ADVANCED CODING AND MODULATION FOR ULTRA-WIDEBAND AND IMPULSIVE NOISES

by

LIBO YANG
B.S.E.E. Nanjing University of Posts & Telecommunications, 2000
M.S.E.E. Nanjing University of Posts & Telecommunications, 2003

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the School of Electrical Engineering and Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2007

Major Professor: Lei Wei

ABSTRACT

The ever-growing demand for higher quality and faster multimedia content delivery over short distances in home environments drives the quest for higher data rates in wireless personal area networks (WPANs). One of the candidate IEEE 802.15.3a WPAN proposals support data rates up to 480 Mbps by using punctured convolutional codes with quadrature phase shift keying (QPSK) modulation for a multi-band orthogonal frequency-division multiplexing (MB-OFDM) system over ultra wideband (UWB) channels. In the first part of this dissertation, we combine more powerful near-Shannon-limit turbo codes with bandwidth efficient trellis coded modulation, i.e., turbo trellis coded modulation (TTCM), to further improve the data rates up to 1.2 Gbps. A modified iterative decoder for this TTCM coded MB-OFDM system is proposed and its bit error rate performance under various impulsive noises over both Gaussian and UWB channel is extensively investigated, especially in mismatched scenarios. A robust decoder which is immune to noise mismatch is provided based on comparison of impulsive noises in time domain and frequency domain.

The accurate estimation of the dynamic noise model could be very difficult or impossible at the receiver, thus a significant performance degradation may occur due to noise mismatch. In the second part of this dissertation, we prove that the minimax decoder in [38], which instead of minimizing the average bit error probability aims at minimizing the worst bit error probability, is optimal and robust to certain noise model with unknown prior probabilities in two and higher dimensions.

Besides turbo codes, another kind of error correcting codes which approach the Shannon capacity is low-density parity-check (LDPC) codes. In the last part of this dissertation,

we extend the density evolution method for sum-product decoding using mismatched noises. We will prove that as long as the true noise type and the estimated noise type used in the decoder are both binary-input memoryless output symmetric channels, the output from mismatched log-likelihood ratio (LLR) computation is also symmetric. We will show the Shannon capacity can be evaluated for mismatched LLR computation and it can be reduced if the mismatched LLR computation is not an one-to-one mapping function. We will derive the Shannon capacity, threshold and stable condition of LDPC codes for mismatched BI-AWGN and BIL noise types. The results show that the noise variance estimation errors will not affect the Shannon capacity and stable condition, but the errors do reduce the threshold. The mismatch in noise type will only reduce Shannon capacity when LLR computation is based on BIL.

To my parents, Mama Hu Qingyun and Papa Yang Xingrong

ACKNOWLEDGMENTS

First, I want to express my great appreciation and gratitude to my supervisor, Prof. Lei Wei for his inspiration, discussion and support during my studies, which made this dissertation possible. Also, I want to thank all other professors in my dissertation committee, Prof. W. Linwood Jones, Prof. Takis C. Kasparis, Prof. Xin Li, and Prof. Jiann S. Yuan for their valuable feedback and help.

I always enjoy working with a group of smart young people in our lab, including Yanxia, Burak, Balaji, etc., from whom I learned a lot through discussion and cooperation during the last few years. I really appreciate the effort of many other faculty, staff and fellow students in the Electrical and Computer Engineering department, who makes here a wonderful place for study. I also want to thank my friends, especially Min Lu and Tao Li, for friendship which makes life in Orlando a happy one.

Finally, my deepest love and gratitude is devoted to my parents, Qingyun Hu and Xingrong Yang. They always extend their love, supporting and encouragement from thousands of miles away. To them, I dedicate this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS/ABBREVIATIONS

3GPP – Third-Generation Partnership Project

ADSL – Asymmetric Digital Subscriber Line

APP – A Posteriori Probability

AR – Auto-Regressive

BEP – Bit Error Probability

BER – Bit Error Rate

BIAWGN – Binary Input Additive White Gaussian Noise

BIL – Binary Input Laplace

BPSK – Binary Phase Shift Keying

BSC – Binary Symmetric Channel

CCSDS – Consultative Committee for Space Data Systems

CP – Cyclic Prefix

CP-OFDM – Cyclic Prefix - Orthogonal Frequency-Division Modulation

CSI – Channel State Information

DAB – Digital Audio Broadcasting

DVB – Digital Video Broadcasting

DMT – Discrete Multi-Tone

DFT – Discrete Fourier Transform

EIRP – Efficient Isotropically-Radiated Power

FCC – Federal Communications Commission

FFT – Fast Fourier Transform

FIR – Finite Impulse Response

GPS – Global Positioning System

ICI – Inter-Channel Interference

IFFT – Inverse Fast Fourier Transform

IR-UWB – Impulse Radio Ultra-WideBand

ISI – Inter-Symbol Interference

LDPC – Low-Density Parity-Check codes

LLR – Log-Likelihood Ratio

LOS – Line-Of-Sight

MB-OFDM – Multi-Band Orthogonal Frequency-Division Modulation

MAP – Maximum A Posteriori

ML – Maximum Likelihood

MPSK – Multi-Phase Shift Keying

MSED – Minimum Squared Euclidean Distance

OFDM – Orthogonal Frequency Division Modulation

PAM – Pulse Amplitude Modulation

PCCC – Parallel Concatenated Convolutional Codes

PCS – Personal Communications Service

PDF – Probability Density Function

QAM – Quadrature Amplitude Modulation

RA – Repeat Accumulate codes

RSC – Recursive Systematic Convolutional codes

SISO – Soft-Input Soft-Output

SNR – Signal-to-Noise Ratio

SPC – Single Parity-Check codes

TCM – Trellis Coded Modulation

TTCM – Turbo Trellis Coded Modulation

WLAN – Wireless Local Area Network

WPAN – Wireless Personal Area Network

UMTS – Universal Mobile Telecommunication Systems

UWB – Ultra-WideBand

ZP-OFDM – Zero-padding Prefix - Orthogonal Frequency-Division Modulation

CHAPTER 1: INTRODUCTION

1.1 Motivations

In 1948, Claude E. Shannon built the mathematical foundation of information theory [1], in which one of the groundbreaking results established bounds for the maximum amount of information that can be transmitted reliably over noisy channels, also called *channel capacity*. Shannon proved the existence of codes that arbitrarily approach capacity for which the probability of error of the optimal maximum likelihood (ML) decoder goes to zero as the block length of the code goes to infinity.

However, Shannon's theorems are non-constructive because random code ensemble is assumed in his proof and thus don't give any method on how to find explicit capacity approaching codes. Though randomly constructed codes have been shown to be good codes with high probability, we need to maintain a codebook describing which information bit sequence (k-tuple) is mapped to which codeword (n-tuple), which is too long to handle in practical systems when block length becomes large. It is also not likely to be practically decodable. For non-random code ensemble, linear codes are equipped with efficient encoding algorithm due to their intrinsic algebraic code structure but so far no polynomial-time algorithm is known under ML decoding algorithms.

A different way is to first find a code ensemble which can be decoded by a sub-optimal algorithm efficiently. Then we find those good (capacity approaching) codes within that ensemble based on further performance analysis. The low-density parity-check (LDPC) codes discovered in the early 1960's by Gallager in his PhD dissertation [2] are actually

within such a kind of code ensemble. LDPC codes are largely forgotten after their invention until recently. The discovery of turbo codes [3] led to a revival of LDPC codes [4] which were shown to achieve capacity using low-complexity decoders. It is by nature equipped with efficient iterative decoding algorithms due to its sparse structure. The task then is to find those codes in the LDPC ensemble that perform very close to Shannon capacity. *Density evolution* [5] is an important tool to evaluate asysmptotic ensemble performance of LDPC codes when iterative decoding is employed and also to find good codes from the code ensemble. The best LDPC codes found by density evolution is within 0.0045 dB away from the Shannon limit [6].

In the digital home of the not-too-distant future, people will be sharing photos, music, video, data and voice among networked consumer electronics, PCs and mobile devices throughout the home and even remotely. Requirements for the digital home include high-speed data transfer for multimedia content, short-range connectivity for transfer to other devices, low power consumption due to limited battery capacity, and low complexity and cost due to market pricing pressures and alternative wired connectivity options. Transfer of video from a camcorder to an entertainment PC is one scenario. The ability to view photos from the user's digital camera on a larger HDTV display is the second. Removing all the wires to the printer, scanner, mass storage devices, and video cameras located in the home office is another possible scenario.

Current technologies, like Bluetooth or wireless local area networks (WLANs), are simply not fast enough for wireless multimedia distribution. A leading candidate for short-range, wireless personal area networks (WPANs) is ultra-wideband (UWB), which features

large capacity, low power, low interference, low probability of detection and excellent multipath immunity. In recent years, UWB is making the transition from laboratories to standardization after the US Federal Communications Commission (FCC) allocated a range of bandwidth from 3.1 GHz to 10.6 GHz for unlicensed use by UWB radios at a limited transmit power of -41.3 dBm/MHz or less [7], which is a key step towards the development of real-world products.

While a traditional carrierless UWB transmitter works by sending billions of pulses across a very wide spectrum of frequency several GHz in bandwidth [8][9][10][11][12][13], a multi-carrier system based on orthogonal frequency division multiplexing (OFDM) is currently under intensive discussion. For highly dispersive channels like UWB, an OFDM scheme is more efficient at capturing multipath energy than an equivalent single-carrier system using the same total bandwidth [14][15][16][17]. OFDM also can eliminate inter-symbol interference (ISI) caused by multipath without using complicated equalization as in the conventional single carrier receiver [18][19]. OFDM systems posses additional desirable properties, such as high spectral efficiency, inherent resilience to narrowband RF interface and spectrum flexibility. One of the candidate proposals for the IEEE 802.15 WPAN project [20] is a multi-band OFDM (MB-OFDM) system over UWB channel supporting data rate up to 480 Mbps. A major drawback of OFDM, however, is that it alone cannot take advantage of the diversity provided by the multipath fading in UWB environment. Appropriate frequency interleaving and coding becomes necessary for OFDM systems [21].

The current MB-OFDM proposal for IEEE 802.15 WPAN suggests using punctured convolutional codes. Modern coding theory provides us more powerful error control codes

with reasonable complexity, such as turbo codes and LDPC codes, which we believe can improve the overall system performance. Motivated by the principle behind both codes, researchers have come up with many compound codes with excellent performance, such as serially concatenated codes [22], parallel concatenated codes [23][24], product code [25], turbo trellis coded modulation (TTCM) [26][27][28][29][30][31], multilevel codes [32][33][34], and parity-concatenated codes [35][36][37]. Among these compound codes, TTCM is an attractive coding scheme for high-data-rate applications, since it combines the impressive near Shannon limit error correcting capability of turbo codes with the high spectral efficiency property of trellis coded modulation (TCM).

In [35][36], it has been shown that single parity check (SPC) codes serially concatenated with convolutional codes can obtain the performance within 0.45 dB of the Shannon limit. Motivated by that, we explore the performance using a serial concatenation of SPC codes with TCM, which is a special case of TTCM. The objective is to develop a novel coding/decoding scheme suitable for WPAN applications, which require superior bit error rate performance over severe channels like UWB, with high bandwidth efficiency and low decoding complexity. Based on this idea, we will build up a TTCM coded MB-OFDM system and investigate its performance over both Gaussian and UWB channels via simulation and density evolution.

Although the assumption of white Gaussian noise is quite appropriate for many applications, it is well known that many noise environments arising in practice are poorly modeled by Gaussian statistics. In particular, man-made electromagnetic interference (or noise), and a great deal of natural interference as well, is basically "impulsive" in nature, i.e.,

it has a highly structured form, characterized by significant probabilities of large interference levels. This impulsive or structured character of the interference can significantly degrade the performance of conventional demodulation systems, which are designed to operate most effectively against the commonly assumed Gaussian background noise processes. By proper system design, this noise structure can often be exploited to yield better performance than would be obtained in the Gaussian case. In this dissertation we will investigate the impact of various impulsive noises on the above mentioned TTCM coded MB-OFDM system over both Gaussian and UWB channels. Performance loss due to mismatch is a major concern in this system.

Further, we investigate a kind of robust decoder under some mismatched cases. The normally optimal maximum *a posterior* (MAP) decoder is not always the best robust decoder when noise is a mixture of several noise types whose percentage is unclear or varying to the decoder. So we want to evaluate the robustness of decoders especially when there is a mismatch between the actual and estimated noise parameters. We find that decoding algorithms based on the minimax rule is robust and optimal for the above mentioned situation.

Since optimum systems for impulsive channels are usually nonlinear, the analysis of such problems is often significantly more difficult than in the conventional Gaussian case. Lastly, we want to do some capacity analysis for another class of important error correcting codes, LDPC codes, via density evolution methods when there is noise mismatch. We want to know the effects of mismatched noises on the capacity of LDPC codes under sum-product decoding algorithm.

1.2 Dissertation Outline

In this dissertation we mainly focus on combining various advanced coding and modulation techniques in ultra-wideband environment for high-speed, short-range applications and investigate the performance and robustness of decoder when impaired with impulsive noise and its mismatch. This dissertation has totally six chapters, which are organized as follows:

Chapter 2 introduces all the techniques involved in this dissertation. It starts with a brief review of the development of ultra-wideband technology covering both carrierless and carrier-modulated UWB, where our MB-OFDM schemes are within the latter category. Another component of our WPAN system, OFDM will be discussed next with applications. Then we present some previous works done related to trellis coded modulation (TCM) scheme for bandwidth-limited communication systems, including Ungerboeck's TCM, up to the recent parity-concatenated TCM codes. Following this, the principle of turbo codes is elaborated including both encoder and decoder structures. Different ways of combining turbo codes and TCM, which lead to different turbo trellis coded modulation (TTCM) schemes, will be described thereafter. Their corresponding decoding algorithms are mentioned as well. This will be the starting point for deriving our own TTCM codes in Chapter 3. Finally we conclude Chapter 2 by an introduction to another class of important correct correcting codes, LDPC codes and especially message-passing algorithm and density evolution techniques, which will be used in performance analysis in Chapter 5.

In Chapter 3 we substitute the punctured convolutional codes suggested by one of the IEEE 802.15.3a WPAN proposals with our new parity-concatenated TCM codes, which

function as TTCM codes. We come up with an MB-OFDM system with frequency hopping, turbo trellis coded modulation, bit as well as symbol level interleaving, OFDM and equalization. We explain the flow of data processing and the advantage compared with traditional OFDM systems. Decoding algorithm works in an iterative way by updating the soft input and soft output information of each component decoder. It is observed that the soft information follows Gaussian distribution with changing means and variances, then density evolution can be used to predict the system performance quite well compared with simulation. The next part is investigation on the effect of a generalized Gaussian noise with various impulsiveness. An optimal branch metric is derived based on the noise structure, which is used to enhance decoding performance. Extensive simulation is done for different cases of noise match and mismatch over both Gaussian and UWB channels. Finally robust solutions for Gaussian and UWB channels are provided by comparing impulsive noises in time domain with those in frequency domain.

Further investigation on robust decoder is presented in Chapter 4. The well known MAP decoder is considered optimal to minimize the average bit error probability by using the knowledge of the noise structure. What if actual noise is a mixture of several types of noise with unknown percentage? In this case, we are unable to design an optimal decoder based on MAP criterion. However, if we look at the problem from an entirely different angle of view, we actually can build a decoder based on minimax criterion, which means minimizing the maximum bit error probability, instead of the average one. A decoder based on minimax rule has been proposed in [38]. In this chapter, we will prove that a minimax decoder is optimal and robust in both 2-dimensional and K-dimensional cases.

Chapter 5 extends the density evolution method for sum-product decoding using mismatched noises. We will prove that as long as the true noise type and the estimated noise type used in the decoder are binary-input memoryless output symmetric channels, e.g. the BIAWGN channel and the BIL channel, the output from mismatched log-likelihood ratio (LLR) computation is also symmetric. We will show the Shannon capacity can be evaluated for mismatched LLR computation. If the mismatched LLR computation is not an one-to-one mapping function, then the capacity can be reduced. We will derive the Shannon capacity, threshold and stable condition of LDPC codes for mismatched BIAWGN and BIL noise types. The results show that the noise variance estimation errors will not affect the Shannon capacity and stable condition, but the errors do reduce the threshold. The mismatch in noise type will only reduce Shannon capacity when LLR computation is based on BIL.

Finally, in Chapter 6 we discuss and summarize the contents of this dissertation followed by future research directions.

1.3 Contributions

The major contributions in this dissertation are listed as follows:

(1) Density evolution method is applied to our parity-concatenated TCM coded MB-OFDM sytem over both Gaussian and UWB channels. From it, we can predict system performance pretty well, which can save a lot of time-consuming simulations. (Chapter 3)

(2) We simulated TTCM coded MB-OFDM systems over both Gaussian and UWB channels impaired by impulsive noises. We discovered impulsive noise in time domain could be

transformed to Gaussian noise under general condition. Thus, the optimal decoder is established when its branch metric computation matched to noise model of its decision statistics. OFDM demodulation typically transforms time domain noise into frequency domain noise. Consequently, a robust decoder can be established when we apply the theory in [38] to noise type in frequency domain. (Chapter 3)

(3) We further investigated equivalence between time domain impulsive noises and frequency domain impulsive noises for other noise scenarios. Thus, the system designer can have a better guidance on how to select metric to achieve robust decoding. (Chapter 3)

(4) We have proved mathematically that the minimax robust decoder in [38] is optimal for finite dimensional constellations (to 2-dimension first and then extended to general K-dimension). That is, the minimax robust decoder in [38] yields minimum worst case bit error probability. (Chapter 4)

(5) We extended the density evolution method for LDPC codes under sum-product decoding using mismatched noises. We proved that as long as the true noise type and the estimated noise type used in the decoder are binary-input memoryless output symmetric channels, the output from mismatched log-likelihood ratio (LLR) computation is also symmetric. (Chapter 5)

(6) We showed the Shannon capacity can be evaluated for mismatched LLR computation. If the mismatched LLR computation is not an one-to-one mapping function, then the capacity can be reduced. (Chapter 5)

(7) We derived the Shannon capacity, threshold and stable condition of LDPC codes for mismatched BIAWGN and BIL noise types. The results show that the noise variance estimation errors will not affect the Shannon capacity and stable condition, but the errors do reduce the threshold. The mismatching in noise type will only reduce Shannon capacity when LLR computation is based on BIL. (Chapter 5)

1.4 Paper List

Journal Papers:

(1) Y. Wang, L. Yang, and L. Wei, "High-speed turbo-TCM-coded orthogonal frequency-division multiplexing ultra-wideband systems," *EURASIP Journal on Wireless Communications and Networking (Special Issue on UWB Communication Systems - Technology and Applications)*, 2006.

(2) L. Yang and L. Wei, "Capacity of LDPC codes under sum-product decoding and noise mismatch," *IEEE Transactions on Communications*, submitted.

(3) L. Yang, Y. Wang, and L. Wei, "Minimax robust turbo-TCM OFDM system on ultra-wideband channels with impulsive noises," in final stage of preparation.

Conference Papers:

(1) Y. Wang, L. Yang, and L. Wei. "High speed turbo TCM coded OFDM system for UWB channels," *IEEE International Symposium on Information Theory (ISIT 2005)*, Adelaide, Australia, Sept.4-9, 2005.

(2) Y. Wang, L. Yang, and L. Wei, "Turbo TCM coded OFDM system for powerline channels," *4th International Symposium on Turbo Codes and Related Topics (Turbo-Coding 2006)*, Munich, Germany, April 3-7, 2006.

(3) L. Yang and L. Wei, "Optimality on probabilistic minimax robust decoder," *International Symposium on Information Theory and Its Applications (ISITA 2006)*, Seoul, Korea, Oct 29 - Nov 1, 2006.

(4) L. Yang and L. Wei, "Capacity of LDPC codes under sum-product decoding and noise mismatch," *IEEE Military Communications Conference (MILCOM 2007)*, submitted.

(5) L. Yang and L. Wei, "Effects of time-domain and frequency-domain impulsive noises on turbo-TCM coded OFDM systems," in final stage of preparation.

CHAPTER 2: BACKGROUND AND RELATED WORK

2.1 Ultra-Wideband Technology

An increased interest from both major chip manufacturing companies and standardization bodies in Ultra-Wideband (UWB) technology has been witnessed in the last five years since the U.S. Federal Communications Commission (FCC) approved the first guidelines allowing the intentional emission of UWB signals contained within specified emission masks [7]. Traditionally, UWB signals have been generated by radiating very short pulses, rather than continuous waveforms. This technique, Impulsive Radio (IR), has been extensively used in radar applications driven by the promise of fine-range resolution that comes with large bandwidth. However, due to technological limitations and commercial pressure for reliable communications at that time, research and development were gradually shifted towards continuous-wave transmissions, and IR remained relatively confined to the radar field up to recent years. A complete survey of IR-UWB research in both radar and communication fields in its early years is presented by [39]. In the late 1980s, a few small companies reintroduced the idea of wireless communications based on the UWB concept and developed UWB technology following the IR paradigm, promoting the transmission of virtually carrierless and extremely short pulses [40]. The pioneering work to develop UWB as a commercial communications technology began at the University of Southern California during the mid-90s [9] [10]. Soon, several start-up companies, like Time Domain and Xtreme Spectrum produced the first chipset solutions. Finally, the FCC allowed UWB technology to be commercialized in April 2002, which is a major change in the regulation of RF emissions,

allowing a significant portion of the RF spectrum, originally allocated in many smaller bands exclusively for specific uses, to be effectively shared with low-power UWB radios.

According to the FCC's definition, a signal is considered UWB if its -10 dB bandwidth is greater than 20% of its center frequency or greater than 500 MHz. Because of its large bandwidth, the UWB waveform must friendly coexist with other waveforms from such as GPS(1.6GHz), PCS(1.9GHz), Bluetooth(2.4GHz), 802.11b(2.4GHz), cordless phone(2.4GHz), microwave ovens(2.4GHz), 802.11a(5.2GHz), etc., which are present in the air interface. The coexistence principle makes the FCC also defines rigid transmission power limits as shown in Figure 2.1, where the frequency range between 3.1 GHz and 10.6 GHz is allowed maximum effective isotropically-radiated power (EIRP), -41.3 dBm/MHz. The FCC's power requirement of -41.3 dBm/MHz, equal to 75 nanowatts/MHz for UWB systems, puts them in the category of unintentional radiators, such as TVs and computer monitors. Such power restriction allows UWB systems to reside below the noise floor of a typical narrowband receiver and enables UWB signals to coexist with current radio services with minimal or no interference. UWB also has some other significant advantages. The huge bandwidth provides the potential for extremely high data rate over short distances. Limit of low transmit power also provides the benefit of very low power consumption and frequency reuse. UWB is particularly well-suited for wireless personal area networks (WPANs). An area of particular interest is for wireless USB and wireless firewire applications, which eliminate cables in communications and computer equipments.

However, the UWB concept defined by the FCC is not limited to pulsed transmission, but can be extended to continuous-like transmission techniques, provided that the occupied

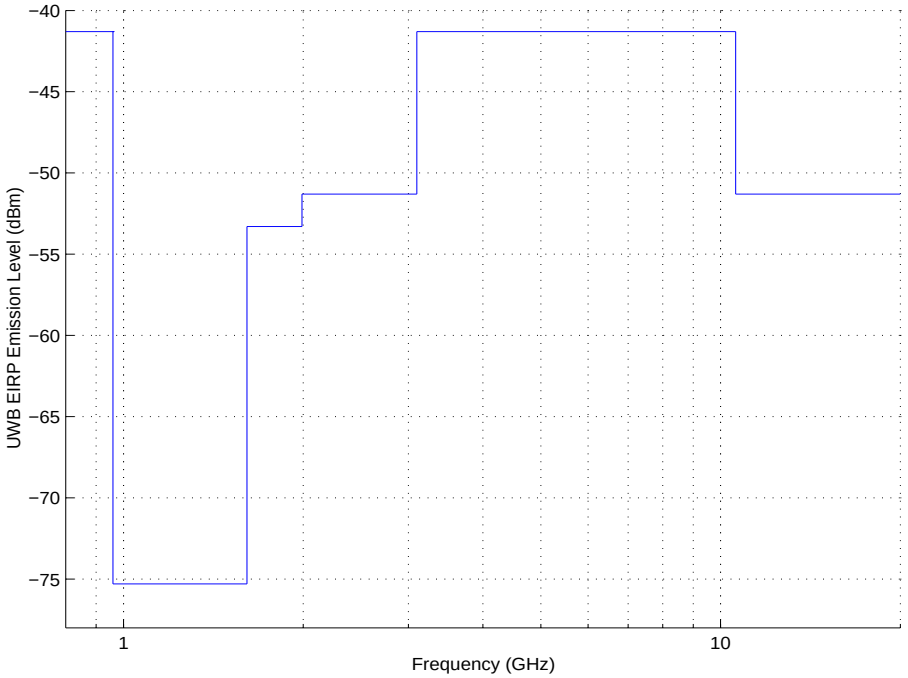

Figure 2.1: FCC emission mask for UWB devices.

bandwidth of the transmitted signal is greater than 500 MHz. Discussions were triggered around the advantages of the original IR scheme vs. the carrier-based continuous transmission alternative. The above lack of agreement is the framework of the IEEE 802.15.3a Task Group. This group was formed in late 2001 with the task of investigating innovative physical layer solutions for the development of very high data rate WPANs for applications that involve imaging and multimedia. Currently, two different merged proposals for a physical layer based on UWB are under consideration: a Multi-Band (MB) approach combining frequency hopping with Orthogonal Frequency Division Multiplexing (OFDM) [20] being backed by the WiMedia Alliance [41], which is a large coalition led by major semiconductor manufacturers like Texas Instruments, Intel, Philips and ST Microelectronics, and a second approach using Direct-Sequence UWB (DS-UWB) preserving the original UWB pulsed nature [42], which is heavily supported by Motorola who acquired Xtreme Spectrum. No conclusion has been

drawn since then and supporters of either proposal will continue their work outside of the IEEE 802.15.3a task group.

To coexist with existing narrowband systems, IR scheme requires notch filters which will distort the received signal. The short duration of the pulses required by IR UWB makes this scheme more susceptible to time jitter, require faster timing acquisition and have more multipath effects. Multi-band UWB on the other hand, can avoid using frequency bands where other wireless systems, e.g. 802.11a are present and it does not use such short pulses, which eases the design of pulse shaping filters and avoids notch filters. It is much easier to do time acquisition, capture multipath energy and eliminate inter-symbol interference (ISI) at low complexity [43] [44]. The spectral flexibility and coexistence features of MB-OFDM also makes it easier to meet the requirements of emerging spectral regulations in other regions of the world [45]. The major difference between MB-OFDM and traditional OFDM is that MB-OFDM symbols are not continually sent on a single frequency band; instead, they are interleaved over different subbands across both time and frequency [20].

Despite the merits mentioned above, the extremely short range, e.g., 10 meters for a data rate of 110 Mbps, puts UWB at an obvious disadvantage when compared to other competitive technologies, such as the soon coming IEEE 802.11n standard [46], which supports a data rate of 200 Mbps for around 50 meters in indoor environments. There is a trade-off between data rates and ranges [47]. Hence, in order to push UWB as an more attractive option for WPAN applications, it is crucial to improve the range limit of UWB devices at a fixed data rate or equivalently increase the data rate at fixed distances. In this dissertation, we will develop a system using more advanced turbo trellis coded modulation (TCM) with

OFDM to achieve much higher data rate within a short range in UWB environment than the current MB-OFDM scheme. UWB can also work with additional networking technology for a larger home coverage, such as a metal backplane (i.e. cable or power line) or mesh networking from room to room. Using multi-antenna technique with UWB to enhance data rates and coverage ranges also gains a lot of interest recently [48]. However, these topics are beyond the scope of this dissertation.

2.2 OFDM Overview

The history of OFDM dates back to the mid 60's, when Chang published a paper on the synthesis of bandlimited signals for multichannel transmission [14]. He presents a principle for transmitting messages simultaneously through a linear bandlimited channel without interchannel (ICI) and intersymbol interference (ISI). Shortly after Chang's paper, Saltzberg performed an analysis of the performance [15], where he concluded that "the strategy of designing an efficient parallel system should concentrate more on reducing crosstalk between adjacent channels than on perfecting the individual channels themselves, since the distortion due to crosstalk seems to dominate". This important conclusion has been proven correct in digital baseband processing that emerges a few years later.

A major contribution to OFDM was presented in 1971 by Weinstein and Ebert [18], who used the *discrete Fourier transform* (DFT) to perform baseband modulation and demodulation. This work did not focus on "perfecting the individual channels", but rather on introducing efficient processing, eliminating the banks of subcarrier oscillators. To combat ISI and ICI they used both a guard space between the symbols and raised-cosine windowing

in the time domain. Their system did not obtain perfect orthogonality between subcarriers over a dispersive channel, but it was still a major contribution to OFDM.

Another important contribution was due to Peled and Ruiz in 1980 [49], who introduced the *cyclic prefix* (CP) or cyclic extension, solving the orthogonality problem. Instead of using an empty guard space, they filled the guard space with a cyclic extension of the OFDM symbol. This effectively simulates a channel performing cyclic convolution, which implies orthogonality over dispersive channels when the CP is longer than the impulse response of the channel. This introduces an energy loss proportional to the length of the CP, but the zero ICI generally motivated the loss.

OFDM systems are usually designed with rectangular pulses, but recently there has been an increased interest in pulse shaping [50][51]. By using pulse other than rectangular, the spectrum can be shaped to be more well-localized in frequency, which is beneficial from an interference point of view.

OFDM is currently used in the European *digital audio broadcasting* (DAB) standard [52]. Several DAB systems proposed for North America are also based on OFDM [53], and its applicability to digital TV broadcasting is currently being investigated [54] [55] [56]. OFDM in combination with multiple-access techniques are subject to significant investigation, see e.g., [57]. OFDM, under the name DMT, has also attracted a great deal of attention as an efficient technology for high-speed transmission on the existing telephone network, see e.g., [19] [58] [59] [60].

One of the major advantages of OFDM is its robustness against multipath propagation. Hence, its typical application are in tough radio environments. OFDM is also suitable

for single frequency networks, since the signals from other transmitters can be viewed as echoes, i.e., multipath propagation. This means that it is favorable to use OFDM in broadcasting applications, such DAB and DVB. The use of OFDM in multiuser systems has gained an increasing interest the last few years. The downlink in those systems is similar to broadcasting, while the uplink puts high demand on e.g., synchronization. The future of OFDM as a transmission technique for multiuser systems depends on how well these problems can be solved.

In wired systems the structure of OFDM offers the possibility of efficient bit loading. By allocating a different number of bits to different subchannels, depending on their individual SNRs, efficient transmission can be achieved. Although other systems have been proposed, OFDM is the dominating technique on e.g. digital subscriber lines. Note that OFDM often goes under the name DMT when used in wired systems with bitloading.

There are also problems associated with OFDM system design. The two main obstacles when using OFDM are the high peak-to-average-power ratio and synchronization. The former puts high demands on linearity on amplifiers. Synchronization errors, in both time and frequency, destroy the orthogonality and cause interference. By using a cyclic prefix, the timing requirements are somewhat relaxed, so the biggest problems are due to high frequency synchronization demands. Degradation due to frequency errors can be caused both by differences in local oscillators and by Doppler shifts. A great deal of effort is therefore spent on designing accurate frequency synchronizers for OFDM.

As in any digital communication system, there are two alternatives for modulation: coherent and differential. The European DAB system uses differential QPSK, while the

proposed scheme for DVB is coherent 64-QAM. Differential PSK is suitable for low data rates and gives simple and inexpensive receivers, which is important for portable consumer products like DAB receivers. However, in DVB the data rate is much higher and low bit-error rates are difficult to obtain with differential PSK. A natural choice for DVB is therefore multiamplitude schemes. Due to the structure in OFDM, it is easy to design efficient channel estimators and equalizers. This is one of the appealing properties of OFDM which could be exploited to achieve high spectral efficiency.

Coding in wireless OFDM doesn't differ much from coding in wireless single-carrier systems. The main difference is that interleaving in OFDM allows symbols to be spread in both time and frequency. The possibility to interleave in frequency overcomes the drawback of not obtaining diversity from the equalizer. Since each subchannel experiences flat fading, code design and performance analysis developed for flat fading channels can be used. Decoding can be performed with a Viterbi decoder, where the metric depends on the (estimated) channel attenuations. This means that symbols are weighted with their respective channel strength. This reduces the effect of errors caused by symbols transmitted during a fade.

Multi-band OFDM (MB-OFDM) is one of the promising candidates for PHY layer for short-range high date-rate UWB communications. It combines OFDM with the above multi-band approach enabling UWB transmission to inherit all the strength of OFDM technique which has already been proven for wireless communications (ADSL, DVB, 802.11a, 802.16a, etc.). The difference between MB-OFDM and conventional OFDM lies in that OFDM symbols are not continually transmitted on the same frequency band. Instead, they are transmitted using one of the subbands in a particular time-slot. The subband selection

at each time-slot is determined by a Time-Frequency Code (TFC). The TFC is used not only to provide frequency diversity in the system but also to distinguish between multiple users. There are many advantages associated with using the MB-OFDM approach. This includes the ability to efficiently capture multipath energy, simplified transceiver architecture, enhanced frequency diversity, increased interference mitigation capability and spectral flexibility to avoid low quality subbands and to cope with local regulations.

2.3 Trellis Coded Modulation

The use of error control coding improves bit error rate performance of a digital communication system at the expense of bandwidth expansion by an amount equal to the reciprocal of the code rate. For power-limited channels, e.g., deep space communication channels, one may trade bandwidth for transmit power to achieve a desired system performance. However, that is not the case for bandwidth-limited channels, such as telephone channels. Traditionally coding is not favored in such channels, instead a bandwidth-efficient signaling technique such as Pulse Amplitude Modulation (PAM), Quadrature Amplitude Modulation (QAM), or Multi-Phase Shift Keying (MPSK) modulation is employed to support a high bandwidth efficiency (in bits/s/Hz).

Instead of considering coding and modulation as two separate parts, Ungerboeck [61] introduced a new method of combining them together to achieve significant coding gain over conventional uncoded multilevel modulation without reducing data rate or increasing bandwidth in 1976. It was subsequently termed as *trellis-coded modulation* (TCM) [62] [63] [64]. To avoid bandwidth expansion, TCM integrates a trellis code with an expanded bandwidth-efficient signal set and utilize the redundancy from it. Although the expansion

of a signal set reduces the distance between the signal points if the average signal energy is kept constant, substantial gain is still obtained by transmitting only constrained sequences of symbols through trellis coding. Typically one can get a coding gain of about 3-6 dB using 4-256 state convolutional codes by sacrificing neither data rate nor bandwidth. Within years of its emergence, TCM has become an industry standard. Modem, ADSL, Ethernet, Wi-Fi and WiMAX are all examples of TCM's successful application around us.

2.3.1 Ungerboeck's Trellis Coded Modulation

Ungerboeck's TCM is composed of a convolutional code followed by a signal set mapping, which is expanded to provide redundancy for coding. The convolutional code and signal mapping function are jointly designed to maximize the minimum squared Euclidean distance (MSED) between coded symbol sequences instead of Hamming distance, which exceeds the minimum distance between uncoded modulation signals at the same symbol rate. The general structure of Ungerboeck's TCM encoder/modulator is shown in Figure 2.2 [64].

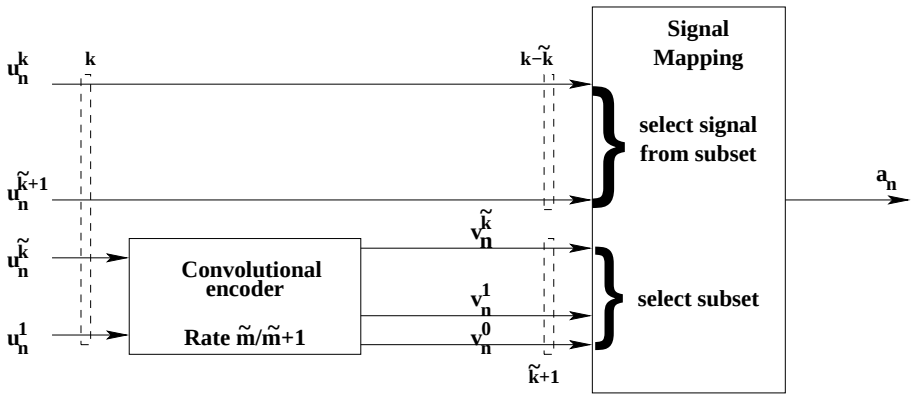

Figure 2.2: General structure of Ungerboeck's TCM encoder/modulator.

Information bits are grouped into blocks of length k as the input to the Ungerboeck's TCM encoder, among which $\tilde{k}$ bits are coded bits passing through a rate $\tilde{k}/(\tilde{k}+1)$ binary convolutional encoder with $\tilde{k}+1$ coded bits. These bits are used to select one of $2^{\tilde{k}+1}$ subsets. The remaining $k-\tilde{k}$ uncoded bits go directly to signal mapping and select one of the $2^{k-\tilde{k}}$ signal points from the selected subset. Combined together they form a symbol in a 2^{k+1} constellation to be transmitted at k bit/s/Hz.

The signal mapping follows a heuristic search technique called "mapping by set partitioning", which partitions a larger constellation successively into smaller subsets with maximally increasing minimum distance $\Delta_0 < \Delta_1 < \Delta_2 < \cdots$ between the signals of these subsets. The partitioning is repeated $\tilde{k}+1$ times until $\Delta_{\tilde{k}+1}$ is equal or greater than the desired MSED. This concept is illustrated in Figure 2.3(a) for 16-QAM modulation.

Figure 2.3(b) shows the trellis representation of a 4-state Ungerboeck code $(h^0, h^1) = (2, 5)$ [62], where there are 4 parallel (uncoded) paths associated with transition from current state to next state because each subset contains 4 signal points. A thicker trajectory is also presented in Figure 2.3(b) to show an error event corresponding to minimum non-parallel free distance. Then, the free Euclidean distance of this TCM code can be expressed as $d_{free} = min\{\Delta_{\tilde{k}+1}, d_{free}(\tilde{k})\}$, where $\Delta_{\tilde{k}+1}$ is the minimum distance between parallel transitions and $d_{free}(\tilde{k})$ denotes the minimum distance between non-parallel paths in the TCM trellis diagram.

Since there exists parallel branches between states, a branch metric is determined first by selecting the point within each subset which is closest to the received signal. This procedure makes the TCM trellis the same as the trellis for a convolutional code. Then, the

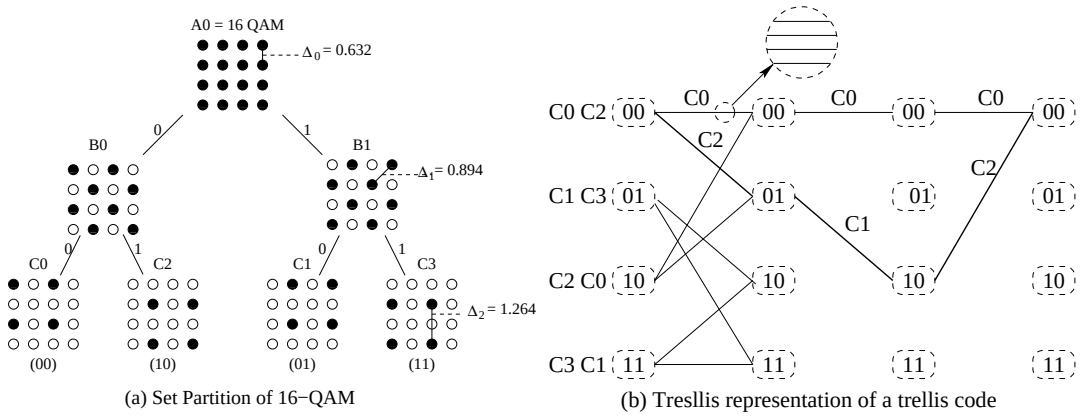

Figure 2.3: Set partition and trellis representation of a 4-state trellis code

Viterbi algorithm [65] can be used to find the mostly likely state sequence that the encoded signal traverses by finding a path with minimized accumulated metric.

2.4 Turbo Codes and Turbo Trellis Coded Modulation

2.4.1 Concatenated Codes

The exponentially increasing decoding complexity for a large code makes it complicated for implementation. A good idea is to divide it into several smaller component codes, each with a lower complexity decoder and concatenate them together to achieve near Shannon capacity. One of the first examples of concatenated codes was put forward by G. D. Forney in [66], where an algebraic block outer code is serially concatenated with a convolutional inner code by an block interleaver as shown in Figure 2.4.

A near-ML decoding algorithm for the inner code is used to achieve a moderate error rate like $10^{-2} - 10^{-3}$ at a code rate as close to capacity as possible, then another decoder for the outer code is applied to drive the error rate further down [67]. With such "separated" decoding scheme, the error rate could be made to decrease exponentially with block length at rate less than capacity, while decoding complexity increases only polynomially [66].

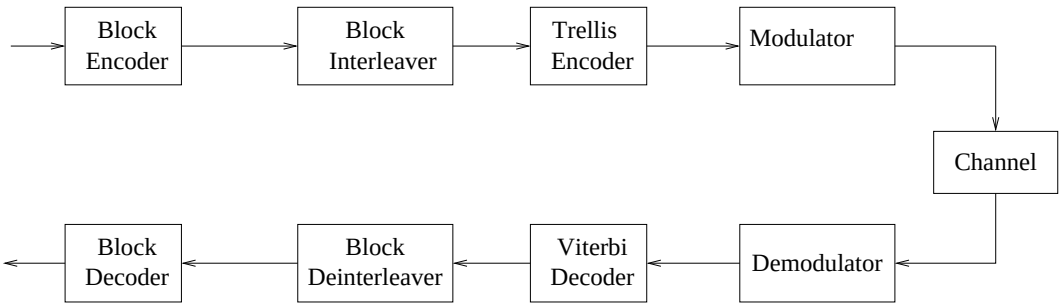

Figure 2.4: Forney's concatenated coding system

2.4.2 Turbo Codes

Turbo codes represent the most important achievement after Ungerboeck's TCM in the coding area [3], whose astonishing performance comes closer to the Shannon limit than any other class of error correcting codes at that time, with a simple encoding structure and a practical, implementable iterative decoding algorithms. Turbo codes yield very large coding gains (10-11 dB) at the expense of a data rate reduction, or bandwidth increase. It once again shows the potential of concatenated codes formed by two or more simple codes.

Since the appearance of turbo codes, there has been a tremendous amount of research in many aspects of turbo codes. As a result, many capacity-approaching turbo-like codes have been introduced [68] [69] [70] [71], and the turbo principle has also been applied to other communication areas, including equalization [72] [73] and detection [74]. Moreover, turbo codes have been selected as the Consultative Committee for Space Data Systems (CCSDS) standard [75] and also one of the universal mobile telecommunication systems (UMTS) standards by the third-generation partnership project (3GPP) [76].

2.4.2.1 Turbo Encoder

Turbo codes are in essence parallel concatenated convolutional codes (PCCC) whose encoder is composed of two or more constituent systematic recursive convolutional (RSC) encoders joined through one or more pseudo-random interleavers. The structure of a rate 1/3 turbo encoder is shown in Figure 2.5. The input bit sequence u_k with block length N feeds the first RSC encoder ENC1 and, after scrambled by an interleaver π with length N, it enters the second RSC encoder ENC2. A codeword of a parallel concatenated code consists of the input bits to the first encoder x_k followed by the parity check bits y_{1k} and y_{2k} from each encoder. Finally, the coded bit sequence $[v_1, v_2, \ldots, v_N]$ is sent to the channel, where $v_k = [x_k, y_{1k}, y_{2k}]$ is the output codeword at time k. We note that the systematic output of ENC2 $\pi(u_k)$ is not sent because it is just a permutation of x_k. Different code rates can be obtained by performing different puncturing on those parity check bits. For example, if y_{1k} and y_{2k} are selected alternatively, the code rate will become 1/2.

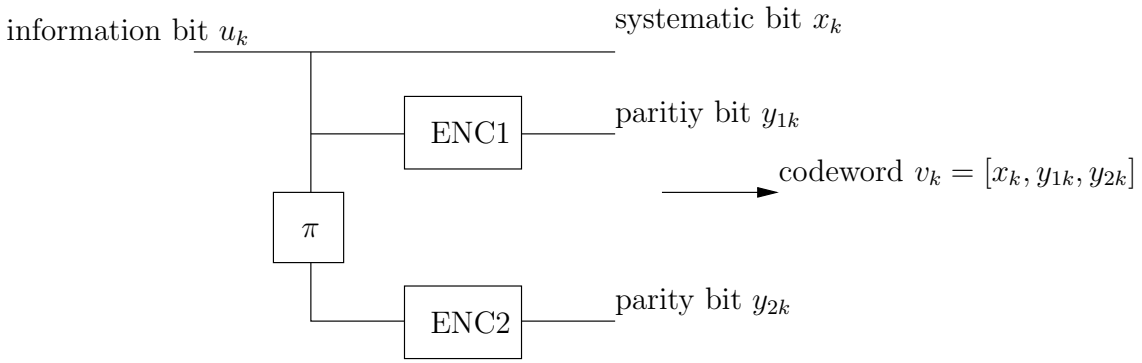

Figure 2.5: A rate 1/3 turbo encoder

2.4.2.2 BCJR Algorithm

The maximum *a posterior* (MAP) decoder is designed for each constituent code using BCJR algorithm [77]. The input will be the received noisy signal sequence $\mathbf{r}_1^N = [r_1, r_2, \ldots, r_N]$ from the channel and the output is the estimated u_k, where $r_k = v_k + n_k$ and n_k is channel noise. The goal of the decoder is to estimate u_k via the maximum *a posterior* (MAP) rule as follows

$$\hat{u}_k = sign[L(u_k)] \tag{2.1}$$

where

$$L(u_k) = log\left(\frac{P(u_k = 1|\mathbf{r}_1^N)}{P(u_k = 0|\mathbf{r}_1^N)}\right) \tag{2.2}$$

is the log-likelihood ratio (LLR). So the key is to calculate the LLR before making decision on u_k.

We can rewrite $L(u_k)$ using the trellis structure as

$$L(u_k) = log\frac{\sum_{S^1} P(s_{k-1} = s', s_k = s, \mathbf{r}_1^N)}{\sum_{S^0} P(s_{k-1} = s', s_k = s, \mathbf{r}_1^N)} \tag{2.3}$$

where s_k is the encoder state at time k and S^i represents all the state transitions from $s_{k-1} = s'$ to $s_k = s$ associated with $u_k = i$.

By using BCJR algorithm, we can obtain

$$P(s_{k-1} = s', s_k = s, \mathbf{r}_1^N) = \alpha_{k-1}(s')\gamma_k(s', s)\beta_k(s) \tag{2.4}$$

where

$$\alpha_{k-1}(s') = P(s_{k-1} = s', \mathbf{r}_1^{k-1}) \tag{2.5}$$

$$\gamma_k(s', s) = P(s_k = s, r_k | s_{k-1} = s') \tag{2.6}$$

$$\beta_k(s) = P(\mathbf{r}_{k+1}^N | s_k = s) \tag{2.7}$$

where $\mathbf{r}_1^{k-1} = [r_1, \ldots, r_{k-1}]$, $\mathbf{r}_{k+1}^N = [r_{k+1}, \ldots, r_N]$ are received signal before and after time k, respectively.

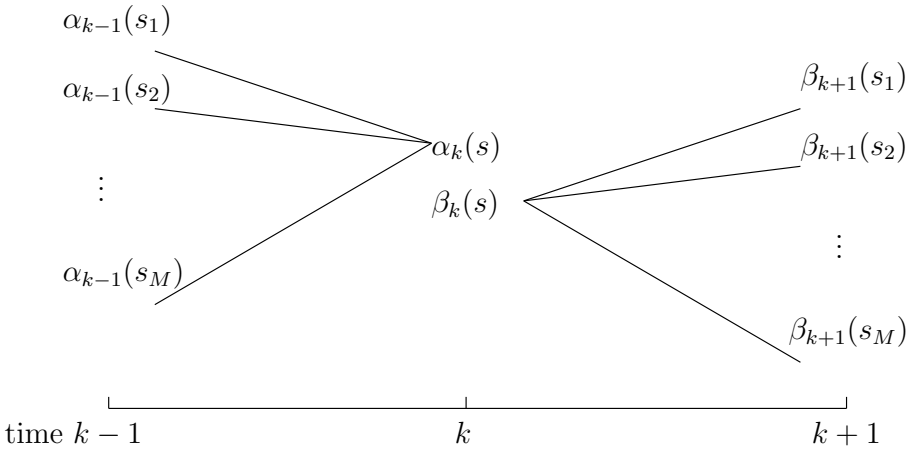

Figure 2.6: A forward and backward recursive calculation for $\alpha_k(s)$ and $\beta_k(s)$.

Further, we can calculate $\alpha_k(s)$ and $\beta_k(s)$ at each time k recursively forwards and backwards as in Figure 2.6, respectively, as

$$\alpha_k(s) = \sum_{s' \in S} \alpha_{k-1}(s') \gamma_k(s', s) \tag{2.8}$$

where $\alpha_0(s) = 1$ if $s = 0$ and 0 otherwise.

$$\beta_k(s') = \sum_{s \in S} \gamma_{k+1}(s', s) \beta_{k+1}(s) \tag{2.9}$$

where $\beta_N(s') = 1$ if $s' = 0$ and 0 otherwise. S is the set of all the possible states.

$\gamma_k(s', s)$ is related to channel as

$$\gamma_k(s', s) = P(s_k = s|s_{k-1} = s')P(r_k|s_{k-1} = s', s_k = s) \tag{2.10}$$

where $P(s_k = s|s_{k-1} = s') = P(u_k)$ and $P(r_k|s_{k-1} = s', s_k = s) = P(r_k|u_k, s_{k-1} = s')$ if there is a valid transition from $s_{k-1} = s'$ to $s_k = s$. Otherwise, $\gamma_k(s', s) = 0$.

2.4.3 Turbo Trellis Coded Modulation

The bandwidth-efficient TCM concept when used in parallel concatenation with an interleaved data stream permits the use of iterative decoding that results in error performance close to turbo codes. This scheme, referred to as turbo trellis coded modulation (TTCM) was initially proposed by LeGoff, Glavieux and Berrou [29]. In their paper, a simple scheme is presented in which a standard turbo code, after puncturing to obtain a desired spectral efficiency, was mapped to a high-density constellation. The constituent encoders encode m-ary information symbols into $(m + 1)$-ary code symbols. The encoders are separated by a symbol interleaver, which means groups of m information bits are interleaved together. The two sequences of coded symbols are then punctured to reduce the overall code rate and to ensure that the systematic outputs are only transmitted once. The puncturing imposes constraints on the symbol interleaver, so that the interleaving gain is reduced. The decoding first computes the log-likelihood ratios of transmitted bits and then uses standard turbo decoding algorithm. Although this system utilizes a bandwidth efficient modulation scheme, the encoder and modulator are not jointly designed as in TCM systems.

Later, turbo codes were embedded in multilevel codes with multistage decoding as shown in Figure 2.7 [34]. The data sequence is first divided into several sub-sequences

which are encoded by individual binary turbo encoders as component encoders and then the combined parallel codewords are mapped to a higher level constellation. Instead of using a maximum likelihood decoding over an enlarged trellis of multiple turbo codes, a suboptimal multistage decoding strategy is used to decode each component code individually using turbo decoders starting from the lowest level.

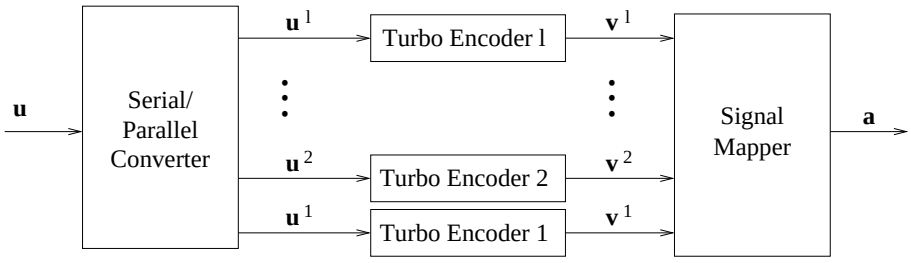

Figure 2.7: Multilevel turbo encoder

Robertson and Worz , in [26]-[28], presented a different approach in which two simple Ungerboeck's TCM codes were concatenated in parallel in the same way as in a turbo code, while the interleaver operates on groups of input bits to the second trellis encoder. Figure 2.8 shows the encoder structure.

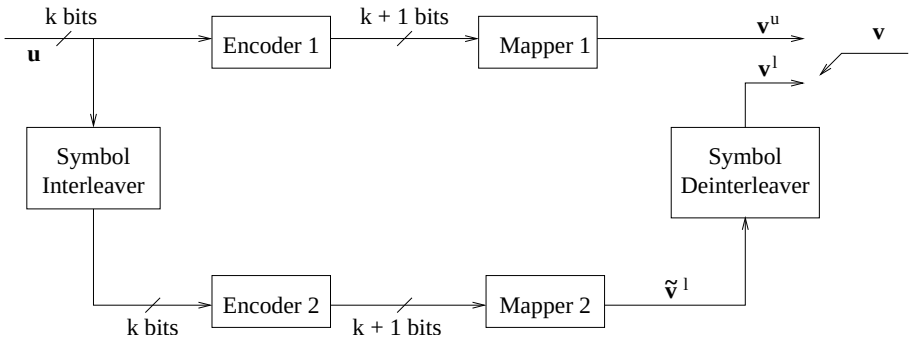

Figure 2.8: Turbo TCM encoder with parity symbol puncturing

It is noted that the interleaver is constrained to interleave symbols. That is, the ordering of k information bits arriving at the interleaver at a particular instant remains unchanged. For the component trellis code, some of the input bits may not be encoded.

In practical implementations these inputs do not need to be interleaved, but are directly used to select the final point in a signal subset. At the receiver, the values of these bits are estimated by set decoding [62]. The output of the second encoder is de-interleaved. This ensures that the k information bits which determine the encoded $(k+1)$ binary digits of both the upper and lower encoder at a given time instant are identical. The selector then alternately connects the upper and lower encoder to the channel. Thus, the parity symbols is alternately chosen from the upper and lower encoder. Each information group appears in the transmitted sequence only once. A generalized decoding algorithm used by standard binary turbo codes was used, where the symbol probability is used as the extrinsic information rather than the bit probability.

A totally different approach was presented by Benedetto et al [30] where two component convolutional codes, with entire information block and its interleaved version as inputs, are used to produce parity bits. The outputs of component codes are then punctured in such a way that alternate half of the information bits (even) are coded by the first constituent encoder and the other alternate half (odd) by the second constituent encoder. The systematic information bits and the parity bits are often mapped to a higher constellation. A 16QAM turbo trellis-coded modulation encoder is given in Figure 2.9.

By using MAP decoding [79] slightly superior performance compared to that of the above two schemes was reported. However, the latter two schemes can be considered as only special cases of the first scheme, which can be obtained from the former by restricting the interleaver structure and the puncturing scheme. Further, the additional performance gain

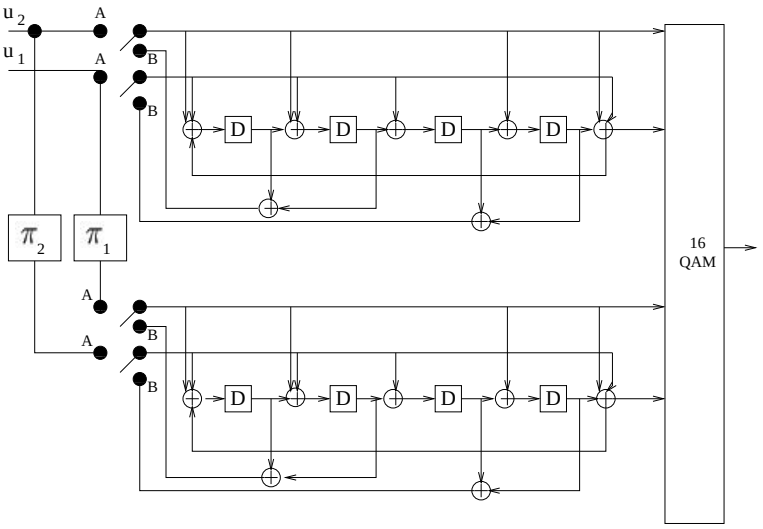

Figure 2.9: Trubo trellis-coded modulation, 16 QAM, 2bits/s/Hz

of the latter two schemes can be attributed to the choice of better component codes and as such the schemes itself can not be concluded to be better than the former.

2.5 Low-Density Parity Check Codes

2.5.1 LDPC Codes and Codes on Graphs

LDPC codes are a subset of linear block codes. We concentrate on binariy codes defined over the field $\mathbb{F}_2$ with two elements $\{0, 1\}$ and assume it has block length n and dimension k. Then the codewords can be expressed as the set of all n-tuple $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ satisfying the parity check equation $\mathbf{x} \cdot H^\top = \mathbf{0}$, where H is the $(n-k) \times n$ parity check matrix and $\top$ denotes transpose. Each row of H induces one parity check constraint on $\mathbf{x}$. For $n = 10$ and $k = 5$, an example of H is shown as follows. The codes are called *low-density* only if they are defined based on a sparse parity check matrix H, i.e. very few ones are

within H.

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \tag{2.11}$$

Another useful way of representing LDPC codes is via a graphical representation called *Tanner graph* [80], see Fig. 2.10. A Tanner graph $\mathcal{G}$ is a bipartite graph with n *variable nodes* on the left side and r *check nodes* on the right side. Variable nodes correspond to codewords $\mathbf{x}$ and also the columns of H. Check nodes correspond to parity check constraints that any valid codeword $\mathbf{x}$ must satisfy, hence they correspond to the rows of H. Edges in the bipartite graph connect variable nodes to check nodes, where an edge indicates the associated code bit participates in the associated parity check equation. Thus, the edges in the graph correspond to the ones in the parity check matrix H. A n-tuple $\mathbf{x}$ associated to the variable nodes is a valid codeword if and only if for each check node the modulo 2 sum of the values of its adjacent variable nodes is 0. Since the number of constraints is r, the rate of the code (defined as k/n) is at least $(n-r)/n$. If each parity check constraint is independent, i.e. H is full-rank, then code dimension $k = n - r$. In this case, the code defined in Fig. 2.10 has a rate 0.5.

Since Tanner generalized Gallager's codes to general bipartite graphs and introduced max-sum and sum-product decoding algorithms [80], the field of codes on graphs becomes active. Wiberg extends Tanner graphs to include state variables as well as other symbol variables, thus turbo codes, LDPC codes and trellis codes are put into a common framework

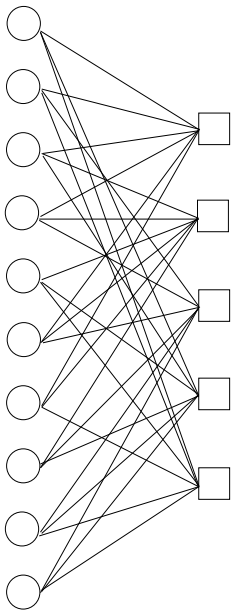

Figure 2.10: Tanner graph with $n = 10$ and $r = 5$, which is corresponding to the above parity check matrix H. Here circles and squares denote variable and check nodes, respectively. It is also a $(3, 6)$ regular LDPC code.

[81]. Further, Kschischang *et al* generalized Wiberg-type graphs to *factor graphs* [82], a natural graphical model of a global function that can be factored into a product of local functions. Forney later introduces *normal graphs* [83], which leads to a clean separation of functions in sum-product decoding. It is because of the sparse graphical model, design of various efficient iterative decoding algorithms based on message-passing via edges becomes convenient, which will be discussed later.

If each variable node has the same degree d_v and each check node has the same degree d_c, it is called (d_v, d_c) *regular* LDPC code. The example given above is actually a $(3, 6)$ regular LDPC code of block length $n = 10$. Otherwise, it is *irregular*. For irregular LDPC codes, the degree of each variable and check node is random, with distribution given

by some degree sequence

$$\lambda(x) = \sum_{i=1}^{d_{v,max}} \lambda_i x^{i-1} \text{ and } \rho(x) = \sum_{i=1}^{d_{c,max}} \rho_i x^{i-1} \tag{2.12}$$

where λ_i is fraction of edges incident to variable node of degree i and ρ_i fraction of edges incident to check node of degree i, i.e. $\sum_{i=1}^{d_{v,max}} \lambda_i = \sum_{i=1}^{d_{c,max}} \rho_i = 1$. Here, $d_{v,max}$ and $d_{c,max}$ is the maximum degree of variable and check nodes, respectively. Regular LDPC code is a special case with $\lambda(x) = x^{d_v - 1}$ and $\rho(x) = x^{d_c - 1}$. It can be easily verified that the code rate is a function of degree distributions as $1 - (\int_0^1 \rho(x)\,dx)/(\int_0^1 \lambda(x)\,dx)$.

2.5.2 Message-Passing Algorithms

LDPC codes can be decoded by a class of efficient iterative algorithms, called *message-passing algorithms*, based on the graphical model. At each round of the algorithm messages are passed from variable nodes to check nodes and from check nodes back to variable nodes. The message sent from a variable node v to a check node c depends on messages sent from channel and all check nodes incident to v, excluding message sent from c. Similarly, the message from c to v depends on messages sent from all variable nodes incident to c, excluding message sent from v. That is, only *extrinsic* information is exchanged, which turns out to be an important property of good message-passing decoders and which also makes the analysis of the decoding algorithm feasible. Fig. 2.11 shows the basic operations in the message-passing algorithm. It works based on an assumption that messages from different incoming edges are conditionally independent, under which convergence is guaranteed.

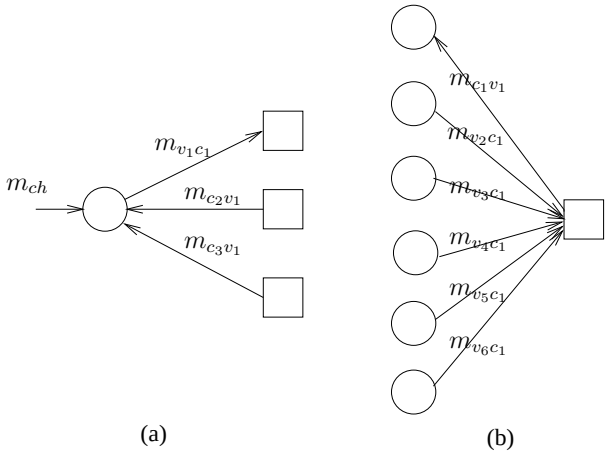

Figure 2.11: Message-passing algorithms (a case for regular $(3, 6)$ LDPC codes). (a) variable node update: outgoing message $m_{v_1c_1}$ is determined by message from channel m_{ch} and incoming messages $m_{c_2v_1}$ and $m_{c_3v_1}$, but not by $m_{c_1v_1}$. (b) check node update: outgoing message $m_{c_1v_1}$ is determined by incoming messages $m_{v_2c_1}, \ldots, m_{v_6c_1}$, but not by $m_{v_1c_1}$.

2.5.2.1 Gallager's algorithm A

The first example of message-passing decoders was given by Gallager [2] for the binary symmetric channel (BSC). The messages passed back and forth are estimated bits. It obeys the following rules:

1) *Initiation*: In the 0-th iteration, the variable nodes send their received value from channel to their incident check nodes directly.

2) *Check node update*: In the first half of the l-th iteration, the message sent from check node c to variable node v is the modulo 2 sum of the message received by c from all incident variable nodes other than v in the previous round.

3) *Variable node update*: In the second half of the l-th iteration, the message sent from variable node v to check node c is b if all the messages received by v from all incident

check nodes other than c are the same and equal to b. If they are not the same, the received value from channel will be sent from v to c.

There is also a *Gallager's algorithm B*, which is slightly different in step 3). See [2] for details. Both these two algorithms belong to the class of hard-decision decoding algorithms. It is simple but performance is not good enough.

2.5.2.2 Belief propagation algorithm

To achieve performance closer to capacity, one needs to use more refined messages. Precisely, the messages are conditional probability (or soft information), i.e. $Pr[x = 0|m_1, \ldots, m_l]$ or $Pr[x = 1|m_1, \ldots, m_l]$, which contains information on hard decision as well as reliability. Here, m_i means current l incoming messages. Usually another equivalent form of message, i.e. conditional *log-likelihood ratios* (LLR) defined as

$$logL(x|m_1, \ldots, m_l) \triangleq log\left(\frac{Pr[x = 0|m_1, \ldots, m_l]}{Pr[x = 1|m_1, \ldots, m_l]}\right), \tag{2.13}$$

is preferred. If $x = 0$ or 1 with certain, we know LLR $= +\infty$ or $-\infty$, respectively. Otherwise, LLR $\in \mathbb{R}$. The sign of LLR gives hard decision and its magnitude tells how reliable the decision is. Message-passing algorithm that uses LLR as message is also called *belief propagation algorithm*, which is already developed in Artificial Intelligence community [84].

Following Gallager's rule, we can similarly set up update equations when messages are in LLR-domain. Let $m_{vc}^{(l)}$ be the message passed from variable node v to check node c in the l-th iteration and $m_{cv}^{(l)}$ in the reverse direction. In the 0-th iteration, $m_{vc}^{(0)} = m_{ch}$ which is the LLR conditioned on its observed value and it is independent of check node c. In the l-th

iteration, "tanh" rule and "sum" rule are followed at check and variable nodes, respectively, due to operation on LLR-domain. For details about so called *LLR-algebra*, see [25].

Then the messages passed between variable nodes and check nodes can be specified by the following update rule at variable and check nodes, respectively

$$m_{vc}^{(l)} = \begin{cases} m_{ch}, & \text{if } l = 0 \\ m_{ch} + \sum_{c' \in C_v \setminus \{c\}} m_{c'v}^{(l-1)}, & \text{if } l \geq 1 \end{cases} \tag{2.14}$$

$$m_{cv}^{(l)} = 2tanh^{-1}\left(\prod_{v' \in V_c \setminus \{v\}} tanh\left(m_{v'c}^{(l)}/2\right)\right) \tag{2.15}$$

where C_v is the set of check nodes incident to variable node v and and V_c the set of variable nodes incident to check node c, respectively. From the above formulation, we find that sum operation is used in variable node update while product operation is used in check node update. So the belief propagation algorithm is also called *sum-product algorithm*, which is the most powerful message-passing algorithm.

CHAPTER 3: NOISE MISMATCH IN TURBO TCM CODED MB-OFDM SYSTEMS

3.1 Turbo TCM Coded MB-OFDM System

We follow the frequency band allocation scheme in the MB-OFDM proposal for IEEE 802.15.3a project [20]. The entire 3.1 - 10.6 GHz UWB bandwidth is divided into 14 frequency subbands with 528 MHz each. Among them, the first twelve subbands are grouped into four band groups each including three bands while the last two subbands are grouped into a fifth band group. Currently, the only band group that is mandatory is the first band group between 3168 MHz to 4752 MHz. The other subbands are allocated for optional use or future expansions since propagation loss severely limits signals at higher frequencies. Information is transmitted using OFDM modulation on each subband and different Time-Frequency Codes (TFC), i.e. frequency hopping patterns, are used for channelization.

Instead of using punctured convolutional codes followed by QPSK modulation as in the MB-OFDM proposal, we use more powerful turbo TCM codes and iterative decoding to achieve higher data rate.

3.1.1 General Structure of a Multi-Band OFDM System

The diagram of a baseband transceiver of our turbo TCM coded MB-OFDM system is illustrated as in Figure 3.1.

Information bit sequences are first encoded and mapped by turbo TCM encoder onto complex points of a higher constellation, such as 16QAM or 64QAM. Then these mapped data are grouped into blocks of length equal to 100. Together with inserted pilot/guard/null

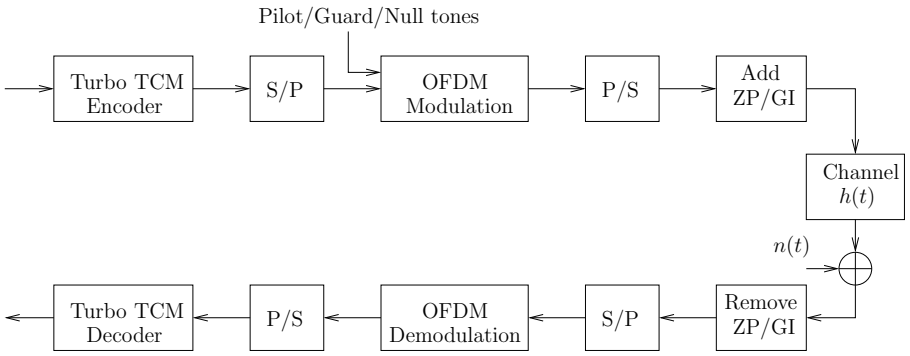

Figure 3.1: A baseband turbo TCM coded MB-OFDM transceiver.

bits, they modulate $N_{FFT} = 128$ subcarriers of OFDM in a selected frequency subband. Since the bandwidth of each subband is 528 MHz, the frequency spacing between adjacent subcarriers is $\Delta_f = 528/128 = 4.125$ MHz. A 128-point complex inverse fast Fourier transform (IFFT) is employed for OFDM modulation to construct the 128 time-domain samples for each subcarrier of an OFDM symbol, which occupies $T_{FFT} = 1/\Delta_f = 242.4$ ns. Of the 128 subcarriers, there are 100 of them dedicated for data from the mapper and 12 of them located at subcarrier numbers 5, 15, 25, 35, 45, 55, 73, 83, 93, 103, 113, 123 are used as pilot tones, which are useful for coherent detection. Another 10 subcarriers located at numbers 57-61 and 67-71 are dedicated for guard tones, which has various purposes such as relaxing the specifications on the transmitter and receiver filters. The other 6 subcarriers at numbers 0 and 62-66 are simply set to be zero and named null tones, which are reserved for future use. The subcarrier allocation for each OFDM symbol is summarized in Figure 3.2.

Unlike CP-OFDM method, ZP-OFDM assumes zero-padding prefix (ZP) of length $T_{ZP} = 60.6$ ns (corresponding to $N_{ZP} = 32$ samples) at the beginning of each OFDM symbol instead of a cyclic prefix (CP) [90]. Also, a guard interval of length $T_{GI} = 9.5$ ns (corresponding to $N_{GI} = 5$ samples) is appended at the end of each OFDM symbol, which

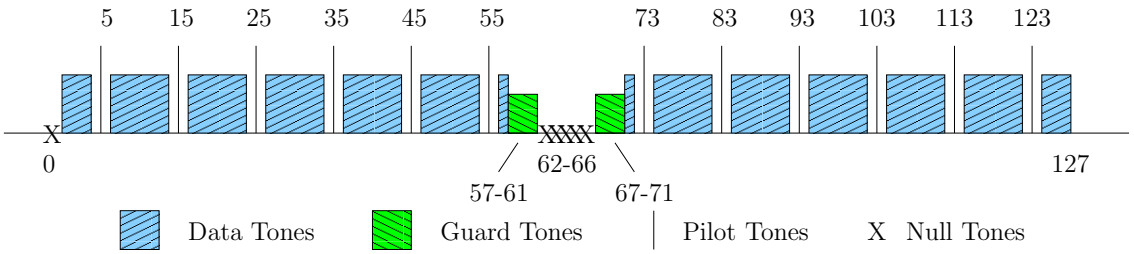

Figure 3.2: OFDM subcarrier allocation.

results in a total duration of $T_{SYM} = T_{FFT} + T_{ZP} + T_{GI} = 312.5$ ns (corresponding to $N_{SYM} = N_{FFT} + N_{ZP} + N_{GI} = 165$ samples) for an OFDM symbol period. A general OFDM symbol structure is shown in Figure 3.3, where the shaded block represents the start of a new OFDM symbol period.

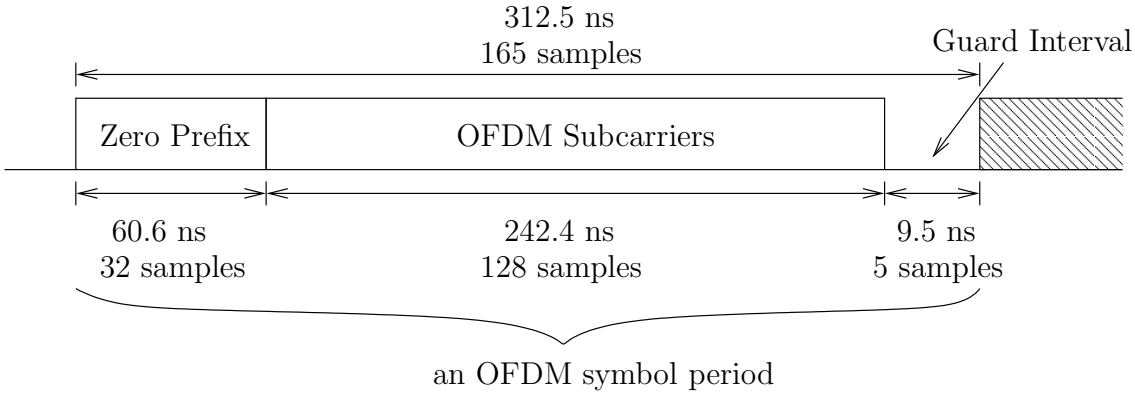

Figure 3.3: A general OFDM symbol structure.

The k-th analog baseband OFDM signal after digital-to-analog converter is

$$s_k(t) = \begin{cases} 0, & t \in [0, T_{ZP}) \\ \sum_{n=0}^{N_{FFT}-1} s_k[n] e^{j2\pi(n-\frac{N_{FFT}}{2})\Delta_f(t-T_{ZP})}, & t \in [T_{ZP}, T_{ZP}+T_{FFT}) \\ 0, & t \in [T_{ZP}+T_{FFT}, T_{SYM}) \end{cases} \tag{3.1}$$

where $s_k[n]$ denotes data/pilot/guard/null tones at the n-th subcarrier during the k-th OFDM symbol period. Further, the complex baseband signal $s_k(t)$ is filtered, up-converted

to an RF signal with a carrier frequency $f_{\Theta(k)}$, which is hopping among various subbands following a time-frequency code (TFC). Figure 3.4 illustrates a time-frequency representation of 6 OFDM symbols, where the carrier frequency index changes following a hopping pattern $\Theta(.)$ defined as

$$\Theta(k) = \begin{cases} 1, & k \bmod 6 = 0 \\ 3, & k \bmod 6 = 1 \\ 2, & k \bmod 6 = 2 \\ 1, & k \bmod 6 = 3 \\ 3, & k \bmod 6 = 4 \\ 2, & k \bmod 6 = 5 \end{cases} \tag{3.2}$$

where $f_1 = 3432$ MHz, $f_2 = 3960$ MHz, and $f_3 = 3388$ MHz represent center frequencies of the three subbands within band group 1, respectively, as shown in Figure 3.4.

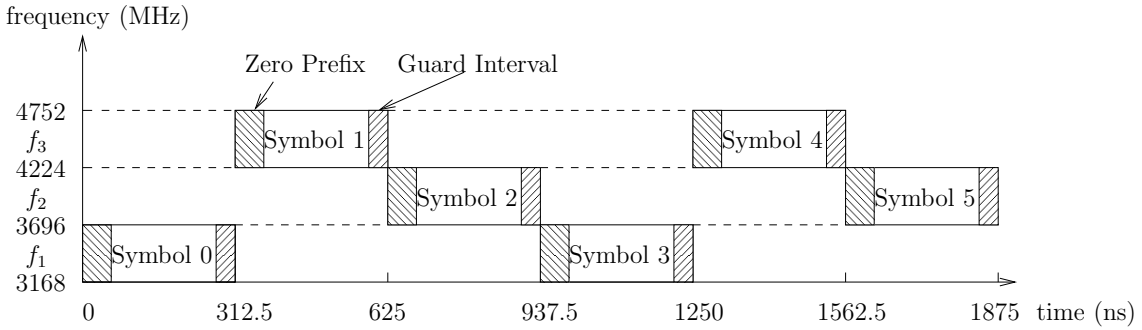

Figure 3.4: Time-frequency code of 6 MB-OFDM symbols over band group 1.

Finally, the transmitted multi-band OFDM signal is given by

$$s(t) = \sum_k \mathrm{Re}\{s_k(t - kT_{SYM})g(t - kT_g)e^{j2\pi f_{\Theta(k)}t}\} \tag{3.3}$$

where $g(t)$ denotes the waveform shaping filter of duration T_g and Re{.} takes the real part of a complex variable.

After passing through a channel with impulse response $h(t)$ with additive noise $n(t)$, the received signal can be represented as

$$r(t) = s(t) \otimes h(t) + n(t) \tag{3.4}$$

where $\otimes$ means convolution. At the receiver, zero-padding prefix and guard interval are removed before OFDM demodulation, which can be implemented efficently as fast Fourier transform (FFT). Then during each OFDM symbol period, 100 data are extracted from 128 subcarriers after a simple one-tap equalizer is employed to offset the effect of channel imperfectness. Demodulated OFDM symbols are further passed to the turbo TCM decoder to recover the information bits.

3.1.2 Turbo TCM Encoder Structure

Now let us take a deep look at the turbo TCM encoder, which is demonstrated in Figure 3.5 step by step.

First, each of two information bit sequences u_1 (or u_2) with length $L-2$ passes through an independent random interleaver and generate its scrambled copy u_3 (or u_4) with length $L-2$, respectively. This step is a parity check code and it is constructed as a repeater followed by an interleaver [92]. Then after appending two tailing bits for each sequence, u_1, u_2, u_3 and u_4 are further encoded by a recursive systematic convolutional (RSC) code with rate $R = 4/3$ after puncturing the systematic output for u_3 and u_4. The output of RSC is named v_0, v_1, v_2. The parity check outer code is serially concatenated to the RSC inner code with 16 states and this functions as a turbo code, whose overall rate is 2/3. The encoder structure of this parity concatenated trellis code is given in Figure 3.6. As proved

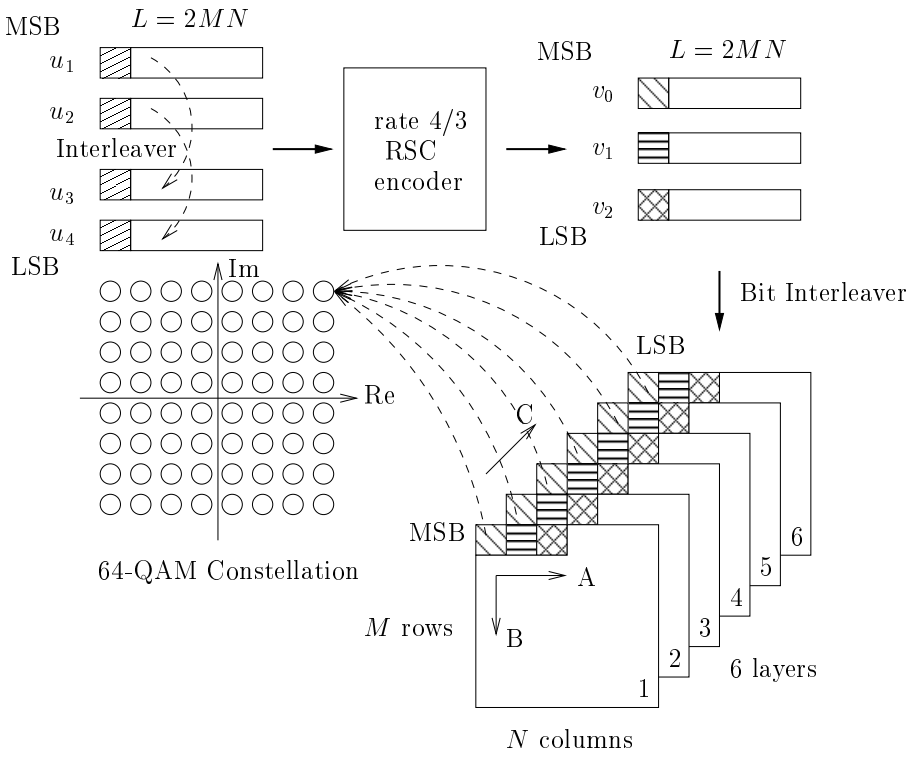

Figure 3.5: Turbo TCM structure in details.

in [93] that this structure is equivalent to Benedetto et al's TTCM [30], the information bits only occur once at the output.

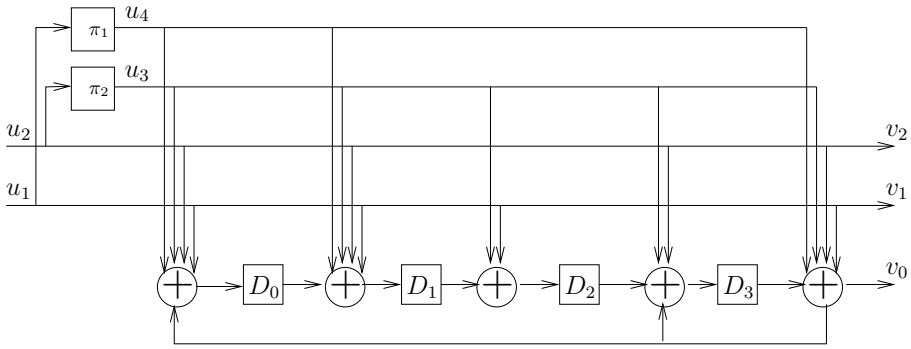

Figure 3.6: Parity concatenated trellis encoder.

We name the 16 states for the trellis of RSC as $S_0 - S_{15}$, where S_0 is the initial state. We will find the tail-bit sequences for (u_1, u_2, u_3, u_4) that terminate the trellis to state S_0 such that a better BER performance is obtained. By examining the RSC trellis,

we know that there are 16 outwards transition branches from each state due to 4-bit input (u_1, u_2, u_3, u_4). Also we find out that actually there exists parallel branches between state transitions. For example, from state S_1 to state S_8, there are two parallel branches, one is $(u_1, u_2, u_3, u_4) = (0, 1, 0, 1)$, the other is $(u_1, u_2, u_3, u_4) = (1, 0, 1, 0)$. For simplicity, we represent each binary 4-tuple (u_1, u_2, u_3, u_4) using a decimal number by assuming u_1 the most significent bit (MSB) while u_4 the least significant bit (LSB). Thus, the above mentioned two parallel branches from state S_1 to state S_8 become 5 and 10 in decimal, respectively.

From the trellis, we notice that for those S_i (i even), there is a direct transition branch from S_i to S_0. However, that is not the case for those S_i (i odd). Generally, there only exists transition branch from S_i (i odd) to S_j ($j \in [8, 15]$) and from S_i (i even) to S_j ($j \in [0, 7]$) and none otherwise. That means in order to go back to state S_0 at time t, the previous state at time $t-1$ must be S_i (i even). Based on this fact, we select S_i at time $t-1$ to be S_0 such that all even states at time $t-2$ can connect through it to go back to S_0 at time t. Also, we select S_i at time $t-1$ to be S_8 such that all odd states at time $t-2$ can connect through it to go back to S_0 at time t. Since the choice is not unique, there are other possible paths to go back to S_0 within two transitions. All the paths based on our selection are shown in Figure 3.7.

For simplicity, we only use a single solid line to represent all parallel transitions between states. The two numbers attached to each line are decimal representation of inputs (u_1, u_2, u_3, u_4) associated with each possible transition branch. From Figure 3.7, we easily know that a tail-bit of 5 or 10 will bring S_1 to S_8 and a second tail-bit of 4 or 11 can further

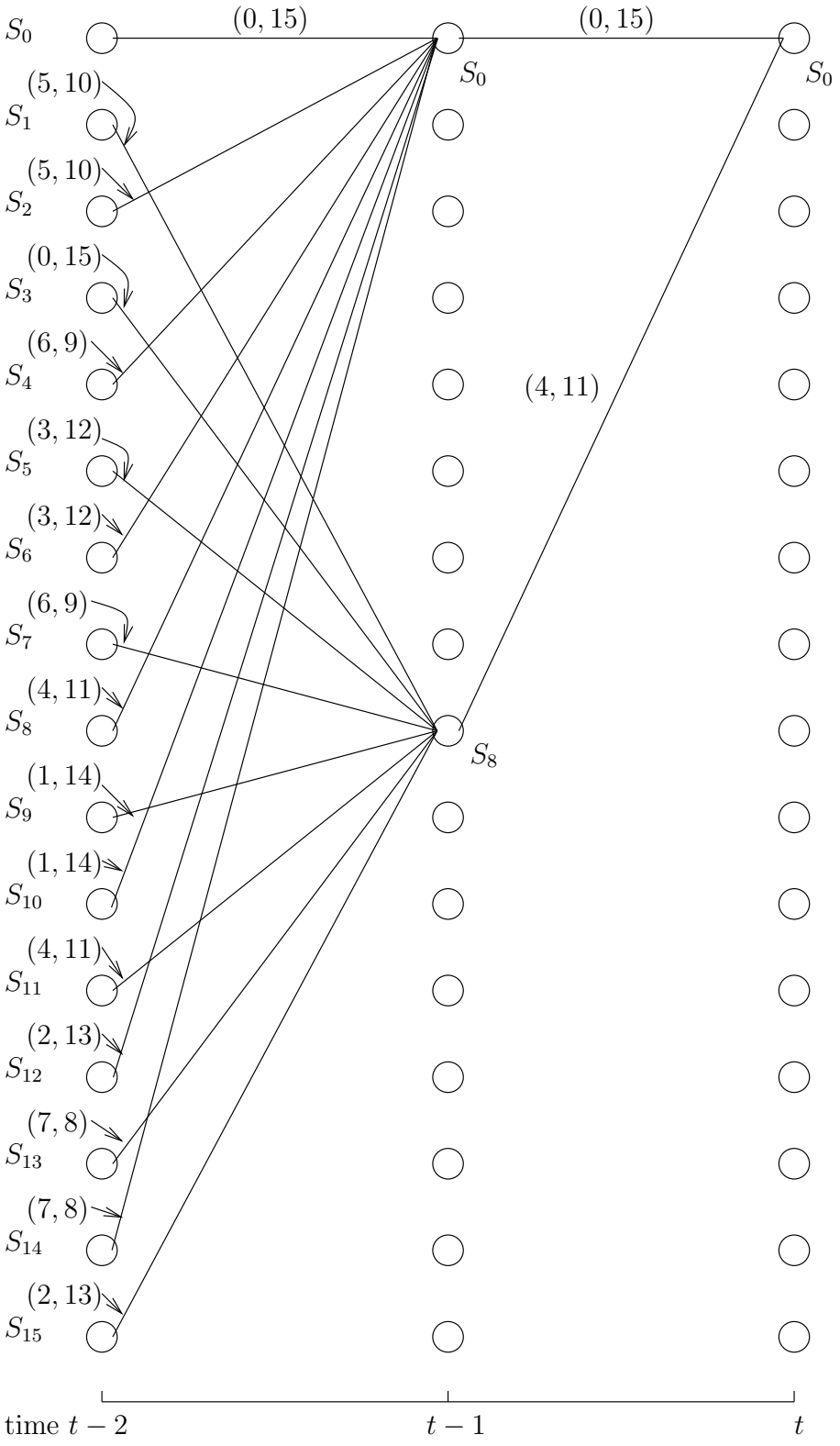

Figure 3.7: Two tail bits to terminate the TTCM trellis.

go to S_0. It is similar for other states to return to S_0. Thus, two tail bits are enough for trellis termination and the input bit sequence length is L.

The multiplexed outputs v_0, v_1, v_2, each with length $L = 2 \times M \times N$, are rearranged one by one into a three-dimensional array, with size $6 \times M \times N$ (six layers, each of which is a matrix with M columns and N rows). It follows the order of row first, column second, and layer last, i.e. the order of directions A, B, and C (see Figure 3.5) in serial. In our system, parameters are set to be $L = 8000, M = 200, N = 20$. After such a bit level block interleaver, every six bits on the same row-and-column position of that matrix are grouped together and mapped onto a 64QAM constellation point, which is a complex number. The 64QAM constellation with gray mapping is shown in Figure 3.8, where the number associated with each point is the decimal representation of the six bits from different layers. Totally we have a matrix with size $M \times N = 200 \times 20$ of such complex numbers to be modulated by OFDM.

Before OFDM modulation, there is another symbol level block interleaver to rearrange the order of these $M \times N = 4000$ complex numbers in order to combat burst errors due to the imperfect channel. They are read out columnwise and shift an additional $step = 10$ rows downwards for each column as described in Table 3.1, where the entries represent the order numbers for transmission, starting from 0 and ending at $M \times N - 1 = 3999$. We define it as $ord \in [0, M \times N - 1]$.

Assume a complex number is located at the I-th row and J-th column within the matrix with size $M \times N$, where $I \in [0, M - 1]$ and $J \in [0, N - 1]$. Then, the symbol level block interleaver in Table 3.1 functions as a mapping from row-column position (I, J) to its

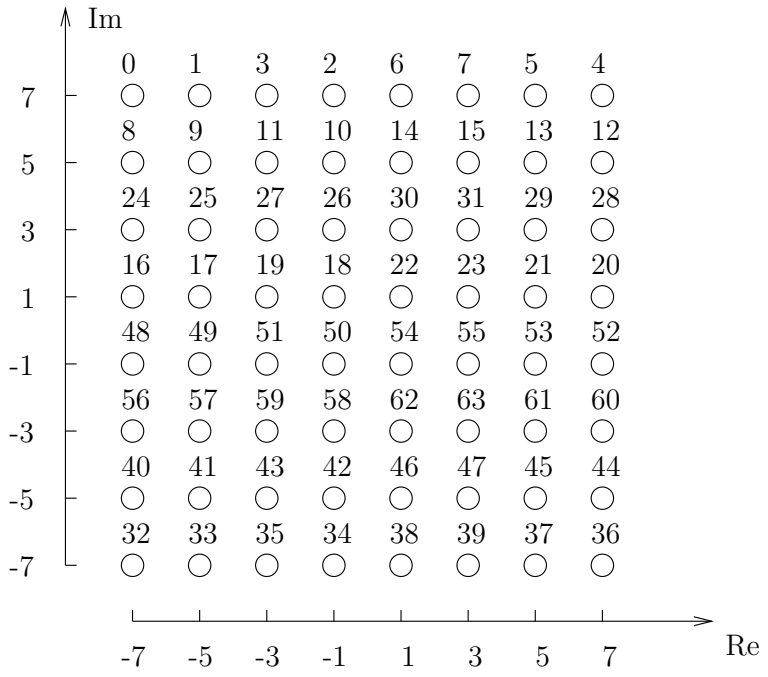

Figure 3.8: 64QAM with gray mapping.

Table 3.1: A symbol block interleaver.

0	390	580		3810
1	391	581		3811
	⋮			
	399		⋮	
	200	⋮		
				⋮
		599	⋮	
⋮		400		
	⋮	401		
		⋮	⋮	
				3999
				3800
				3801
				⋮
199	389	579		3809

200 rows

$step = 10$

20 columns

order number ord as

$$ord = J \times M + (I - step \times J)_M \tag{3.5}$$

where $(.)_M$ denotes (.) mod M.

Once the order number ord is obtained, we can easily know that it is corresponding to the $n_{sub}-$th data subcarrier of the $n_{OFDM}-$th OFDM symbol via

$$n_{sub} = (ord)_{N_{FFT}} \tag{3.6}$$

$$n_{OFDM} = \frac{ord - n_{sub}}{N_{FFT}} \tag{3.7}$$

where $n_{sub} \in [0, N_{FFT} - 1], n_{OFDM} \in [0, \lceil \frac{M \times N}{N_{FFT}} \rceil - 1]$ and here $\lceil . \rceil$ means the ceiling function. For our system $N_{FFT} = 100$, thus we group the 4000 complex points from the mapper into 40 blocks each with 100 data.

By inserting 12 pilot bits, 10 guard bits and 6 null bits for each block following the subcarrier allocation rule (see Figure 3.2), we form the samples with length 128, which will be used to modulate 128 OFDM subcarriers in frequency domain. It is easy to get its time domain OFDM symbol via IFFT.

3.2 Characteristics of UWB Channel and Impulsive Noise

3.2.1 UWB Channel Modeling

After OFDM symbols are transmitted out, it goes through a channel with impulsive response $h(t)$ and is further corrupted by an additive noise $n(t)$. We are especially interested in a UWB indoor channel for home environment and an additive noise with impulsive characteristics.

Previous research on signal propagation measurements in home environment has been done over various bandwidth ranges and transmitter-receiver separation distances in two different ways, time domain sounding [94][95][96] vs. frequency domain sounding [97][98][99][100]. In time domain techniques, the channel is excited by a short pulse and a receiver records samples of the channel response captured by its antenna. However, it requires the generation of ultra-short pulses and the non-ideal transmit pulses will distort the observed impulse response. In frequency domain techniques, the bandwidth centered around the frequency of interest is scanned by a vector network analyzer through discrete frequencies. For each frequency point, a known sinusoidal signal is transmitted and both magnitude and phase of the received signal are measured. Thus we get the transfer function of the channel in frequency domain and the corresponding impulse response in time domain can be obtained via inverse Fourier transform [101].

Once we get the measurements, we can parameterize a mathematical model to best fit the data. One method is to model the channel impulse response in time domain [94][102][103]. Several statistical models based on a time domain tapped-delay line model with clusters [104][105] have been found to be effective in modeling UWB channel with dense multipath. It assumes multipath components arrive in clusters, thus the channel impulse response is expressed as

$$h(t) = \sum_{l=0}^{N_c-1} \sum_{m=0}^{N_e-1} \alpha_{m,l} \delta(t - T_l - \tau_{m,l}) e^{j\theta_{m,l}} \tag{3.8}$$

where N_c is the number of clusters of scatterers, N_e is the number of echoes in each cluster, $\alpha_{m,l}$ and $\theta_{m,l}$ are the amplitude and phase shift of the m-th echo of the l-th cluster, respectively, T_l and $\tau_{m,l}$ are the arrival time of the l-th cluster and the time delay of the m-th

echo in that cluster with respect to first arriving echo, respectively, and $\delta(.)$ is the Dirac delta function. A modified S-V model based on [104] is finally taken as the IEEE 802.15.3a standard channel model due to its accuracy for capturing the important characteristics of the UWB channel, such as RMS delay spread, power decay profile and number of multipath components [105].

The main disadvantage of channel models based on the description of the channel impulse response is that they involve a large number of parameters, when a satisfactory description of a given UWB propagation scenario is required. In [106], it shows that a second order auto-regressive (AR) model in frequency domain can provide a channel description closely matching experimental data. The channel model is

$$\hat{H}(f_i, t; d) + a_1\hat{H}(f_{i-1}, t; d) + a_2\hat{H}(f_{i-2}, t; d) = n_i \tag{3.9}$$

where d denotes the transmitter-receiver separation, $\hat{H}(f_i, t; d)$ is the channel frequency response which does not exhibit significant variability in time, n_i is the excitation for the model which is independent identically Gaussian distributed and a_i is the model coefficient. From equation (3.9), we know that the channel frequency response can be generated by passing a Gaussian noise through an Infinite Impulse Response (IIR) filter. Since this model is generative and allows simple simulation, we will generate our UWB channel based on this AR model in this chapter.

3.2.2 Impulsive Noise Modeling

In the real wireless communications systems besides AWGN there are impulsive man-made noise from ignition of automobile or other sources such as power line which affect

the performance of the system. The impulsive noise is an additive disturbance that arises primarily from the switching electric equipment and their statistical characteristics are much different from those of Gaussian noise.

Analytical characterization of impulsive noise is quite difficult task. Several analytical and empirical impulsive noise models were proposed in a literature in last twenty years. Some of these models characterize impulsive noise from particular sources in certain conditions with high degree of accuracy. However, in many practical situations it is very difficult to predict statistical properties of impulsive interference. On the other hand, impulsive noise from any source in any system has two common properties: (a) impulsive noise energy is concentrated into short periods; (b) impulsive noise energy is much higher than that of background (thermal) noise.

In this dissertation, we consider a generalized Gaussian noise model [38] as:

$$p(n_k; A_k) = \frac{\gamma_k^C}{2\sigma\sqrt{a(\gamma_k^C)\Gamma(\frac{1}{\gamma_k^C})}} e^{-\frac{|n_k|^{\gamma_k^C}}{[a(\gamma_k^C)]^{\gamma_k^C/2}\sigma_k^{\gamma_k^C}}} \tag{3.10}$$

where the parameter $A_k = (\sigma_k^2, \gamma_k^C)$, n_k is the noise added on transmitted symbol, $\Gamma(.)$ denotes the Gama function, σ_k^2 is the variance of the noise and $a(\gamma_k^C) = \frac{\Gamma(\frac{1}{\gamma_k^C})}{\Gamma(\frac{3}{\gamma_k^C})}$. When $\gamma_k^C = 2$, $p(n)$ is the Gaussian distribution function. When $\gamma_k^C = 1$, $p(n)$ becomes the Laplace distribution function. And when $\gamma_k^C = 0.5$, $p(n)$ is the Sqrt noise mentioned in [38].

For OFDM system, the longer OFDM symbol duration provides an advantage in a presence of impulse noise, because impulsive noise energy is spread out among simultaneously transmitted OFDM subcarriers. However, it has been recently recognized that this advantage turns into a disadvantage if the impulsive noise energy exceeds certain threshold. When

OFDM system is applied to UWB channel, the impulsive noise mentioned above will cause severe degradation of the system performance.

Therefore, a powerful channel decoder is required for such applications, that is robust against impulsive noise in order to minimize its impact. We have studied the performance of conventional iterative bit MAP decoder which is designed for Gaussian noise and reported that, under impulsive noise environment, the bit MAP decoding algorithm for TTCM must be modified to match the noise statistical characteristics. The code performance is examined when applied to OFDM systems in non-Gaussian channel environment.

3.3 Impact of Impulsive Noise on MB-OFDM Receivers

At the receiver, the received RF signal is first down-converted to baseband by a fast switching local oscillator, which uses the same hopping frequency as the TFC used by the transmitter. Further it goes through a typical matched filter to obtain discrete signal samples. Since the transmission over different subbands is independent one with respect to another, the performance loss due to channel and additive noise can be considered almost the same as the loss of a system using only one subband. Therefore in the following part, it is sufficient to analyze the impact of channel and noise on a MB-OFDM receiver without considering frequency hopping. Also as we know that OFDM system is efficient for ISI effects, only a very simple one-tap equalizer is used for eliminating the channel imperfectness before OFDM demodulation, which is implemented by FFT.

The MB-OFDM receiver structure is given is Figure 3.9. If the duration of multipath channel $h(t)$ is shorter than the guard interval, it has been shown in [90] that the ZP-OFDM

receiver achieves a good performance (close to that of a CP-OFDM receiver) by using simple overlap-and-add (OLA) operation, FFT and one-tap equalizer.

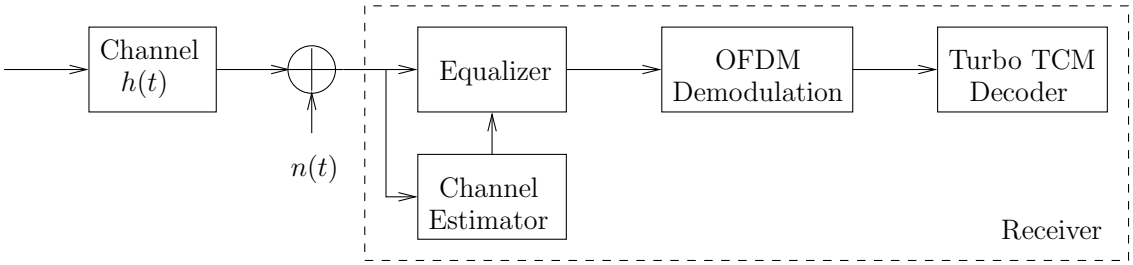

Figure 3.9: The MB-OFDM receiver.

An equivalent receiver model with noise added directly at the input of turbo TCM decoder is given as in Figure 3.10, where $n'(t)$ is the output of $n(t)$ passing through the equalizer and OFDM demodulator, which is basically an FFT operation.

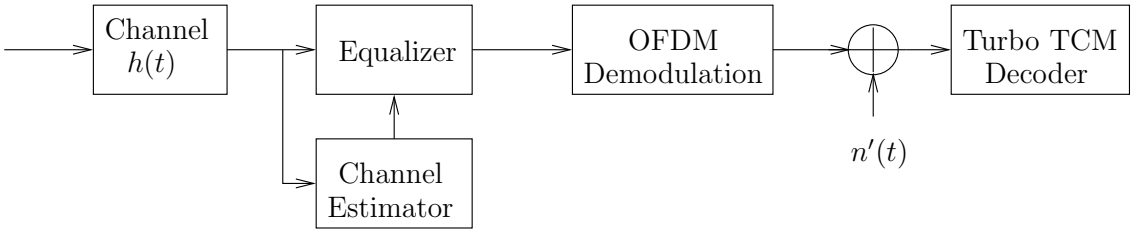

Figure 3.10: An equivalent MB-OFDM receiver with noise added the input of turbo TCM decoder.

Since OFDM demodulation is implemented via FFT, which means $n(t)$ is impulsive noise in time-domain while $n'(t)$ is impulsive noise in frequency-domain. Then we have the relationship before $n(t)$ and $n'(t)$ as follows:

$$n'[m] = \sum_{k=0}^{N_{FFT}-1} n[k]\hat{C}_k e^{j2\pi mk/N_{FFT}} \tag{3.11}$$

where $n[k]$ is k-th discrete sample of $n(t)$ at the input of equalizer with one tap coefficient equal to $\hat{C}_k$ to compensate the channel imperfectness. The noise $n[k]$ is generally non-Gaussian in time domain, however after passing through an N_{FFT}-point complex FFT the

output represents the noise in frequency domain, which could be Gaussian according to central limit theorem. We will compare impulsive noises in time domain and frequency domain in details using computer simulation.

3.4 Decoding Algorithms for Impulsive Noise

The demodulated data symbols will be fed to the bit MAP decoder for turbo TCM. The decoder works in an iterative way, first decoding trellis codes then by parity check codes, separated by de-interleaver. The output of decoder is probability information instead of bit decision, which means soft. It will be used to pass to the next stage for further decoding. The interaction between two component decoders will improve performance of overall system, and approach maximum likelihood performance. The advantage of iterative decoding over a single maximum likelihood decoder is to reduce the complexity with little loss in performance.

We still apply the standard turbo iterative decoding scheme except making certain modifications to match the statistical characteristics of the channel impulsive noise. The branch metric in the basic MAP algorithm is modified according to the distribution of the impulsive noise.

Under Gaussian noise environment, the channel transition probability is based on Gaussian PDF and the Euclidian distance between the received signal and the candidates of the transmitted signal, which can be described by (3.12):

$$p(n_k) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{n_k^2}{2\sigma^2}} \tag{3.12}$$

where

$$|n_k| = \sqrt{(\mathrm{Re}\{Y\} - \mathrm{Re}\{X\})^2 + (\mathrm{Im}\{Y\} - \mathrm{Im}\{X\})^2} \tag{3.13}$$

where Re{.} and Im{.} represent in-phase and quadrature component of the transmitted symbol X within constellation and received noisy symbol Y, respectively. And $|n_k|$ is the Euclidian distance between X and Y. So the conventional MAP decoding is optimized for Gaussian noise by selecting the symbol which has minimum Euclidean distance from transmitted one.

For channels with impulsive noise, the channel noise can not be approximated through Gaussian PDF any more. Therefore, the channel transition probability should be modified accordingly and take into account the statistical distribution of the channel noise. The modified channel probability is given in (3.10) by substituting

$$|n_k| = (|\mathrm{Re}\{Y\} - \mathrm{Re}\{X\}|^{\gamma_k^C} + |\mathrm{Im}\{Y\} - \mathrm{Im}\{X\}|^{\gamma_k^C})^{1/\gamma_k^C} \tag{3.14}$$

where $|n_k|$ is the modified distance between constellation point X and received symbol Y, which is exactly the same as (3.13) if $\gamma_k^C = 2$.

If we assume γ_k^C and γ_k^D are the actual and estimated noise parameter, respectively. Obviously, the optimal decoder requires that $\gamma_k^D = \gamma_k^C$. Otherwise, due to channel variation or estimation error, an additional mismatched error will occur to increase the total error probability at the decoder.

3.5 Performance Evaluation

Coherent detection and ideal Channel State Information (CSI) are assumed in our computer simulation. We compare performance over a UWB channel and memoryless Gaussian channel assuming $h(t) = \delta(t)$.

3.5.1 Density Evolution for Gaussian and UWB Channels

Density evolution is used to determine the threshold, which is the minimum SNR for the decoder to converge assuming infinite block length as well as the extrinsic information exchanged between component decoders is independent Gaussian. In density evolution chart, as long as the SNR is above the threshold, these two constituent transfer curves will never intersect, which means convergence in the limiting case.

In Figure 3.11, we show density evolution for OFDM systems using 16QAM on Gaussian and UWB channels. For Gaussian channels, we find the threshold is $2.6dB$ and show the EXIT chart for $Eb/No = 2.8dB$. On UWB channels, we found that if we take average over all 2000 channels, then EXIT chart shows the clear case of convergence (see curves with solid squares in figure 3.11). However, if we run EXIT over each individual channel instance, then some channel instances require much large SNRs to allow iterative decoding to converge to correct codewords. For example, at $Eb/No = 5.5dB$, about 2% of the channels are difficult to converge (see curves with crosses in figure 3.11). We call them "bad" channels. When Eb/No is small, the percentage of worst channels increases significantly. For example, when $Eb/No = 4.5dB$, about 20% of channels are bad. Good performance can only be achieved unless the interleave can fully randomize the extrinsic information over all channels. If the

bits of a packet are interleaved over a number of channels containing significant amount of "bad" channels, then the performance will be much poorer.

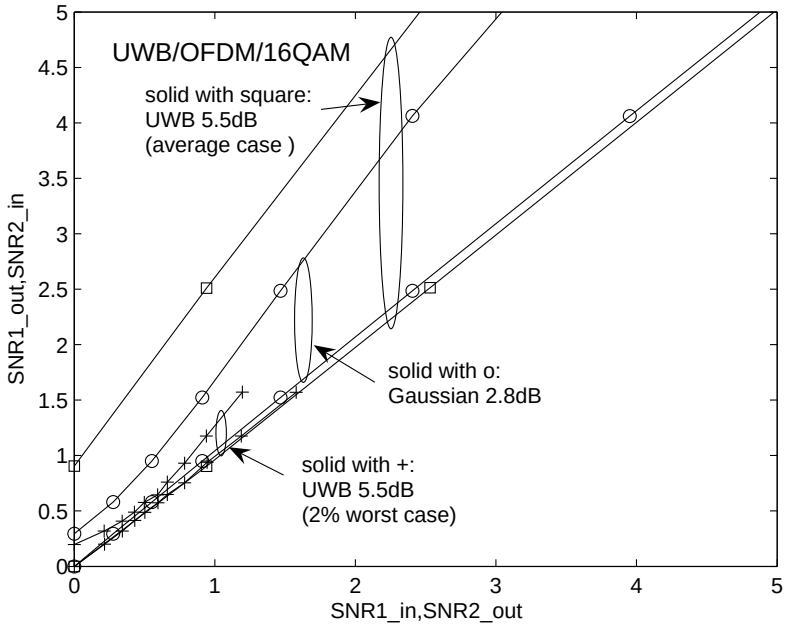

Figure 3.11: Density evolution for 16QAM/OFDM on AWGN and UWB channels.

Figure 3.12 presents the density evolution analysis for OFDM system using 64QAM on Gaussian and UWB channels. For Gaussian channels, we find the threshold is $3.7dB$ and show the EXIT chart for $Eb/No = 4.2dB$. For UWB channels, when we set $Eb/No = 9.2dB$, about 2% of the channels are bad.

3.5.2 Impulsive Noise Mismatch

Based on two different settings in Figures 3.9 and 3.10, we do simulation for both Gaussian and UWB channels with impulsive noise inserted either before or after the FFT block within OFDM demodulation.

1) Case I: AWGN channel before FFT.

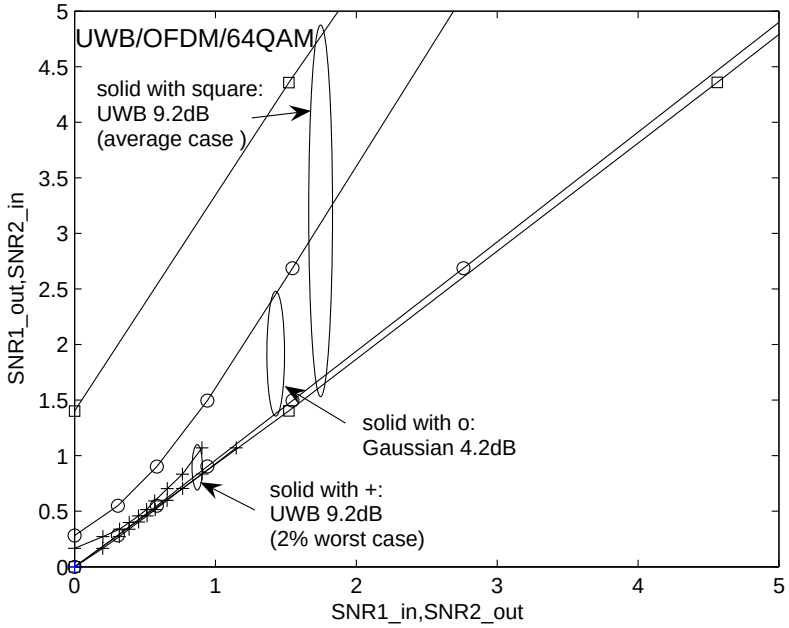

Figure 3.12: Density evolution for 64QAM/OFDM on AWGN and UWB channels.

We generate impulsive noise $n(t)$ following the generalized Gaussian model with various γ^C as in Figure 3.9. Normalized distribution of the equivalent noise $n'(t)$ after FFT for each different γ^C is shown in Figure 3.13-3.15. It is more like Gaussian distribution (corresponding to $\gamma^D = 2.0$) than Sqrt (corresponding to $\gamma^D = 0.5$) or Laplace distribution (corresponding to $\gamma^D = 1.0$) no matter what actual value γ^C is.

BER performance vs. SNR curves are presented in Figure 3.16-3.19 for both matched and mismatched cases. We see that in this setting decoder matched to $n(t)$ is not necessarily better than those mismatched cases. The results show that $\gamma^D = 2.0$ is always better than $\gamma^D = 0.5$ or 1.0 no matter what γ^C is, since $\gamma^D = 2.0$ is always matched to the distribution of the equivalent noise $n'(t)$ shown above. So we suggest Gaussian decoder with $\gamma^D = 2.0$ in this setting, which is always matched to $n'(t)$ and is quite robust to the impulsive noise mismatch.

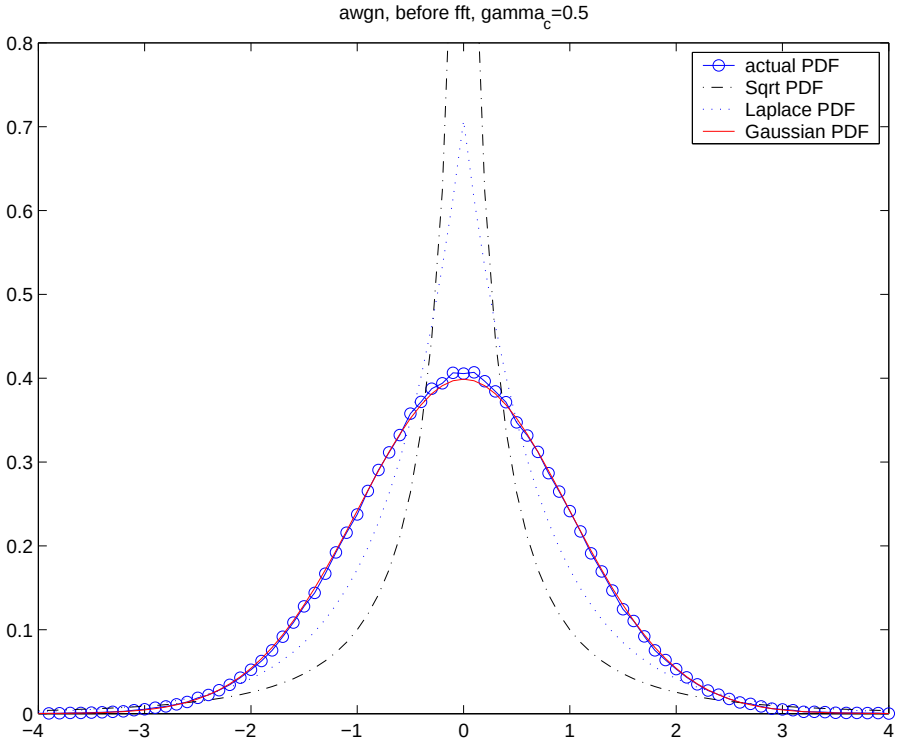

Figure 3.13: Normalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 0.5$)

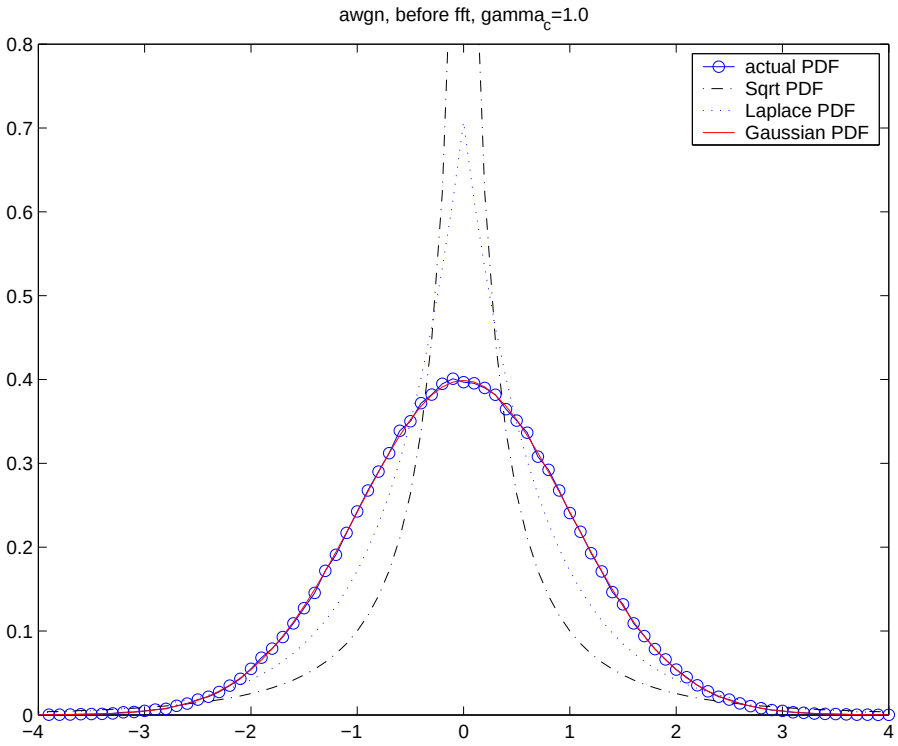

Figure 3.14: Normalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 1.0$)

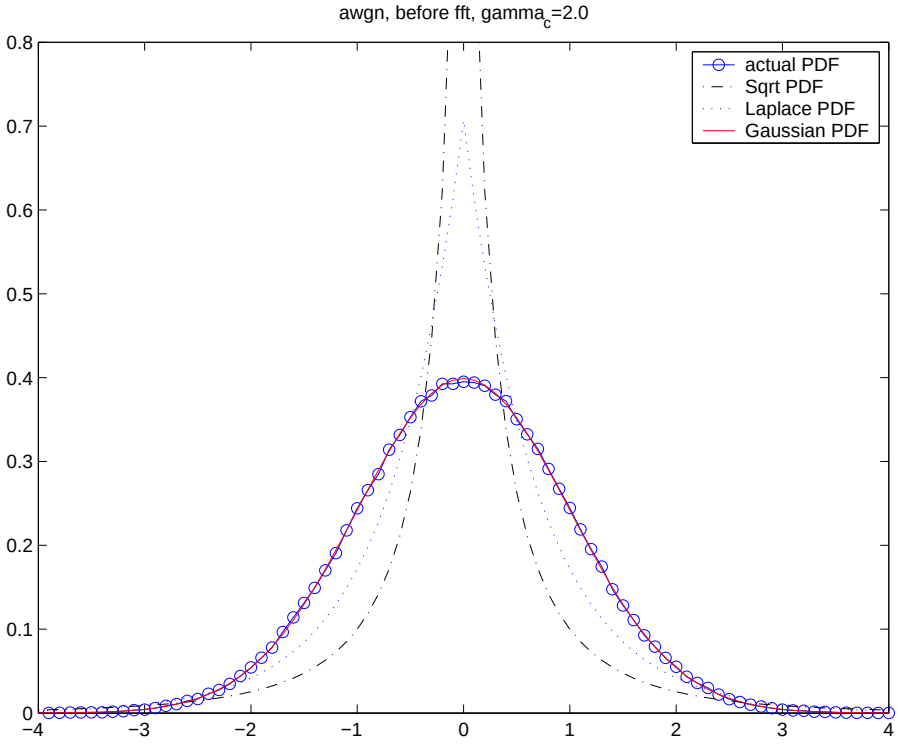

Figure 3.15: Normalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 2.0$)

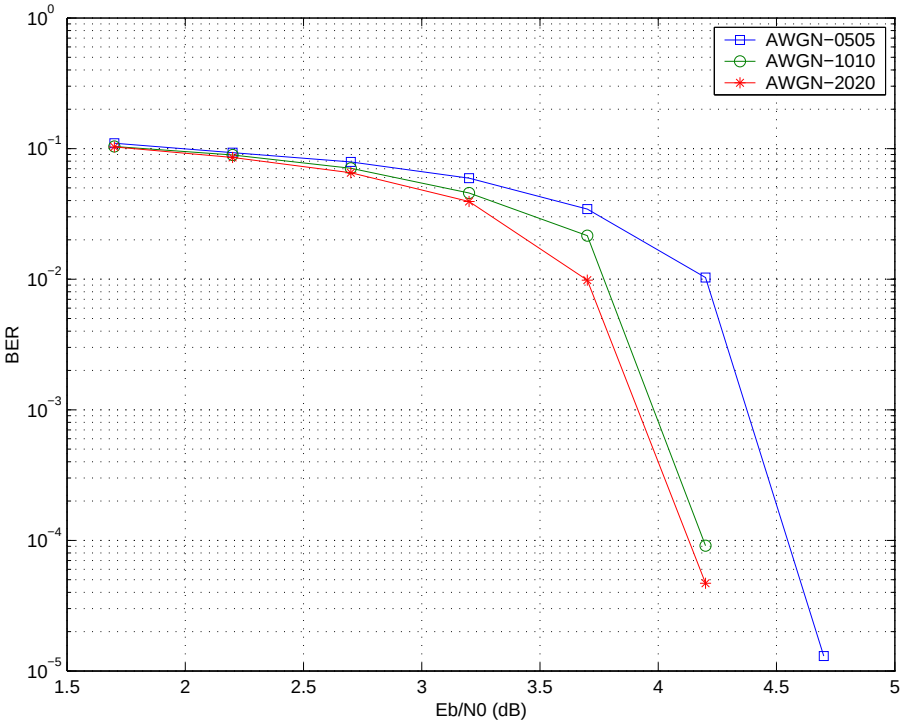

Figure 3.16: BER vs. SNR: AWGN before FFT. ($\gamma^C = \gamma^D = 0.5, 1.0, 2.0$, respectively)

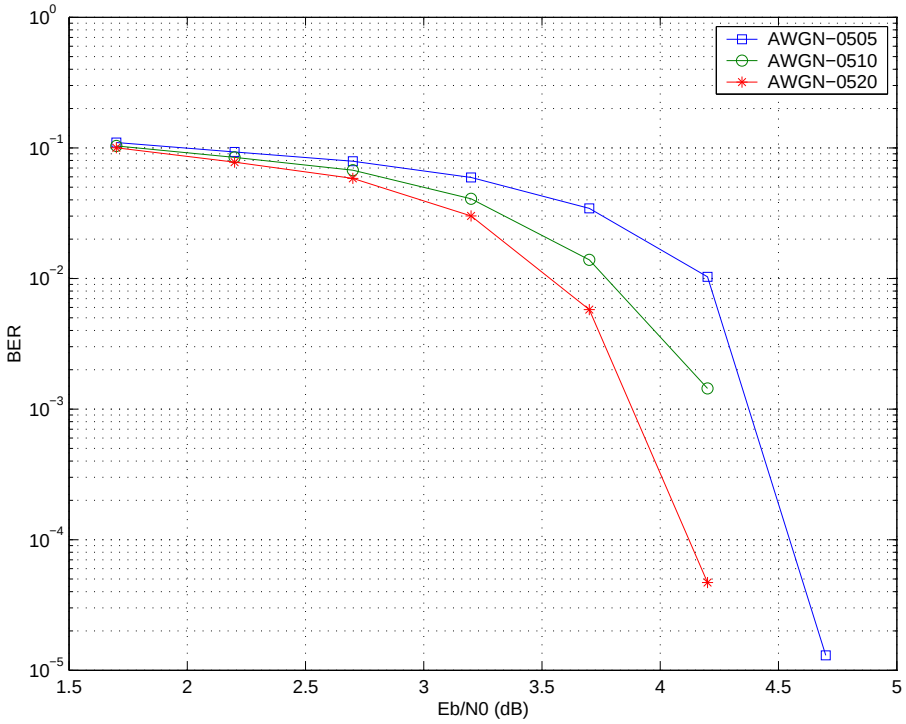

Figure 3.17: BER vs. SNR: AWGN before FFT. ($\gamma^C = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

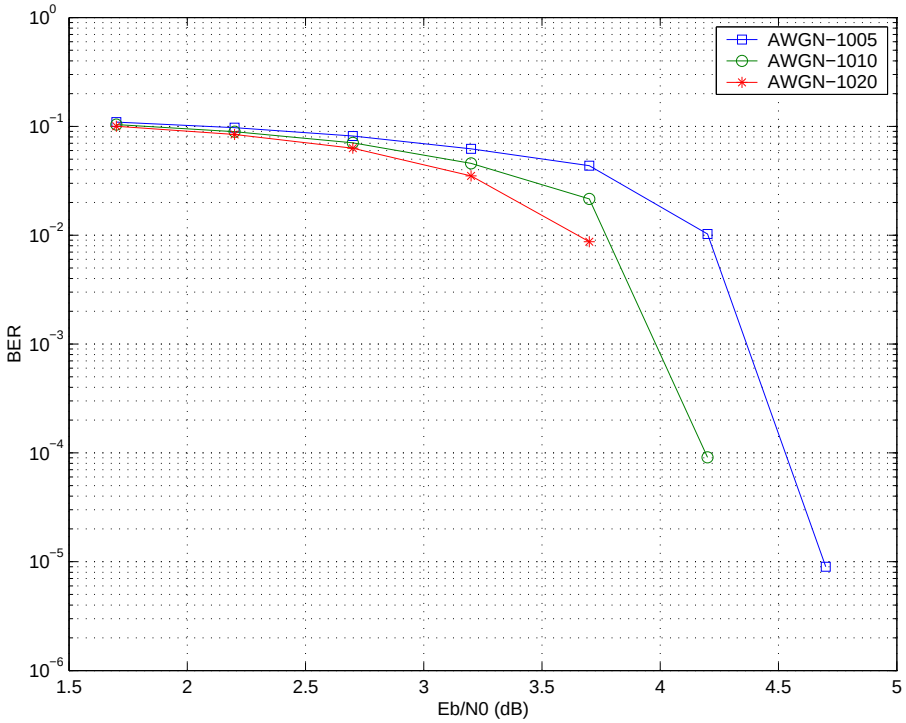

Figure 3.18: BER vs. SNR: AWGN before FFT. ($\gamma^C = 1.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

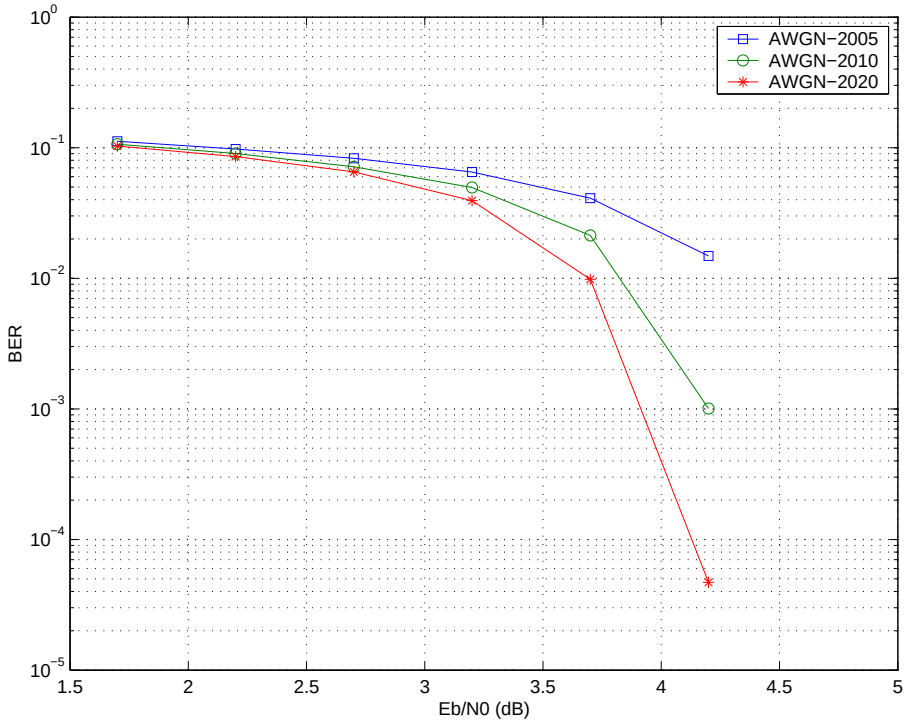

Figure 3.19: BER vs. SNR: AWGN before FFT. ($\gamma^C = 2.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

2) Case II: UWB channel before FFT.

We again generate impulsive noise $n(t)$ following the generalized Gaussian model with various γ^C as in Figure 3.9. Normalized distribution of the equivalent noise $n'(t)$ after FFT for each different γ^C is shown in Figure 3.20 - 3.22. It is more like Sqrt distribution (corresponding to $\gamma^D = 0.5$) than Laplace (corresponding to $\gamma^D = 1.0$) or Gaussian distribution (corresponding to $\gamma^D = 2.0$) no matter what actual value γ^C is.

BER performance vs. SNR curves are presented in Figure 3.23- 3.34 for both matched and mismatched cases. Again, We see that decoder matched to $n(t)$ is not necessarily better than mismatched cases. The results show that $\gamma^D = 0.5$ is always better than $\gamma^D = 1.0$ or 2.0 no matter what γ^C is, since $\gamma^D = 0.5$ is always matched to the distribution of the

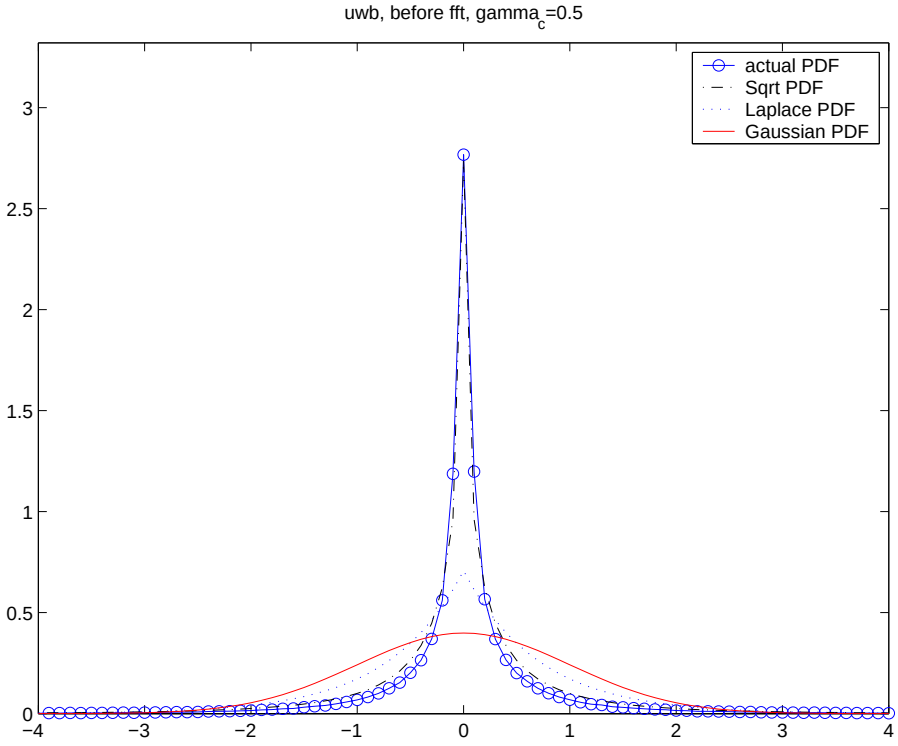

Figure 3.20: Normalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 0.5$)

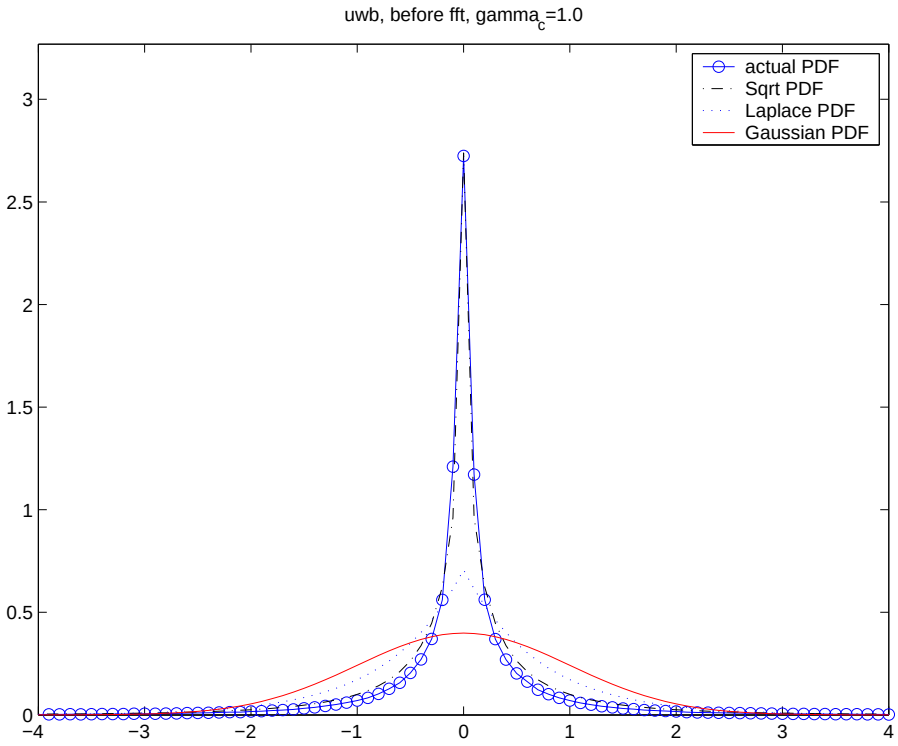

Figure 3.21: Normalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 1.0$)

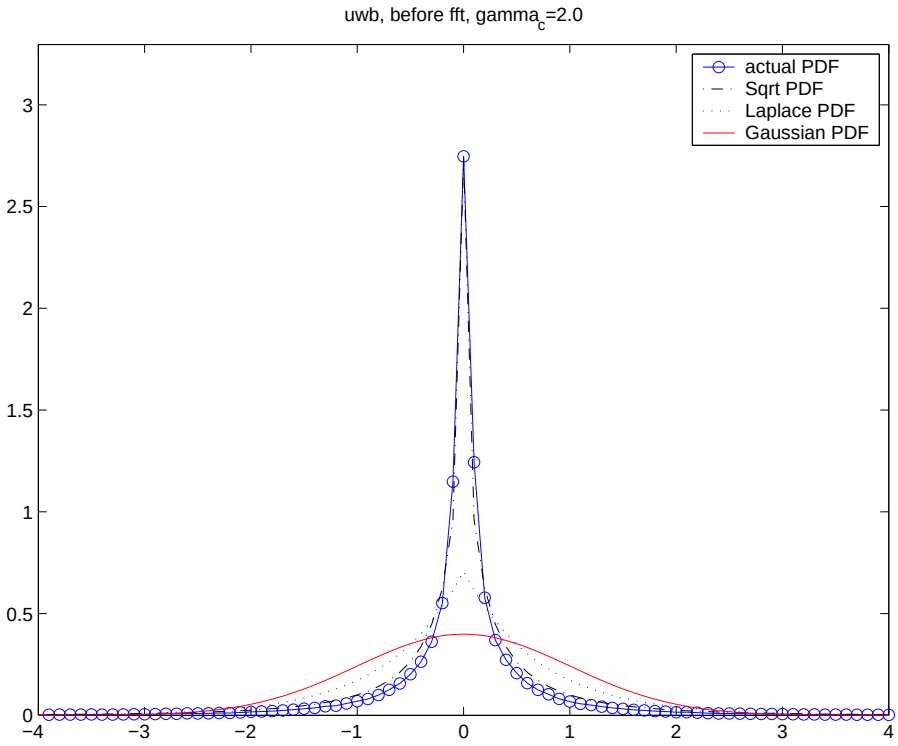

Figure 3.22: NNormalized noise $n'(t)$ distribution after FFT. ($\gamma^C = 2.0$)

equivalent noise $n'(t)$ shown above. So we suggest Sqrt decoder with $\gamma^D = 0.5$ in this setting, which is always matched to $n'(t)$ and is quite robust to the impulsive noise mismatch.

3) Case III: AWGN channel after FFT.

This time we generate impulsive noise $n'(t)$ following the generalized Gaussian model as in Figure 3.10. BER performance vs. SNR curves are presented in Figure 3.27 - 3.30 for matched and mismatched cases. The simulation results confirm that decoder matched to the distribution of $n'(t)$ is always better than mismatched case.

4) Case IV: UWB channel after FFT.

Again we generate impulsive noise $n'(t)$ following the generalized Gaussian model as in Figure 3.10. BER performance vs. SNR curve is presented in Figure 3.31 - 3.34 for both

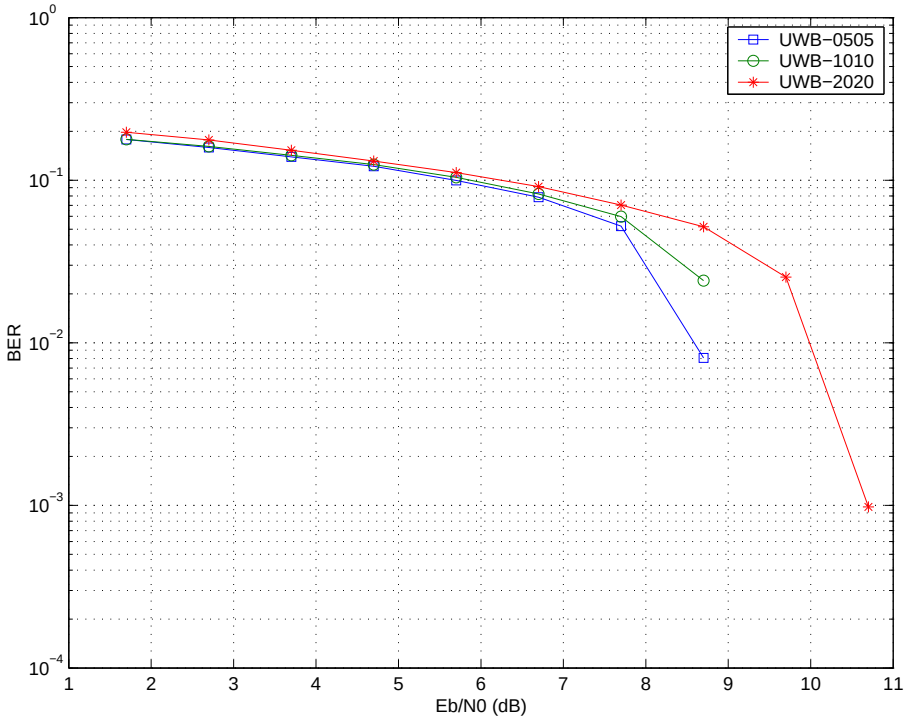

Figure 3.23: BER vs. SNR: UWB before FFT. ($\gamma^C = \gamma^D = 0.5, 1.0, 2.0$, respectively)

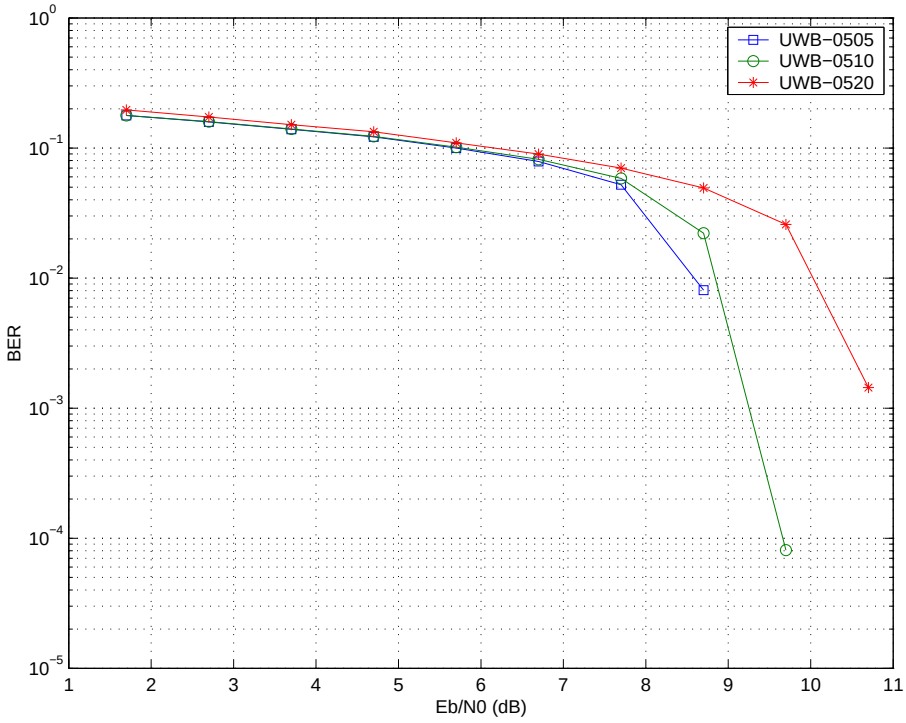

Figure 3.24: BER vs. SNR: UWB before FFT. ($\gamma^C = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

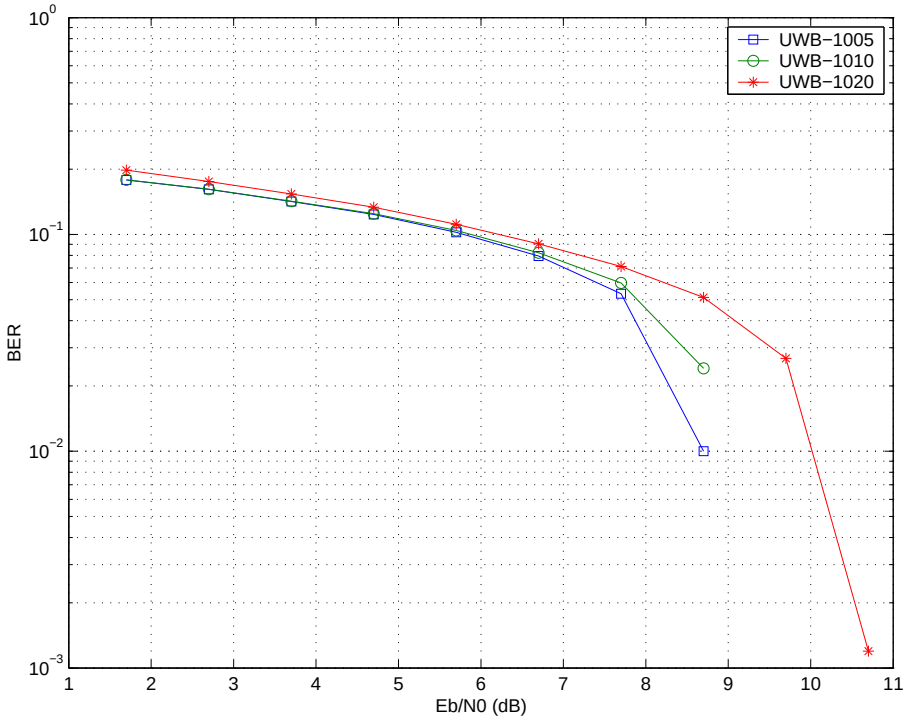

Figure 3.25: BER vs. SNR: UWB before FFT. ($\gamma^C = 1.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

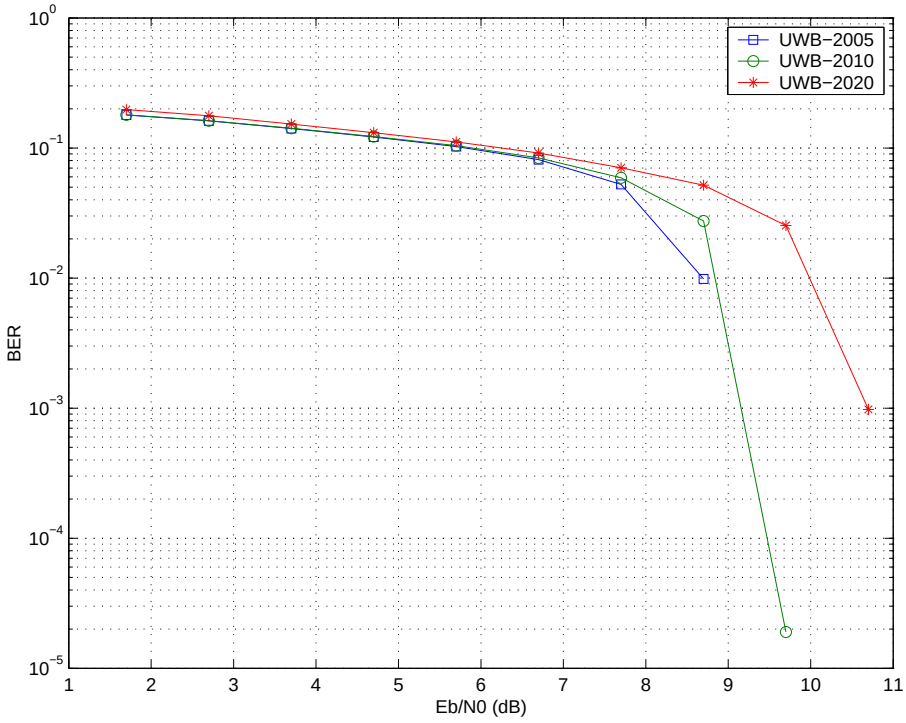

Figure 3.26: BER vs. SNR: UWB before FFT. ($\gamma^C = 2.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

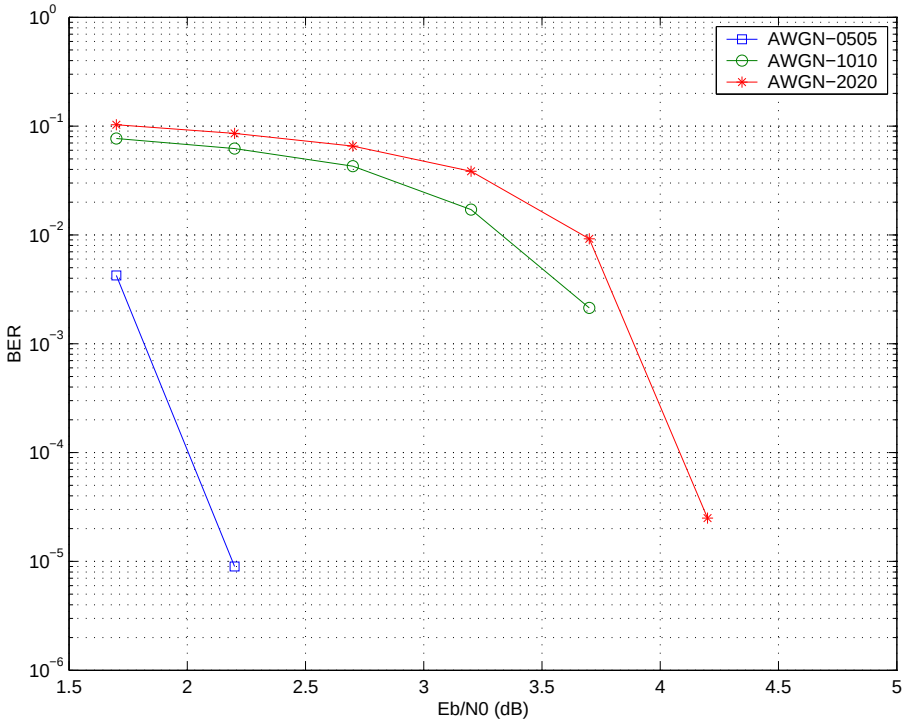

Figure 3.27: BER vs. SNR: AWGN after FFT. ($\gamma^C = \gamma^D = 0.5, 1.0, 2.0$, respectively)

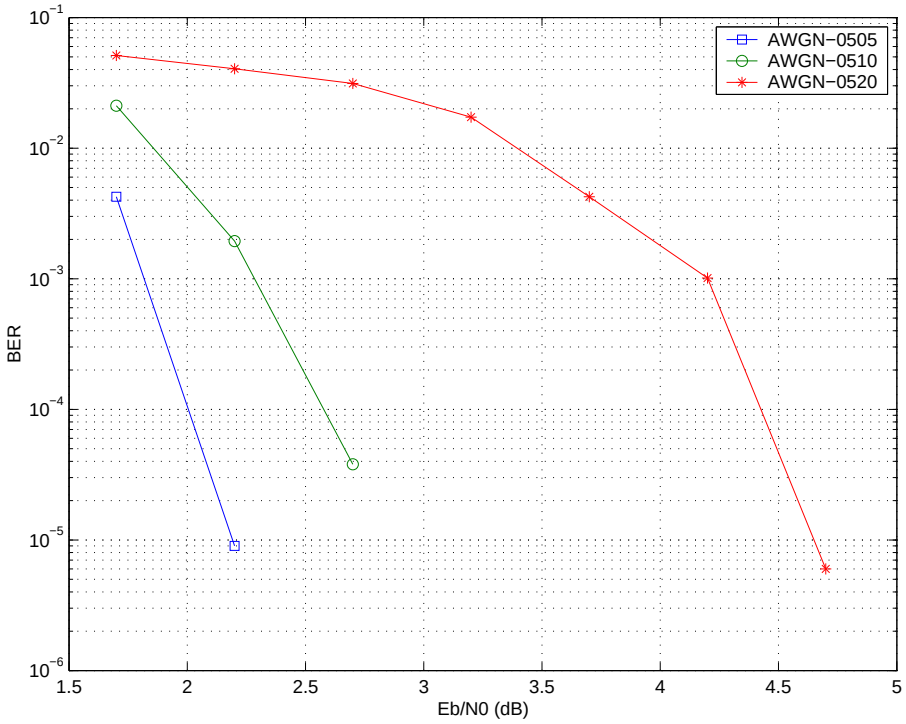

Figure 3.28: BER vs. SNR: AWGN after FFT. ($\gamma^C = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

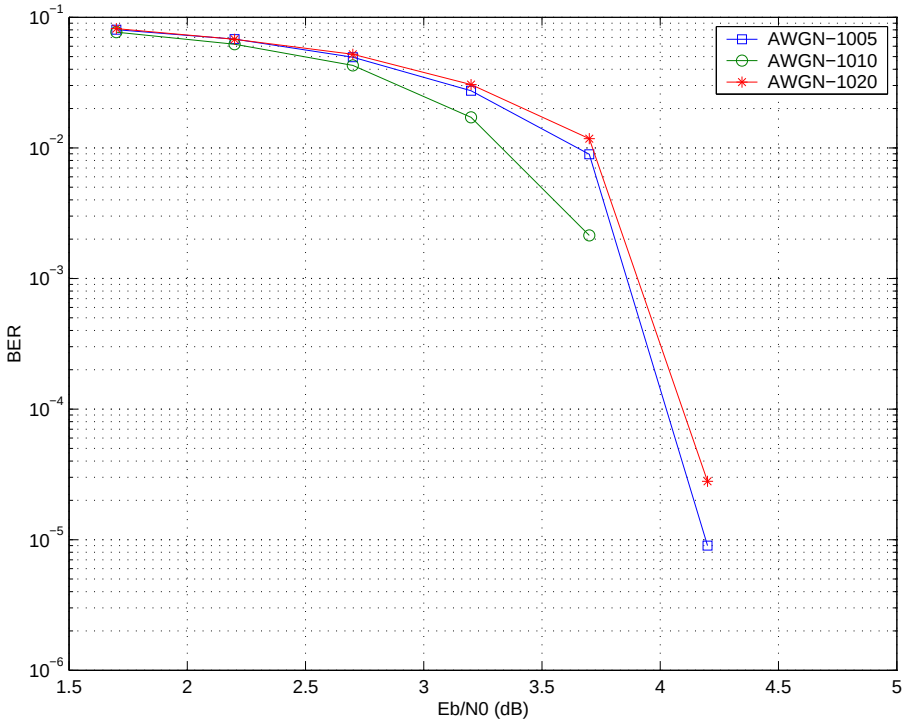

Figure 3.29: BER vs. SNR: AWGN after FFT. ($\gamma^C = 1.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

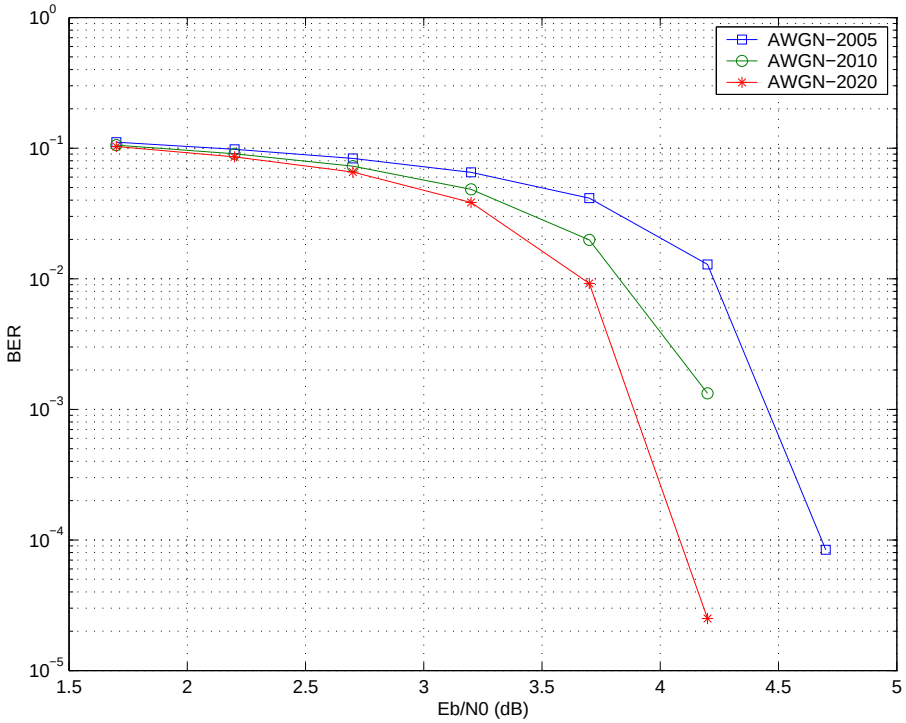

Figure 3.30: BER vs. SNR: AWGN after FFT. ($\gamma^C = 2.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

matched and mismatched cases. The simulation results confirm that decoder matched to the distribution of $n'(t)$ is always better than mismatched case.

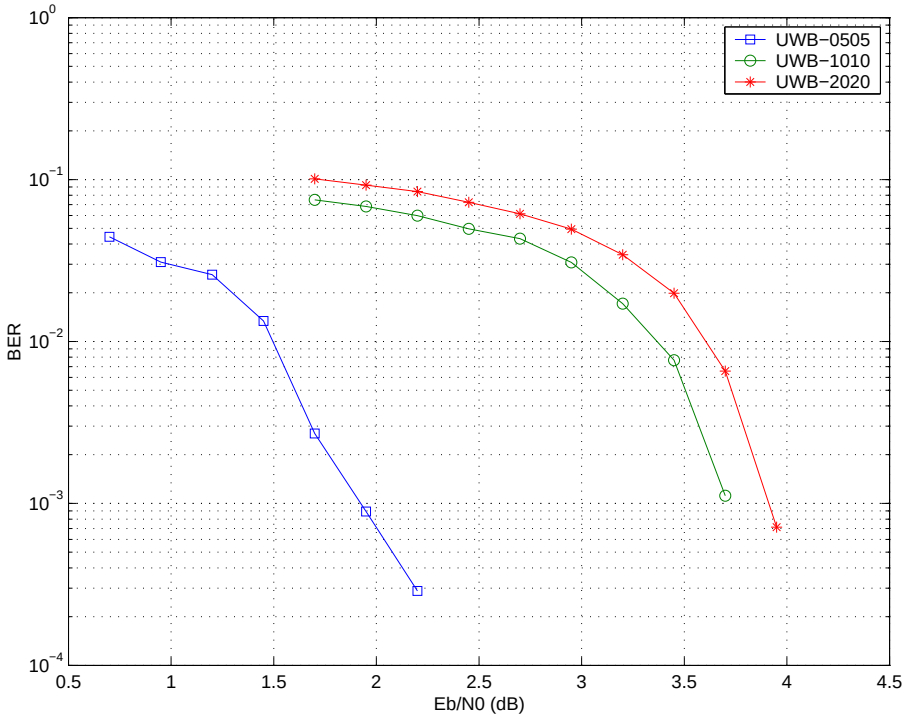

Figure 3.31: BER vs. SNR: UWB after FFT. ($\gamma^C = \gamma^D = 0.5, 1.0, 2.0$, respectively)

5) Case V: AWGN channel plus mixed Gaussian noise with $\epsilon = 0.5$.

We generate impulsive noise $n(t)$ as in Figure 3.9 following the mixed Gaussian model:

$$p(n_k) = \epsilon p_1(n_k) + (1 - \epsilon) p_2(n_k) \tag{3.15}$$

where $p_1(n_k), p_2(n_k)$ are two Gaussian PDF with zero mean and different variances equal to 1 and 100, respectively, and $p(n_k)$ is the mixed Gaussian PDF for the impulsive noise $n(t)$ as shown in Figure 3.35. We choose coefficient $\epsilon = 0.5$. Normalized distribution for equivalent noise $n'(t)$ after FFT is shown here in Figure 3.36 together with the BER performance simulation results in Figure 3.37. It is more like Gaussian distribution (corresponding to

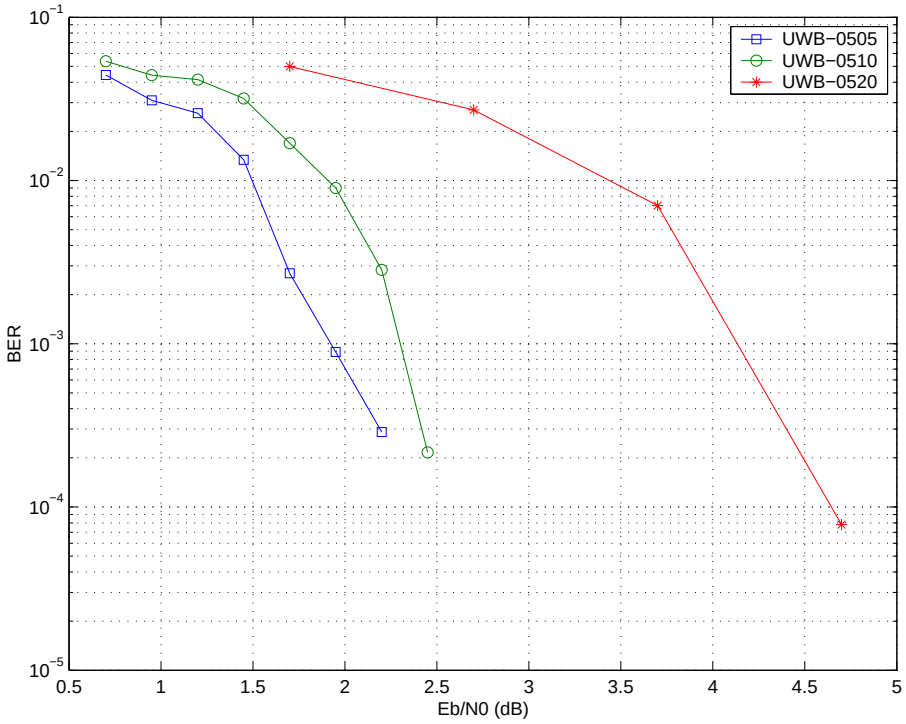

Figure 3.32: BER vs. SNR: UWB after FFT. ($\gamma^C = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

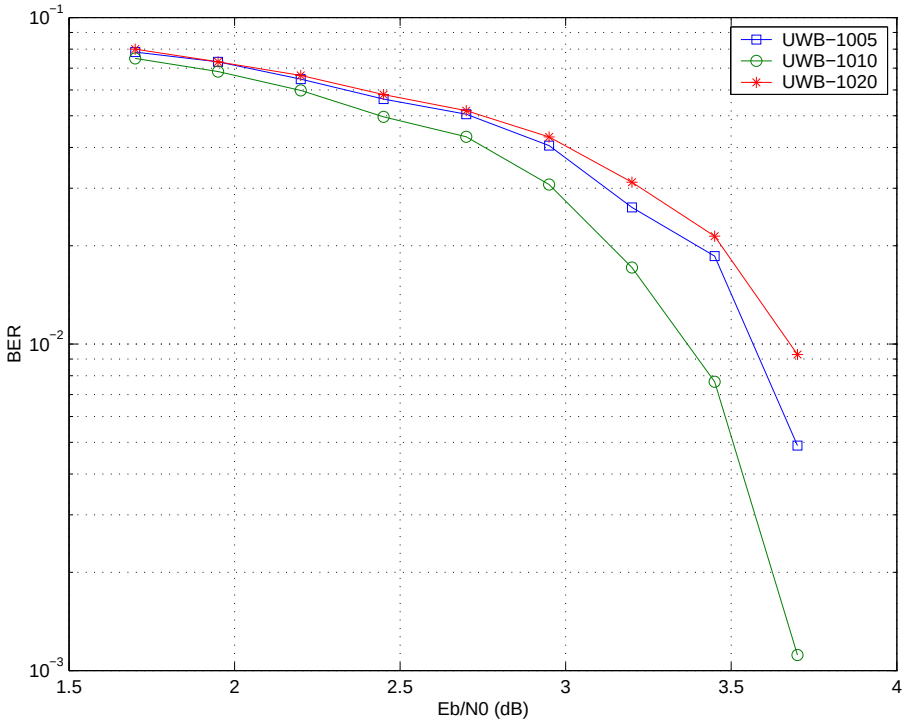

Figure 3.33: BER vs. SNR: UWB after FFT. ($\gamma^C = 1.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

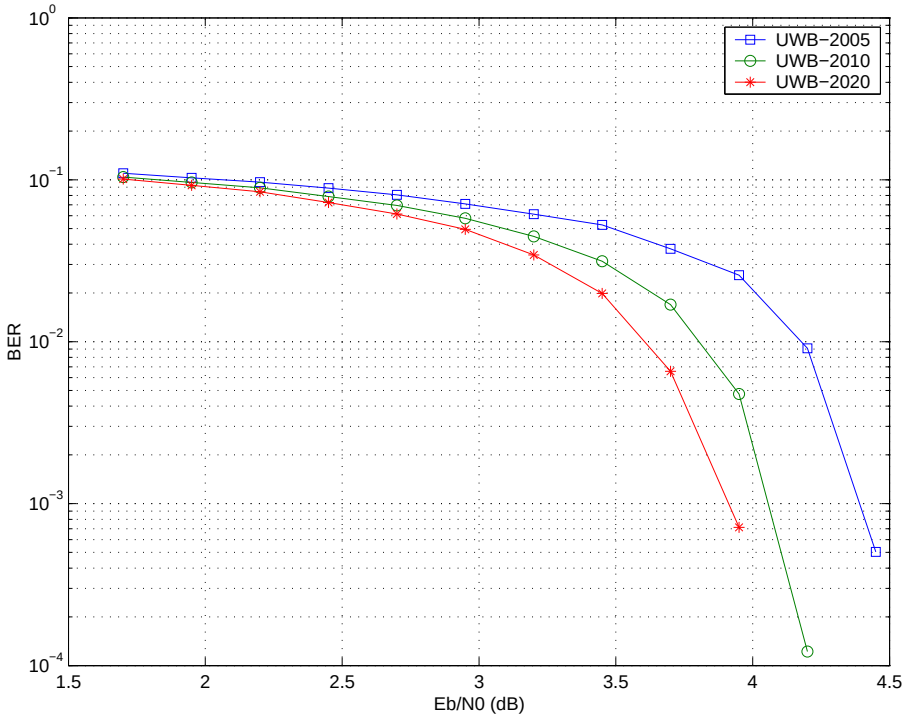

Figure 3.34: BER vs. SNR: UWB after FFT. ($\gamma^C = 2.0, \gamma^D = 0.5, 1.0, 2.0$, respectively)

$\gamma^D = 2.0$) than Sqrt (corresponding to $\gamma^D = 0.5$) or Laplace distribution (corresponding to $\gamma^D = 1.0$). Then, $\gamma^D = 2.0$ achieves robust performance as we predicted.

6) Case VI: UWB channel plus mixed Gaussian noise with $\epsilon = 0.5$.

Again we generate impulsive noise $n(t)$ as in Figure 3.9 following the same mixed Gaussian model with coefficient $\epsilon = 0.5$ shown in Figure 3.35. Normalized distribution for equivalent noise $n'(t)$ after FFT is shown here in Figure 3.38 together with the BER performance simulation results in Figure 3.39. It is more like Sqrt distribution (corresponding to $\gamma^D = 0.5$) than Laplace (corresponding to $\gamma^D = 1.0$) or Gaussian distribution (corresponding to $\gamma^D = 2.0$). Then, $\gamma^D = 0.5$ achieves robust performance as we predicted.

7) Case VII: AWGN channel with narrowband interference (NBI).

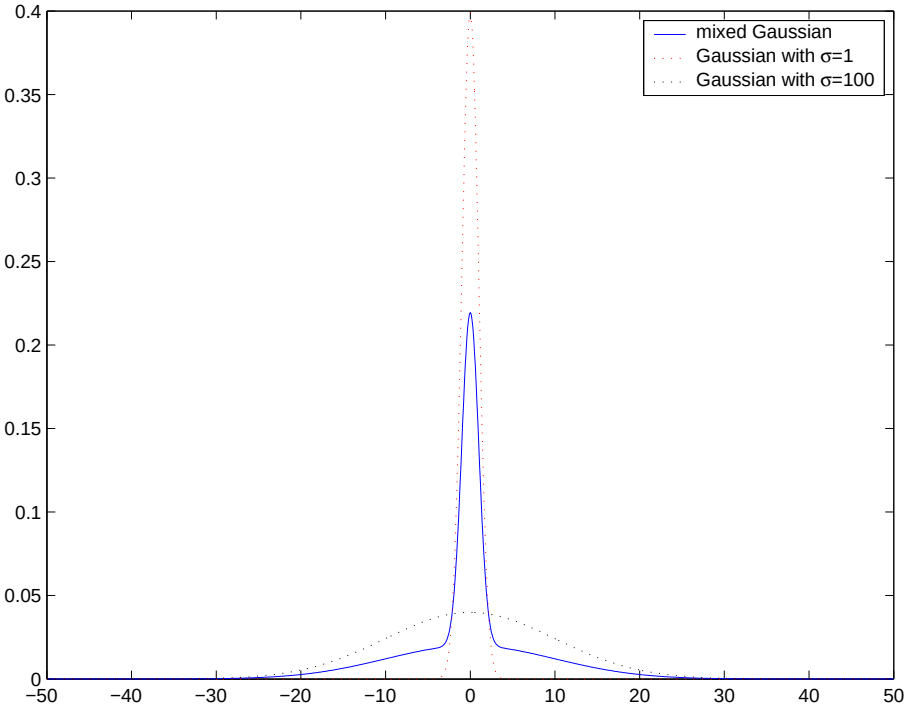

Figure 3.35: Noise distribution $n(t)$ before FFT. ($\epsilon = 0.5$)

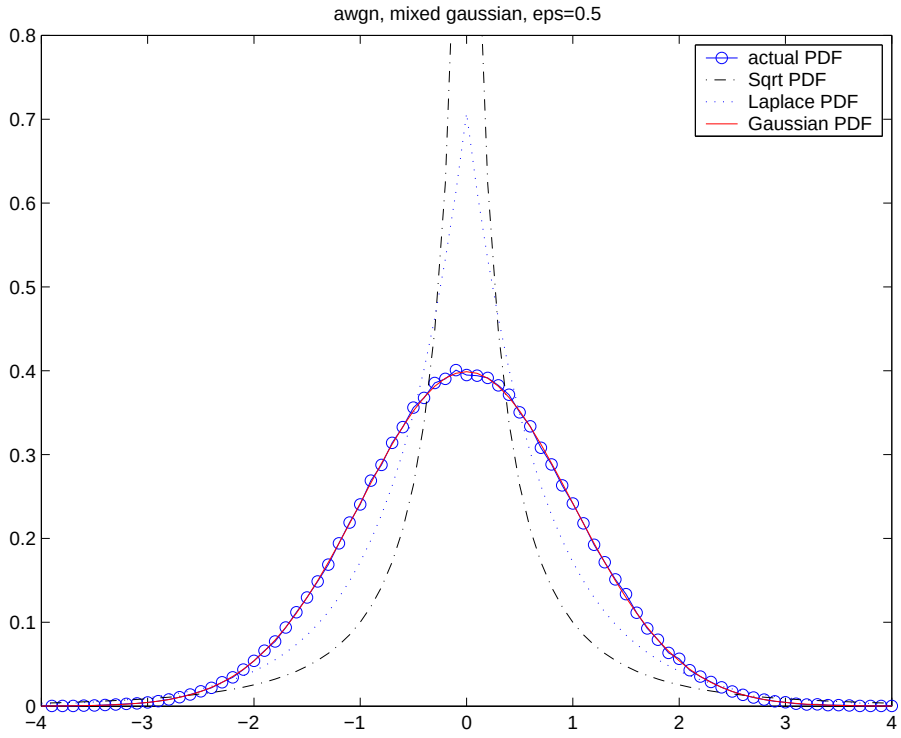

Figure 3.36: Normalized noise distribution $n'(t)$ after FFT. ($\epsilon = 0.5$)

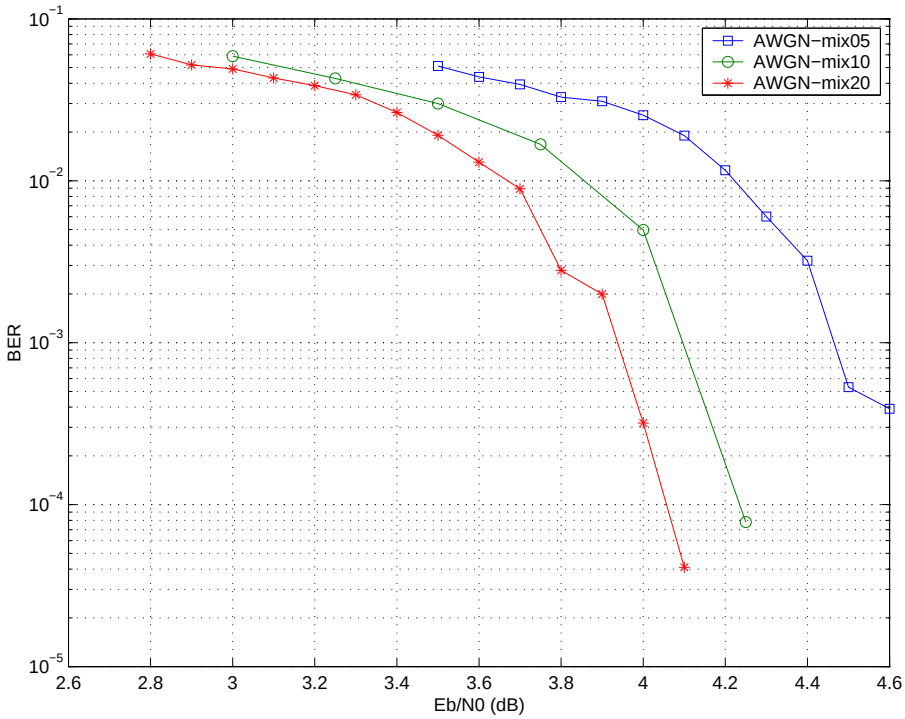

Figure 3.37: BER vs SNR: AWGN plus mixed Gaussian. ($\epsilon = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

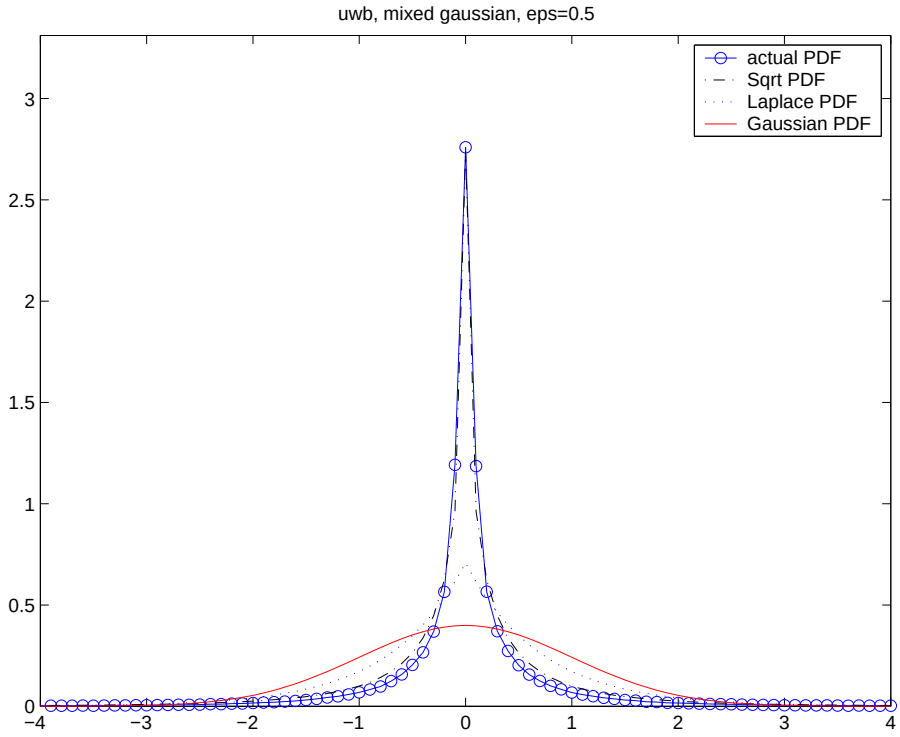

Figure 3.38: Normalized noise $n'(t)$ distribution after FFT. ($\epsilon = 0.5$)

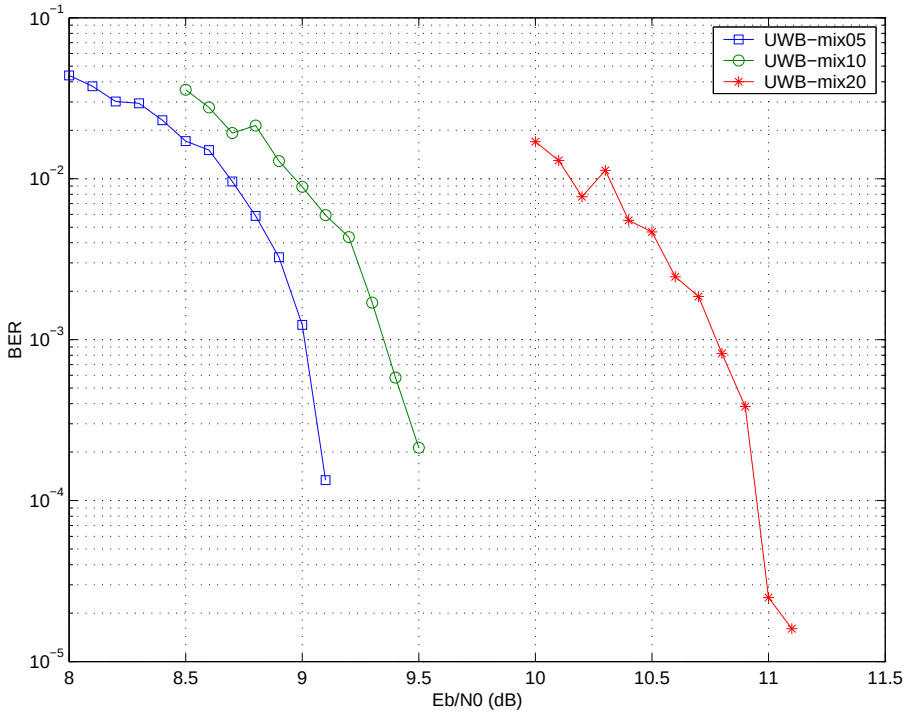

Figure 3.39: BER vs SNR: UWB plus mixed Gaussian. ($\epsilon = 0.5, \gamma^D = 0.5, 1.0, 2.0$, respectively)

We model impulsive noise $n(t)$ as in Figure 3.9 to be a Gaussian noise plus some narrowband interference. Here the Gaussian noise is zero mean and unit variance while the narrowband interference is assumed to be a complex sinusoid with fixed frequency but various average power. We want to investigate the BER performance vs. signal-to-interference-plus-noise (SINR) for different sinusoid interference to Gaussian noise power ratio.

When this ratio is low, e.g. 0.01. We observe that the distributions of $n(t)$ and $n'(t)$ are shown in Figure 3.40 and Figure 3.41, respectively. The BER performance vs. SINR simulation result is given in Figure 3.42. Since $n'(t)$ is more like Gaussian distribution (corresponding to $\gamma^D = 2.0$) than Sqrt (corresponding to $\gamma^D = 0.5$) or Laplace distribution (corresponding to $\gamma^D = 1.0$). Then, $\gamma^D = 2.0$ achieves robust performance as we predicted.

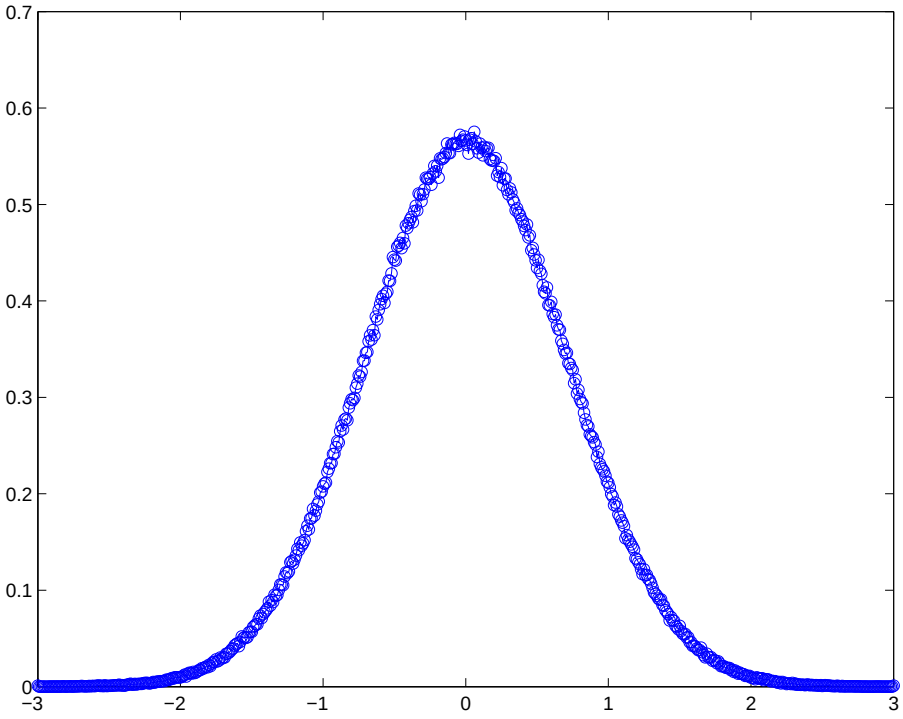

Figure 3.40: Noise distribution $n(t)$ before FFT. (power ratio is 0.01)

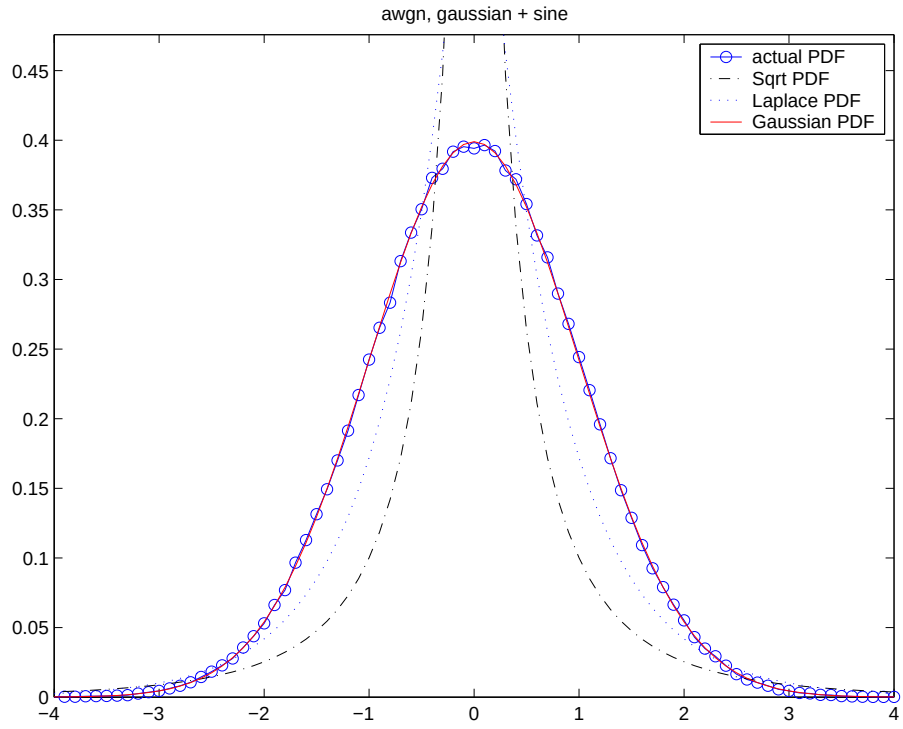

Figure 3.41: Normalized noise distribution $n'(t)$ after FFT. (power ratio is 0.01)

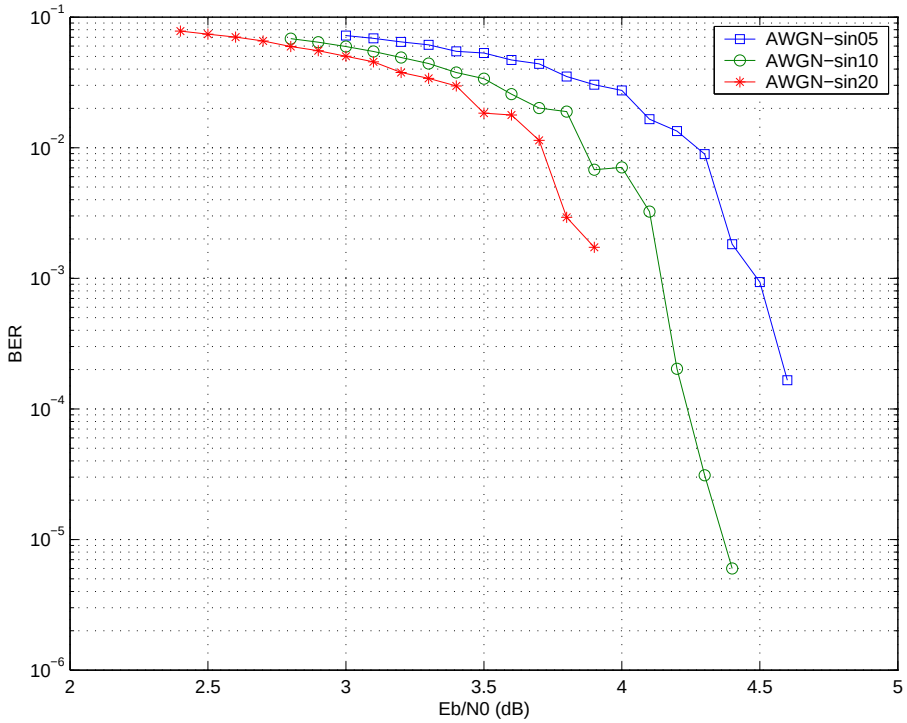

Figure 3.42: BER vs SINR: AWGN with NBI. (power ratio is 0.01, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

When this ratio is high, e.g. 10. We observe that the distributions of $n(t)$ and $n'(t)$ are shown in Figure 3.43 and Figure 3.44, respectively. The BER performance vs. SINR simulation result is given in Figure 3.45. Since $n'(t)$ is more like Sqrt distribution (corresponding to $\gamma^D = 0.5$) than Laplace (corresponding to $\gamma^D = 1.0$) or Gaussian distribution (corresponding to $\gamma^D = 2.0$). Then, $\gamma^D = 0.5$ achieves robust performance as we predicted.

When this ratio is medium, e.g. 0.5. We observe that the distributions of $n(t)$ and $n'(t)$ are shown in Figure 3.46 and Figure 3.47, respectively. The BER performance vs. SINR simulation result is given in Figure 3.48. Since $n'(t)$ is more like Laplace distribution (corresponding to $\gamma^D = 1.0$) than Sqrt (corresponding to $\gamma^D = 0.5$) or Gaussian distribution (corresponding to $\gamma^D = 2.0$). Then, $\gamma^D = 1.0$ achieves robust performance as we predicted.

8) Case VIII: UWB channel with narrowband interference (NBI).

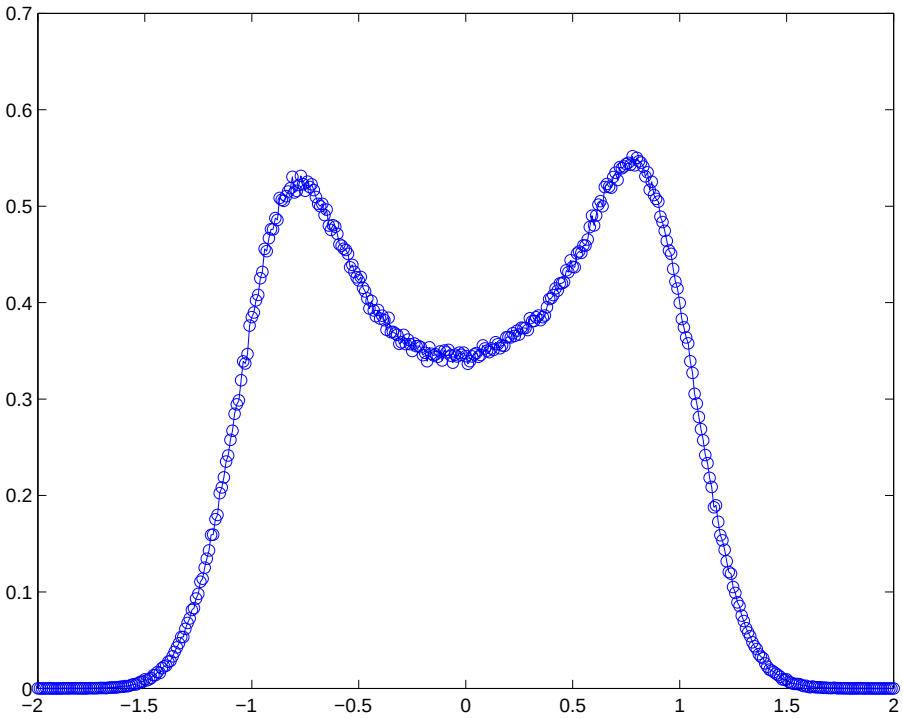

Figure 3.43: Noise distribution $n(t)$ before FFT. (power ratio is 10)

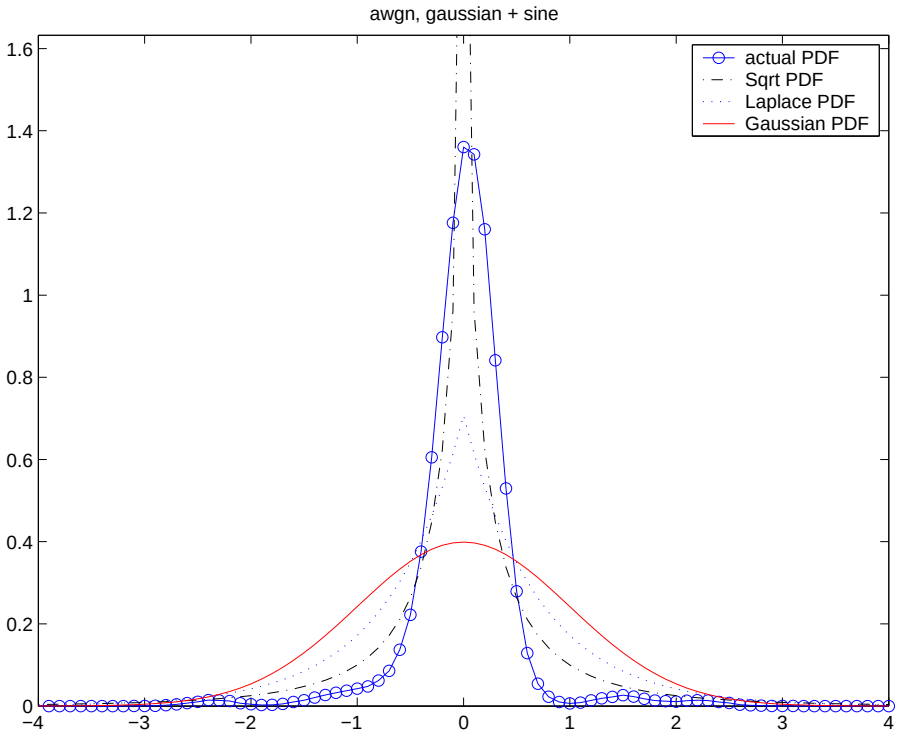

Figure 3.44: Normalized noise distribution $n'(t)$ after FFT. (power ratio is 10)

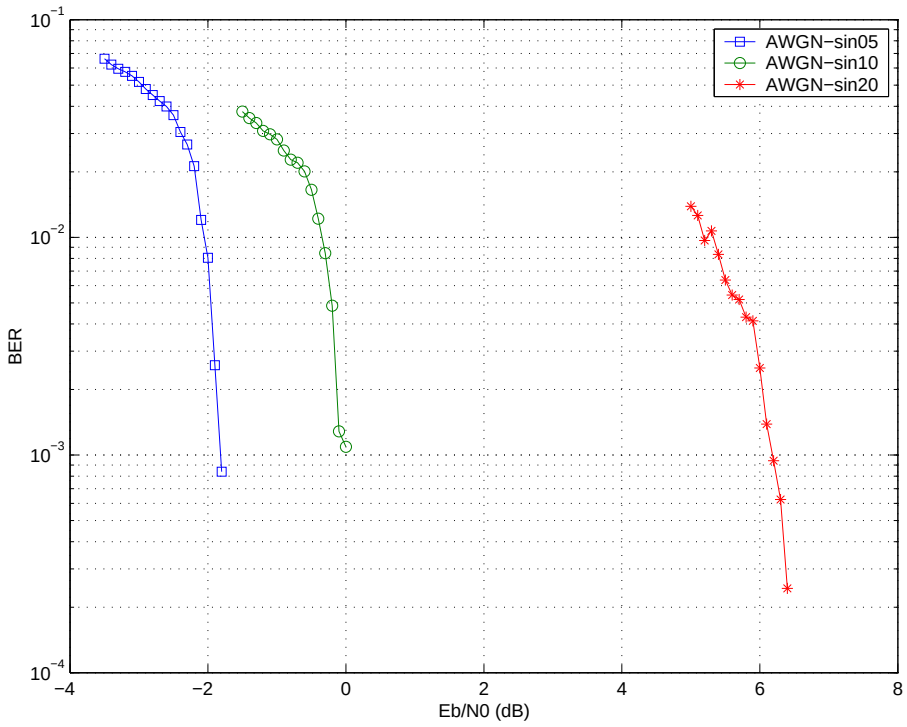

Figure 3.45: BER vs SINR: AWGN with NBI. (power ratio is 10, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

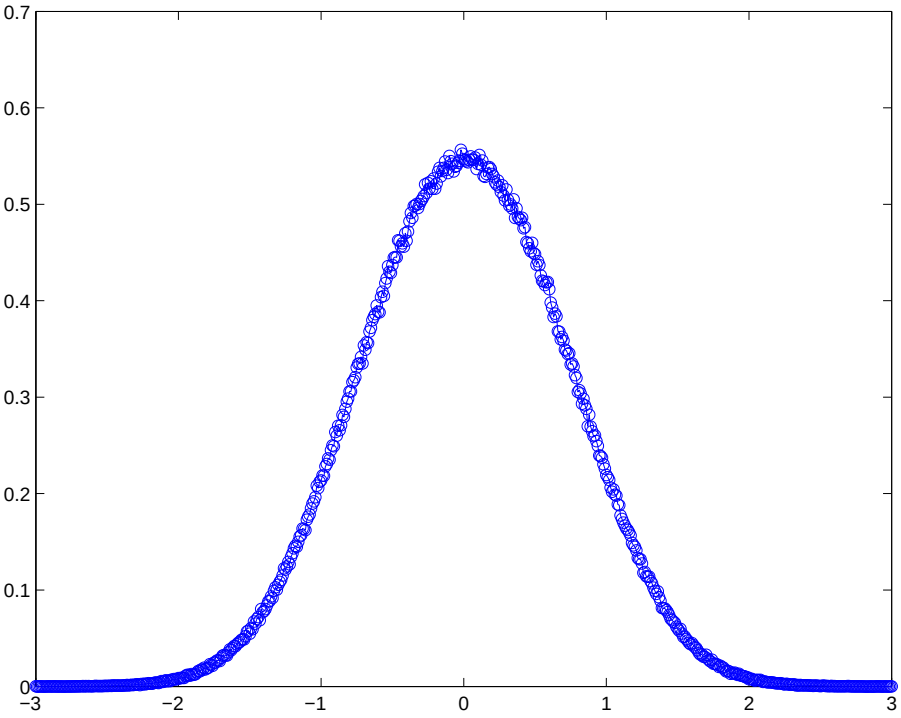

Figure 3.46: Noise distribution $n(t)$ before FFT. (power ratio is 0.5)

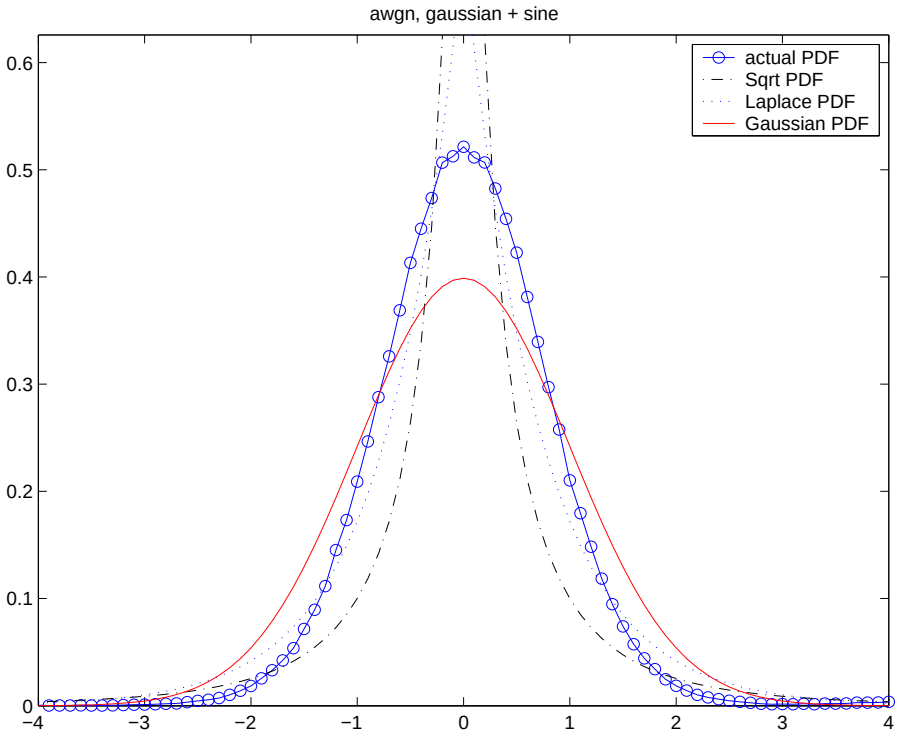

Figure 3.47: Normalized noise distribution $n'(t)$ after FFT. (power ratio is 0.5)

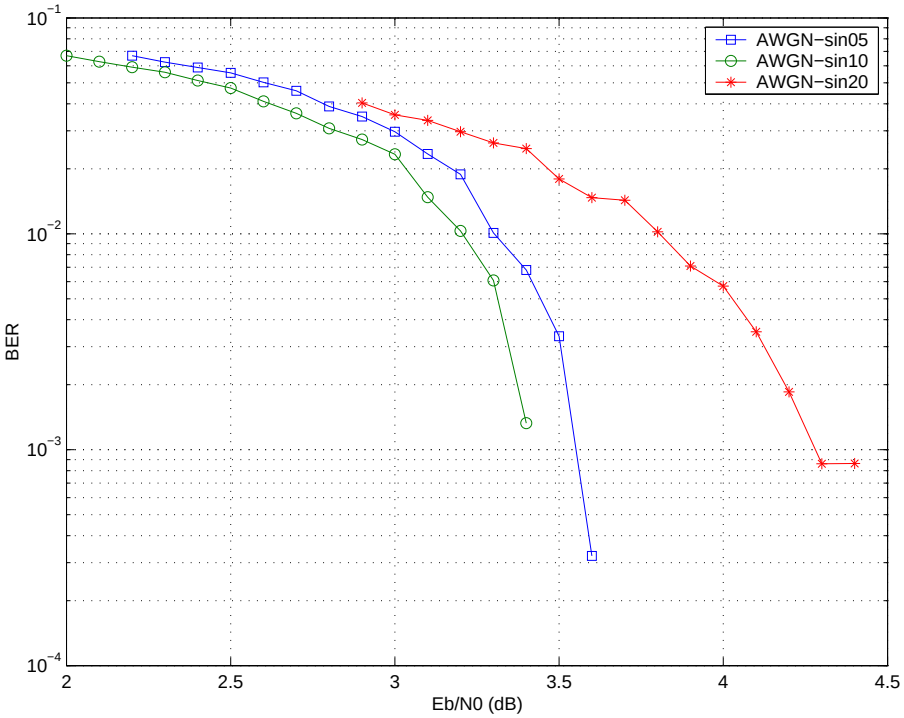

Figure 3.48: BER vs SINR: AWGN with NBI. (power ratio is 0.5, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

Similarly we model impulsive noise $n(t)$ as in Figure 3.9 to be a Gaussian noise plus some narrowband interference.

However no matter what the sinusoid interference to Gaussian noise power ratio is, we observe that the distribution of $n'(t)$ is always like the way shown as in Figure 3.49, which is more like Sqrt distribution (corresponding to $\gamma^D = 0.5$) than Laplace (corresponding to $\gamma^D = 1.0$) or Gaussian distribution (corresponding to $\gamma^D = 2.0$). The BER performance vs. SINR simulation result is given in Figure 3.50-3.52 for power ratio 0.01, 10 and 0.5, respectively. Then, $\gamma^D = 0.5$ achieves robust performance as we predicted.

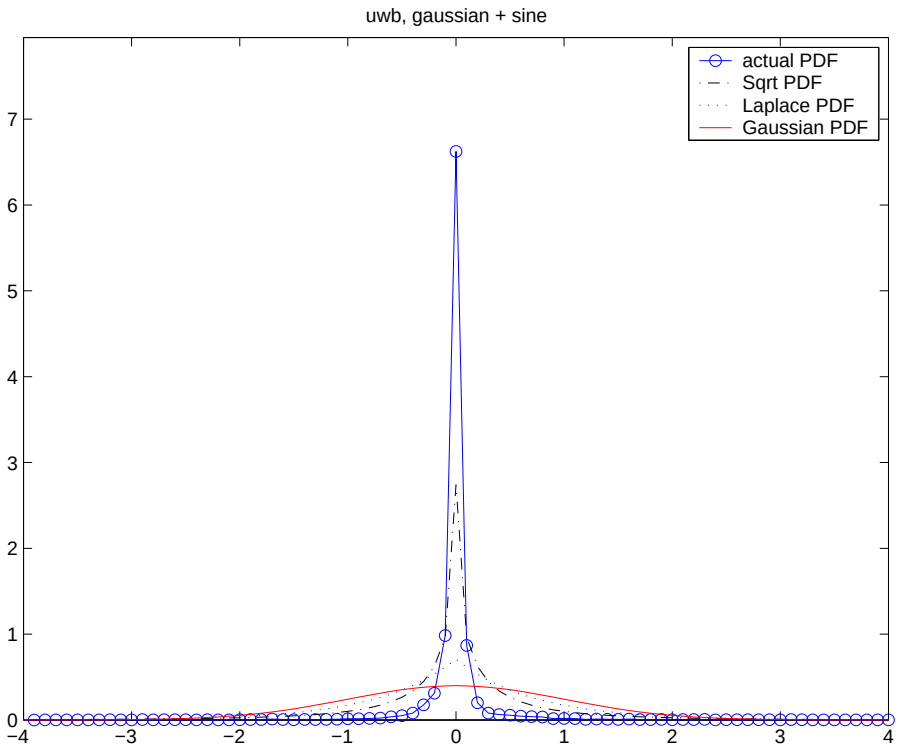

Figure 3.49: Normalized noise distribution $n'(t)$ after FFT. (No matter power ratio is 0.01, 10, or 0.5)

Based on the above observation in Case VII and VIII, we are interested in how the sinusoid interference to Gaussian noise power ratio affects the BER vs. SINR performance over both AWGN and UWB channels. We summarize their relationship via Figure 3.53.

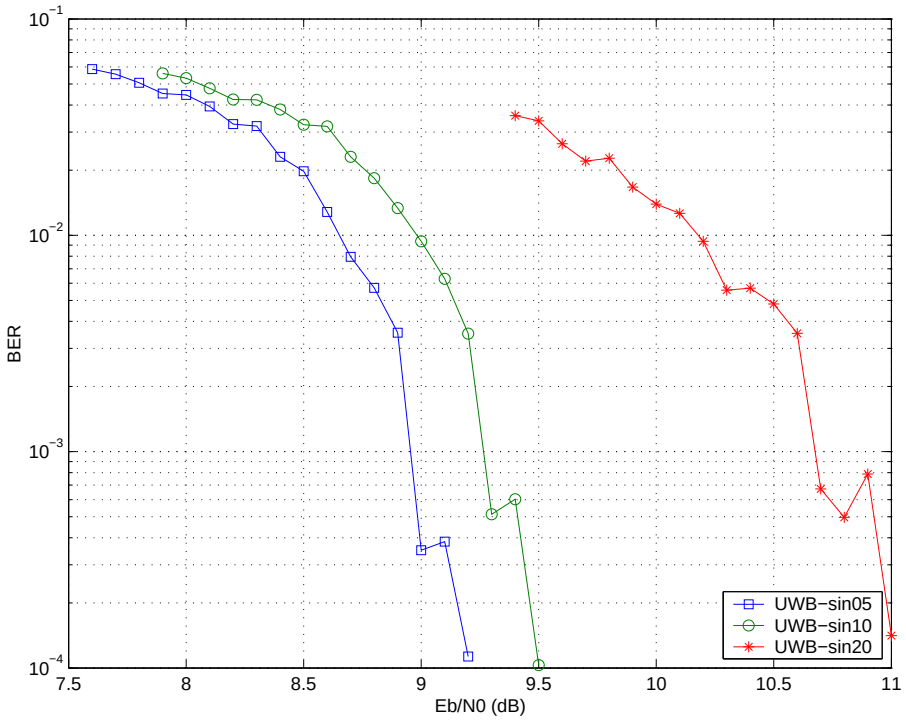

Figure 3.50: BER vs SINR: UWB with NBI. (power ratio is 0.01, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

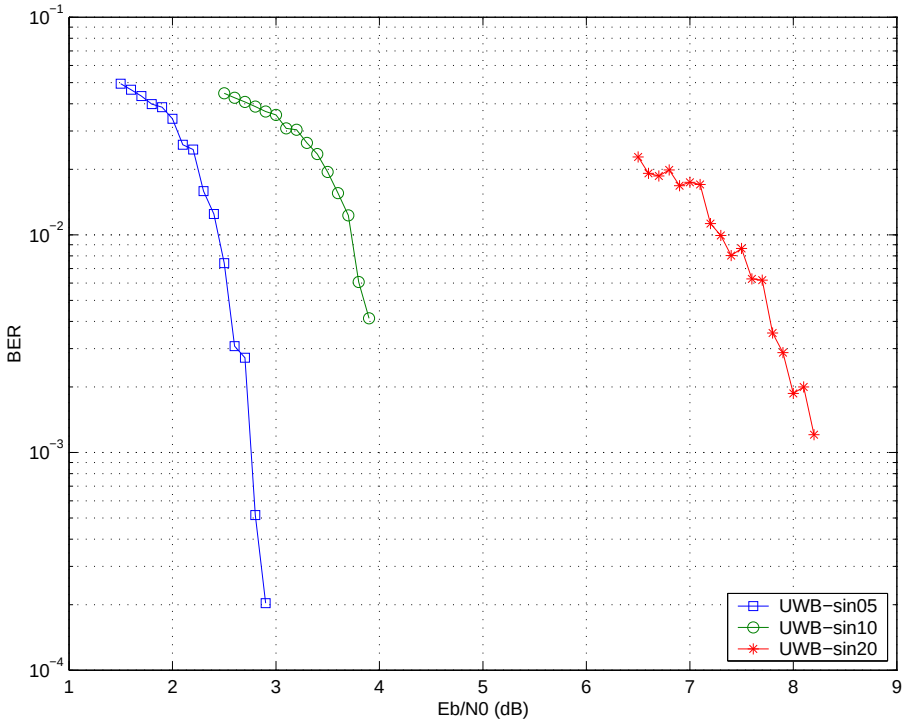

Figure 3.51: BER vs SINR: UWB with NBI. (power ratio is 10, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

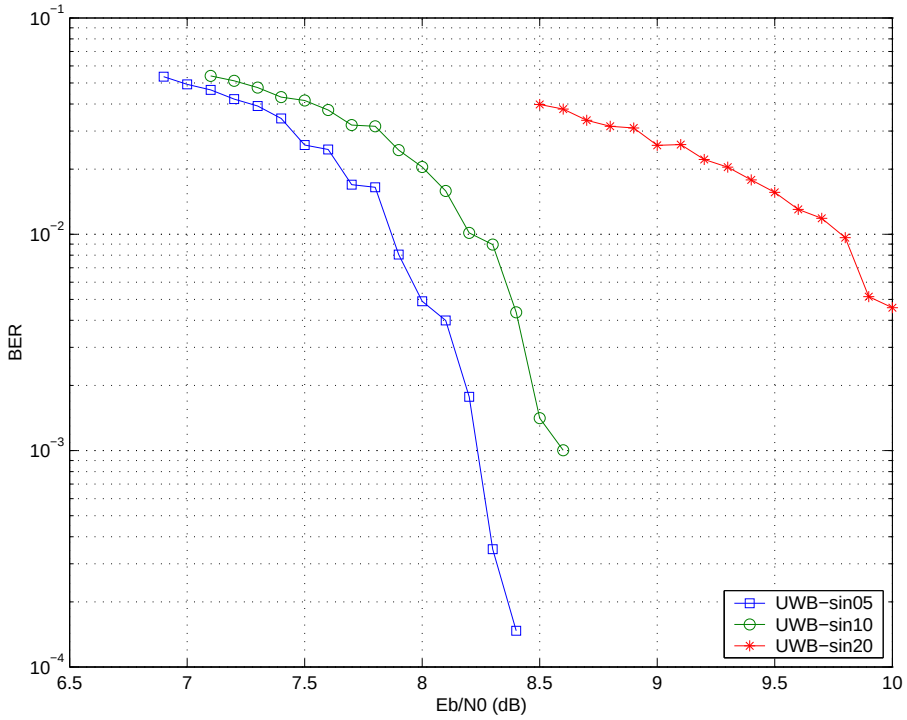

Figure 3.52: BER vs SINR: UWB with NBI. (power ratio is 0.5, $\gamma^D = 0.5, 1.0, 2.0$, respectively)

For AWGN channel, we find out that when the power ratio is less than 0.1, receiver based on $\gamma^D = 2.0$ is better than 1.0, which is further better than 0.5. When the power ratio is between 0.1 and 0.8, receiver based on $\gamma^D = 1.0$ is better than 0.5, which is further better than 2.0. When the power ratio is larger thn 0.8, receiver based on $\gamma^D = 0.5$ is better than 1.0, which is further better than 2.0. This result is consistent with our BER vs. SINR simulation in Case VII for power ratio equal to 0.01, 10 and 0.5.

For UWB channel, we find out that no matter what power ratio is, receiver based on $\gamma^D = 0.5$ is better than 1.0, which is further better than 2.0. This result is also consistent with our BER vs. SINR simulation in Case VIII for power ratio equal to 0.01, 10 and 0.5.

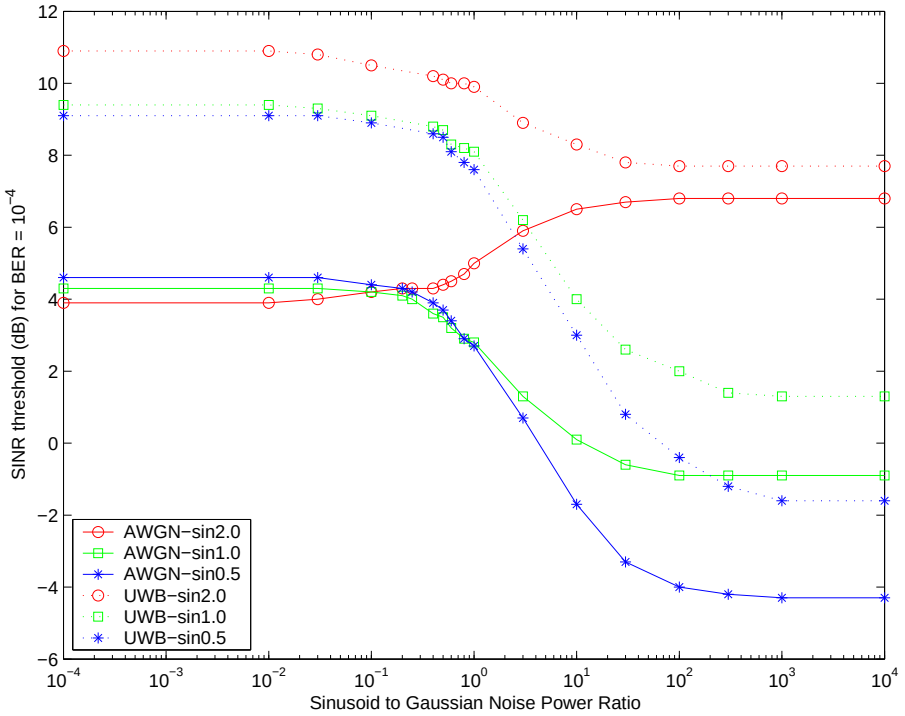

Figure 3.53: SINR threshold at BER $= 10^{-4}$ vs. sinusoid inference to Gaussian noise power ratio over both AWGN and UWB channels. ($\gamma^D = 0.5, 1.0, 2.0$, respectively)

3.6 Conclusions

In this chapter, we demonstrated an MB-OFDM system using powerful parity-concatenated TCM schemes for IEEE 802.15.3a WPAN applications. Frequency hopping, turbo trellis coded modulation, bit as well as symbol level interleaving, OFDM and equalization were described step by step. An modified MAP decoder for TTCM, whose branch metrics were matched to general Gaussian noise structure, was provided to enhance decoding performance. Extensive simulation was done for different cases of noise match and mismatch over both Gaussian and UWB channels. Finally robust solutions for Gaussian and UWB channels were provided by comparing impulsive noises in time domain with those in frequency domain.

CHAPTER 4: OPTIMALITY ON PROBABILISTIC MINIMAX ROBUST DECODER

Since the invention of turbo codes [3] [107] and the rediscovery of LDPC codes [2] [80] [85] [108], there has been renewed interest in these two classes of codes which can approach asymptotically the Shannon limit while still having a low-complexity decoder. Both classes of codes rely on the application of the iterative soft-decision decoder where *a posteriori* probability (APP) decoding plays an important role.

As we know, the Viterbi [65] and BCJR algorithms [77] are optimal if the noise distribution is perfectly known to the decoder. Otherwise, a significant performance degradation may occur due to mismatch. However, the probability density function (PDF) of the noise could change within a short frame in an unpredictable manner. This could be caused by either a natural phenomenon such as lighting or man-made noise such as automotive noise and power-line noise [124]. The accurate estimation of the dynamic noise model could be very difficult or impossible in practice. A robust decoder which is immune to unknown noise model is very useful. In [38] we introduced a probabilistic minimax decoder and showed its robustness by extensive simulation. In this chapter, we will demonstrate its optimality in terms of minimizing the worst case bit error probability for channels with two-type noise model with unknown prior probabilities. Further, we extend some results to more general K-type noise model.

This chapter is organized as follows. In Section 1, we present the signal and system. In Section 2, we show the optimal decoder for known noise. In Section 3 we prove the optimality of our minimax robust decoder for channels with two-type noise model with unknown prior

probabilities. More general K-dimensional case is discussed in Section 4. Section 5 gives a simple example and Section 6 concludes this chapter.

4.1 Signals and Systems

Let $[\mathbf{u}]_1^l = [u_1, u_2, \ldots, u_{kl}]$ and $[\mathbf{v}]_1^l = [v_1, v_2, \ldots, v_{nl}]$ be the input and output sequences of a rate k/n convolutional encoder, respectively, where l is the length of sequence, $[u_{ik-k+1}, \ldots, u_{ik}]$ and $[v_{in-n+1}, \ldots, v_{in}]$ are the k-tuple input and the n-tuple output at time i, respectively, and $u_i, v_i \in \{0, 1\}$. By BPSK modulation, the transmitted sequence is $[\mathbf{s}]_1^l = [s_1, s_2, \ldots, s_{nl}]$, where $s_i = \sqrt{E_b k/n}(2v_i - 1)$ and E_b is the signal power per source bit. The received signal $[\mathbf{r}]_1^l = [r_1, r_2, \ldots, r_{nl}]$ can be expressed as $r_i = s_i + n_i$, where n_i is the channel noise whose PDF may change from slot to slot.

The decoding problem is to recover sequence $[\mathbf{u}]_1^l$ from sequence $[\mathbf{r}]_1^l$ based on certain design criterion. We assume that the noise variables n_i for time slots $i = 1, \ldots, nl$ are mutually independent so that decoding can be done on a slot-by-slot basis. It forms a binary hypothesis detection problem: given $r_i \in \mathbb{R}$, the decoder accepts $u_i = 0$ (hypothesis H_0) or $u_i = 1$ (hypothesis H_1). The noise model for time slot i can be described as the noise PDF $p_{n_i}(n_i; \mathbf{A}_i)$, or equivalently, a pair of conditional PDFs, $p_{r_i|H_0}(r_i|H_0; \mathbf{A}_i)$ if "0" is sent and $p_{r_i|H_1}(r_i|H_1; \mathbf{A}_i)$ if "1" is sent, where $\mathbf{A}_i$ is a time-varying parameter vector defining the noise type. For a Gaussian PDF, $\mathbf{A}_i = (\mu_i, \sigma_i^2)$ where μ_i is the mean and σ_i^2 is the variance.

Let $\mathbf{A}_i^c$ be the actual model parameter vector for the channel and $\mathbf{A}_i^d$ be its estimation at the decoder. Later on we will just omit the subscript i for simplicity.

4.2 Optimal Decoder for Known Noise

The decoder partitions the entire observation space $\mathbb{R}$ into two non-overlap regions Z_0 and Z_1 ($Z_0 \cap Z_1 = \varnothing$ and $Z_0 \cup Z_1 = \mathbb{R}$) and make decisions following this rule: $u = j$ if $r \in Z_j$ ($j = 0, 1$). One of the most reasonable design criteria is to minimize the average bit error probability (BEP). The decoder design and therefore its performance clearly depends on how much information the receiver knows about the noise. It can be classified into four cases of different levels:

i) noise type known for each time slot;

ii) *a priori* probability of each noise type known;

iii) nothing else known except possible types of noise;

iv) nothing known at all about noise.

Naturally, in case iv) the receiver has the worst performance. However, with some knowledge about the noise model we can perform better. In the following two subsections, we will first address cases i) and ii), respectively, which will be useful to case iii) later. Case iii) will be studied in Section 4, which is the main focus of this chapter.

4.2.1 Single-Type Noise Model

For case i), we assume that the noise model is exactly known, which means the parameter vector estimated by the decoder is equal to the actual parameter vector, i.e.

$\mathbf{A}^d = \mathbf{A}^c$. Then, the BEP can be defined as

$$R(\mathbf{A}^c) = \sum_{j=0}^{1} P(H_j) \int_{Z_{1-j}} p_{r|H_j}(r|H_j; \mathbf{A}^c)\, dr \tag{4.1}$$

where $P(H_j)$ is the probability of $u = j$. The decoder can find the optimal decision regions $\{Z_0, Z_1\}$ via the following minimization problem:

$$\min_{\{Z_0,Z_1\}} \; R(\mathbf{A}^c) \tag{4.2}$$

Finally, we can get the following decision rule:

$$\frac{p_{r|H_1}(r|H_1; \mathbf{A}^c)}{p_{r|H_0}(r|H_0; \mathbf{A}^c)} \underset{Z_0}{\overset{Z_1}{\gtrless}} \frac{P(H_0)}{P(H_1)} \tag{4.3}$$

which is the *likelihood ratio test*. If we know $P(H_j)$, from equation (4.3) we get the optimal $\{Z_0^*, Z_1^*\}$ and then the minimum BEP from equation (4.1).

For a standard Viterbi decoder which is optimized towards Gaussian noise model, we just need to change the branch metrics from $|r - \hat{s}|^2$ to $-\log\left(p_n(r - \hat{s}; \mathbf{A}^c)\right)$. For a standard BCJR decoder, we simply replace the Gaussian PDF by the known $p_n(r - \hat{s}; \mathbf{A}^c)$ in branch metric calculation.

4.2.2 Two-Type Noise Model with Known Prior Probabilities

As for case ii), we consider two noise models (denoted by $\mathbf{A}^1$ and $\mathbf{A}^2$) occur with *a priori* probabilities $P(\mathbf{A}^1)$ and $P(\mathbf{A}^2)$, respectively, and the decoder knows these probabilities. The BEP is further averaged over two noises as

$$R(\mathbf{A}^1, \mathbf{A}^2) = \sum_{k=1}^{2} P(\mathbf{A}^k) R(\mathbf{A}^k) \tag{4.4}$$

where $R(\mathbf{A}^k)$ is defined as in (4.1) by substituting $\mathbf{A}^c$ by $\mathbf{A}^k$.

The decoder will find the optimal decision regions $\{Z_0, Z_1\}$ via the following minimization problem:

$$\min_{\{Z_0,Z_1\}} \; R(\mathbf{A}^1, \mathbf{A}^2) \tag{4.5}$$

Similar to the procedure in the previous subsection, we can get the following decision rule:

$$\frac{\sum_{k=1}^{2} P(\mathbf{A}^k) p_{r|H_1}(r|H_1; \mathbf{A}^k)}{\sum_{k=1}^{2} P(\mathbf{A}^k) p_{r|H_0}(r|H_0; \mathbf{A}^k)} \underset{Z_0}{\overset{Z_1}{\gtrless}} \frac{P(H_0)}{P(H_1)} \tag{4.6}$$

which is a generalized version of (4.3). Thus we can get $\{Z_0^*, Z_1^*\}$ and then minimum BEP from equation (4.4).

For a standard Viterbi decoder, similarly we just need to change the branch metrics from $|r - \hat{s}|^2$ to $-\log\left(\sum_{k=1}^{2} P(\mathbf{A}^k) p_n(r - \hat{s}; \mathbf{A}^k)\right)$. For a standard BCJR decoder, we simply replace the Gaussian PDF by the averaged $\sum_{k=1}^{2} P(\mathbf{A}^k) p_n(r - \hat{s}; \mathbf{A}^k)$ in branch metric calculation.

4.3 Minimax Decoder for Noise with Unknown Prior Probabilities

As we know from Section 3, the PDF $p_{r|H_j}(r|H_j; \mathbf{A})$ is used to derive various branch metrics for optimal trellis-based decoding algorithms such as standard Viterbi and BCJR algorithms. If the true noise model is known, so that $\mathbf{A}^c = \mathbf{A}^d$, then the Viterbi and BCJR decoders are optimal. When the true noise model is unknown, we have $\mathbf{A}^c \neq \mathbf{A}^d$, and the standard Viterbi and BCJR decoders are no longer optimal due to mismatch errors. However, it is often difficult to know the exact noise PDFs because of the existence of the so-called man-made noise such as power-line noise, automobile noise, etc. These man-made noise types are often impulsive, which could result in mixed type of noise even in one packet

(about 200 bits). The impulsive noise often lasts a short period of time, which makes noise estimation a very difficult task.

4.3.1 Minimax Criterion

With the knowledge of noise as described in case i) and ii), Section 3 actually follows the *maximum a posteriori probability* (MAP) criterion if $P(H_j)$ is known. It is equivalent to the *maximum likelihood* (ML) criterion if we assume $P(H_0) = P(H_1)$.

If we don't know anything about the noise, nothing we can do better than guessing. However, if we can know all the possible types of noise according to prior knowledge though their probabilities are not clear, which is exactly case iii), we can do better than guessing. Instead of minimizing the average BEP, we try to minimize the maximum BEP for each possible noise type.

Again we consider a model with two noise types (denoted by $\mathbf{A}^1$ and $\mathbf{A}^2$) but with unknown priori probabilities. By minimax criterion, we do optimization as follows:

$$\min_{\{Z_0,Z_1\}} \ \max_{k\in\{1,2\}} R(\mathbf{A}^k) \tag{4.7}$$

where $R(\mathbf{A}^k)$ is defined the same way as before.

4.3.2 Minimax Decoder is Robust

In this subsection we only discuss a two-type noise model where the model prior probabilities are unknown to the decoder. We want to investigate the following problem: if the actual *a priori* probability $P(\mathbf{A}^1) = p$, but somehow we use an "optimal" decoder matched to $P(\mathbf{A}^1) = q$, what will be the average BEP $R(p; q)$ of this mismatched case? And

what is the best we can do if p is unknown. Here the "optimal" decoder is in the MAP sense based on the optimization in Subsection 4.2.2.

From the case discussed in Subsection 4.2.2, we know that if the actual probability p is known, we can design the optimal decoder matched to $q = p$, then the average BEP $R(p;p)$ will be the optimum. We can plot a curve C as in Figure 4.1. It forms a lower bound for the average BEP performance, which is achieved only if the design probability matches the actual probability, i.e. $q = p$. For $q \neq p$, which is usual for the unknown noise model as in case iii), there is performance degradation due to mismatch.

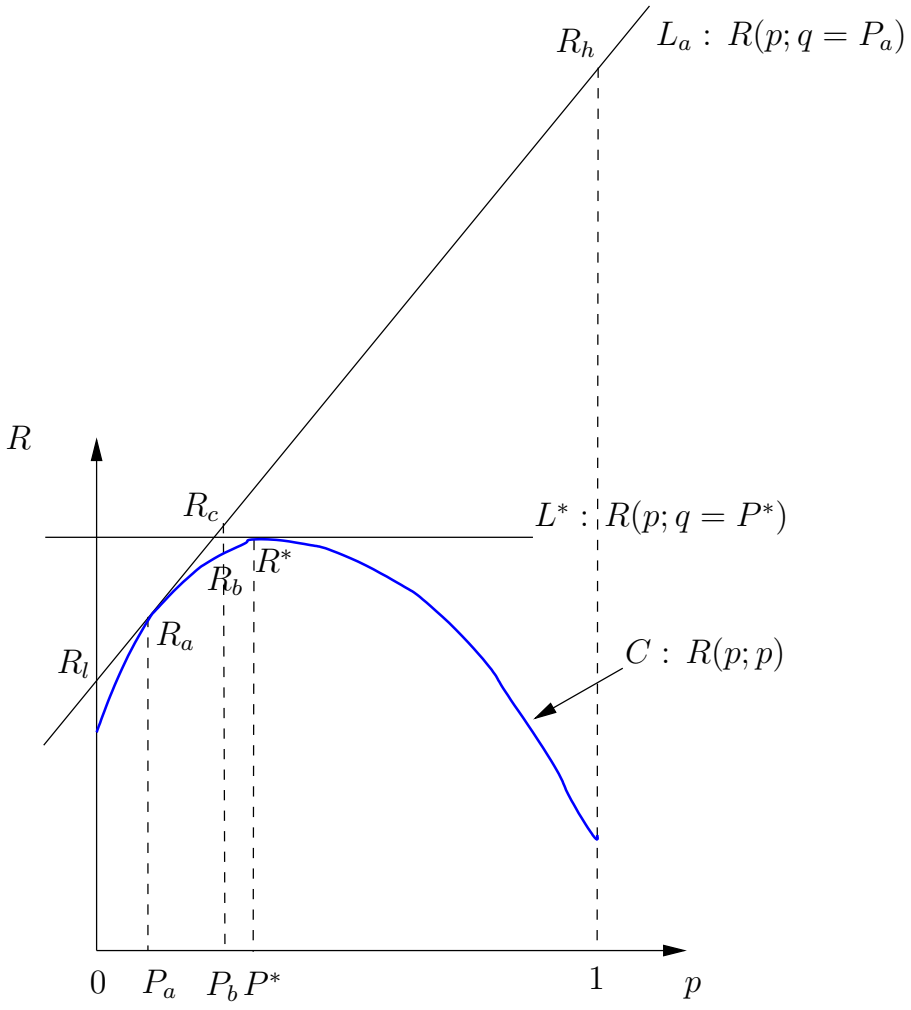

Figure 4.1: Average BEP R versus actual probability p.

We find some interesting properties of the minimax decoder as follows.

Property 1. *If the design probability $q = q_0$ is fixed, the average BEP $R(p; q_0)$ is a* linear *function of actual probability p. And it is a straight line tangent to the curve $R(p; p)$ at point $p = q_0$.*

Proof: We design our decoder based on $q = q_0$. The *mismatched* decision regions $\{Z_0^{\triangle}, Z_1^{\triangle}\}$ are obtained via solving equation (4.6) by substituting $P(\mathbf{A}^1) = q_0$ and $P(\mathbf{A}^2) = 1 - q_0$.

For noise type $\mathbf{A}^1$, we have *probability of miss* $P_M(\mathbf{A}^1) = \int_{Z_0^{\triangle}} p_{r|H_j}(r|H_1; \mathbf{A}^1)\, dr$ and *probability of false-alarm* $P_F(\mathbf{A}^1) = \int_{Z_1^{\triangle}} p_{r|H_j}(r|H_0; \mathbf{A}^1)\, dr$, both of which are fixed when $\{Z_0^{\triangle}, Z_1^{\triangle}\}$ are determined.

Thus, the BEP for noise type $\mathbf{A}^1$, i.e. $R(\mathbf{A}^1) = P(H_1)P_M(\mathbf{A}^1) + P(H_0)P_F(\mathbf{A}^1)$, is also fixed.

Similarly, the BEP for noise type $\mathbf{A}^2$, i.e. $R(\mathbf{A}^2)$, is fixed, too.

Then, $R(p; q_0) = P(\mathbf{A}^1)R(\mathbf{A}^1) + P(\mathbf{A}^2)R(\mathbf{A}^2) = pR(\mathbf{A}^1) + (1-p)R(\mathbf{A}^2) = R(\mathbf{A}^2) + p[R(\mathbf{A}^1) - R(\mathbf{A}^2)]$ is a linear function of p only.

Easily we know that $p = q_0$ is a solution to equation $R(p; q_0) = R(p; p)$, which means the straight line $R(p; q_0)$ intersects the curve $R(p; p)$ at point $p = q_0$. Actually, the curve is always below the straight line where the gap between them is the mismatch BEP loss. So the straight line $R(p; q_0)$ is tangent to the curve $R(p; p)$ at point $p = q_0$. $\square$

For example, we design a decoder based on $q = P_a$, then the BEP for actual probability p can be described as a straight line $L_a : R(p; q = P_a)$, where $R_a = R(p = P_a; q = P_a)$

means matched case. If $p = P_b$ instead, the BEP becomes $R_c = R(p = P_b; q = P_a)$ which will be greater than $R_b = R(p = P_b; q = P_b)$ as if we design the decoder based on $q = P_b$. The BEP loss due to mismatch is $R_c - R_b$ as you can see in Figure 4.1.

Property 2. *The curve $R(p;p)$ is concave and has a maximum R^* at $p = P^*$.*

Proof: This follows from the fact that for any p, $R(p;q) \geq R(p;p)$. That means the curve C is always below the straight line tangent to it.

Since $P(\mathbf{A}^1) = p$, by solving

$$\frac{p\ p_{r|H_1}(r|H_1;\mathbf{A}^1) + (1-p)p_{r|H_1}(r|H_1;\mathbf{A}^2)}{p\ p_{r|H_0}(r|H_0;\mathbf{A}^1) + (1-p)p_{r|H_0}(r|H_0;\mathbf{A}^2)} \underset{Z_0}{\overset{Z_1}{\gtrless}} \frac{P(H_0)}{P(H_1)},$$

we get the optimal decision regions $\{Z_0^*, Z_1^*\}$ which minimize the average BEP according to Subsection 4.2.2. And its minimum is exactly $R(p;p)$.

If we instead get a set of different decision regions $\{Z_0^\triangle, Z_1^\triangle\}$ by solving

$$\frac{q\ p_{r|H_1}(r|H_1;\mathbf{A}^1) + (1-q)p_{r|H_1}(r|H_1;\mathbf{A}^2)}{q\ p_{r|H_0}(r|H_0;\mathbf{A}^1) + (1-q)p_{r|H_0}(r|H_0;\mathbf{A}^2)} \underset{Z_0}{\overset{Z_1}{\gtrless}} \frac{P(H_0)}{P(H_1)},$$

we know that the corresponding BEP is $R(p;q)$, which is not necessarily the minimum. Thus, $R(p;q) \geq R(p;p)$.

Since $R(p;p)$ is concave, it has a maximum either at a point $p = P^*$ with zero slope or at one of the endpoints $p = 0$ or 1. $\square$

Property 3. *The minimax criterion defined in equation (4.7) is equivalent to the following criterion:*

$$max_p R(p;p) \tag{4.8}$$

Proof: It is not difficult to understand that the minimax criterion (4.7) actually means

$$\min_q \max_p R(p;q) \tag{4.9}$$

which minimizes the maximum BEP. Once the design probability q is given, the BEP performance is decided by the actual probability p only and has its maximum achieved at certain value of p. We try to choose a q such that the maximum BEP is as small as possible. It takes maximization of $R(p;q)$ over p by fixing q first, then takes minimization over q.

We can also do minimization over q by fixing p first, and get a function of p only, which is exactly $R(p;p) = min_q R(p;q)$. Then we can do maximization over p. This is exactly what equation (4.8) does.

So equation (4.8) is equivalent to equation (4.9), which means equation (4.8) is an alternative representation to the minimax criterion. □

Property 4. *Suppose $P^* = arg\ max_p R(p;p)$ and the slope of the line $L^* : R(p; q = P^*)$ is zero, then the decoder with design probability P^* is robust to unknown noise prior probabilities.*

Proof: From Property 1, we know that $R(p;q) = R(\mathbf{A}^2) + p[R(\mathbf{A}^1) - R(\mathbf{A}^2)]$. When $q = P^*$, we get a zero slope, i.e. $R(\mathbf{A}^1) = R(\mathbf{A}^2) \triangleq R^*$.

Given any actual probability $P(\mathbf{A}^1) = p$, we can calculate average BEP from equation (4.4): $pR(\mathbf{A}^1) + (1-p)R(\mathbf{A}^2) = R^*$, which is a constant independent of unknown *a priori* probability p over the entire range $p \in [0,1]$. In this sense, the minimax decoder is "robust". It is the optimal solution to our unknown noise model for case iii). □

4.4 K-Dimensional Cases

For K-dimensional cases, there are totally K noise models with unknown priors. Let $p_k = P(\mathcal{A}_k)$ for $k = 1, 2, \ldots, K$ where $\sum_{k=1}^{K} p_k = 1$.

The bit error probability (BEP) is defined as follows:

$$R(A_1, A_2, \ldots, A_K) = \sum_{k=1}^{K} p_k R(A_k)$$

where the conditional risk when noise model k appears can be described as

$$R(A_k) = P(H_0) \int_{Z_1} p_{r|H_0}(r|H_0; A_k)\, dr + P(H_1) \int_{Z_0} p_{r|H_1}(r|H_1; A_k)\, dr$$

Since $p_K = 1 - \sum_{k=1}^{K-1} p_k$, BEP can be expressed as a function of $K-1$ variables, $p_1, p_2, \ldots, p_{K-1}$. Let $\mathbf{p} = [p_1, p_2, \ldots, p_{K-1}]^T$, which means a vector in a $(K-1)$-dimensional space.

If the prior distribution for $\mathbf{p}$ is known to be $\mathbf{p}_0$, then according to Section 4.2.2, we can get the optimum BEP R^* by minimizing $R(A_1, A_2, \ldots, A_K)$, i.e. by solving the inequality

$$\frac{\sum_{k=1}^{K} p_{0k} p_{r|H_1}(r|H_1; A_k}{\sum_{k=1}^{K} p_{0k} p_{r|H_0}(r|H_0; A_k)} \underset{Z_0}{\overset{Z_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}$$

For each particular distribution $\mathbf{p}$, we can get the above mentioned optimum. Thus we can name such an optimum as a function of distribution, $R^*(\mathbf{p})$.

Property 5. *$R^*(\mathbf{p})$ is a lower bound of BEP for any rule.*

Proof. Given any valid $\mathbf{p} \in [0, 1]^K$, R^* is the minimum BEP we can get by minimizing $R(A_1, A_2, \ldots, A_K)$. □

However, this lower bound is not achievable when $\mathbf{p}$ is unknown. A implementable rule is to design a rule independent of unknown $\mathbf{p}$.

Property 6. *For every rule which is independent of* $\mathbf{p}$*, BEP is a linear function of* $\mathbf{p}$*.*

Proof. Once a rule is chosen, $\{Z_0, Z_1\}$ is then determined also independent of $\mathbf{p}$, such that the conditional risk $R(A_k)$ is determined independent of $\mathbf{p}$. By substituting $p_K = 1 - \sum_{k=1}^{K-1} p_k$, $R(A_1, A_2, \ldots, A_K) = \mathbf{a}^T\mathbf{p} + b$, where $\mathbf{a} = [R(A_1) - R(A_K), R(A_2) - R(A_K), \ldots, R(A_{K-1}) - R(A_K)]^T$ and $b = R(A_K)$. □

Minimax is a very reasonable rule for unknown prior distributions $\mathbf{p}$, which minimizes the maximum conditional risks as follows

$$\min \max\{R(A_1), R(A_2), \ldots, R(A_K)\}$$

Corollary 1. *A minimax rule is also a linear function of* $\mathbf{p}$*.*

Property 7. *A rule, which is independent of* $\mathbf{p}$*, cannot be a solution for minimax rule if it is not tangent to the curve* $R^*(\mathbf{p})$*.*

Proof. Since the rule is independent of $\mathbf{p}$, it can be described as the linear function $R = \mathbf{a}^T\mathbf{p} + b$. If it is not tangent to $R^*(\mathbf{p})$, there must exist another parallel line tangent to $R^*(\mathbf{p})$ which has the form $R' = \mathbf{a}^T\mathbf{p} + b'$, where $b' < b$. We know that for every $\mathbf{p}$, $R' < R$. Then $\max R' < \max R$. It means R cannot be a solution to $\min \max R$. □

Actually a rule tangent to $R^*(\mathbf{p})$ at $\mathbf{p}_0$ means that though we don't know the actual distribution, we design the decoder based on the distribution $\mathbf{p}_0$. The corresponding BEP is a function of the actual distribution $\mathbf{p}$ as $R(\mathbf{p}; \mathbf{p}_0)$. Then we can see that the previously

defined $R^*(\mathbf{p}) = R(\mathbf{p};\mathbf{p})$. According to the above property, the solution for a minimax rule must be within the family of straight lines tangent to the lower bound $R^*(\mathbf{p})$, i.e. $R(\mathbf{p};\mathbf{q})$. We can do minimax operation over these candidates only instead of over the entire rule space.

Property 8. *Minimax, i.e.* $min_{\mathbf{q}}max_{\mathbf{p}}R(\mathbf{p};\mathbf{q})$ *is equivalent to Maximin, i.e.* $max_{\mathbf{p}}min_{\mathbf{q}}R(\mathbf{p};\mathbf{q}) = max_{\mathbf{p}}R(\mathbf{p};\mathbf{p})$.

Proof. Similar to proof of Property 3. □

It means if we want to find a minimax solution, we can actually a maximum of $R(\mathbf{p};\mathbf{p})$ or equivalently $R^*(\mathbf{p})$. The only question left is the existence of a maximum of $R(\mathbf{p};\mathbf{p})$, which will be guaranteed by the following property.

Property 9. $R^*(\mathbf{p})$ *is concave, which guarantees the existence of a maximum solution of* $R(\mathbf{p};\mathbf{p})$*, and then the existence of a minimax solution of* $min_{\mathbf{q}}max_{\mathbf{p}}R(\mathbf{p};\mathbf{q})$.

Proof. For any $\mathbf{p}_a, \mathbf{p}_b \in \underbrace{[0,1] \times \cdots \times [0,1]}_{K-1 \text{ times}}$ and any $\lambda \in [0,1]$, let

$$\mathbf{p}_c = \lambda \mathbf{p}_a + (1-\lambda)\mathbf{p}_b$$

that is, $p_{ci} = \lambda p_{ai} + (1-\lambda)p_{bi}$ for $i = 1, 2, \ldots, K-1$.

Then

$$\frac{\mathbf{p}_c - \mathbf{p}_a}{\mathbf{p}_b - \mathbf{p}_c} = \frac{1-\lambda}{\lambda}$$

which means

$$\frac{\mathbf{p}_b - \mathbf{p}_a}{\mathbf{p}_c - \mathbf{p}_a} = \frac{1}{1-\lambda}$$

Since we have

$$\frac{R(\mathbf{p}_b;\mathbf{p}_c)-R(\mathbf{p}_a;\mathbf{p}_c)}{\mathbf{p}_b-\mathbf{p}_a}=\frac{R^*(\mathbf{p}_c)-R(\mathbf{p}_a;\mathbf{p}_c)}{\mathbf{p}_c-\mathbf{p}_a}$$

we get

$$\frac{R(\mathbf{p}_b;\mathbf{p}_c)-R(\mathbf{p}_a;\mathbf{p}_c)}{R^*(\mathbf{p}_c)-R(\mathbf{p}_a;\mathbf{p}_c)}=\frac{\mathbf{p}_b-\mathbf{p}_a}{\mathbf{p}_c-\mathbf{p}_a}=\frac{1}{1-\lambda}$$

that is

$$R^*(\mathbf{p}_c)=\lambda R(\mathbf{p}_a;\mathbf{p}_c)+(1-\lambda)R(\mathbf{p}_b;\mathbf{p}_c)$$

However, $R(\mathbf{p}_a;\mathbf{p}_c)\geq R(\mathbf{p}_a;\mathbf{p}_a)=R^*(\mathbf{p}_a)$ and $R(\mathbf{p}_b;\mathbf{p}_c)\geq R(\mathbf{p}_b;\mathbf{p}_b)=R^*(\mathbf{p}_b)$. Thus, we finally get

$$R^*(\mathbf{p}_c)\geq\lambda R^*(\mathbf{p}_a)+(1-\lambda)R^*(\mathbf{p}_b)$$

which means $R^*(\mathbf{p})$ is concave. □

4.5 Numerical Results

We use some numerical examples to confirm our results. In [38] (Fig. 2), a two-type noise model is introduced. Their PDFs can be described as follows:

$$p(r|H_1;\mathbf{A}^1)=\begin{cases}\frac{1}{4}(r+1), & \text{if } -1\leqslant r<\frac{1}{3};\\ \frac{1}{3}, & \text{if } \frac{1}{3}\leqslant r<\frac{2}{3};\\ -\frac{1}{4}(r-2), & \text{if } \frac{2}{3}\leqslant r<1;\\ -\frac{1}{4}(r-3), & \text{if } 1\leqslant r<\frac{7}{3};\\ \frac{1}{6}, & \text{if } \frac{7}{3}\leqslant r<\frac{8}{3};\\ \frac{1}{4}(r-2), & \text{if } \frac{8}{3}\leqslant r<3;\\ 0, & \text{otherwise.}\end{cases}$$

$$p(r|H_1;\mathbf{A}^2)=\begin{cases}\frac{1}{6}, & \text{if } -1\leqslant r<1;\\ \frac{1}{3}, & \text{if } 1\leqslant r<3;\\ 0, & \text{otherwise.}\end{cases}$$

and

$$p(r|H_0;\mathbf{A}^1)=p(r+2|H_1;\mathbf{A}^1)$$

$$p(r|H_0;\mathbf{A}^2)=p(r+2|H_1;\mathbf{A}^2)$$

If only $\mathbf{A}^1$ type noise exists, the optimal decoder should follow the results in Subsection 4.2.1. The optimal decision regions are $Z_0=(-\infty,0)$ and $Z_1=[0,+\infty)$ and the minimum BEP is $\frac{23}{144}$. Similarly, if only $\mathbf{A}^2$ type noise exists, the optimal decision regions are $Z_0=(-\infty,1)$ and $Z_1=[1,+\infty)$ and the minimum BEP is $\frac{1}{6}$.

When both noises exist and the actual probability p is known, we can always design a decoder matched to p and get the best BEP performance. The optimal decision region will rely on both received r and the model probability p. It is easy to verify that $r \in (-\infty, 0] \subset Z_0$ and $r \in [1, +\infty) \subset Z_1$ for any p. For $r \in (0, 1)$, we get the optimal decision regions for each pair of (r, p) as shown in Figure 4.2.

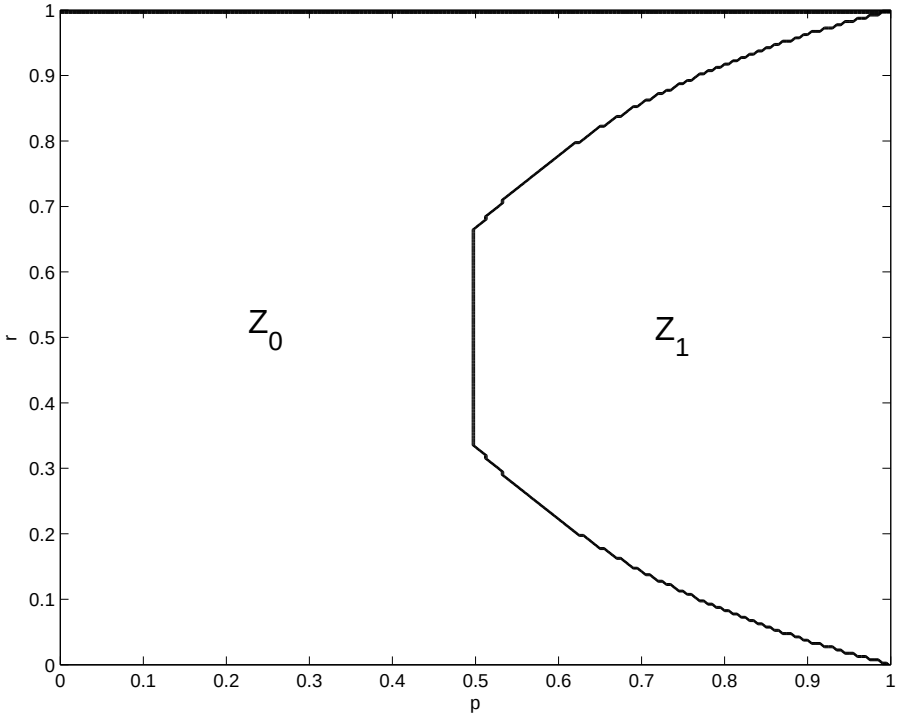

Figure 4.2: Optimal decision regions for $r \in (0, 1)$.

When both noises exist but the *a priori* probability $P(\mathbf{A}^1)$ is unknown to the decoder, there will be mismatch if we use the optimal decoder matched to either type of noise. For an actual probability p, the straight line $L_1 : R = -\frac{13}{144}p + \frac{1}{4}$ in Figure 4.3 denotes the BEP when the decoder matches $\mathbf{A}^1$ only, and the straight line $L_2 : R = \frac{7}{144}p + \frac{1}{6}$ denotes the BEP when the decoder matches $\mathbf{A}^2$ only. We can use the minimax criterion to design our robust decoder and the optimal BEP is $\frac{55}{288}$, which is corresponding to the horizontal line $L_3 : R = \frac{55}{288}$. All the results we obtain here are consistent with those in [38].

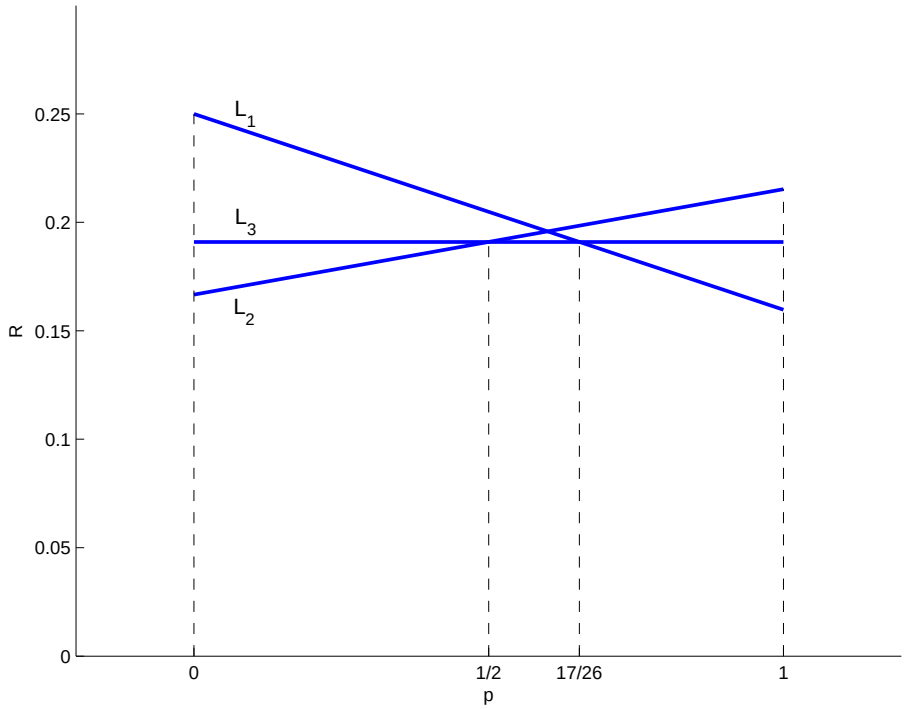

Figure 4.3: Average BEP R versus actual probability p. (L_1: decoder matched to $\mathbf{A}^1$ only; L_2: decoder matched to $\mathbf{A}^2$ only; L_3: minimax robust decoder.)

4.6 Conclusions

In this chapter, we discussed the decoder design problem under different noise models. After reviewing the optimal MAP decoder under known noise model, we formulated a minimax decoder for channels with two-type noise model with unknown prior probabilities. We generated an alternative optimization problem which actually is equivalent to minimax criterion. Further, we proved that minimax decoder is robust to unknown noise model and it is an optimal solution to our problem. Numerical examples further confirmed our results.

CHAPTER 5: CAPACITY OF LDPC CODES UNDER SUM-PRODUCT DECODING AND NOISE MISMATCH

A decade after the break-through of turbo codes [3] [107], many coding schemes can now closely approach the Shannon limit. These coding schemes include turbo codes [110] [111] [112], LDPC codes [2] [6], "repeat-accumulate" (RA) codes [113], and several concatenated coding schemes using simple parity-check codes [114] [115] [116] [117] [118]. Density evolution and the Tanner graph have played an essential role in code optimization and decoding performance analysis [85] [108] [80].

Both the Viterbi [65] and BCJR algorithms [77] algorithms are special cases of min-sum and sum-product algorithms [81]. In [119] Forney provided a detailed description of the two-way algorithms as well as a review of their rich history. Both algorithms are optimal if the decoders know the exact probability density function (**PDF**) of the channel noises; otherwise, a significant performance degradation may occur. In other words, we need to know the exact noise PDF of the channel to design the optimal decoders. In practice, if the noise PDF changes slowly or are constant, then it can be estimated [120] [121] [122]. However, in the real world, the noise could change within a short time frame in an uncertain manner. This could be caused by either natural phenomenon such as lighting or man-made noise such as automotive noise and power-line noise [123] [124]. The characteristic of one type of man-made noise is impulsive with a typical rate of 10-500 impulses/second [123]. For a mobile phone with a data rate of 10 Kbits/s, it could experience up to one impulse every 200 bits (roughly every speech packet will be affected). For a HF radio with a data rate of 1 Kbits/s, the situation will be worse. Some types of man-made noise can still be

approximated by Gaussian noise, while others might only be modeled as other types of noise (for example, impulsive Laplace noise). Furthermore, even in Gaussian channels the sum-product decoder needs to estimate the noise variance. The accurate estimation of these noise parameters could be very difficult if not impossible in practice. Alternatively, we can design decoders robust to mismatched noises and noise parameter estimation errors [125] [38].

LDPC codes and turbo codes have been included in many practical application recently. Impulsive noise has become a major focus in application of turbo and LDPC codes in practice [126] [127] [128] [129]. However, it is unclear whether the main analytical results of density evolution in [5] and [130] can be extended to mismatched cases. In other words, can we analyze the threshold for LDPC codes with sum-product decoder designed to match the AWGN noise under the effect of impulsive Laplace noises? Does the symmetric condition hold? Is the Shannon capacity affected by the mismatched noise used in the decoder? How is the stable condition affected by the mismatched noise? This chapter is dedicated to answer these questions. In this chapter, we will focus on binary input AWGN (BIAWGN) channels and binary input Laplace (BIL) channels, as example. The general results can be applied to many other mismatched noise cases.

5.1 Capacity for BIAWGN and BIL Channels with Mismatched Noise

Figure 5.1 shows the block diagram of the system. Let $x_D \in \{\pm 1\}$ denote the binary input information source. Let x denote the binary coded bit with probability $q = Prob(x = 1)$ (here $Prob(.)$ means probability, not PDF), which is passed through an additive channel and the output is $y = x + n$, where n is the independent noise component. $L(\cdot)$ denotes the

computation of log-likelihood ratio (LLR) given y, channel type and channel parameters. The initial log-likelihood ratio value is passed to the sum-product decoder. In order to compute the log-likelihood ratio value, we need to know the channel type and estimate the channel parameters such as the variance. In practice, we may use wrong type of channels or an inaccurate channel parameters in the decoder. For clarification, we use subscript "c" denotes the true channel type and channel parameters, and subscript "r" to denote the estimated channel type and channel parameters used in the decoders.

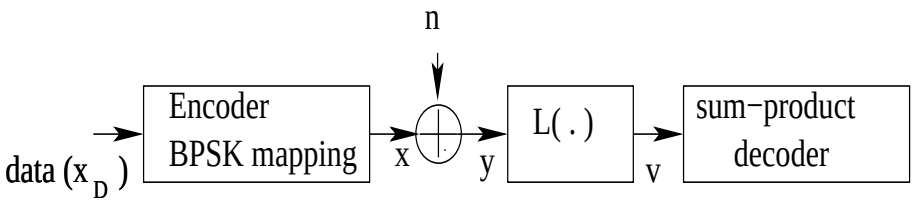

Figure 5.1: Block Diagram of the System

Property 10. *Mismatch between true channel (type and parameters) and the estimated channel (type and parameters) used in the decoder will only affect the performance of sum-product decoder through the log-likelihood ratio computation for any binary-input memoryless output symmetric channels.*

Proof. At first, we prove that the channel symmetric condition holds as long as both the true channel and the channel used in the receiver are binary-input memoryless output symmetric. Presumably, we have $p_c(y|x_c = 1) = p_c(-y|x_c = -1)$ and $p_r(y|x_r = 1) = p_r(-y|x_r = -1)$ where x_c means true bit sent at the transmitter while x_r means estimated bit at the receiver, then we have $p(v|x_c = 1) = p(y|x_c = 1)|dy|/|dv|$ and $p(v|x_c = -1) = p(y|x_c = -1)|dy|/|dv|$, where $v = log\frac{p_r(y|x_r=1)}{p_r(y|x_r=-1)} \triangleq f(y)$. Since $-v = -log\frac{p_r(y|x_r=1)}{p_r(y|x_r=-1)} = log\frac{p_r(-(-y)|x_r=-1)}{p_r(-(-y)|x_r=1)} = log\frac{p_r(-y|x_r=1)}{p_r(-y|x_r=-1)} = f(-y)$, we obtain $p(-v|x_c = -1) = p(-y|x_c = -1)|dy|/|dv| = p(y|x_c =$

$1)|dy|/|dv| = p(v|x_c = 1)$. Thus, the output from mismatched LLR computation is also symmetric. One shall note that for the mismatched case, we do have $p(v) \neq e^v p(-v)$.

Secondly, it is easily verified that the symmetric conditions on check node and variable nodes are maintained regardless of mismatch or not. Thus, we can analyze the behavior of sum-product decoder from its behavior assuming transmission of the all-one codeword [85].

Finally, the computation of discrete density evolution of sum-product decoder is related to outputs of channel only through the initial message density of the likelihood ratios. Thus, regardless matched or mismatched channel type or parameters, we can use density evolution to evaluate its performance as long as we find appropriate initial message density. □

Now, let us take two channels as examples, one is binary input additive white Gaussian noise (BIAWGN) channels and another is binary input Laplace (BIL) channels. The results for matched cases can be found at [5], [85] and [130]. Now, we will focus on mismatched cases.

For BIAWGN channels, n is Gaussian distribution of zero mean and variance σ_c^2 and

$$p_{c,BIAWGN}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{(y-x)^2}{2\sigma_c^2}} \tag{5.1}$$

For BIL channels, n is Laplace distribution of zero mean and variance $2\lambda_c^2$ and

$$p_{c,BIL}(y|x) = \frac{1}{2\lambda_c} e^{-\frac{|y-x|}{\lambda_c}} \tag{5.2}$$

where $p(\cdot)$ denotes the probability density function. One can replace subscript "c" by subscript "r" to obtain the estimated channel noise PDF used in the decoders as

$$p_{r,BIAWGN}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_r} e^{-\frac{(y-x)^2}{2\sigma_r^2}} \tag{5.3}$$

$$p_{r,BIL}(y|x) = \frac{1}{2\lambda_r} e^{-\frac{|y-x|}{\lambda_r}} \tag{5.4}$$

The signal to noise ratio (SNR) is $\frac{1}{\sigma_c^2}$ and $\frac{1}{2\lambda_c^2}$ for BIAWGN and BIL channels respectively. Bit energy to noise density ratio $\frac{E_b}{N_o}$ is then related to SNR as $\frac{E_b}{N_o} = \frac{SNR}{2R}$, where R denotes the code rate.

$L(\cdot)$ computation for the BIAWGN channel is

$$v = log \frac{p_{r,BIAWGN}(y|x=1)}{p_{r,BIAWGN}(y|x=-1)} = \frac{2y}{\sigma_r^2} \tag{5.5}$$

It is a linear operation.

$L(\cdot)$ computation for the BIL channel is

$$v = log \frac{p_{r,BIL}(y|x=1)}{p_{r,BIL}(y|x=-1)} = \begin{cases} -\frac{2}{\lambda_r}, & y \leq -1 \\ \frac{2y}{\lambda_r}, & -1 < y \leq 1 \\ \frac{2}{\lambda_r}, & y > 1 \end{cases} \tag{5.6}$$

It is a non-linear operation.

The initial message density given all one-codeword, i.e., the PDF of v given $x = 1$, or $p(v|x=1) = p_c(y|x=1)|dy|/|dv|$ as

$$\frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{(y-1)^2}{2\sigma_c^2}} \cdot \frac{\sigma_r^2}{2}\bigg|_{y=v\sigma_r^2/2} = \frac{\sigma_r^2}{2\sqrt{2\pi}\sigma_c} e^{-\frac{(v\sigma_r^2-2)^2}{8\sigma_c^2}} \tag{5.7}$$

and

$$\frac{1}{2\lambda_c} e^{-\frac{|y-1|}{\lambda_c}} \cdot \frac{\sigma_r^2}{2}\bigg|_{y=v\sigma_r^2/2} = \frac{\sigma_r^2}{4\lambda_c} e^{-\frac{|\frac{v\sigma_r^2}{2}-1|}{\lambda_c}} \tag{5.8}$$

for BIAWGN and BIL channels, respectively, where the decoder uses the BIAWGN noise type, i.e., $L(\cdot)$ is computed by (5.5).

When the decoder uses the BIL to calculate v via (5.6), due to its nonlinearity, we can calculate $p(v|x=1)$ separately as follows.

$$Prob(v=-\frac{2}{\lambda_r}|x=1)=Prob(y\leq -1|x=1)=\int_{-\infty}^{-1}p_c(y|x=1)dy\triangleq F_1 \tag{5.9}$$

$$Prob(v=\frac{2}{\lambda_r}|x=1)=Prob(y>1|x=1)=\int_{1}^{\infty}p_c(y|x=1)dy\triangleq F_2 \tag{5.10}$$

When $v=\frac{2y}{\lambda_r}\in(-\frac{2}{\lambda_r},\frac{2}{\lambda_r})$, its PDF is

$$p(v|x=1)=p_c(y|x=1)\cdot\frac{\lambda_r}{2}\bigg|_{y=\frac{v\lambda_r}{2}}\triangleq F(v) \tag{5.11}$$

Also, we know that $p(v|x=1)=0$ when $v<-\frac{2}{\lambda_r}$ or $v>\frac{2}{\lambda_r}$.

Overall, the PDF of v is

$$p(v|x=1)=F_1\delta(v+\frac{2}{\lambda_r})+F_2\delta(v-\frac{2}{\lambda_r})+F(v)\Pi_v(\frac{2}{\lambda_r}) \tag{5.12}$$

where the rectangular pulse function is $\Pi_v(x)=1$ if $|v|<x$, 0 elsewhere.

Therefore, we have $p(v|x=1)$ as

$$Q(\frac{2}{\sigma_c})\delta(v+\frac{2}{\lambda_r})+\frac{1}{2}\delta(v-\frac{2}{\lambda_r})+\frac{\lambda_r}{2\sqrt{2\pi}\sigma_c}e^{-\frac{(\frac{v\lambda_r}{2\sigma_c}-\frac{1}{\sigma_c})^2}{2}}\Pi_v(\frac{2}{\lambda_r}) \tag{5.13}$$

and

$$\frac{1}{2}e^{-\frac{2}{\lambda_c}}\delta(v+\frac{2}{\lambda_r})+\frac{1}{2}\delta(v-\frac{2}{\lambda_r})+\frac{\lambda_r}{4\lambda_c}e^{-|\frac{v\lambda_r}{2\lambda_c}-\frac{1}{\lambda_c}|}\Pi_v(\frac{2}{\lambda_r}) \tag{5.14}$$

for BIAWGN and BIL channels, respectively, where the decoder uses the BIL noise type, i.e., $L(\cdot)$ is computed by (5.6) and $Q(x)=\int_x^\infty\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$.

By replacing v by $-v$ in (5.7), (5.8), (5.13) and (5.14), we obtain corresponding $p(v|x=-1)$ as

$$\frac{\sigma_r^2}{2\sqrt{2\pi}\sigma_c}e^{-\frac{(v\sigma_r^2+2)^2}{8\sigma_c^2}} \tag{5.15}$$

$$\frac{\sigma_r^2}{4\lambda_c} e^{-\frac{|\frac{v\sigma_r^2}{2}+1|}{\lambda_c}} \tag{5.16}$$

$$Q(\frac{2}{\sigma_c})\delta(v-\frac{2}{\lambda_r}) + \frac{1}{2}\delta(v+\frac{2}{\lambda_r}) + \frac{\lambda_r}{2\sqrt{2\pi}\sigma_c} e^{-\frac{(\frac{v\lambda_r}{2\sigma_c}+\frac{1}{\sigma_c})^2}{2}} \Pi_v(\frac{2}{\lambda_r}) \tag{5.17}$$

and

$$\frac{1}{2}e^{-\frac{2}{\lambda_c}}\delta(v-\frac{2}{\lambda_r}) + \frac{1}{2}\delta(v+\frac{2}{\lambda_r}) + \frac{\lambda_r}{4\lambda_c} e^{-|\frac{v\lambda_r}{2\lambda_c}+\frac{1}{\lambda_c}|} \Pi_v(\frac{2}{\lambda_r}) \tag{5.18}$$

Property 11. *(a) The Shannon capacity for the matched channels is given by*

$$C_c = \max_{Prob(x)} I(X;Y) \tag{5.19}$$

or for the mismatched case is

$$C_v = \max_{Prob(x)} I(X;V) \tag{5.20}$$

where the mutual information between X and Z (where Z can be either Y or V) is

$$I(X;Z) = H(Z) - H(Z|X) \tag{5.21}$$

$$H(Z) = \int_Z -p(z) log_2 p(z) dz \tag{5.22}$$

$$H(Z|X) = \int_Z \int_X -p(z,x) log_2 p(z|x) dz dx \tag{5.23}$$

$$p(z) = p(z|x=1)q + p(z|x=-1)(1-q) \tag{5.24}$$

$$p(z,x) = p(z|x) Prob(x) \tag{5.25}$$

(b) If $L(\cdot)$ is a linear function, then mismatch does not reduce the Shannon capacity. If $L(\cdot)$ is a non-linear function, then mismatch may reduce the Shannon capacity.

Proof. The proof is based on the well-known fact that the mutual information between X and Y is identical to the mutual information between X and $f(Y)$ as long as $f(\cdot)$ is a one-to-one mapping function. □

Linear transform in (5.5) does not reduce the capacity, but non-linear transform in (5.6) may reduce the capacity. For parameter mismatched cases, we can convert it to matched parameters by a linear scaling, for example, multiplying (5.5) by σ_r^2/σ_c^2 or multiplying (5.6) by λ_r/λ_c. However, a non-linear transform given in (5.6) could result in a capacity reduction $C_v < C_c$ since we will not be able to map v back to y once we map the values of y into v.

Let the first letter (either G or L) in the subscript denote the channel types (BIAWGN or BIL, respectively) and the second letter (either G or L) denote the decoder type used in LLR computation. Then for BIAWGN and BIL channels and decoder matched to BIAWGN and BIL channels, we will obtain four capacities, $C_{G,G}$, $C_{L,G}$, $C_{G,L}$, $C_{L,L}$. Although $C_{G,G}$ and $C_{L,L}$ use the same type of noise, mismatching could come from the parameter mismatch, for example, σ_r^2 instead of σ_c^2 and λ_r instead of λ_c. In the following property, we show that for the LLR computation matched to BIAWGN will not reduce the capacity. We also show the mismatched parameters will not reduce the capacity for $C_{L,L}$ and $C_{G,L}$, even though we did not find close-form solution for $C_{G,L}$.

Property 12. *(a)*

$$C_{G,G} = \int_Y -p(y)log_2p(y)dy - \frac{1}{2}log_2(2\pi e\sigma_c^2) \tag{5.26}$$

(b)

$$C_{L,G} = C_{L,L} = \frac{4\tan^{-1}(e^{-\frac{1}{\lambda_c}}) - \pi}{2e^{\frac{1}{\lambda_c}}\ln 2} - log_2\frac{1+e^{-\frac{2}{\lambda_c}}}{2} \tag{5.27}$$

(c)

$$\begin{aligned} C_{G,L} &= Q(\tfrac{2}{\sigma_c})log_2\tfrac{4Q(2/\sigma_c)}{2Q(2/\sigma_c)+1} + \tfrac{1}{2}log_2\tfrac{2}{2Q(2/\sigma_c)+1} \\ &\quad -\tfrac{1}{\sqrt{2\pi}}\int_0^{2/\sigma_c} e^{-\alpha^2/2}log_2(\tfrac{1+e^{\frac{2\alpha}{\sigma_c}-\frac{2}{\sigma_c^2}}}{2})d\alpha \end{aligned} \tag{5.28}$$

Proof. According to Property 11, we have the capacity $C_{G,G}$ and $C_{L,G}$ are equal to the capacity of BIAWGN and BIL channels, regardless the noise parameter mismatches, which are the right hand sides of Eqns. (5.26) (5.27), respectively [85] [130].

For the $C_{G,G}$ case, it is very straightforward from capacity of BIAWGN channel as

$$C_{G,G} = I(X;Y)|_{q=0.5} = H(Y) - H(Y|X)|_{q=0.5} \tag{5.29}$$

where $H(Y|X)|_{q=0.5} = \frac{1}{2}log_2(2\pi e\sigma_c^2)$.

Now let us prove for the $C_{L,L}$ case. Given $q = 0.5$ and v in (5.14), we have

$$C_{L,L} = I(X;V)|_{q=0.5}$$

$$= \frac{1}{2}\int_R p(v|x=1)log_2\frac{2p(v|x=1)}{p(v|x=1)+p(v|x=-1)}$$

$$+\frac{1}{2}\int_R p(v|x=-1)log_2\frac{2p(v|x=-1)}{p(v|x=1)+p(v|x=-1)}$$

$$= \frac{1}{2}e^{-\frac{2}{\lambda_c}}log_2\frac{2e^{-\frac{2}{\lambda_c}}}{e^{-\frac{2}{\lambda_c}}+1} + \frac{1}{2}log_2\frac{2}{e^{-\frac{2}{\lambda_c}}+1}$$

$$+\frac{\lambda_r}{8\lambda_c}e^{-\frac{1}{\lambda_c}}\int_{-2/\lambda_r}^{2/\lambda_r} e^{\frac{v\lambda_r}{2\lambda_c}} log_2 \frac{2e^{v\lambda_r/(2\lambda_c)}}{e^{v\lambda_r/(2\lambda_c)}+e^{-v\lambda_r/(2\lambda_c)}} dv$$

$$+\frac{\lambda_r}{8\lambda_c}e^{-\frac{1}{\lambda_c}}\int_{-2/\lambda_r}^{2/\lambda_r} e^{-\frac{v\lambda_r}{2\lambda_c}} log_2 \frac{2e^{-v\lambda_r/(2\lambda_c)}}{e^{v\lambda_r/(2\lambda_c)}+e^{-v\lambda_r/(2\lambda_c)}} dv$$

$$= \frac{1}{2}e^{-\frac{2}{\lambda_c}} log_2 \frac{2e^{-\frac{2}{\lambda_c}}}{e^{-\frac{2}{\lambda_c}}+1} + \frac{1}{2} log_2 \frac{2}{e^{-\frac{2}{\lambda_c}}+1}$$

$$+\frac{1}{2}e^{-\frac{1}{\lambda_c}}\int_{e^{-1/\lambda_c}}^{e^{1/\lambda_c}} log_2 \frac{2\alpha^2}{\alpha^2+1}\, d\alpha$$

$$= \frac{4\tan^{-1}(e^{-\frac{1}{\lambda_c}})-\pi}{2e^{1/\lambda_c}\ln 2} - log_2\left(\frac{1+e^{-2/\lambda_c}}{2}\right) \tag{5.30}$$

where $\alpha = e^{v\lambda_r/(2\lambda_c)}$.

Next let us prove for the $C_{L,G}$ case. Given $q = 0.5$ and v in (5.8), we have

$$C_{L,G} = I(X;V)|_{q=0.5}$$

$$= \frac{1}{2}\int_R p(v|x=1) log_2 \frac{2p(v|x=1)}{p(v|x=1)+p(v|x=-1)}$$

$$+\frac{1}{2}\int_R p(v|x=-1) log_2 \frac{2p(v|x=-1)}{p(v|x=1)+p(v|x=-1)}$$

$$= \int_{-\infty}^{-\frac{2}{\sigma_r^2}} \frac{\sigma_r^2}{4\lambda_c} e^{\frac{v\sigma_r^2-2}{2\lambda_c}} log_2 \left(\frac{2e^{\frac{v\sigma_r^2-2}{2\lambda_c}}}{e^{\frac{v\sigma_r^2-2}{2\lambda_c}}+e^{\frac{v\sigma_r^2+2}{2\lambda_c}}}\right) dv$$

$$+\int_{\frac{2}{\sigma_r^2}}^{\infty}\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{v\sigma_r^2-2}{2\lambda_c}}log_2\left(\frac{2e^{-\frac{v\sigma_r^2-2}{2\lambda_c}}}{e^{-\frac{v\sigma_r^2-2}{2\lambda_c}}+e^{-\frac{v\sigma_r^2+2}{2\lambda_c}}}\right)dv$$

$$+\int_{-\frac{2}{\sigma_r^2}}^{\frac{2}{\sigma_r^2}}\frac{\sigma_r^2}{4\lambda_c}e^{\frac{v\sigma_r^2-2}{2\lambda_c}}log_2\left(\frac{2e^{\frac{v\sigma_r^2-2}{2\lambda_c}}}{e^{\frac{v\sigma_r^2-2}{2\lambda_c}}+e^{-\frac{v\sigma_r^2+2}{2\lambda_c}}}\right)dv$$

$$=\int_{-\infty}^{-\frac{2}{\sigma_r^2}}\frac{\sigma_r^2}{4\lambda_c}e^{\frac{v\sigma_r^2-2}{2\lambda_c}}log_2\frac{2}{1+e^{2/\lambda_c}}dv$$

$$+\int_{\frac{2}{\sigma_r^2}}^{\infty}\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{v\sigma_r^2-2}{2\lambda_c}}log_2\frac{2}{1+e^{-2/\lambda_c}}dv$$

$$+\int_{-\frac{2}{\sigma_r^2}}^{\frac{2}{\sigma_r^2}}\frac{\sigma_r^2}{4\lambda_c}e^{\frac{v\sigma_r^2-2}{2\lambda_c}}log_2\frac{2}{1+e^{-v\sigma_r^2/\lambda_c}}dv$$

$$=\frac{1}{2}e^{-\frac{2}{\lambda_c}}log_2\frac{2}{1+e^{2/\lambda_c}}+\frac{1}{2}log_2\frac{2}{1+e^{-2/\lambda_c}}$$

$$+\frac{1}{2}e^{-\frac{1}{\lambda_c}}\int_{e^{-1/\lambda_c}}^{e^{1/\lambda_c}}log_2\frac{2\alpha^2}{\alpha^2+1}\,d\alpha$$

$$=\frac{4\tan^{-1}(e^{-\frac{1}{\lambda_c}})-\pi}{2e^{1/\lambda_c}\ln 2}-log_2\left(\frac{1+e^{-2/\lambda_c}}{2}\right)=C_{L,L} \tag{5.31}$$

where $\alpha=e^{v\sigma_r^2/(2\lambda_c)}$. So we also verified that $C_{L,G}=C_{L,L}$ which is the result of Property 11.

Following the similar procedure, we can prove the $C_{G,L}$ case. Given $q=0.5$ and v in (5.13), we have

$$C_{G,L} = I(X;V)|_{q=0.5}$$

$$= \frac{1}{2}\int_R p(v|x=1) log_2 \frac{2p(v|x=1)}{p(v|x=1)+p(v|x=-1)}$$

$$+\frac{1}{2}\int_R p(v|x=-1) log_2 \frac{2p(v|x=-1)}{p(v|x=1)+p(v|x=-1)}$$

$$= Q(\frac{2}{\sigma_c}) log_2 \frac{4Q(2/\sigma_c)}{2Q(2/\sigma_c)+1} + \frac{1}{2} log_2 \frac{2}{2Q(2/\sigma_c)+1}$$

$$+\frac{\lambda_r}{2\sqrt{2\pi}\sigma_c}\int_{-2/\lambda_r}^{2/\lambda_r} e^{-\frac{(v\lambda_r-2)^2}{8\sigma_c^2}} log_2 \frac{2}{1+e^{-v\lambda_r/\sigma_c^2}} dv$$

$$= Q(\frac{2}{\sigma_c}) log_2 \frac{4Q(2/\sigma_c)}{2Q(2/\sigma_c)+1} + \frac{1}{2} log_2 \frac{2}{2Q(2/\sigma_c)+1}$$

$$-\frac{1}{\sqrt{2\pi}}\int_0^{2/\sigma_c} e^{-\alpha^2/2} log_2 \left(\frac{1+e^{\frac{2\alpha}{\sigma_c}-\frac{2}{\sigma_c^2}}}{2}\right) d\alpha \tag{5.32}$$

where $\alpha = \frac{2-v\lambda_r}{2\sigma_c}$.

Clearly, $C_{G,L}$ is the only capacity reduced, i.e., $C_{G,L} < C_{G,G}$. □

The following Property states the behavior of $\frac{E_b}{N_o}$ as the code rate approaches to zero, i.e., $R \to 0$, or equivalently $\sigma_c \to \infty$ for BIAWGN and $\lambda_c \to \infty$ for BIL.

Property 13. *(a)* $\lim_{R\to 0}(E_b/N_0)_{G,G} = ln2$,

(b) $\lim_{R\to 0}(E_b/N_0)_{G,L} = \pi\ln(2)/2$,

(c) $\lim_{R\to 0}(E_b/N_0)_{L,L} = \lim_{R\to 0}(E_b/N_0)_{L,G}$=ln(2)/2

Proof. (a) $\lim_{R\to 0}(E_b/N_0)_{G,G} = \lim_{\sigma_c\to\infty}\frac{1}{2\sigma_c^2 C_{G,G}}$,

$$\lim_{\sigma_c\to\infty} 2\sigma_c^2 C_{G,G}$$

$$= -\lim_{\sigma_c\to\infty} 2\sigma_c^2\left\{\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_c}e^{-\frac{(y-1)^2}{2\sigma_c^2}}log_2\left[\frac{2e^{-\frac{(y-1)^2}{2\sigma_c^2}}}{e^{-\frac{(y-1)^2}{2\sigma_c^2}}+e^{-\frac{(y+1)^2}{2\sigma_c^2}}}\right]dy\right.$$

$$\left.+\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_c}e^{-\frac{(y+1)^2}{2\sigma_c^2}}log_2\left[\frac{2e^{-\frac{(y+1)^2}{2\sigma_c^2}}}{e^{-\frac{(y+1)^2}{2\sigma_c^2}}+e^{-\frac{(y-1)^2}{2\sigma_c^2}}}\right]dy\right\}$$

$$= -\lim_{\sigma_c\to\infty}\frac{\sigma_c}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{(y-1)^2}{2\sigma_c^2}}log_2\frac{1+e^{-\frac{4y}{2\sigma_c^2}}}{2}dy$$

$$-\lim_{\sigma_c\to\infty}\frac{\sigma_c}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{(y+1)^2}{2\sigma_c^2}}log_2\frac{1+e^{\frac{4y}{2\sigma_c^2}}}{2}dy$$

$$= \lim_{\beta\to 0}\frac{-4\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}log_2(\frac{(1+e^{-\beta^2}+e^{y\beta-\frac{\beta^2}{2}}+e^{-y\beta-\frac{\beta^2}{2}})}{4})dy}{\sqrt{2\pi}\beta^2}$$

$$= 2log_2(e) - \lim_{\beta\to 0}\frac{log_2(e)\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}(ye^{y\beta-\frac{\beta^2}{2}} - ye^{-y\beta-\frac{\beta^2}{2}})dy}{2\sqrt{2\pi}\beta}$$

$$= 2log_2(2) - \frac{log_2(e)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy = log_2(e) \tag{5.33}$$

where $\beta = \frac{2}{\sigma_c}$.

Thus, $\lim_{R\to 0}(E_b/N_0)_{G,G} = 1/log_2(e) = ln2$.

(b) $\lim_{R\to 0}(E_b/N_0)_{G,L} = \lim_{\sigma_c\to\infty} \frac{1}{2\sigma_c^2 C_{G,L}}$,

$$\lim_{\sigma_c\to\infty} 2\sigma_c^2 C_{G,L}$$

$$= \lim_{\sigma_c\to\infty} 2\sigma_c^2 \left\{ Q(\frac{2}{\sigma_c}) log_2 \frac{4Q(2/\sigma_c)}{2Q(2/\sigma_c)+1} + \frac{1}{2} log_2 \frac{2}{2Q(2/\sigma_c)+1} \right.$$

$$\left. - \frac{1}{\sqrt{2\pi}} \int_0^{2/\sigma_c} e^{-\alpha^2/2} log_2(\frac{1+e^{\frac{2\alpha}{\sigma_c} - \frac{2}{\sigma_c^2}}}{2}) d\alpha \right\}$$

$$= \lim_{\beta\to 0} \frac{8Q(\beta)log_2(4Q(\beta)) + 4 - 4(2Q(\beta)+1)log_2(2Q(\beta)+1)}{\beta^2}$$

$$= \lim_{\beta\to 0} \frac{-4log_2(4Q(\beta)) + 4log_2(2Q(\beta)+1)}{\beta\sqrt{2\pi}}$$

$$= \lim_{\beta\to 0} \frac{2log_2(e)}{\pi} \left(\frac{1}{Q(\beta)} - \frac{2}{2Q(\beta)+1} \right)$$

$$= \frac{2log_2(e)}{\pi} \tag{5.34}$$

where $\beta = \frac{2}{\sigma_c}$.

Thus, $\lim_{R\to 0}(E_b/N_0)_{G,L} = \frac{1}{2log_2(e)/\pi} = \pi \ln(2)/2$.

(c) $\lim_{R\to 0}(E_b/N_0)_{L,L} = \lim_{R\to 0}(E_b/N_0)_{L,G} = \lim_{\lambda_c\to\infty} \frac{1}{4\lambda_c^2 C_{L,L}}$

$$\lim_{\lambda_c\to\infty} 4\lambda_c^2 C_{L,L}$$

$$= \lim_{\lambda_c\to\infty} 4\lambda_c^2 \left[\frac{4\tan^{-1}(e^{-\frac{1}{\lambda_c}}) - \pi}{2e^{1/\lambda_c}\ln 2} - log_2\left(\frac{1+e^{-2/\lambda_c}}{2}\right)\right]$$

$$= \lim_{\beta\to 0} \frac{8e^{-\beta}\tan^{-1}(e^{-\beta}) - 2\pi e^{-\beta}}{\beta^2\ln(2)} - \frac{4log_2(\frac{1+e^{-2\beta}}{2})}{\beta^2}$$

$$= \lim_{\beta\to 0} \frac{-4e^{-\beta}\tan^{-1}(e^{-\beta}) + \pi e^{-\beta}}{\beta\ln(2)}$$

$$= \lim_{\beta\to 0} \frac{4e^{-\beta}\tan^{-1}(e^{-\beta}) - \pi e^{-\beta} + \frac{4e^{-2\beta}}{1+e^{-2\beta}}}{\ln(2)}$$

$$= 2log_2(e) \tag{5.35}$$

where $\beta = \frac{1}{\lambda_c}$.

Thus, $\lim_{R\to 0}(E_b/N_0)_{L,L} = \lim_{R\to 0}(E_b/N_0)_{L,G} = \frac{1}{2log_2(e)} = \ln(2)/2$.

□

5.2 Capacity and Stability for LDPC Codes Under Sum-Product Decoding with Mismatched Noises

In this section, we will use density evolution to evaluate the capacity for infinite length of LDPC codes and also evaluate the stability of sum-product decoding with mismatched noises. The details of technics used here can be found in [5] [130].

A LDPC code can be represented by a Tanner graph. Typically, the variable nodes are drawn at the left-hand side and the check nodes are listed at the right-hand side. A variable node can only connect to check nodes and a check node can only connect to variable nodes. The degree of a node denotes the number of other type of nodes that it connects to. We call the graph a regular LDPC code if all degrees are the same for variable nodes and also if all degrees are the same for check nodes, and an irregular LDPC code if not. For example, in the Tanner graph for a regular (3,6) LDPC code each variable node connects to 3 check nodes and each check node connects to 6 variable nodes.

Let $\lambda(x)$ and $\rho(x)$ denote the degree distributions of variable and check nodes (often called the left and right degree), where $\lambda(x) = \sum_{i=2}^{d_l} \lambda_i x^{i-1}$ and $\rho(x) = \sum_{j=2}^{d_r} \rho_j x^{j-1}$. Here, λ_i is the fraction of edges that belong to degree-i variable nodes, ρ_j is the fraction of edges that belong to degree-j check nodes, d_l denotes the maximum variable degree, and d_r is the maximum check degree. Then, the code rate is $R = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx}$. If $\lambda_{d_v} = 1$ and $\rho_{d_c} = 1$, then it is reduced to regular (d_v, d_c) LDPC codes.

We apply the discrete density evolution method [130] to analyze the performance of LDPC codes under sum-product decoding with mismatched noises. In short, we use log-likelihood ratio for messages, i.e., $w = log\frac{p(z|x=1)}{p(z|x=-1)}$, where x is the bit value of the node and z denotes all the messages available to the node up to the present iteration obtained from edges other than one carrying w. For the variable nodes, we have [130]

$$p_u^{(l+1)} = \rho(p_v^{(l)}) \tag{5.36}$$

$$p_v^{(l+1)} = p_{u_0} * \lambda(p_u^{(l)}) \tag{5.37}$$

where $p_{u_0} = p(v|x = 1)$ given in (5.7), (5.8), (5.13) and (5.14), $\lambda(p) = \sum_{i=2}^{d_l} \lambda_i \otimes^{i-1} p$ and $\rho(p) = \sum_{j=2}^{d_r} \rho_j \mathcal{R}^{j-1} p$, $\otimes$ denotes discrete convolution, $\mathcal{R}^{j-1} p = \mathcal{R}(p, \mathcal{R}(p, ..., \mathcal{R}(p, p)))$, and

$$\mathcal{R}(p_1, p_2) = 2 \tanh^{-1} \left(\Pi_{i=1}^{2} \tanh(p_i) \right) \tag{5.38}$$

For four different cases, the initial density functions (p_{u_0}) are given in (5.7), (5.8), (5.13) and (5.14), respectively. These four types of initial density functions satisfies the channel symmetric condition [85]. Thus, substituting these initial density functions into (5.36) and (5.37), we can obtain the thresholds for LDPC codes. Let $\gamma = \frac{\lambda_r}{\lambda_c} = \frac{\sigma_r}{\sigma_c}$, where $2\lambda^2 = \sigma^2$. The numerical examples can be found in Table 5.2-5.5.

In [85] and [130] a stability function $S(\mathcal{N})$ is introduced to study the behavior of density evolution near the zero probability of error, where $\mathcal{N}$ denotes the channel parameter and

$$S(\mathcal{N}) = \inf_{s<0} \int_{\mathcal{R}} e^{su} p_{u_0} du. \tag{5.39}$$

In detail, if $\lambda'(0)\rho'(1) < S^{-1}(\mathcal{N})$, then the error probability will converge to zero if it is initially small enough, and if $\lambda'(0)\rho'(1) > S^{-1}(\mathcal{N})$, then the error probability is bounded away from zero. Since $\lambda'(0)\rho'(1) = \lambda_2 \sum_{j=2}^{d_r} \rho_j (j-1)$ the limitation is mainly on degree 2 variable nodes. If $S(\mathcal{N}) \to \infty$, then no degree 2 variable node is allowed.

For mismatched cases, we have the following property on stability condition.

Property 14. *(a)*

$$S_{G,G}(\sigma_c, \gamma) = e^{\frac{-1}{2\sigma_c^2}} \tag{5.40}$$

(b)

$$S_{L,G}(\lambda_c, \gamma) = \frac{e^{1-\sqrt{1+1/\lambda_c^2}}}{2\lambda_c \sqrt{1+\lambda_c^2} - 2\lambda_c^2} \tag{5.41}$$

(c)

$$\begin{aligned} S_{G,L}(\sigma_c,\gamma) &= \inf_{s<0} 0.5e^{\frac{2\sqrt{2}s}{\gamma\sigma_c}} + Q(\tfrac{2}{\sigma_c})e^{\frac{-2\sqrt{2}s}{\gamma\sigma_c}} \\ &+ \left(Q(-\tfrac{2\sqrt{2}s}{\gamma} - \tfrac{2}{\sigma_c}) - Q(-\tfrac{2\sqrt{2}s}{\gamma})\right) e^{\frac{4s^2}{\gamma^2}+\frac{2\sqrt{2}s}{\gamma\sigma_c}} \end{aligned} \tag{5.42}$$

(d)

$$S_{L,L}(\lambda_c,\gamma) = (1+\frac{1}{\lambda_c})e^{-\frac{1}{\lambda_c}} \tag{5.43}$$

Proof. (a) Substituting (5.7) into (5.39) we have

$$S_{G,G}(\sigma_c,\gamma) = \inf_{s<0}\int_{\mathcal{R}} e^{su}\frac{\sigma_r^2}{2\sqrt{2\pi}\sigma_c}e^{-\frac{(u\sigma_r^2-2)^2}{8\sigma_c^2}}du$$

$$= \inf_{s<0}\int_{\mathcal{R}} \frac{\sigma_r^2}{2\sqrt{2\pi}\sigma_c}e^{-\frac{u^2\sigma_r^4-4u(\sigma_r^2+2\sigma_c^2 s)+4}{8\sigma_c^2}}du$$

$$= \inf_{s<0}\int_{\mathcal{R}} \frac{\sigma_r^2}{2\sqrt{2\pi}\sigma_c}e^{-\frac{\left(u\sigma_r^2-\frac{2(\sigma_r^2+2\sigma_c^2 s)}{\sigma_r^2}\right)^2+4-\frac{4(\sigma_r^2+2\sigma_c^2 s)^2}{\sigma_r^4}}{8\sigma_c^2}}du$$

$$= \inf_{s<0} e^{\frac{(\sigma_r^2+2\sigma_c^2 s)^2-\sigma_r^4}{2\sigma_r^4\sigma_2^2}}$$

$$= \inf_{s<0} e^{\frac{(\gamma^2+2s)^2\sigma_c^2-\gamma^4}{2\gamma^4\sigma_c^2}} \tag{5.44}$$

where $\gamma = \frac{\sigma_r}{\sigma_c}$.

By solving $\frac{\partial}{\partial s}\left(e^{\frac{(\gamma^2+2s)^2\sigma_c^2-\gamma^4}{2\gamma^4\sigma_c^2}}\right) = 0$, we get $s^* = -\frac{\gamma^2}{2}$. Substituting it back to the above equation, we finally get $S_{G,G}(\sigma_c,\gamma) = e^{-\frac{1}{2\sigma_c^2}}$.

(b) Substituting (5.8) into (5.39) we have

$$S_{L,G}(\lambda_c,\gamma)=\inf_{s<0}\int_{\mathcal{R}}e^{su}\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{|\frac{u\sigma_r^2}{2}-1|}{\lambda_c}}du$$

$$=\inf_{s<0}\left\{\int_{-\infty}^{\frac{2}{\sigma_r^2}}\frac{\sigma_r^2}{4\lambda_c}e^{\frac{\frac{u\sigma_r^2}{2}-1}{\lambda_c}}e^{su}du+\int_{\frac{2}{\sigma_r^2}}^{\infty}\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{\frac{u\sigma_r^2}{2}-1}{\lambda_c}}e^{su}du\right\}$$

$$=\inf_{s<0}\left\{\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{1}{\lambda_c}}\int_{-\infty}^{\frac{2}{\sigma_r^2}}e^{u(\frac{\sigma_r^2}{2\lambda_c}+s)}du+\frac{\sigma_r^2}{4\lambda_c}e^{\frac{1}{\lambda_c}}\int_{\frac{2}{\sigma_r^2}}^{\infty}e^{u(-\frac{\sigma_r^2}{2\lambda_c}+s)}du\right\}$$

$$=\inf_{s<0}\left\{\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{1}{\lambda_c}}\frac{1}{\frac{\sigma_r^2}{2\lambda_c}+s}e^{u(\frac{\sigma_r^2}{2\lambda_c}+s)}\Big|_{-\infty}^{\frac{2}{\sigma_r^2}}+\frac{\sigma_r^2}{4\lambda_c}e^{\frac{1}{\lambda_c}}\frac{1}{-\frac{\sigma_r^2}{2\lambda_c}+s}e^{u(-\frac{\sigma_r^2}{2\lambda_c}+s)}\Big|_{\frac{2}{\sigma_r^2}}^{\infty}\right\}$$

$$=\inf_{-\frac{\sigma_r^2}{2\lambda_c}<s<0}\left\{\frac{\sigma_r^2}{4\lambda_c}e^{-\frac{1}{\lambda_c}}\frac{1}{\frac{\sigma_r^2}{2\lambda_c}+s}e^{\frac{2}{\sigma_r^2}(\frac{\sigma_r^2}{2\lambda_c}+s)}-\frac{\sigma_r^2}{4\lambda_c}e^{\frac{1}{\lambda_c}}\frac{1}{-\frac{\sigma_r^2}{2\lambda_c}+s}e^{\frac{2}{\sigma_r^2}(-\frac{\sigma_r^2}{2\lambda_c}+s)}\right\}$$

$$=\inf_{-\frac{\sigma_r^2}{2\lambda_c}<s<0}\left\{\frac{(\frac{\sigma_r^2}{2\lambda_c})^2}{(\frac{\sigma_r^2}{2\lambda_c})^2-s^2}e^{\frac{2s}{\sigma_r^2}}\right\}$$

$$=\inf_{-\gamma^2\lambda_c<s<0}\left\{\frac{(\lambda_c\gamma^2)^2}{(\lambda_c\gamma^2)^2-s^2}e^{\frac{s}{\lambda_c^2\gamma^2}}\right\}\tag{5.45}$$

where $\gamma=\frac{\sigma_r}{\sqrt{2\lambda_c}}$.

By solving $\frac{\partial}{\partial s}\left(\frac{1}{(\lambda_c\gamma^2)^2-s^2}e^{\frac{s}{\lambda_c^2\gamma^2}}\right)=0$, we get

$$\frac{e^{\frac{s}{\lambda_c^2\gamma^2}}\frac{1}{\lambda_c^2\gamma^2}}{(\lambda_c\gamma^2)^2-s^2}+\frac{e^{\frac{s}{\lambda_c^2\gamma^2}}2s}{[(\lambda_c\gamma^2)^2-s^2]^2}=0$$

or simply

$$s^2 - 2\lambda_c^2\gamma^2 s - \lambda_c^2\gamma^4 = 0$$

Taking the negative solution of the above equation, we get $s^* = \lambda_c^2\gamma^2(1 - \sqrt{1 + 1/\lambda_c^2})$ which we can easily verify that $-\gamma^2\lambda_c < s^* < 0$. Substituting it back to the original equation (5.45), we finally get $S_{L,G}(\lambda_c, \gamma) = \frac{e^{1-\sqrt{1+1/\lambda_c^2}}}{2\lambda_c\sqrt{1+\lambda_c^2}-2\lambda_c^2}$.

(c) Substituting (5.13) into (5.39) we have

$$S_{G,L}(\sigma_c, \gamma) = \inf_{s<0} \int_{\mathcal{R}} e^{su} \left[Q(\frac{2}{\sigma_c})\delta(u + \frac{2}{\lambda_r}) \right.$$

$$\left. + \frac{1}{2}\delta(u - \frac{2}{\lambda_r}) + \frac{\lambda_r}{2\sqrt{2\pi}\sigma_c} e^{-\frac{(\frac{u\lambda_r}{2\sigma_c} - \frac{1}{\sigma_c})^2}{2}} \Pi(\frac{2}{\lambda_r}) \right] du$$

$$= \inf_{s<0} Q(\frac{2}{\sigma_c}) e^{\frac{-2s}{\lambda_r}} + 0.5e^{\frac{2s}{\lambda_r}} + \int_{-\frac{2}{\lambda_r}}^{\frac{2}{\lambda_r}} \frac{e^{su}\lambda_r}{2\sqrt{2\pi}\sigma_c} e^{-\frac{(\frac{u\lambda_r}{2\sigma_c} - \frac{1}{\sigma_c})^2}{2}} du$$

$$= \inf_{s<0} Q(\frac{2}{\sigma_c}) e^{\frac{-2s}{\lambda_r}} + 0.5e^{\frac{2s}{\lambda_r}} + e^{\frac{(\lambda_r + 2\sigma_c^2 s)^2 - \lambda_r^2}{2\lambda_r^2\sigma_c^2}} \int_{-\frac{2\sigma_c s}{\lambda_r} - \frac{2}{\sigma_c}}^{-\frac{2\sigma_c s}{\lambda_r}} e^{-\frac{u^2}{2}} du$$

$$= \inf_{s<0} Q(\frac{2}{\sigma_c}) e^{\frac{-2\sqrt{2}s}{\gamma\sigma_c}} + 0.5e^{\frac{2\sqrt{2}s}{\gamma\sigma_c}}$$

$$+ \left(Q(-\frac{2\sqrt{2}s}{\gamma} - \frac{2}{\sigma_c}) - Q(-\frac{2\sqrt{2}s}{\gamma}) \right) e^{\frac{4s^2}{\gamma^2} + \frac{2\sqrt{2}s}{\gamma\sigma_c}} \tag{5.46}$$

where $\gamma = \frac{\sqrt{2}\lambda_r}{\sigma_c}$.

No close-form solution has been found in this case. Later we will give an upper bound instead.

(d) Substituting (5.14) into (5.39) we have

$$S_{L,L}(\lambda_c,\gamma)=\inf_{s<0}\int_{\mathcal{R}}e^{su}\left[\frac{1}{2}e^{-\frac{2}{\lambda_c}}\delta(u+\frac{2}{\lambda_c\gamma})\right.$$

$$\left.+\frac{1}{2}\delta(u-\frac{2}{\lambda_c\gamma})+\frac{\gamma}{4}e^{-|\frac{u\gamma}{2}-\frac{1}{\lambda_c}|}\Pi(\frac{2}{\lambda_c\gamma})\right]du$$

$$=\inf_{s<0}\frac{1}{2}e^{-\frac{2}{\lambda_c}-\frac{2s}{\lambda_c\gamma}}+\frac{1}{2}e^{\frac{2s}{\lambda_c\gamma}}+\int_{-\frac{2}{\lambda_c\gamma}}^{\frac{2}{\lambda_c\gamma}}\frac{\gamma e^{su}}{4}e^{\frac{u\gamma}{2}-\frac{1}{\lambda_c}}du$$

$$=\inf_{s<0}\frac{1}{2}e^{-\frac{2}{\lambda_c}-\frac{2s}{\lambda_c\gamma}}+\frac{1}{2}e^{\frac{2s}{\lambda_c\gamma}}+\frac{\gamma}{4(s+\gamma/2)}\left(e^{\frac{2s}{\lambda_c\gamma}}-e^{\frac{-2s}{\lambda_c\gamma}-\frac{2}{\lambda_c}}\right) \tag{5.47}$$

where $\gamma=\frac{\lambda_r}{\lambda_c}$.

By solving $\frac{\partial}{\partial s}\left[\frac{1}{2}e^{-\frac{2}{\lambda_c}-\frac{2s}{\lambda_c\gamma}}+\frac{1}{2}e^{\frac{2s}{\lambda_c\gamma}}+\frac{\gamma}{4(s+\gamma/2)}\left(e^{\frac{2s}{\lambda_c\gamma}}-e^{\frac{-2s}{\lambda_c\gamma}-\frac{2}{\lambda_c}}\right)\right]=0$, we get

$$(\frac{1}{\lambda_c\gamma}+\frac{\gamma}{(2s+\gamma)^2})(e^{\frac{2s}{\lambda_c\gamma}}-e^{-\frac{2(s+\gamma)}{\lambda_c\gamma}})+\frac{1}{\lambda_c(2s+r)}(e^{\frac{2s}{\lambda_c\gamma}}+e^{-\frac{2(s+\gamma)}{\lambda_c\gamma}})=0$$

or simply

$$[(2s+\gamma)^2+\gamma(2s+\gamma)-\lambda_c\gamma^2]e^{\frac{2(2s+\gamma)}{\lambda_c\gamma}}=(2s+\gamma)^2-\gamma(2s+\gamma)-\lambda_c\gamma^2$$

We find that $s^*=-\frac{\gamma}{2}$ is the solution to the above equation. Substituting it into the original equation (5.47), we can get $S_{L,L}(\lambda_c,\gamma)=e^{-\frac{1}{\lambda_c}}+\lim_{s\to-\frac{\gamma}{2}}\frac{\gamma(e^{\frac{2s}{\lambda_c\gamma}}-e^{-\frac{2(s+\gamma)}{\lambda_c\gamma}})}{2(2s+\gamma)}=(1+\frac{1}{\lambda_c})e^{-\frac{1}{\lambda_c}}$.

□

For the (G, L) case, although no close-form solution has been found for equation (5.46), we can obtain an upper bound on $S_{G,L}(\sigma_c, \gamma)$ by substituting $s = -\frac{\gamma}{2\sqrt{2}\sigma_c}$ into the right-hand side of (5.46), i.e.,

$$\begin{aligned} S_{G,L}(\sigma_c, \gamma) &\leq \tfrac{1}{2} e^{-\frac{1}{\sigma_c^2}} + Q(\tfrac{2}{\sigma_c}) e^{\frac{1}{\sigma_c^2}} + (1 - 2Q(\tfrac{1}{\sigma_c})) e^{-\frac{1}{2\sigma_c^2}} \\ &\leq (\tfrac{1}{2} + \tfrac{\sigma_c}{2\sqrt{2\pi}}) e^{-\frac{1}{\sigma_c^2}} + (1 - 2Q(\tfrac{1}{\sigma_c})) e^{-\frac{1}{2\sigma_c^2}}. \end{aligned} \tag{5.48}$$

where the second inequality follows $Q(x) \leq \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}$.

5.3 Numerical Results

We first evaluate the Shannon capacity loss due to mismatched noises. We then examine the threshold for regular LDPC codes due to mismatched noises. After that, stability conditions are compared between matched and mismatched cases. Finally, we examine the threshold for two irregular LDPC codes under mismatched noises.

The results in Table 5.1 show: (a) The non-linear transformation in (5.6) will reduce the capacity mainly for $R \leq 1/2$. It has marginal effects on higher rate. (b) The capacity (in term of E_b/N_o) of a decoder with mismatched noise density given in (5.13) will decrease to -0.46 dB as rate reduces from 0.9 to 1/8, then increases to 0.369 dB as rate further decreases. (c) Laplace channel has a higher capacity (i.e., lower E_b/N_o) for rate less than 2/3, reverses for the higher rate.

It is worth-mentioning here we assume that the impulsive Laplace noise and non-impulsive Gaussian noise have the same power level, i.e., $\sigma_c^2 = 2\lambda_c^2$. In practice, when the impulsive noise happens, it often has much higher power levels than the non-impulsive noise.

Table 5.1: Capacity of channels with and without mismatched noises

C	$\sigma_{G,G}$, $\frac{E_b}{N_o}$ dB	$\sigma_{G,L}$, $\frac{E_b}{N_o}$ dB	$\lambda_{L,L}$, $\frac{E_b}{N_o}$ dB
0	∞, -1.59	∞, 0.369	∞, -4.602
1/8	2.30, -1.21	2.11, -0.46	2.03, -3.12
1/4	1.55, -0.79	1.47, -0.36	1.30, -2.26
1/3	1.29, -0.49	1.25, -0.20	1.05, -1.70
1/2	0.979, 0.19	0.964, 0.32	0.752, -0.54
2/3	0.767, 1.06	0.763, 1.10	0.552, 0.90
3/4	0.677, 1.63	0.675, 1.65	0.469, 1.80
4/5	0.625, 2.04	0.623, 2.05	0.422, 2.44
7/8	0.545, 2.84	0.544, 2.86	0.354, 3.61
9/10	0.516, 3.20	0.515, 3.20	0.323, 4.25

Table 5.2: Threshold of regular LDPC codes on BIAWGN under sum-product decoding of BIAWGN with mismatched noise variance. Let $\sigma_\alpha = \sigma_{G,G}(\gamma = \alpha)$

C	d_v	d_c	σ_{opt}	$\sigma_{1.0}$	$\sigma_{1.4}$	$\sigma_{1.2}$	$\sigma_{0.83}$	$\sigma_{0.71}$
1/4	3	4	1.55	1.26	1.11	1.21	1.20	1.10
1/2	3	6	0.979	0.88	0.80	0.85	0.86	0.84
2/3	3	9	0.767	0.70	0.64	0.69	0.69	0.69
3/4	3	12	0.677	0.63	0.58	0.61	0.63	0.62
4/5	3	15	0.625	0.58	0.54	0.57	0.58	0.57
7/8	3	24	0.545	0.52	0.48	0.50	0.52	0.51
9/10	3	30	0.516	0.49	0.46	0.48	0.48	0.48

In Tables 5.2-5.5, we list thresholds of regular LDPC codes with different rates for various mismatched cases. In the computation we set the maximum number of iterations to 1000, the bit error rate limit to 10^{-10}.

Tables 5.2 and 5.3 show that if we use the BIL decoder for BIAWGN noises, then the reductions on E_b/N_o are 0.57 dB, 0.31 dB, 0.13 dB, 0.14 dB, 0.0 dB, 0.17 dB, 0.18 dB for rate 1/4, 1/2, 2/3, 3/4, 7/8, 9/10, respectively. Tables 5.4 and 5.5 show that if we use the BIAWGN decoder for BIL noises, then the reductions on E_b/N_o are 0.81 dB, 0.55 dB, 0.18 dB, 0.21 dB, 0.23 dB, 0.27 dB, 0.29 for rate 1/4, 1/2, 2/3, 3/4, 7/8, 9/10, respectively. On BIAWGN channels, the gap between capacities of LDPC codes and channels are from 1.8 dB for rate 1/4 to 0.45 dB for rate 9/10. Similarly, on BIL channels, the gap between capacities of LDPC codes and channels are from 2.19 dB for rate 1/4 to 0.64 dB for rate 9/10.

Table 5.3: Threshold of regular LDPC codes on BIAWGN under sum-product decoding of BIL with mismatched noise variance. Let $\sigma_\alpha = \sigma_{G,L}(\gamma = \alpha)$

C	d_v	d_c	σ_{opt}	$\sigma_{1.0}$	$\sigma_{1.4}$	$\sigma_{1.2}$	$\sigma_{0.83}$	$\sigma_{0.71}$
1/4	3	4	1.47	1.18	1.17	1.19	1.15	1.12
1/2	3	6	0.964	0.85	0.82	0.84	0.85	0.84
2/3	3	9	0.763	0.69	0.66	0.68	0.69	0.69
3/4	3	12	0.675	0.62	0.58	0.60	0.62	0.62
4/5	3	15	0.623	0.58	0.54	0.56	0.58	0.58
7/8	3	24	0.544	0.51	0.46	0.49	0.51	0.51
9/10	3	30	0.515	0.48	0.43	0.46	0.48	0.48

Table 5.4: Threshold of regular LDPC codes on BIL under sum-product decoding of BI-AWGN with mismatched noise variance. Let $\lambda_\alpha = \lambda_{L,G}(\gamma = \alpha)$

C	d_v	d_c	λ_{opt}	$\lambda_{1.0}$	$\lambda_{1.4}$	$\lambda_{1.2}$	$\lambda_{0.83}$	$\lambda_{0.71}$
1/4	3	4	1.30	0.92	0.80	0.88	0.86	0.83
1/2	3	6	0.752	0.61	0.57	0.61	0.59	0.58
2/3	3	9	0.552	0.47	0.45	0.48	0.45	0.45
3/4	3	12	0.469	0.41	0.41	0.41	0.40	0.40
4/5	3	15	0.422	0.37	0.37	0.38	0.36	0.37
7/8	3	24	0.354	0.31	0.31	0.32	0.31	0.31
9/10	3	30	0.323	0.29	0.30	0.29	0.29	0.29

Table 5.5: Threshold of regular LDPC codes on BIL under sum-product decoding of BIL with mismatched noise variance. Let $\lambda_\alpha = \lambda_{L,L}(\gamma = \alpha)$

C	d_v	d_c	λ_{opt}	$\lambda_{1.0}$	$\lambda_{1.4}$	$\lambda_{1.2}$	$\lambda_{0.83}$	$\lambda_{0.71}$
1/4	3	4	1.30	1.01	0.96	1.00	1.00	0.96
1/2	3	6	0.752	0.65	0.62	0.64	0.64	0.63
2/3	3	9	0.552	0.48	0.47	0.48	0.48	0.46
3/4	3	12	0.469	0.42	0.41	0.42	0.42	0.42
4/5	3	15	0.422	0.38	0.37	0.38	0.38	0.37
7/8	3	24	0.354	0.32	0.31	0.32	0.32	0.31
9/10	3	30	0.323	0.30	0.29	0.29	0.29	0.29

Figs. 5.2 and 5.3 show the results of Property 14. We conclude (a) the stability functions are independent of γ; (b) for mismatched decoders the limitation on degree 2 variable nodes from stability functions are marginally tightened than their matched counterparts.

Lastly, we consider two rate 1/2 irregular LDPC codes. The first one is from Table I in [85] with $\lambda_2 = 0.25105$, $\lambda_3 = 0.30938$, $\lambda_4 = 0.00104$, $\lambda_{10} = 0.43853$, $\rho_7 = 0.63676$, and $\rho_8 = 0.36324$. The second code is from Table 8.1 in [130] with $\lambda_2 = 0.234029$, $\lambda_3 = 0.212425$, $\lambda_6 = 0.146898$, $\lambda_7 = 0.102840$ $\lambda_{20} = 0.303808$, $\rho_8 = 0.71875$, and $\rho_9 = 0.28125$.

The degradation for the (G,L) case compared with the (G,G) case is 0.34 and 0.41 dB for codes 1 and 2, respectively. The degradation for the (L,G) case compared to the (L,L) case is 0.53 and 1.1 dB for codes 1 and 2, respectively. 20% under and over estimating the noise variance will have little effect on threshold of LDPC codes with and without noise mismatches.

Table 5.6: Threshold of rate 1/2 irregular LDPC code 1

	opt	$\gamma = 1.0$	$\gamma = 1.4$	$\gamma = 1.2$	$\gamma = 0.83$	$\gamma = 0.71$
$\sigma_{G,G}$	0.979	0.955	0.843	0.917	0.928	0.894
$\sigma_{G,L}$	0.964	0.918	0.899	0.915	0.918	0.905
$\lambda_{L,G}$	0.752	0.648	0.608	0.649	0.641	0.637
$\lambda_{L,L}$	0.752	0.689	0.684	0.686	0.686	0.685

Table 5.7: Threshold of rate 1/2 irregular LDPC code 2

	opt	$\gamma = 1.0$	$\gamma = 1.4$	$\gamma = 1.2$	$\gamma = 0.83$	$\gamma = 0.71$
$\sigma_{G,G}$	0.979	0.965	0.845	0.923	0.913	0.885
$\sigma_{G,L}$	0.964	0.921	0.904	0.917	0.919	0.886
$\lambda_{L,G}$	0.752	0.624	0.611	0.623	0.624	0.623
$\lambda_{L,L}$	0.752	0.706	0.696	0.704	0.706	0.702

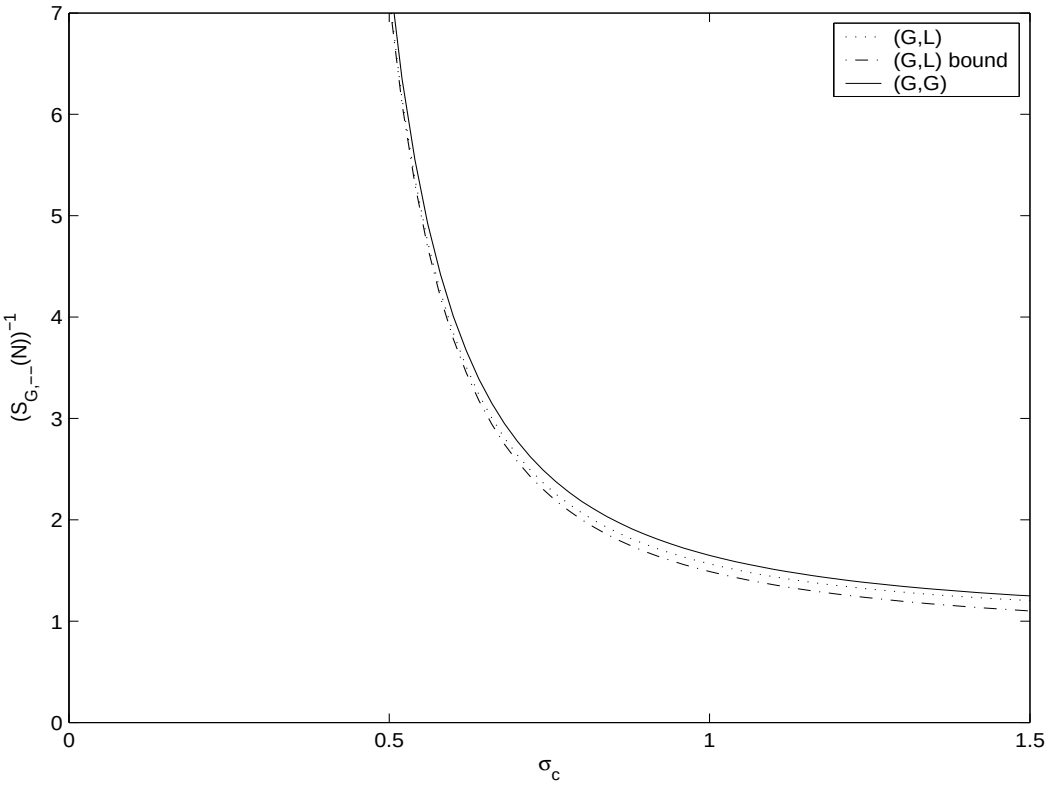

Figure 5.2: Stability as a function of γ for the (G, G) and (G, L) cases.

5.4 Discussion and Conclusion

In this chapter, we proved that as long as the true noise type and the estimated noise type used in the decoder are binary-input memoryless output symmetric channels, the output from mismatched log-likelihood ratio (LLR) computation is also symmetric. Therefore, we can analyze the threshold of LDPC codes under sum-product decoder using mismatched noises by the density evolution method. We have also showed the Shannon capacity can be evaluated for mismatched LLR computation. If the mismatched LLR computation is not an one-to-one mapping function, then the capacity could be reduced. We derived the Shannon capacity, threshold and stable condition of LDPC codes for mismatched BIAWGN and BIL noise types. The results show that the noise variance estimation errors will not

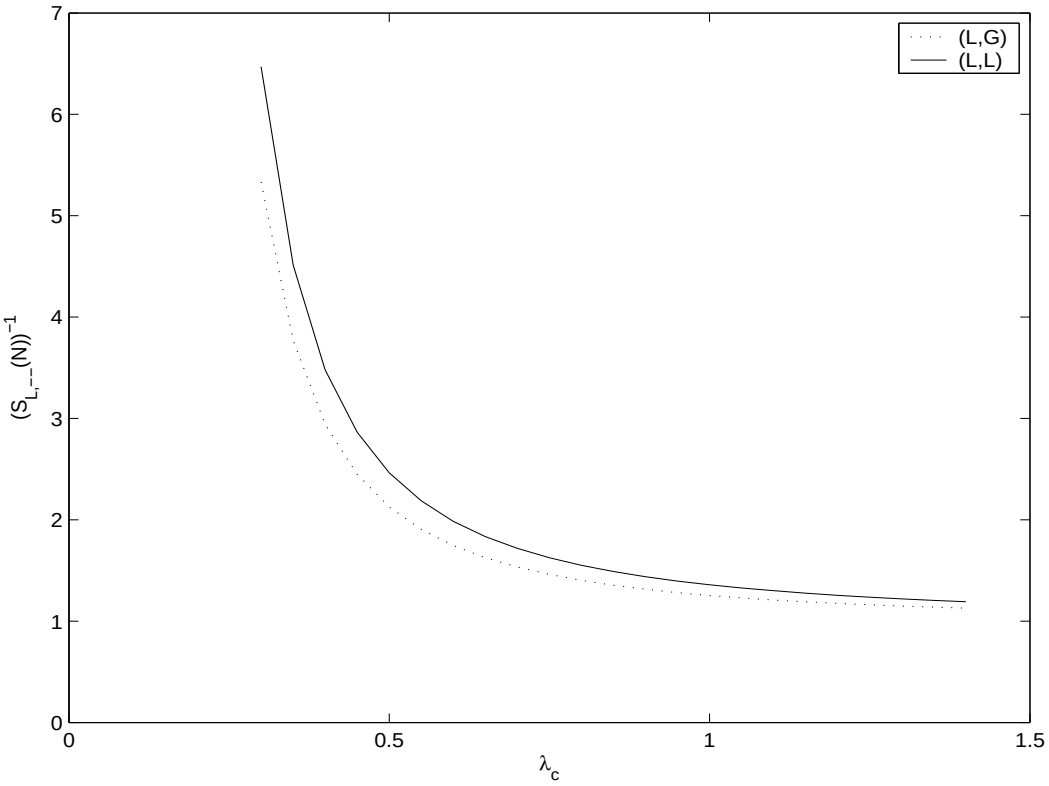

Figure 5.3: Stability as a function of γ for the (L, L) and (L, G) cases.

affect the Shannon capacity and stable condition, but the errors do reduce the threshold. The mismatching in noise type will only reduce Shannon capacity when LLR computation is based on BIL. The stable condition is marginally affected by the noise mismatches.

Although we only analyzed two types of noises, BIAWGN and BIL, the work can be applied to many symmetric channels in practice. These theoretic results are important for studying robust decoding for turbo and LDPC codes.

CHAPTER 6: SUMMARY AND FUTURE WORKS

6.1 Summary

An entire MB-OFDM system over UWB channel for WPAN applications was presented in the first part of this dissertation, which included frequency hopping, TTCM encoder, bit and symbol interleaver, OFDM and equalizer. We used a simple parity-check code as the outer codes and trellis coded modulation as the inner codes, which has less complexity compared to traditional TTCM codes. Excellent performance of this system were verified by density evolution. An iterative decoder with modified branch metric computation for this TTCM coded MB-OFDM system was proposed to optimize the bit error performance, which was extensively examined under various impulsive noises over both Gaussian and UWB channel, especially in mismatched scenarios. Impulsive noises at time domain were compared with their counterparts in frequency domain after passing through equalizer and OFDM demodulator. A robust decoder which is immune to noise mismatch was provided based on comparison of impulsive noises in time domain and frequency domain.

In the second part of this dissertation, we proved that the minimax decoder proposed in [38], which instead of minimizing the average bit error probability aims at minimizing the worst bit error probability, is optimal and robust to certain noise model with unknown prior probabilities. We extended some results from two dimensions to higher dimensions.

In the last part of this dissertation, density evolution was extended for sum-product decoding using mismatched noises. We showed that the output from mismatched log-likelihood ratio (LLR) computation is also symmetric, as long as the true noise type and the

estimated noise type used in the decoder are both binary-input memoryless output symmetric channels. We used BIAWGN and BIL channels to derive our conclusions. We demonstrated that the Shannon capacity can be evaluated for mismatched LLR computation and it can be reduced if the mismatched LLR computation is not an one-to-one mapping function. We further derived the Shannon capacity, threshold and stable condition of LDPC codes for mismatched BIAWGN and BIL noise types. Based on our investigation, the noise variance estimation errors will not affect the Shannon capacity and stable condition, although the threshold can be reduced.

6.2 Future Works

(1) This dissertation only investigated simple parity-concatenated TCM codes. More powerful outer code such as LDPC codes could provide more error correcting capacity and get considerable coding gains if complexity issue is well considered.

(2) Only two-dimensional constellation with gray was considered in this dissertation. Employing different mapping schemes with multi-dimensional constellation to achieve better error performance improvement is an attractive subject as well.

(3) Other impulsive noise cancelation methods in OFDM system, such as clipping, blanking, etc. need to be combined with our decoding algorithm to get additional performance benefit. Iterative noise cancelation is also of interest to be included inside our iterative decoder.

(4) Minimax robust decoder could be applied to achieve better performance in our MB-OFDM system. Both numerical simulation and theoretical analysis need to be done.

LIST OF REFERENCES

[1] C. E. Shannon, "A methematical theory of communication," *Bell System Technical Journal*, Vol. 27, pp.379-423 (Part I), 623-656(Part II), Jul. 1948.

[2] R. G. Gallager, *Low Density Parity-Check Codes*. MIT Press, Cambridge, MA, 1963.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error correcting coding and decoding: turbo-codes," *Proc. ICC'93*, Geneva, Switzerland, May 1993, pp.1064-1070.

[4] D. J. C. Mackay, "Good error-correcting codes based on very sparse matrices", *IEEE Trans. Inform. Theory*, Vol.45, no.2, pp.399-431, Mar. 1999.

[5] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, Vol.47, pp.599-618, Feb. 2001.

[6] S.-Y. Chung, G. D. Forney, Jr., T. J. Richardson, R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Communications Letters*, vol. 5, no. 2, pp. 58-60, Feb 2001.

[7] Federal Communications Commission, "Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems: First Report and Order," *Technical Report FCC 02-48*, April 2002.

[8] R. A. Scholtz, "Multiple access with time-hopping impulsive modulation," *Proc. MILCOM Conf.*, Boston MA, Oct. 1993, pp.447-450.

[9] M. Z. Win and R. A. Scholtz, "Impulse Radio: How It Works," *IEEE Communication Letters*, vol.2, no.2, pp.10C12, Feb 1998.

[10] M. Z. Win and R. A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications," *IEEE Transactions on Communications*, vol.48, no.4, pp.679C691, April 2000.

[11] M. L. Welborn, "System considerations for ultra-wideband wireless networks," *IEEE Radio Wireless Conf.*, Aug. 2001, pp.5C8.

[12] J. R. Foerster, "The performance of a direct-sequence spread ultrawideband system in the presence of multipath, narrowband interference, and multiuser interference," *IEEE Conf. Ultra Wideband Systems Tech.*, May 2002, pp.87C91.

[13] N. Boubaker and K. B. Letaief, "Ultra wideband DSSS for multiple access communications using antipodal signaling," *IEEE Int. Conf. Commun.*, vol.3, May 2003, pp.11C15.

[14] R. W. Chang, "High-speed multichannel data transmission with bandlimited orthogonal signals," *Bell System Technical Journal*, vol. 55, pp.1175-1796, Dec. 1966.

[15] B. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. on Commun.*, vol.15, no.6, pp.805-811, Dec. 1967.

[16] K.-B. Png, X. Peng, and F. Chin, "Performance Studies of a Multi-band OFDM System Using a Simplified LDPC Code," *Joint UWBST & IWUWBS. 2004*, pp.376-380, May 2004.

[17] S. C. Cho, J. U. Kim, K. T. Lee and K. R. Cho; "Convolutional Turbo Coded OFDM/TDD Mobile Communication System for High Speed Multimedia Services," *Proc. Telecommun. - Adv. Industr. Conf. Telecommun./Serv. Assur. Part. Int. Res. Conf./E-Learn. Tel. (Workshop AICT/SAPIR/ELETE 2005),* pp.244-248, Jul. 2005.

[18] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp.628-634, Oct. 1971.

[19] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5-14, May 1991.

[20] A. Batra, et al., "Multi-band OFDM Physical Layer Proposal for IEEE P802.15 Task Group 3a," *IEEE P802.15-03/268r2*, Nov. 2003.

[21] W. Y. Zou, and Y. Wu, "COFDM: An Overview," *IEEE Trans. on Broadcasting*, vol. 41, no. 1, pp.1-8, Mar 1995.

[22] S. Benedetto and G. Montorsi, "Serial concatenation of block and convolutional codes," *Electronics Letters*, vol.32, no.10, pp.887-888, May 1996.

[23] S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Trans. Commun.*, vol.44, no.5, pp.591-600, May 1996.

[24] S. Benedetto and G. Montorsi, "Unveiling turbo-codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, Vol.42, no.2, pp.409-428, Mar. 1996.

[25] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol.42, pp.429-445, Mar. 1996.

[26] P. Robertson and T. Woerz, "A novel bandwidth-efficient coding scheme employing turbo-codes," *Proc. IEEE Int. Conf. Comm.*, ICC'96, pp.962-967.

[27] P. Robertson and T. Worz, "Extensions of turbo trellis-coded modulation to high bandwidth efficiency," *IEEE Int. Conf. on Commun. (ICC' 97)*, vol.3, no.2, pp.1251-1255, Jun. 1997.

[28] P. Robertson and T. Worz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. on select. Areas in Commun.*, vol.16, no.2, pp.206-218, Feb. 1998.

[29] S.LeGoff, A. Glavieux, and C. Berrou, "Turbo codes and high spectral efficiency modulation," *Proceedings of IEEE ICC'94,* May 1-5, 1994, New Orleans, LA.

[30] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Parallel concatenated trellis-coded modulation," *Proceeding of IEEE ICC'96*, pp.974-978, 1996.

[31] C. Fragouli and D. Wesel, "Turbo-encoder design for symbol-interleaved parallel concatenated trellis-coded modulation," *IEEE Trans. on Commun.*, vol.49, no.3, pp.425-435, Mar. 2001.

[32] G. D. Forney, M. D. Trott,and S.-Y. Chung "Approaching AWGN channel capacity with coset codes and multilevel coset codes," *IEEE Trans. Inform. Theory*, Vol. 46, pp.820-850, May 2000.

[33] G. D. Forney,Jr. and L. F. Wei, "Multidimensional Constellation - Part I: Introduction, figure os merit, and generalized cross constellation," *IEEE J. on select. Areas in Commun.*, Vol. 7, pp.877-892, Aug. 1989.

[34] L. U. Wachsmann and J. Huber, "Power and bandwidth efficient digital communications using turbo codes in multilevel codes," *Eur. Trans. on Telecomm*, vol.6, no.5, pp.557-567, Sept./Oct. 1995.

[35] Q. Wang, L.Wei and R.A. Kennedy, "Iterative Viterbi decoding, trellis shaping, and multilevel structure for high-rate parity-check TCM," *IEEE Trans. on Commun.*, vol.50, no.1, pp.48-55, Jan. 2002.

[36] Q. Wang, *Near Optimal Decoding for Trellis-coded Modulaiton.* PhD Disertation, Australian National University, Canberra, Australia, 2000.

[37] Q. Wang and L.Wei, "Graphic-based iterative decoding algorithms for parity-concatenated trellic codes," *IEEE Trans. Infor. Theory*, vol.47, no.3, pp.1062-1074, Mar. 2001.

[38] L. Wei, Z. Li, M. R. James, I. R. Petersen, "A minimax robust decoding algorithm," *IEEE Trans. on Information Theory*, vol.46, no.3, pp. 1158–1167, Mar. 2000.

[39] T. W. Barrett, "History of ultrawideband (UWB) radar and communications: pioneers and innovators," *Proc. Progress in Electromagnetics Symposim 2000 (PIERS2000)*, Cambridge, MA, July 2000.

[40] Robert A. Scholtz, David M. Pozar and Won Namgoong, "Ultra-Wideband Radio," *EURASIP Journal on Applied Signal Processing,*, 2005:3, 252C272.

[41] WiMedia Alliance. [Online]. Available: http://www.wimedia.org

[42] R. Roberts, XtremeSpectrum CFP Document, *IEEE P802.15-03/315r3,* July 2003.

[43] V. S. Somayazulu, J. R. Foerster, and S. Roy, "Design challenges for very high data rate UWB systems," *36th Asilomar Conference on Signal, Systems and Computers*, pp.717-721, Nov. 2002.

[44] K. Mandke, H. Nam, L. Terramneni, C. Zuniga, and T. Rappaport, "The evolution of ultra wide band radio for wireless personal area networks," *High Frequency Electronics*, pp.22-32, Sept 2003.

[45] Stephen Wood, "UWB standards," *WiMedia Alliance White Paper*, June 2006.

[46] IEEE 802.11n report, http://www.ieee802.org/11.

[47] L. Williams, D. Wu, E. Staggs, and A. Yen, "Ultra-wideband radio design for multi-band OFDM 480 Mb/s wireless USB," DesignCon 2005.

[48] W. P. Siriwongpairat, W. Su, M. Olfat, and K. J. R. Liu, "Multiband-OFDM MIMO coding framework for uwb communication systems," *IEEE Trans. on Signal Processing*, vol.54, no.1, pp.214-224, Jan. 2006.

[49] A. Peled and A. Ruiz, "Frequency domaindata transmission using reduced computational complexity algorithms," *Proc. IEEE ICASSP*, pp.964-967, Denver, CO, 1980.

[50] B. Le Flch, M. Alard, and C. Berrou, "Coded orthogonal frequency-division multiplexing," *Proc. IEEE*, vol. 83, no. 6, pp. 982-996, June 1995.

[51] A. Vahlin and N. Holte, "Optimal finite duration pulses for OFDM," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 10-14, Jan. 1996.

[52] ETSI, "Digital audio broadcasting (DAB) to mobile, portable and fixed receivers," Feb. 1995.

[53] T. Keller, et al, "Report on digital audio radio laboratory tests," *Technical report, Electronic Industries Association*, May 1995.

[54] T. de Couasnon, R. Monnier, and J. B. Rault, "OFDM for digital TV broadcasting," *Signal Processing*, vol.39, no.1-2, pp.1-32, Sept. 1994.

[55] M. Moeneclaey and M. van Bladel, "Digital HDTV broadcasting over the CATV distribution system," *Signal Processing: Image Commun.*, vol.5, no.5-6, pp.405-415, Dec. 1993.

[56] P. J. Tourtier, R. Monnier, and P. Lopez, "Multicarrier modem for digital HDTV terrestrial broadcasting," *Signal Processing: Image Commun.*, vol.5, no.5-6, pp.379-403, Dec. 1993.

[57] K. Fazel, "Performance of convolutionally coded CDMA/OFDM in a frequency-time selective fading channel and its near-far resistance," *Proc. ICC*, pp.1438-1442, New Orleans, LA, Nov. 1994.

[58] P. S. Chow, *Bandwidth Optimized Digital Transmission Techniques for Spectrally Shaped Channels with Impulse Noise*. PhD thesis, Stanford University, CA, May 1993.

[59] J. C. Tu, *Theory, Design and Application of Multi-Channel Modulation for Digital Communications*. PhD thesis, Stanford University, CA, June 1991.

[60] G. Young, K. T. Foster, and J. W. Cook, "Broadband multimedia delivery over copper," *Bell System Technical Journal*, vol.13, no.4, pp.78-96, Oct. 1995.

[61] G. Ungerboeck and I. Csajka, "On improving data-link performance by increasing channel alphabet and introducing sequence decoding," *Int. Sym. Inform. Theory,* Ronneby, Sweden, June 1976.

[62] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol.IT-28, pp.55-67, 1982.

[63] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets - part I: Introduction," *IEEE Communications Magazine*, vol.25, no.2, Feb. 1987.

[64] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets - part II: State of the art," *IEEE Communications Magazine*, Vol. 25, no. 2, Feb. 1987.

[65] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory,* vol.13, pp.260-269, Apr. 1967.

[66] G. D. Forney, Jr., *Concatenated Codes*. MIT Press, Cambridge, Mass, 1963.

[67] G. D. Forney, Jr. and G, Undergoeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inform Theory,* vol.IT-44, no.6, pp.2389-2415, Oct. 1998.

[68] D. Divsalar and F. Pollara, "Multiple turbo codes for deep-space communications," Jet Propulsion Laboratory, *Tech. Rep. JPL TDA Progress Report 42-121*, pp. 66- 77, May 1995.

[69] J. Andersen, "Turbo codes extended with outer BCH code," *Electron. Lett.*, vol. 32, no. 22, pp. 2059-2060, Oct. 1996.

[70] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 909-926, May 1998.

[71] O. Takeshita, O. Collins, P. Massey, and D. Costello, "A note on asymmetric turbocodes," *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 69-71, March 1999.

[72] C. Douillard, M. Jezequel, C. Berrou, A. Piccart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 507-511, Sept.-Oct. 1995.

[73] M. Tuchler, R. Koetter, and A. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754-767, May 2002.

[74] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046-1061, July 1999.

[75] Consultative Committee for Space Data Systems, "Recommendations for space data system standards: telemetry channel coding," *CCSDS 101.0-B-4*. Blue Book. May 1999.

[76] The 3rd Generation Partnership Project, 2003, 3GPP TS 25.212, http://www.3gpp.org.

[77] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284-287, Mar. 1974.

[78] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output maximum a posteriori (MAP) module to decode parallel and serial concatenated codes," *JPL TDA Progress Report*, vol. 42-127, Nov. 1996.

[79] J. Hagenauer, P. Robertson, and L. Papke, "Iterative ('Turbo') decoding of systematic convolutional codes with the MAP and SOVA algorithm," *Proc. of 1994 ITG Conference on Source and Channel Coding,* Munich, pp.1-9, Oct. 1994.

[80] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory,* Vol.27, pp.533-547, Sept. 1981.

[81] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation Dept. Elec. Engg., U. Linkŏping, Sweden, Apr. 1996.

[82] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory,* Vol.47, pp.498-519, Feb. 2001.

[83] Jr. G. D. Forney, "Codes on graphs: normal realizations," *IEEE Trans. Inform. Theory,* Vol.47, pp.520-548, Feb. 2001.

[84] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Inc., 1988.

[85] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory,* Vol.47, pp.619-637, Feb. 2001.

[86] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inform. Theory,* Vol.47, pp.657-670, Feb. 2001.

[87] D. Divsalar, S. Dolinar, and F. Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE JSAC,* Vol.19, no.5, pp.891-907, May 2001.

[88] H. El Gamal and A. R. Hammons, Jr., "Analyzing the turbo decoder using the Gaussian approximation," *Proc. Int. Symp. Inform. Theory,* Sorrento, Italy, p.319, June 2000.

[89] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.,* Vol.49, pp.1727-1737, Oct. 2001.

[90] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefix or zero padding OFDM for wireless multicarrier transmission," *IEEE Trans. on Commun.*, vol.50, pp.2136-2148, Dec. 2002.

[91] B. R. Vojcic and R. L. Pickholtz, "Direct sequence code division multiple access for ultra-wide bandwidth impulse radio," *IEEE MILCOM Conf.*, pp.898-902, 2003.

[92] D. Divasalar, H. Jin, and R. J. McEliece, "Coding theorem for 'turbo-like' codes", *Proc. Allerton Conference*, pp.210, 1998.

[93] Y. Wang, L. Yang, and L. Wei, "High-speed turbo-TCM-coded orthogonal frequency-division multiplexing ultra-wideband systems," accepted to *EURASIP Journal on Wireless Communications and Networking (Special Issue on UWB Communication Systems - Technology and Applications)*, Feb 2006.

[94] D. Cassioli, M. Z. Win, and A. R. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE JSAC*, vol.20, no.6, pp.1247-1257, Aug. 2002.

[95] S. Yano, "Investigating the ultra-wideband indoor wireless channel," *Proc. IEEE VTC'02 Spring Conf.*, pp.1200-1204, Oct. 2002.

[96] J. R. Foerster, M. Pendergrass, and A. Molisch, "A channel model for ultrawideband indoor communication," *MERL report TR-2003-73*, Nov. 2003.

[97] M. Z. Win, R. A. Scholtz, and M. A. Barnes, "Ultra-wide bandwidth signal propagation for indoor wireless communications," *Proc. IEEE ICC'97*, pp.56-60, June 1997.

[98] J. Kunisch and J. Pamp, "Measurement results and modeling aspects for UWB radio channel," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies*, pp.19-24, May 2002.

[99] J. Keignart and N. Daniele, "Subnanosecond UWB channel sounding in frequency and temporary domain," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies*, pp.25-30, May 2002.

[100] S. S. Ghassemzadeh, R. Jana, C. W. Rice, W. Turin, and V. Tarokh, " A statistical path loss model for in-home UWB channels," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies*, pp.59-64, May 2002.

[101] Z. Irahhauten, H. Nikookar, and G. J. M. Janssen, "An overview of ultra wide band indoor channel measurements and modeling," *IEEE Microwave and Wireless Components Letters*, vol.14, no.8, pp.386-388, Aug. 2004.

[102] H. Hashemi, "Impulse response modeling of indoor radio propagation channels," *IEEE JSAC*, vol.11, no.7, Sept. 1993.

[103] A. Alvarez, G. Valera, M. Lobeira, R. Torres, and J. L. Garcia, "New channel impulsive response model for UWB indoor system simulations," *IEEE VTC Sring Conf.*, pp.22-25, Apr. 2003.

[104] A. Saleh, R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE JSAC*, vol.5, no.2, pp.128-137, Feb. 1987.

[105] J. Foerster, "Channel modeling sub-committee report final," *IEEE 802.15.3a Tech. Report*, no.02/490r0-SG3a, 2002.

[106] S. S. Ghassemzadeh, R. Jana, C. W. Rice, W. Turin, and V. Tarokh, " Autoregressive modeling of an indoor UWB channel," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies*, pp.71-74, May 2002.

[107] C. Berrou and A. Glavieux, "Near Shannon Limit Error Correcting Coding and Decoding: Turbo-codes," *Proc. IEEE Trans. Commun.*, vol.44, no.10, pp.1261-1271, Oct. 1996.

[108] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi and D. A. Spielman, "Improved low density parity check codes using irregular graphs," *IEEE Trans. Inform. Theory*, vol. 47, pp.585-598, Feb. 2001.

[109] W.C.Y. Lee, *Mobile Communications Design Fundamentals*, John Wiley & Sons Inc., 1993.

[110] J. Boutros et al. "Turbo Code at 0.03 dB from Capacity Limit," *Proc. ISIT*, Lausanne, Switzerland, June-July 2002, pp.56.

[111] S. ten Brink, "Rate one-half code for approaching the Shannon limit by 0.1 dB," *Electron. Lett.* Vol.26, no.15, pp. 1293-1294, July 2000.

[112] D. Divsalar, S. Dolinar, F. Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE JSAC*, Volume: 19 Issue: 5, May 2001 Page(s): 891-907.

[113] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. 1998 Allerton Conf. Communication, Control and Computers*, Allerton, IL, Sept. 1998, pp.201-210.

[114] L. Ping and K. Y. Wu, "Concatenated tree codes: A low complexity, high performance approach," *IEEE Trans. Inform. Theory*, Vol 47, pp.791-800, Feb. 2001.

[115] J. S. Tee and D. P. Taylor, "Multiple parallel concatenated single parity check codes," *Proceedings of IEEE ICC*, Vol 1, pp.60-64, June 2001.

[116] L. Wei, "On bootstrap iterative Viterbi algorithm," *Proceedings of IEEE ICC'99*, Vancourver, Canada, June 6-9, 1999, pp.1187-1192

[117] L. Wei, "High performance iterative Viterbi algorithm for conventional serial concatenated codes," *IEEE Trans. Inform. Theory*, Vol 46, pp.1759-1771, Jul. 2002.

[118] L. Wei, "Near optimum serial concatenation of single parity codes with convolutional codes," *IEE Proceeding on Communications*, Volume 152, Issue 4, August 2005 pp.397 - 403.

[119] G. D. Forney, Jr., "On iterative decoding and the two-way algorithm," *Proc. Intl. Symp. Turbo Codes and Related Topics*, Brest, France, Sept. 1997.

[120] T. A. Summers and S. G. Wilson "SNR mismatch and online estimation in turbo decoding," *IEEE Trans. Commun.*, vol.46, no.4, April, 1998.

[121] M. C. Reed and J. Asenstorfer "A novel variance estimator for turbo-code decoding," *International Conference on Telecommunications*, April, 1997.

[122] X. L. Huang and Nam Phamdo "Turbo decoders which adapt to noise distribution mismatch," *IEEE Communication Letters*, Nov., 1998.

[123] J.D. Parsons and J.G. Gardiner, *Mobile Communications System*, USA Halsted Press.

[124] W.C.Y. Lee, *Mobile Communications Design Fundamentals*, John Wiley & Sons Inc., 1993.

[125] S. Verdú and V. H. Poor, "On Minimax Robustness: A General Approach and Applications," *IEEE Trans. on Information Theory*, vol. 30, no. 2, pp. 328–340, March 1984.

[126] T. C. Chuah, "Decoding of low-density parity-check codes in non-Gaussian channels," *IEE Proc-Commun.*, vol. 152, no. 6, pp. 1086–1097, Dec. 2005.

[127] T. Faber, T. Scholand and P. Jung, "Turbo decoding in impulsive noise environments," *Electronics Letters*, Vol.39, No. 14, July 2003, pp.1069-1071

[128] H. Nakagawa, D. Umehara, S. Denno, and Y. Morihiro, "A Decoding for Low Density Parity Check Codes over Impulsive Noise Channels," *2005 International Symposium on Power Line Communications and Its Applications*, 6-8 Apr. 2005, pp.85 - 89.

[129] R. Hormis, I. Berenguerx, and X. Wang, "A simple baseband transmission scheme for power line channels," *IEEE JSAC* Vol. 24, Iss. 7, Jul. 2006, pp. 1351- 1363.

[130] S.-Y. Chung, *On the construction of some capacity-approaching coding schemes* Ph. D. dissertation, MIT, Cambridge, MA, 2000.