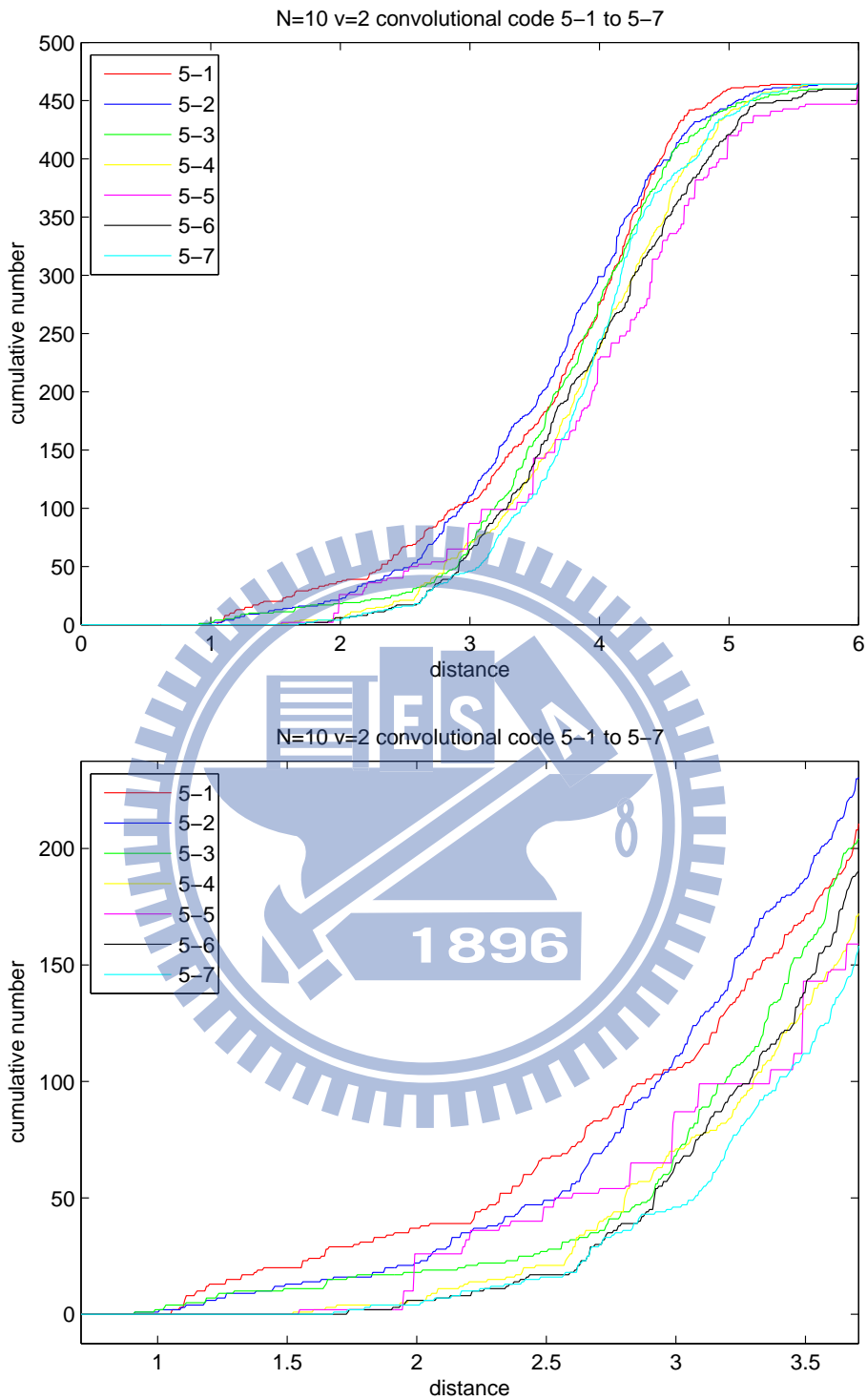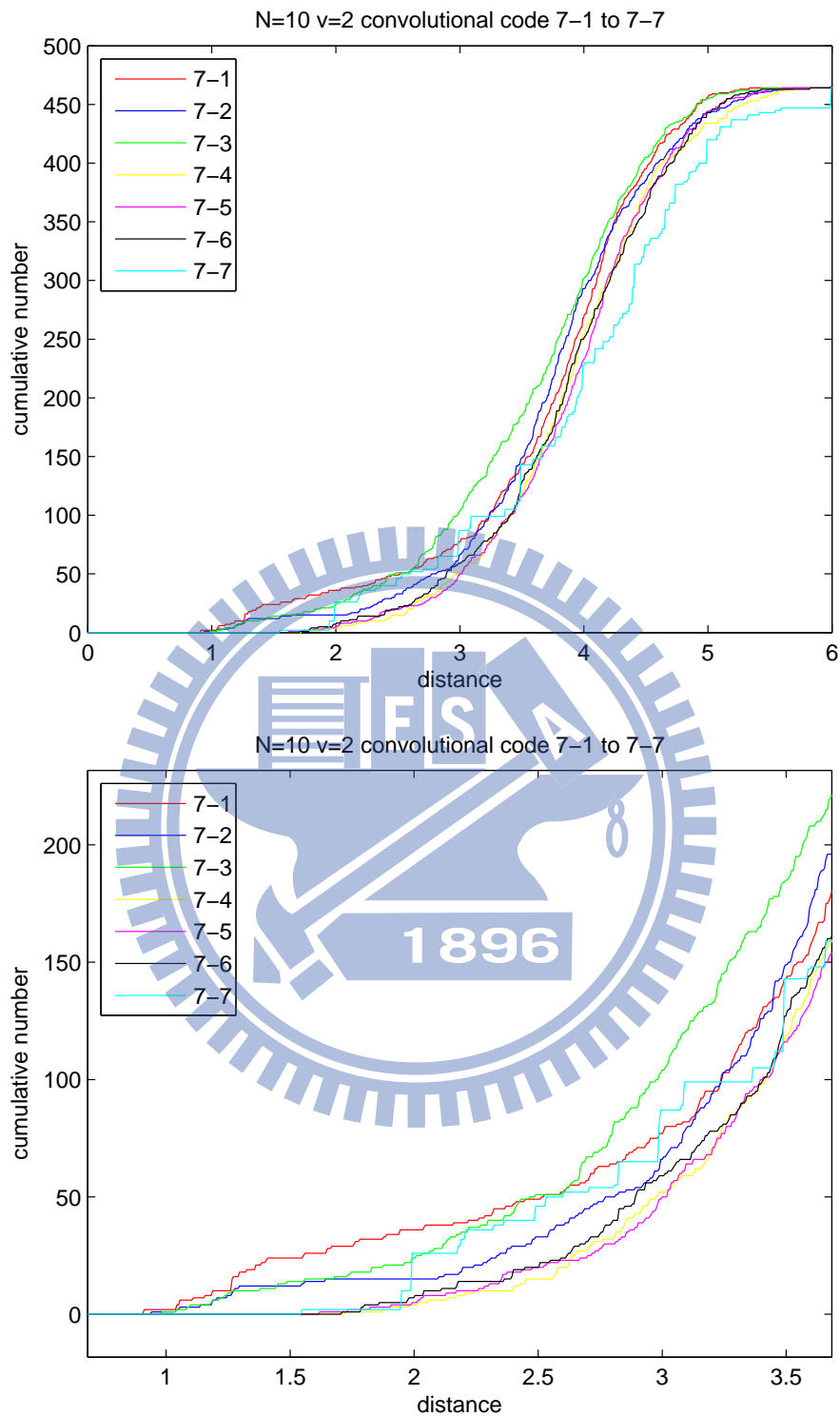
Figure 4.4: Cumulative distance function (CDF) of the convolutional codes under test. Here, we provide two plots that have different x-axis ranges to facilitate our interpretation of the results. The code rate is 1/2.
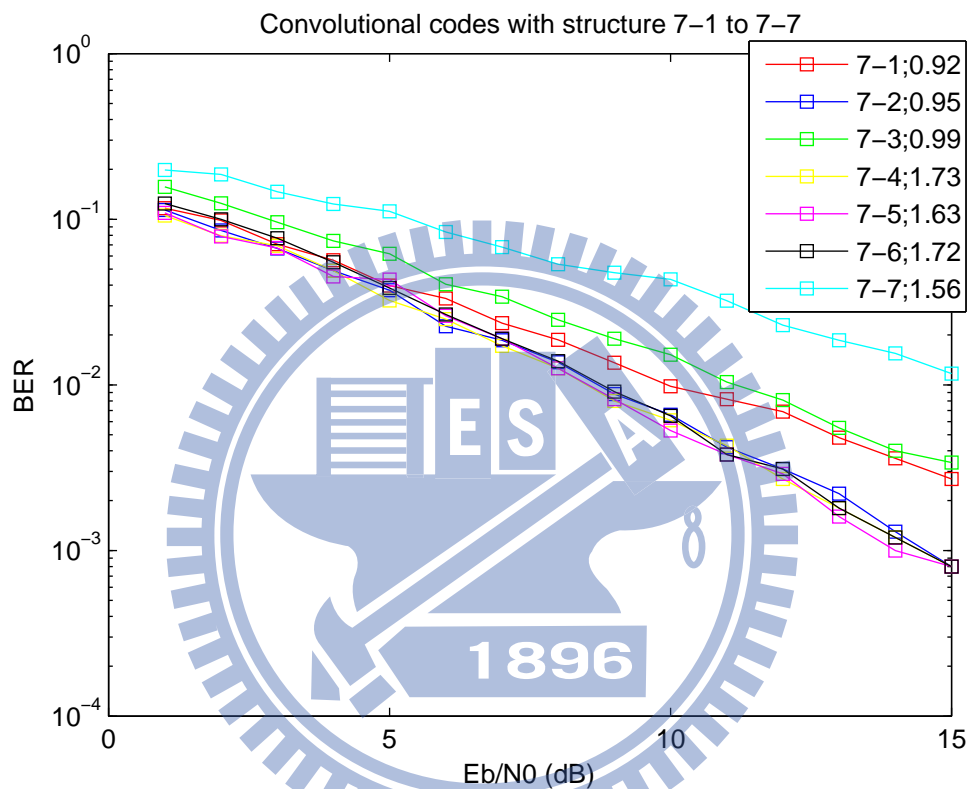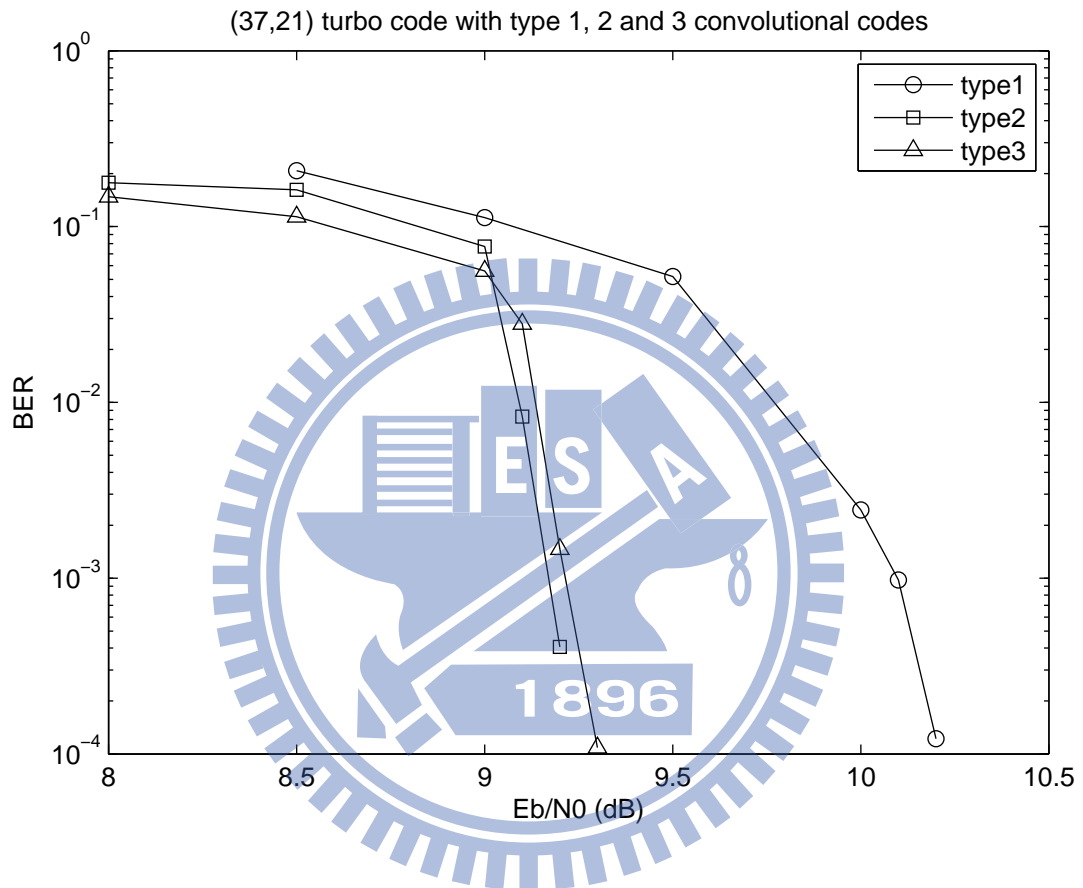
Figure 4.5: Performances of the concatenated coding system. The turbo code used is the (37, 21) code proposed by Berrou and Glavieux in [7]. The codeword lengths of the type 1, 2 and 3 convolutional codes are ten, eight and six, respectively. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3.
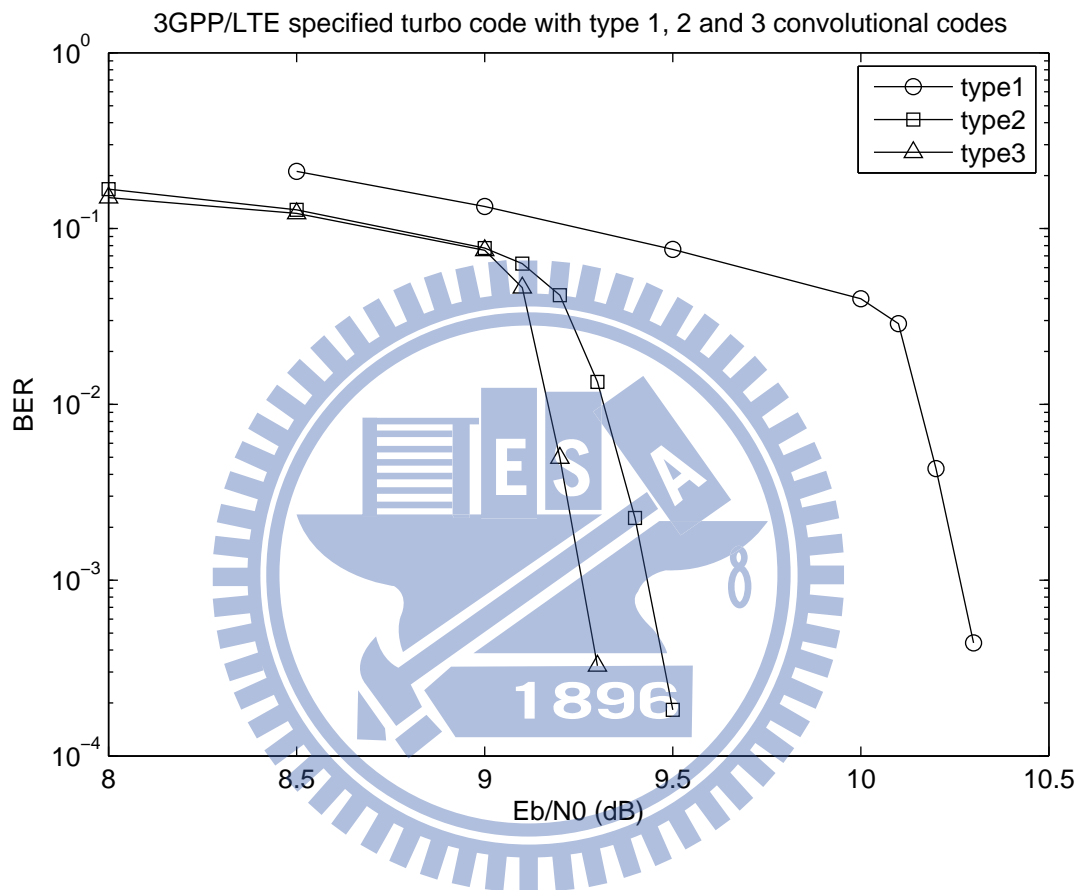
Figure 4.6: Performances of the concatenated coding system. The turbo code used is the one specified in 3GPP/LTE. The codeword lengths of the type 1, 2 and 3 convolutional codes are ten, eight and six, respectively. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3.
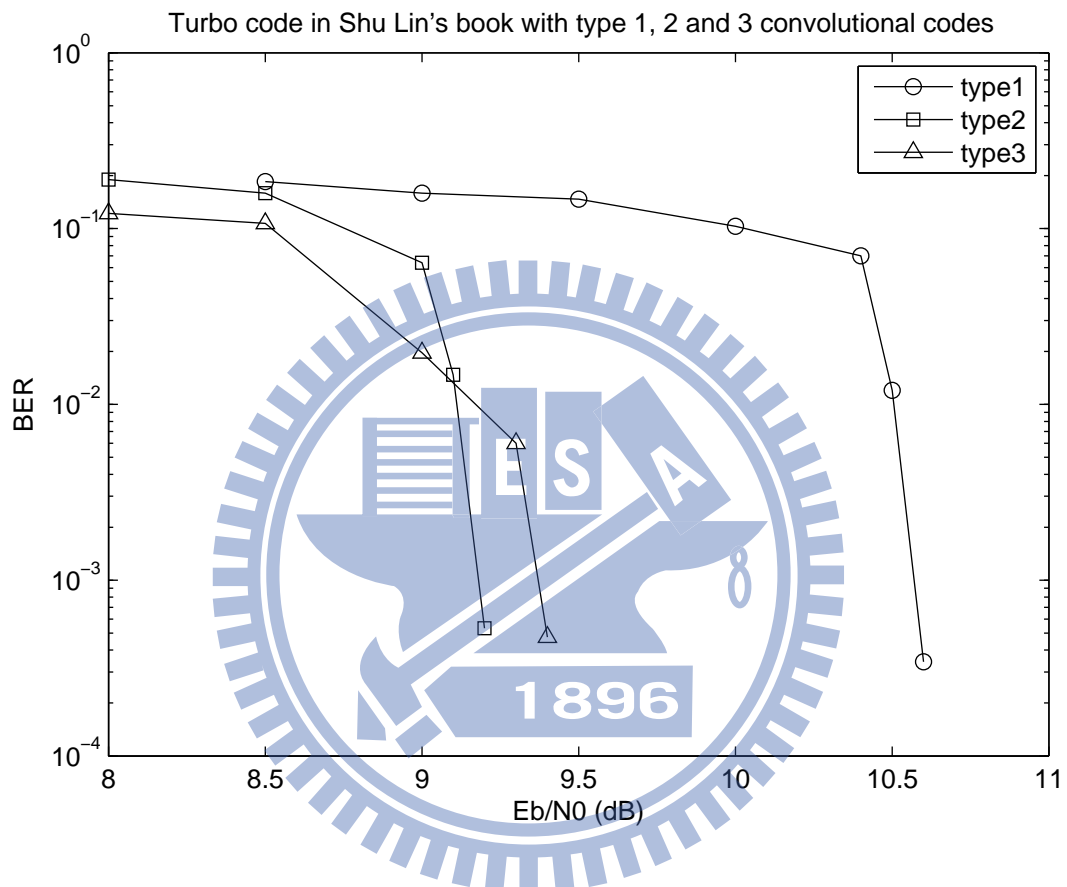
Figure 4.7: Performances of the concatenated coding system. The turbo code used is the one specified in [11]. The codeword lengths of the type 1, 2 and 3 convolutional codes are ten, eight and six, respectively. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3.

Figure 4.8: Performances of a conventional system that employs the $(37, 21)$ turbo code [7] with a least square (LS) channel estimator instead of our non-coherent convolutional code. Here, type 1, type 2 and type 3 similarly denote the sizes of the LS estimator, i.e., 10, 8 and 6. The inner LS channel estimator rate is $2/3$ and the outer turbo code rate is $1/3$. The curves in Figure 4.5 are also illustrated for comparison.
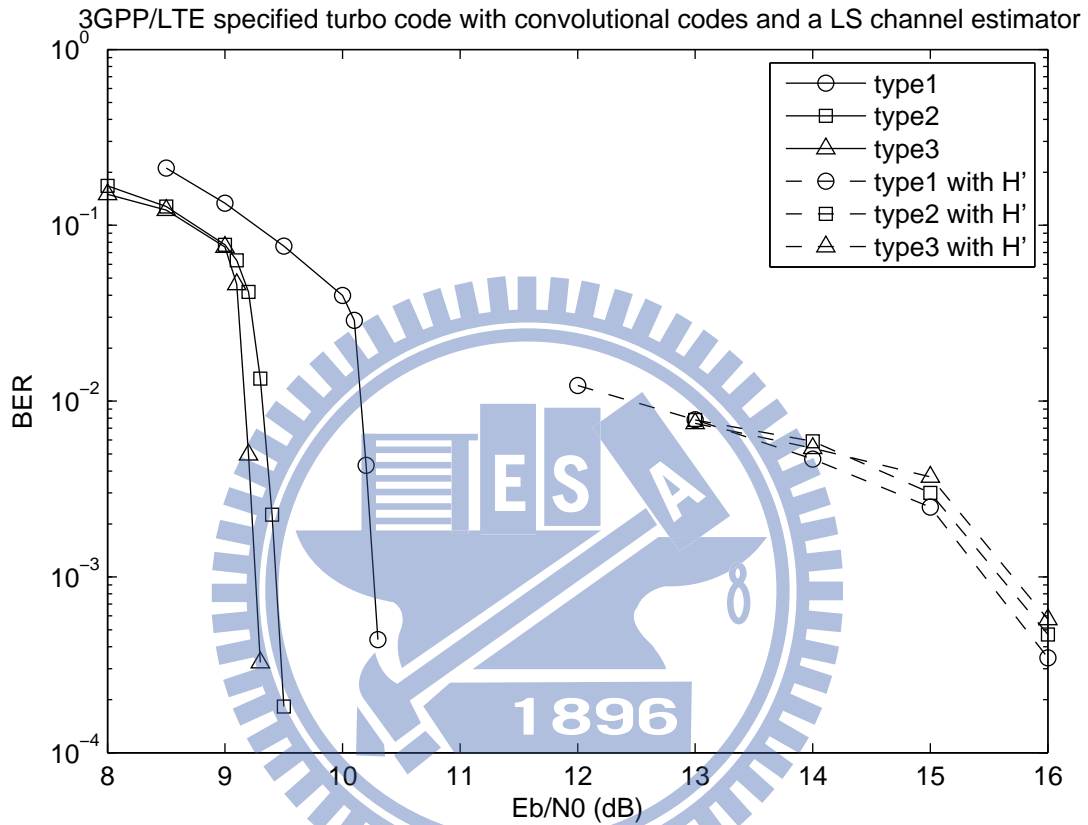
Figure 4.9: Performances of a conventional system that employs the 3GPP/LTE specified turbo code [12] with a least square (LS) channel estimator instead of our non-coherent convolutional code. Here, type 1, type 2 and type 3 similarly denote the sizes of the LS estimator, i.e., 10, 8 and 6. The inner LS channel estimator rate is 2/3 and the outer turbo code rate is 1/3. The curves in Figure 4.6 are also illustrated for comparison.
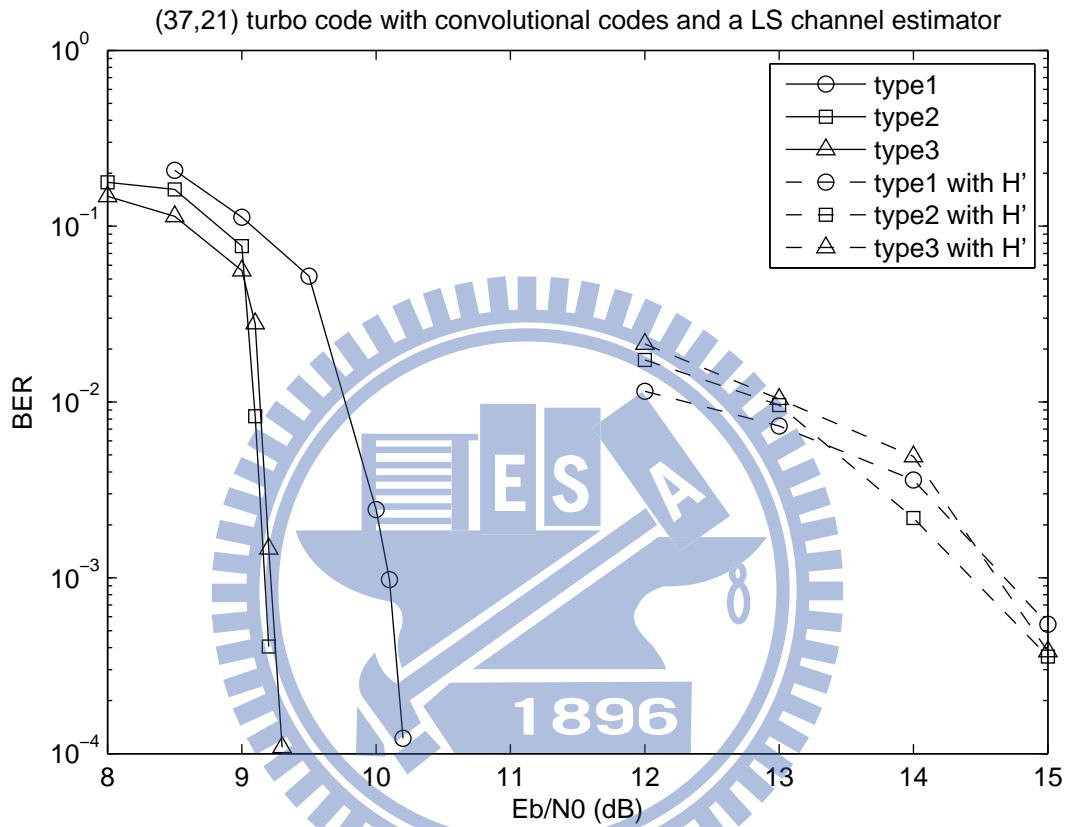
Figure 4.10: Performances of a conventional system that employs the turbo code from [11] with a least square (LS) channel estimator instead of our non-coherent convolutional code. Here, type 1, type 2 and type 3 similarly denote the sizes of the LS estimator, i.e., 10, 8 and 6. The inner LS channel estimator rate is 2/3 and the outer turbo code rate is 1/3. The curves in Figure 4.7 are also illustrated for comparison.
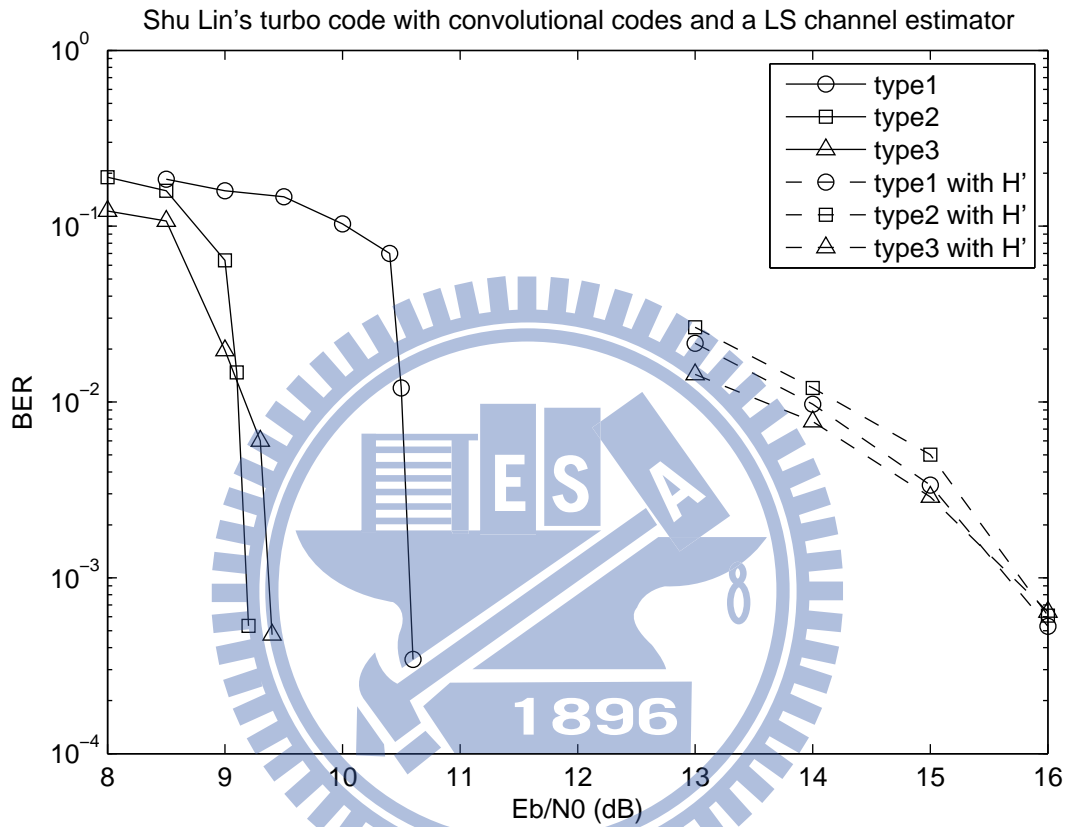
4.2 Remarks

In this subsection, we will remark on the simulation results presented in the previous sub-section.

As having been pointed out, we can observe from Figures 4.1 and 4.3 that the minimum pairwise non-coherent distance does not seem to be a decisive factor for codes' performances. On the one hand, Figure 4.1 hints that the code with the largest minimum pairwise non-coherent distance, i.e., "5-5", gives the worst performance. On the other hand, Figure 4.3 shows the opposite as the code with larger minimum pairwise non-coherent distances usually perform better. A better criterion that can well-connect the pairwise non-coherent distances and the codes' performances should be adopted. So we turn to the so-called cumulative distance function, based on which the method to select non-coherent convolutional codes is proposed (see (3.2)).

We then apply the convolutional code selection method defined in (3.2) for different codeword lengths, varying from $N = 6$ to $N = 12$. The selection results, together with the simulated optimal designs, are listed in Table 4.1. Notably, it is sometimes hard to identify the optimal code design from simulations because their performance curves are close to each other, and one code may outperform the other code in one region but worse in the other region. If such occurs, we will list two or even three as the optimal code designs. There however exist certain code designs that exhibit "strange" CDF patterns like "5-5" and "7-7", of which the performances are very bad. Particularly, these two codes have many sudden increases in their CDFs.

Next, from Figures 4.5, 4.6 and 4.7, we observed that type 1 code always perform the worst among the tested three types, which seems to hint that our design favors a shorter code. However, there is no apparent winner between type 2 and type 3 codes. To be specific, type

Table 4.1: Convolutional codes identified by the proposed method in (3.2) and the optimal convolutional code designs obtained from simulations. The codeword lengths vary from $N = 6$ to $N = 12$. The code rate is $1/2$.

type	codeword length	code selected	optimal code
3	$N = 6$	7-5	5-6 or 7-5
2	$N = 8$	5-6	4-7 or 5-6 or 7-5
1	$N = 10$	7-4	7-4
0	$N = 12$	7-4	5-7 or 7-4 or 7-5

Table 4.2: Convolutional codes identified by the proposed method in (3.2). The codeword lengths vary from $N = 6$ to $N = 12$. The code rate is $2/3$. Since there are many choices for rate $2/3$ codes, to identify the optimal code structures via simulations turns out to be infeasible. Hence, we did not show the optimal code design in this table.

type	codeword length	code selected
3	$N = 6$	7-6; 7-5
2	$N = 8$	5-6; 7-5
1	$N = 10$	5-6; 7-5
0	$N = 12$	7-5; 7-6

2 code performs better than type 3 code when they are concatenated with the 3GPP/LTE specified turbo code; however, the winner changes when they are concatenated with the $(37, 21)$ turbo code and the turbo code in [11].

Although we infer that the concatenated coding system that we propose favors a shorter non-coherent convolutional code, the results on type 2 and 3 codes do not support this inference. Hence, we conduct another set of simulations by adding type 0 code, of which the codeword length is 12.

As such, in this series of simulations, four types of codeword lengths are used. They are type 0, type 1, type 2 and type 3, which respectively correspond to codeword lengths 12, 10, 8 and 6. The results are summarized in Figures 4.11, 4.12 and 4.13. These simulations then clearly show that the longest code, i.e., type 0, has the worst performance. However, by

following this trend, type 3 code should always have the best performance but it sometimes performs worse than the type 2 code. An possible cause could be that the the code selected by our method may not be the one with the optimal performance.

In order to verity this interpretation, we additionally perform the following simulations. We replace the non-coherent convolutional codes selected by our method with the optimal non-coherent convolutional codes obtained from simulations. The simulation results are illustrated in Figures 4.14, 4.15 and 4.16. Here, we only focus on codes of lengths 6 and 8. These figures clearly indicate that the shorter type 3 codes are always superior than the longer type 2 codes, thereby confirming our interpretation.

Finally, we remark from Figures 4.8, 4.9 and 4.10 that the performances using our non-coherent convolutional code together with the turbo code are seemingly better than using the LS estimators with the same turbo code. This implies that when the estimation window size of the LS estimator is short such as 6 and 8, it may be advantageous to consider replacing the conventional LS estimator with our non-coherent convolutional code.
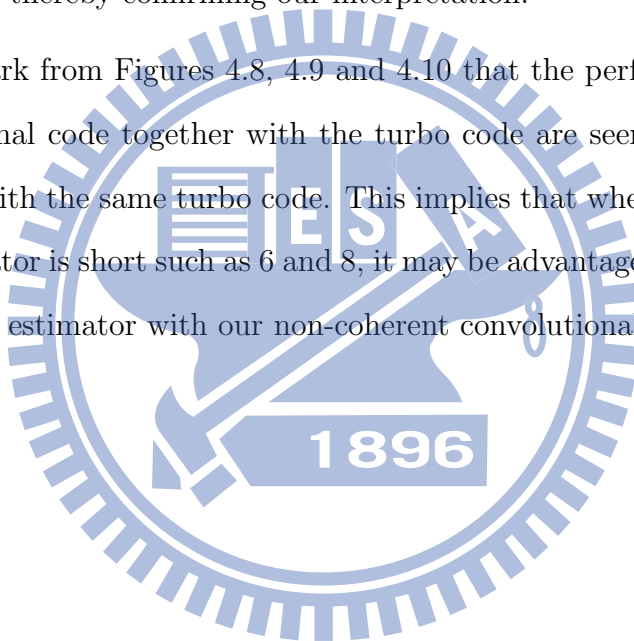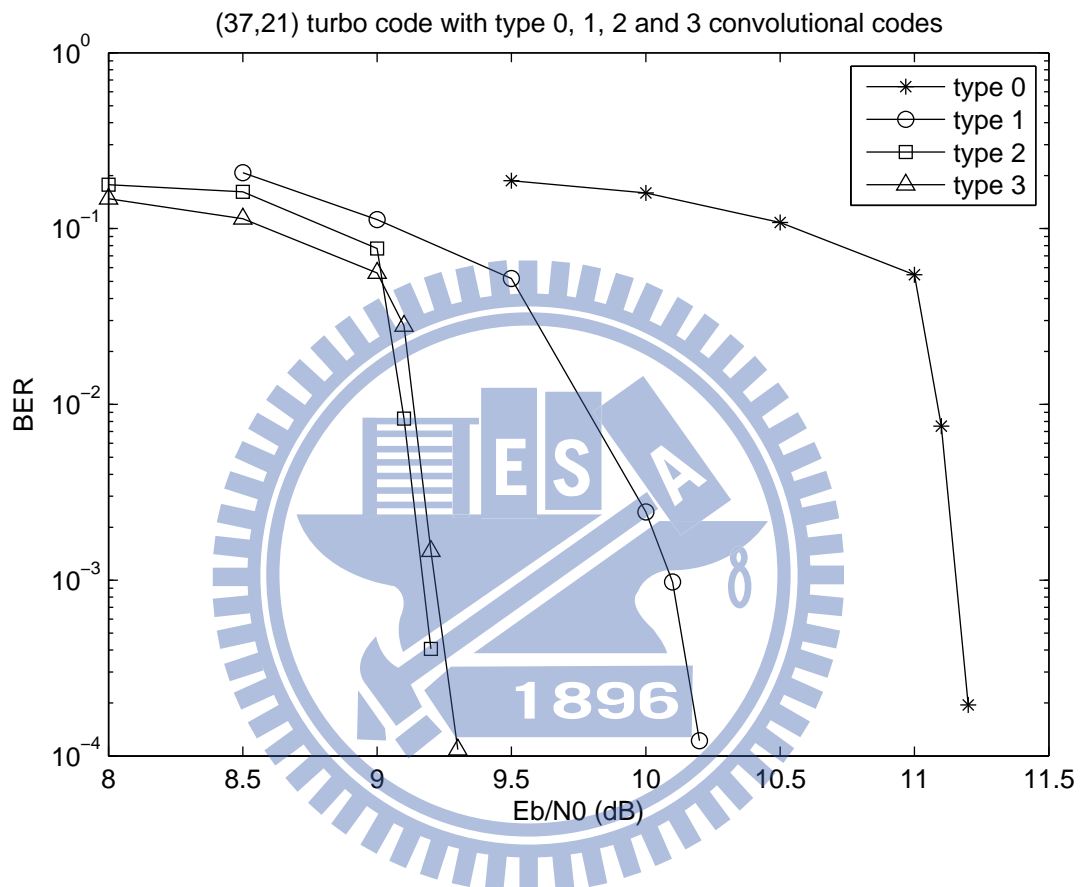
Figure 4.11: Performances of the concatenated system that employs the $(37, 21)$ turbo code [7] concatenated with the non-coherent convolutional codes of type 0, 1, 2, 3. Here, the codeword lengths at type 0, 1, 2, 3 codes are respectively 12, 10, 8 and 6. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3
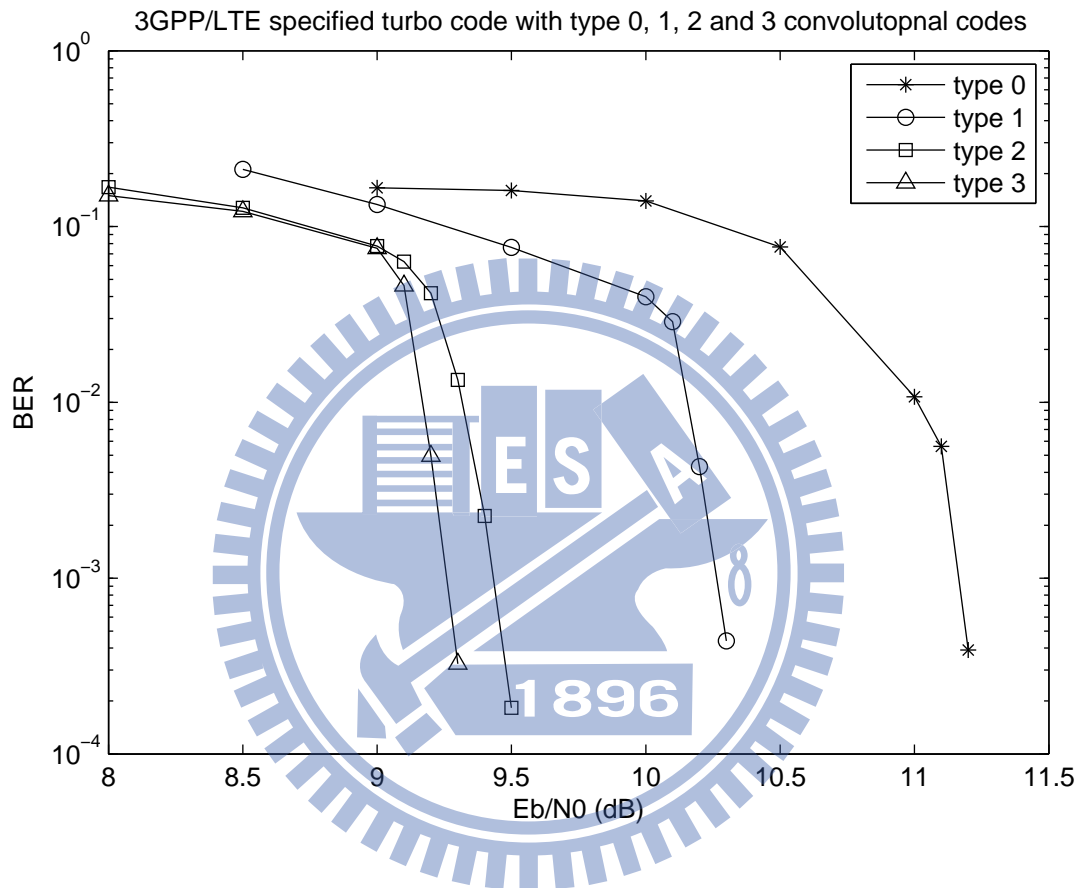
Figure 4.12: Performances of the concatenated system that employs the 3GPP/LTE specified turbo code [12] concatenated with the non-coherent convolutional codes of type 0, 1, 2, 3. Here, the codeword lengths at type 0, 1, 2, 3 codes are respectively 12, 10, 8 and 6. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3
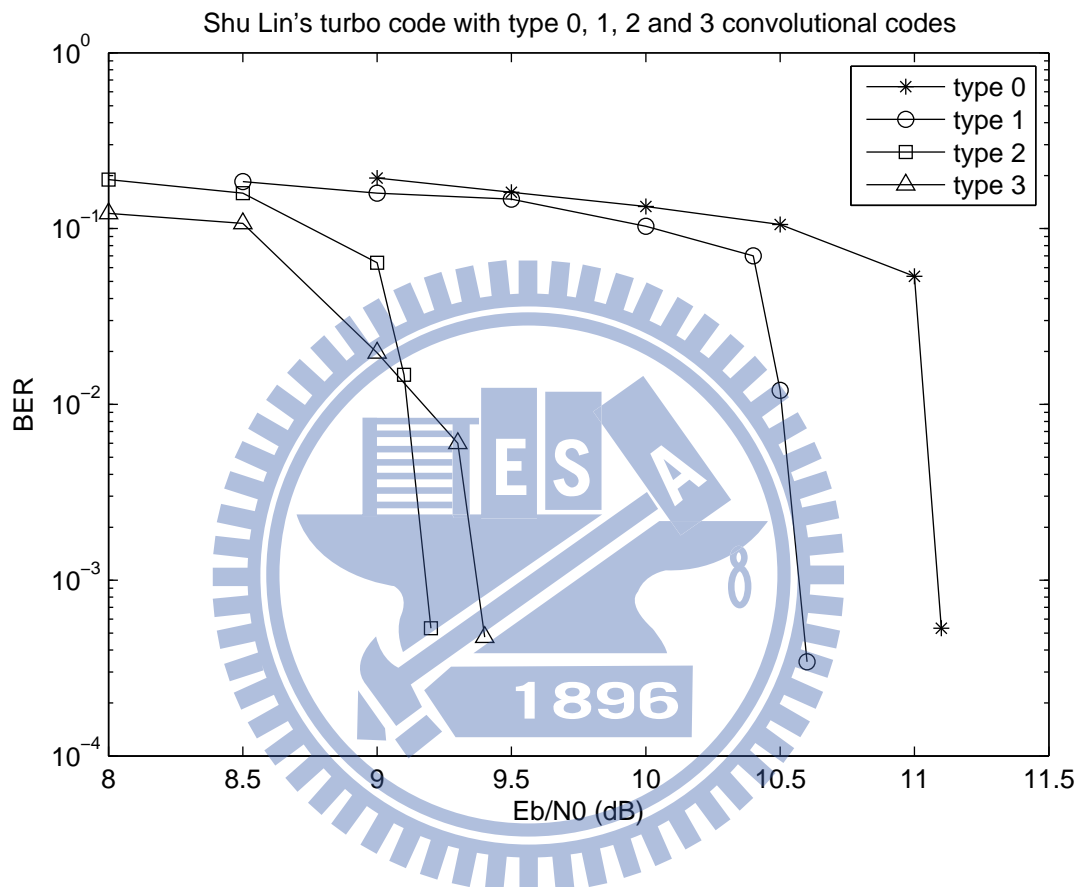
Figure 4.13: Performances of the concatenated system that employs the turbo code from [11] concatenated with the non-coherent convolutional codes of type 0, 1, 2, 3. Here, the codeword lengths at type 0, 1, 2, 3 codes are respectively 12, 10, 8 and 6. The inner convolutional code rate is 2/3 and the outer turbo code rate is 1/3
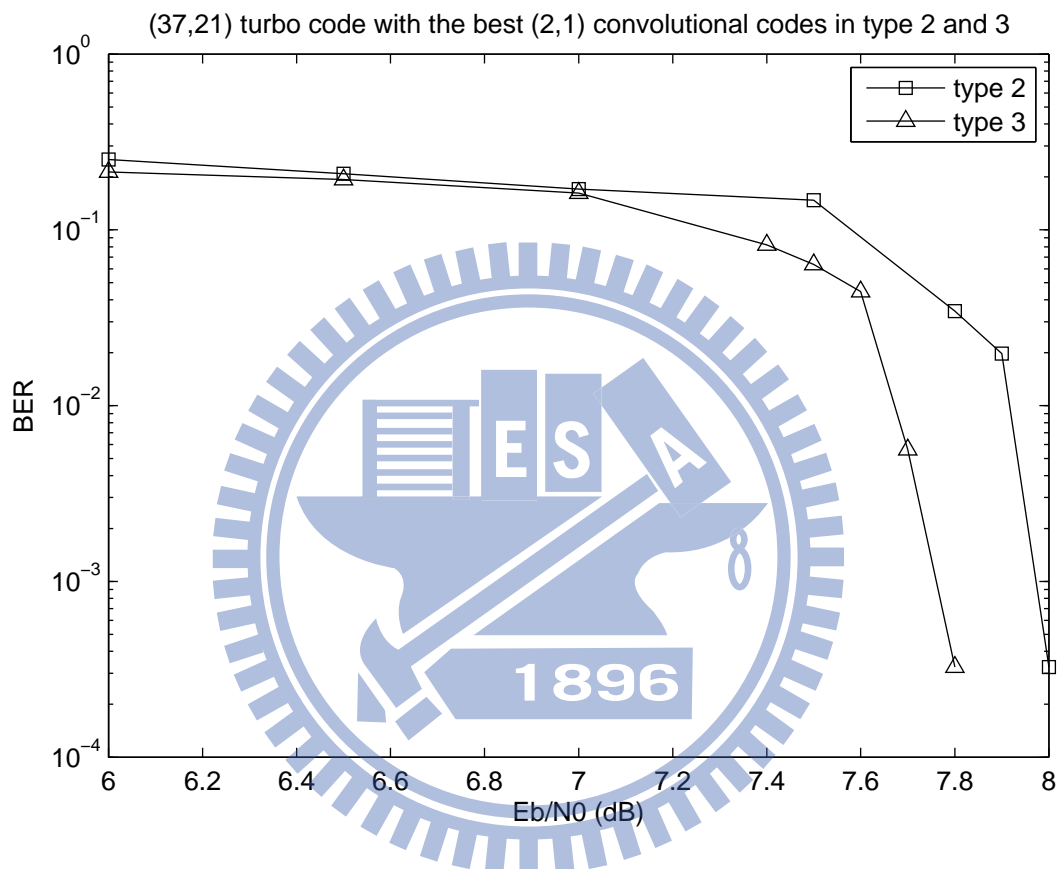
Figure 4.14: Performances of the concatenated system that employs the $(37, 21)$ turbo code [7] concatenated with the optimal non-coherent convolutional codes of type 2 and 3. The inner convolutional code rate is $1/2$ and the outer turbo code rate is $1/3$

Figure 4.15: Performances of the concatenated system that employs the 3GPP/LTE specified turbo code [12] concatenated with the optimal non-coherent convolutional codes of type 2 and 3. The inner convolutional code rate is 1/2 and the outer turbo code rate is 1/3

Figure 4.16: Performances of the concatenated system that employs the turbo code from [11] concatenated with the optimal non-coherent convolutional codes of type 2 and 3. The inner convolutional code rate is 1/2 and the outer turbo code rate is 1/3
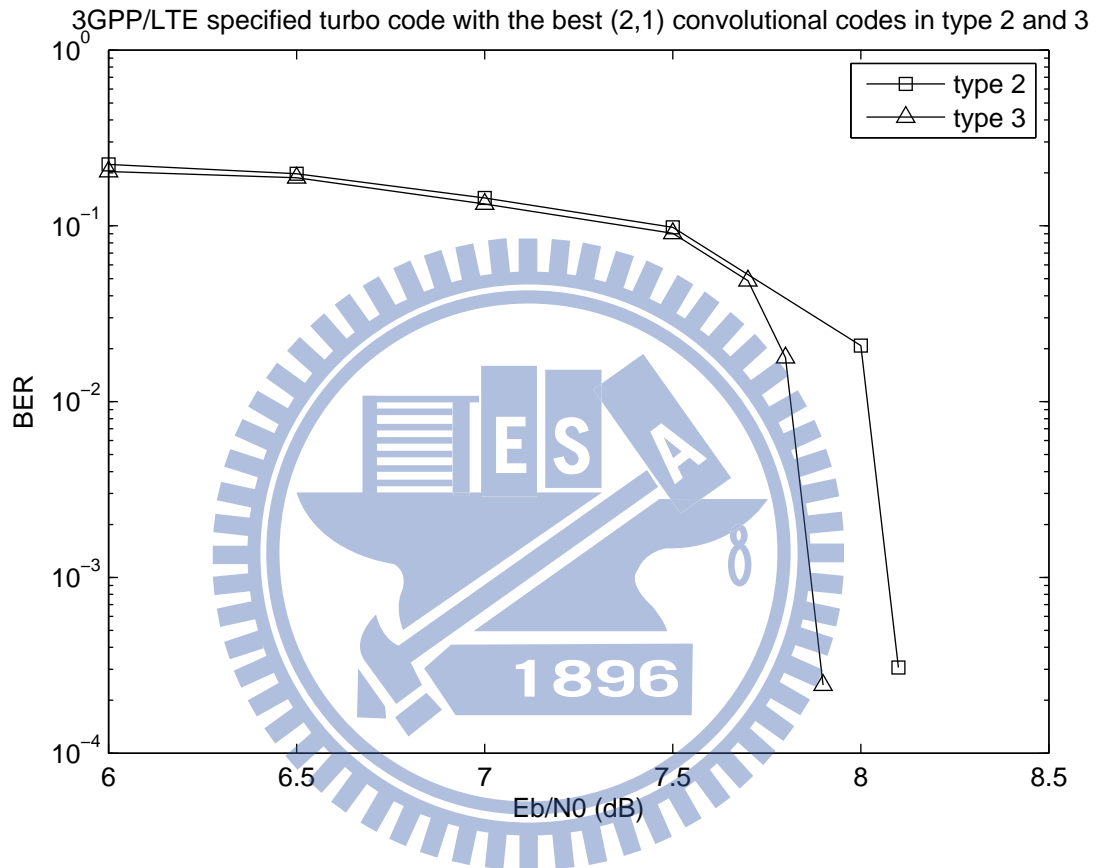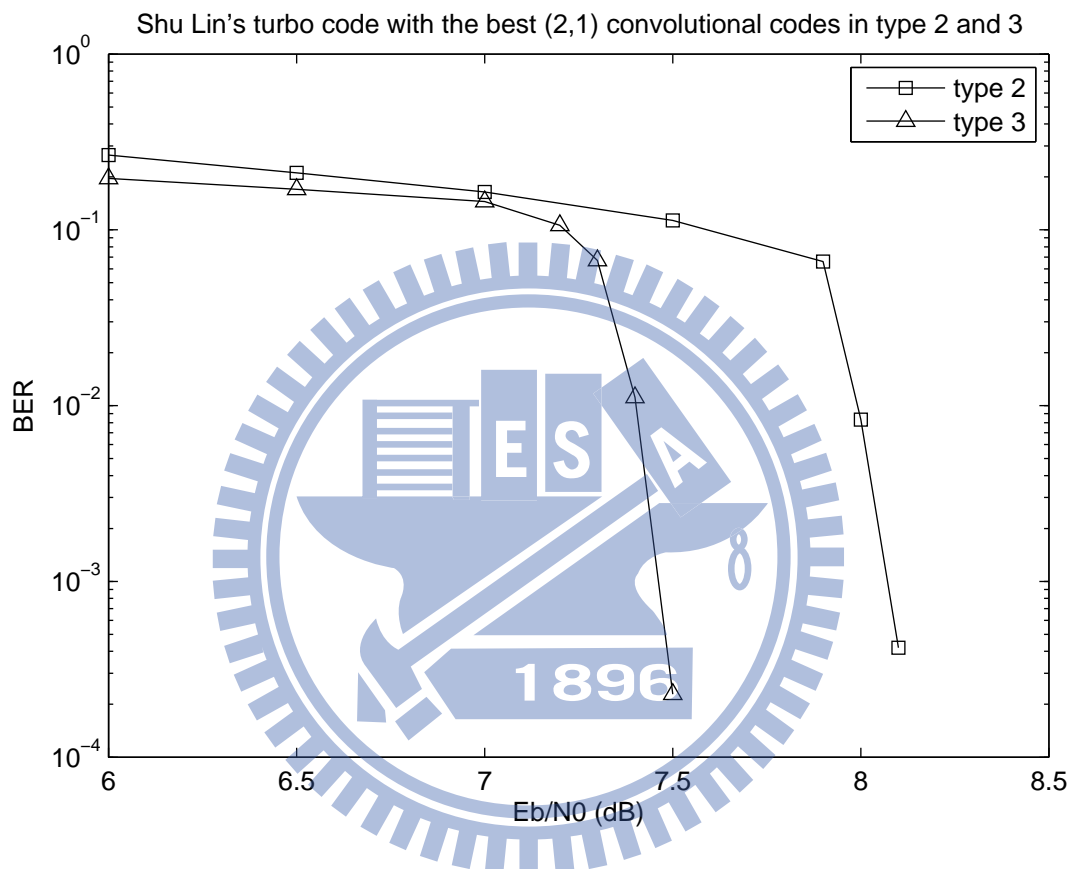
Chapter 5

Conclusion and Future Work

In this thesis, we proposed a simple non-coherent-distance-based method to select a good convolutional code to co-work with the outer turbo code in an unknown frequency-selective fading environment. By using the selected inner non-coherent convolutional code and the corresponding GLRT decoding criterion, a better performance than the conventional LS estimator can be obtained when the codeword length/estimation window size is small. When the code rate is half, the codes that our method locates are the ones with the simulated optimal performance. When the code rate increases to 2/3, however, our method may locate the second-best convolutional code structure in the sense of simulated performance. A side observation from our simulations that the proposed concatenated system prefers using a shorter inner convolutional code. This could be due to that in a dynamic fading environment, it may be better to perform frequent channel estimation based on a smaller window and resort the error protection task to the long outer turbo code.

Our experiments hint that when code rates increase, our method may not select the best code in performance. So some modification may be needed when the code rates further increase. This shall be a useful future research direction. In addition, we restrict the modulation schemes to BPSK in our design and simulations. It shall be more practical to extend our design to a higher order modulations like QPSK and 16-QAM.

Bibliography

[1] O. Shayevitz and M. Feder, "Universal decoding for frequency-selective fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2770 - 2790, August 2005.

[2] Haris Gacanin, Mika Salmela and Fumiyuki Adachi, "Performance analysis of analog network coding with imperfect channel estimation in a frequency-selective fading channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp.742 - 750, February 2012.

[3] Christos Komninakis, Christina Fragouli, Ali H. Sayed and Richard D. Wesell, "Channel estimation and equalization in fading," *The Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1159 - 1163, 24-27 October 1999.

[4] Dilip Warrier and Upamanyu Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651 - 668, March 2002.

[5] Chia-Long Wu, Po-Ning Chen, Yunghsiang S. Han and Ming-Hsin Kao, "Maximum-likelihood priority-first search decodable codes for combined channel estimation and error correction," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4191 - 4203, September 2009.

[6] T. Chu, "Joint data detection and channel estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 670 - 679, April 2006.

[7] Claude Berrou and Alain Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261 - 1271, October 1996.

[8] Jochen Giese and Mikael Skoglund, "Space-time code design for unknown frequency-selective channels," *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, pp. III-2413 - III-2416, 13-17 May 2002.

[9] S. M. Kay, *Fundamental of Statistical Signal of Processing, vol. 1 : Estimation Theory,* Prentice Hall PTR, 1993.

[10] S. M. Kay, *Fundamental of Statistical Signal of Processing, vol. 2 : Detection Theory,* Prentice Hall PTR, 1998.

[11] S. Lin and D. J. Costello, *Error Control Coding,* 2nd edition, Prentice Hall, 2004.

[12] *3GPP Technical Specification: Group Radio Access Network, Evolved Universal Terrestrial Radio Access, Multiplexing and Channel Coding (Release 8),* TS 36.212 v8.3.0, May 2007.