University of Alberta

Robust Signal Detection in Non-Gaussian Noise Using Threshold System and Bistable System

by

Gencheng Guo

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Digital Signals and Image Processing

Department of Electrical and Computer Engineering

To my parents
Wenxu Guo and Yuxiang Guo,

to my in-laws
Bentong Chen (late) and Diyun Xue (late),

to my family
Xiuling, Yijun, and Yijia,

for their endless love throughout years.

Abstract

Signal detection in non-Gaussian noise is fundamental to design signal processing systems like decision making or information extraction. The optimal/near-optimal detector for this problem is the likelihood ratio test (LRT) or generalized LRT (GLRT). However, since the noise is non-Gaussian, sometimes has unknown pdf, the LRT or GLRT suffers high implementation cost, low robustness, and possible poor performance. In this thesis, to deal with these challenges, we investigate two techniques. One is to propose simple and robust detectors using threshold system (TS) and bistable system (BS). The other is to exploit the noise-enhanced effect, to improve performance by adding noise to the observation, for suboptimal detectors.

For the detector design using TS or BS, first, we propose binary TS based detector (TD) under Neyman-Pearson (NP) criterion to detect a known DC signal in known non-Gaussian noise. The optimal TS's, including simple binary TS and composite binary TS, are derived analytically. Secondly, we propose a TD for detecting any known signal in independent non-Gaussian noise whose pdf is unknown but is symmetric and unimodal. The optimality of the proposed TD is proved. It is shown that even without the knowledge of the noise pdf, the proposed TD has close performance to the optimal detector designed with precise noise pdf information. The practical implementation and robustness of the proposed TD are also investigated. Third, we investigate the BS based detector (BD) for watermark extraction. There is no existing efficient and systematic BS design method except exhaustive search. We propose to use the cross-correlation of the watermark signal and the BS output as the criterion. Based on this, we develop a practical BS parameter optimization method, which leads to a BS adaptive to various watermark extraction scenarios. The extraction performance based on the adaptive BD is compared with the white Gaussian noise (WGN) based maximum likelihood (ML) detector and other BDs used in watermark extraction.

For the noise-enhanced effect, we focus on the general binary hypothesis test problem using a binary TD. We adopt the AUC, which refers to the area under receiver operating characteristic (ROC) curve, as the performance measure for its simplicity and robustness. The optimal TS design that maximizes the AUC has been derived. For a given binary TS, the optimal noise pdf that maximizes the AUC is shown to be a delta function. Properties of the derived results and comparisons with other designs are presented.

Acknowledgments

I would like to start by thanking Prof. Mrinal Mandal, my supervisor in these year's study and research at University of Alberta (UofA), for his guidance and invaluable support. No matter how busy he was, he was always there to help. My research works have been significantly benefited from his encouragement, insight, and hard work. I also feel lucky due to his consideration during the years, so that I can enjoy the study without too much pressure, and enjoy my family life as well.

I also would like to express my deepest gratitude to my co-supervisor, Dr. Yindi Jing. She has an uncanny ability to extract a tractable model from a complex problem and dig out an incredibly neat solution. I learned from her not only in research, but also in teaching and presentation. She is knowledgeable, kind, and considerate. I am fortunate to meet such excellent mentor. My thesis would not have been done like this without her help.

I thank Prof. Ben Liang at University of Toronto, and the other members in my defense committee and candidacy committee, Prof. Hong Zhang, Dr. Hai Jiang, Dr. Majid Khabbazian, Prof. Chintha Tellambura, Dr. Sergiy Vorobyov, for taking their valuable time to review my thesis and giving suggestions for my research.

In the long journey to arrive here, many people offered their help generously. Before I came to UofA, I must thank Prof. Qinghu Meng for his guidance and inspiration. Thanks also go to China Scholarship Council for financial support. It was the collaboration with Dr. Adita Saha and Mr. Ming Lu, which leads me to this research topic. Valuable discussion and help come from collaborators, faculty & staff members, classmates, labmates, and friends, who are Dr. Xinwei Yu, Dr. Dangzhi Zhao, Dr. Rongfei Fan, Dr. Dileepan Joseph, Dr. Nilanjan Ray, Dr. Zhengang Chen, Dr. Yu Fu, Dr. Gongpu Wang, Mr. Bo Hu, Dr. Qingfang Dai, Mr. Xin Wang, Ms. Pinder Bains, Dr. Meghna Singh, Mr. Ming Lu, Mr. Guan Pan, Mr. Shugen Li, Dr. Yunfei Wang, Dr. Guiling Wang, Miss. Tingting Liu, Dr. Ming Ying, Miss. Min Liu, Dr. Hengzhen Gao, Mr. Tao Xu, Mr. Cheng Lu, Mr. Chao Feng, Mr. Yanan Fu, Ms. Yue Li, Miss. Xiaohui Gong, Mr. Qiang Guo, Ms. Qian Cao, to list a few. Though I cannot mention all of them, whenever I look back and think of these nice people, deep warmth arises from my heart. I would still mention Chinese Scholarship Receiptor Association at UofA. As the founding chair, I appreciate the good opportunity to practice "it is more blessed to give than to receive".

Finally, my warmest gratitude goes to my beloved parents, my wife, my daughter, and my son, for their endless love over years.

Table of Contents

1 Introduction **1**

1.1 Signal Detection and Binary Hypothesis Testing Problem 1

1.2 BHTP with Gaussian Noise 3

1.3 BHTP with Non-Gaussian Noise 4

1.4 Research Goals and Methodology 6

2 Background, Literature Review, and Summary of Contributions **10**

2.1 Background 10

2.1.1 Threshold System 10

2.1.2 Bistable System 12

2.1.3 Stochastic Resonance, Stochastic Resonance Detection, and Noise-enhanced Detection 14

2.1.3.1 Stochastic Resonance 14

2.1.3.2 Stochastic Resonance Detection and Noise-Enhanced Detection . . 15

2.2 Literature Review 16

2.2.1 Detector Design Using Threshold System 16

2.2.2 Detector Design Using Bistable System 17

2.2.3 Noise-Enhanced Detection 19

2.3 A Note on SR Detection 21

2.3.1 Review on SR Detection 21

2.3.2 Understanding of SR Detection 22

2.3.3 Relationship of SR Detector, Noise-Enhanced Detector, and Detector Design 23

2.4 Thesis Contributions 24

2.5 Summary 26

3 Optimal Design of Threshold Detector for a DC Signal in Non-Gaussian Noise **27**

3.1 Introduction 27

3.2 An Indicator of P_D . . . 29
3.2.1 Calculations of P_D . . . 29
3.2.2 A Detectability Indicator Δp . . . 31
3.3 Optimal TS Design Using Δp . . . 35
3.3.1 Optimal Design of Simple Binary TS . . . 36
3.3.2 Optimal Design of Composite Binary TS . . . 37
3.4 Discussion on Detection with Unknown Noise pdf . . . 38
3.5 Simulation Results . . . 39
3.5.1 Mean-Shift Unimodal Gaussian Mixture Detection . . . 39
3.5.2 Mean-Shift Bimodal Gaussian Mixture Detection . . . 41
3.6 Conclusions . . . 43

4 A Robust Detector of Known Signal in Non-Gaussian Noise Using Threshold Systems 44
4.1 Introduction . . . 44
4.2 Problem Statement and Proposed Detector Structure . . . 46
4.2.1 Problem Statement . . . 46
4.2.2 Proposed TD Structure . . . 46
4.2.3 Optimality of the Proposed TD . . . 47
4.3 Investigation on Detection Probability . . . 48
4.3.1 Detection Probability of the Proposed TD . . . 48
4.3.2 Detection Probability of the MF and the LO Detector . . . 49
4.3.3 Simulation Results . . . 50
4.4 Discussions . . . 53
4.4.1 Robustness . . . 54
4.4.2 Implementation Complexity . . . 57
4.4.3 Region of Validity . . . 58
4.5 Conclusions . . . 61

5 Optimal Design and Noise-Enhanced Effect for Binary Threshold Detector under AUC Measure 67
5.1 Introduction . . . 67
5.2 Problem Statement and AUC Measure . . . 68
5.2.1 Detection Problem and Detector Structure . . . 68
5.2.2 AUC Measure and Design Problems . . . 70
5.3 Optimal Binary TS Dedsign . . . 71
5.4 Optimal Noise-Enhanced Effect . . . 72
5.5 An Example . . . 76

5.6 Conclusions 78

6 An Adaptive Bistable System Based Detector for Watermark Extraction 79

6.1 Introduction 80

6.2 Review of Watermark Embedding and Extraction Using BD 81

6.2.1 Embedding Algorithm 81

6.2.2 Extraction Algorithm 83

6.3 Design Problems of BD in Watermark Extraction 84

6.3.1 Detection Problem in Watermark Extraction and Optimal ML Detector for WGN 84

6.3.2 Investigation on Noise Distribution and Motivation of Using BD 85

6.3.3 Investigation on Mean and Variance of Watermark Bit Bearing Segments and Motivation of Using Adaptive BS 89

6.3.4 BS Design Problems 92

6.4 Proposed BS Design 93

6.4.1 Cross-Correlation as the Design Criterion 93

6.4.2 Observations on the Connection between Cross-Correlation and $(a, b, \Delta t)$ . . 94

6.4.3 Observations on the Connection between Cross-Correlation and Watermark Signal and Noise 95

6.4.4 Adaptive BS 97

6.5 Experimental Results 98

6.6 Conclusions 103

7 Conclusions and Future Work 104

7.1 Conclusions 104

7.2 Future Work 105

References 107

List of Tables

1.1 Comparison of different strategies in detector design. 8

2.1 Comparison of reviewed TDs. 18

6.1 BERs (in %) of the extracted watermark for the four watermarked images attacked by JPEG compression with different qualities. 101

6.2 BERs (in %) of the extracted watermark for the four watermarked images attacked by Gaussian noise with different variances. 102

List of Figures

1.1 Schematic of the matched filter (replica-correlator). 3
1.2 Schematic of the quadrature matched filter. 4
1.3 Schematic of the LO detector for known signal in non-Gaussian noise. 5
1.4 Schematic of the LO detector for unknown sinusoidal signal detection in non-Gaussian noise. 6
1.5 Research methodology: design of TD and BD, and noise-enhanced effect. 9

2.1 Bistable potential. 12
2.2 Illustration of SR. 15
2.3 General model of SR detector. 15
2.4 SR effect via adding noise in SR detector. 22
2.5 SR effect via tuning parameters in BS based detector. 23
2.6 Relationship of SR detection, noise-enhanced detection, and detector design. 24
2.7 Overview of the thesis contributions. 25

3.1 Schematic of the proposed TD. 28
3.2 Contour of P_D with respect to (p_1, p_0). 35
3.3 P_D with respect to τ and P_{FA} for detection of a DC signal $A = 0.1$ in GM noise with $\alpha = 0.9$, $\beta = 5$, and $\sigma = 1$. 40
3.4 ROCs from LO detector, MF and the proposed TD for detection of a DC signal $A = 0.1$ in GM noise with $\alpha = 0.9$, $\beta = 5$, and $\sigma = 1$. 41
3.5 ROCs from LO detector, MF and the proposed TD in the detection of a DC signal $A = 0.1$ in the bimodal GM noise with $\mu = 3$, $\sigma = 1$, and $N = 100$. 42

4.1 Proposed TD structure. 46
4.2 P_D versus ENR of the proposed TD, the MF, and the LO detector in GM noise with $\alpha = 0.3$, $\beta = 5$, and $\sigma = 1$ for $N = 100$ and $P_{FA} = 0.01$. 51
4.3 Comparison between theoretical and simulated ROCs for the proposed TD. GM is with $\alpha = 0.3, \beta = 5, \sigma = 1$; GG is with $\beta = 0.9, \sigma = 1$; and $N = 100$. 52

4.4 Comparison of the ROCs obtained from Monte Carlo simulations in detecting the sinusoid signal (4.13) in GM and GG noise, $N = 100$. 53
4.5 Comparison of the ROCs in detecting the random signal (4.14) in GM and GG noise, $N = 100$. 53
4.6 Comparison of robustness measure. 57
4.7 ROCs under inaccurate estimations of the noise pdf and the signal. 58
4.8 The validity regions for different detection problem with GM noise. 61
4.9 P_D^{TD} versus β for the sinusoidal signal detection in GG noise. 62
4.10 ROCs of the TS with different τ. 63

5.1 Noise-enhanced binary TS based detector. 69
5.2 ROC's obtained from the optimal TS, the non-optimal TS, adding optimal noise $V_{\rm opt}$ under AUC criterion, adding optimal WGN, adding $V_{\rm opt}^{NP}$ under NP criterion for $P_{FA} = 0.1$ and $P_{FA} = 0.5$, respectively. 78

6.1 Schematic of the BD based watermarking scheme. 82
6.2 Embedded watermark "W" and watermarked Lena image. 83
6.3 Histograms of two segments in $X_w[n]$ of the watermarked image. 85
6.4 pdfs of two segments in $X_w[n]$ and Gaussian pdfs. 86
6.5 Histograms and pdfs of two segments in $X_w[n]$ of the JPEG compressed watermarked image with quality of 50. 87
6.6 Histograms and pdfs of two segments in $X_w[n]$ of the watermarked image added Gaussian noise with variance of 0.03. 88
6.7 BERs of the extracted watermark from watermarked Lena image after JPEG compression with quality 30 and 50 using BD with fixed $a, \Delta t$, and different b's. 89
6.8 Histograms of the means and variances of the 256 segments in $X_w[n]$ of the original watermarked Lena image. 90
6.9 Histograms of the means and variances of the 256 segments in $X_w[n]$ of the watermarked Lena image after JPEG compression with quality of 50. 90
6.10 Histograms of the means and variances of the 256 segments in $X_w[n]$ of the watermarked Lena image after adding zero-mean Gaussian noise with variance 0.03. . . . 91
6.11 BERs of the extracted watermark from watermarked Lena image after JPEG compression with quality 50 and after adding Gaussian noise with variance of 0.03 using BD with fixed $a, \Delta t$, and different b's. 92
6.12 Cross-correlation versus b and numerous fixed $(a, \Delta t)$ for WGN and GM noise. . . . 94
6.13 Cross-correlation versus b for different PAM signals with different amplitudes and different noises with different variances. 96

6.14 Schematic of adaptive bistable system. 97
6.15 Host images used in watermark experiments. 99

List of Abbreviations

Acronyms	**Definition**
ARE	asymptotic relative efficiency
AUC	area under the ROC curve
BD	bistable system based detector
BER	bit error rate
BHTP	binary hypothesis testing problem
BS	bistable system
cdf	cumulative distribution function
DCT	discrete cosine transform
ENR	energy-to-noise ratios
GLRT	generalized likelihood ratio test
i.i.d.	independent and identically distributed
LO	locally optimal
LRT	likelihood ratio test
MF	matched filter
ML	maximum likelihood
MLE	maximum likelihood estimation
MSE	mean squared error
NP	Neyman-Pearson
PAM	pulse amplitude modulation
pdf	probability density function
pmf	probability mass function
PSNR	peak signal-to-noise ratio
ROC	receiver operating characteristic
SNR	signal-to-noise ratio
SR	stochastic resonance
TD	threshold system based detector
TS	threshold system

WGN	white Gaussian noise

List of Symbols

$\mathbb{E}$	Mathematical expectation
H_0	Null hypothesis, noise only hypothesis
H_1	Alternative hypothesis, signal and noise hypothesis
$f_{\mathbf{X}}(\mathbf{x}; H_0)$	pdf of $\mathbf{X}$ under hypothesis H_0
$f_{\mathbf{X}}(\mathbf{x}; H_1)$	pdf of $\mathbf{X}$ under hypothesis H_1
$\mathcal{N}(z; \mu, \sigma^2)$	Random variable Z whose pdf is Gaussian with mean μ and variance σ^2
P_D	probability of detection
P_{FA}	probability of false alarm
$\mathcal{O}(h^n)$	truncation error of the order of h^n
$Q(\cdot)$	complementary cdf of standard normal pdf
Q^{-1}	inverse function of $Q(\cdot)$
$\mathbb{R}^N$	N dimensional real vector
$T(\mathbf{x})$	test statistic
$\phi(\mathbf{x})$	critical function (decision function)
η	decision threshold in critical function $\phi(\mathbf{x})$

Chapter 1

Introduction

1.1 Signal Detection and Binary Hypothesis Testing Problem

Signal detection, also called hypothesis testing, is to decide from an observation which event of interest occurs [1]–[3]. One example is the radar system, where the goal is to determine the presence or absence of an approaching aircraft based on the received waveform (observation) of the radar. More applications can be found in communication system, sonar, image processing, control system, to list a few. Detection problems in these applications can be classified into two types: one is to decide between two hypotheses; the other is to decide among more than two hypotheses. Correspondingly, the former is termed as binary hypothesis testing problem (BHTP) and the latter is termed as multiple hypothesis testing problem. Since BHTP is basic and essential, we will focus on BHTP in this thesis.

In general, a BHTP can be formulated as follows:

$$\begin{cases} H_0: & \mathbf{X} = \mathbf{W} \\ H_1: & \mathbf{X} = \mathbf{s} + \mathbf{W} \end{cases}, \tag{1.1}$$

where H_0 and H_1 represent the two hypotheses. $\mathbf{X}, \mathbf{s}, \mathbf{W}$ are N-dimensional real vectors, i.e., $\mathbf{X}, \mathbf{s}, \mathbf{W} \in \mathbb{R}^N$, and sometimes we represent them as $x[n], s[n], w[n], n = 0, 1, ..., N-1$. $\mathbf{X}$ is the observation. $\mathbf{W}$ represents the noise, which in this thesis, is assumed to be independent and identically distributed (i.i.d.). $\mathbf{s}$ represents the discrete-time signal. Hence H_0 refers to the noise only hypothesis (or null hypothesis) and H_1 refers to the signal and noise hypothesis (or alternative hypothesis).

For BHTPs, we wish to use the received data ($\mathbf{x}$) to make a decision (H_0 or H_1) as reliable or costless as possible. The decision making process is called a detector or a test. It is composed of a test statistic $T(\mathbf{x})$ and a decision threshold η, shown as

$$T(\mathbf{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{1.2}$$

The test statistic $T(\mathbf{x})$ is a function of the observation $\mathbf{x}$, which is typically a numerical summary of the observation that reduces the observation data to one scalar value [1]–[3]. It is common to use a critical function (decision function) $\phi(\mathbf{x})$ to completely characterize the detector as follows [3].

$$\phi(\mathbf{x}) = \begin{cases} 1: & T(\mathbf{x}) > \eta \\ \nu: & T(\mathbf{x}) = \eta \\ 0: & T(\mathbf{x}) < \eta \end{cases}, \tag{1.3}$$

where $0 \leq \nu \leq 1$.

Let the probability density functions (pdf)s of $\mathbf{X}$ under H_0 and H_1 be respectively $f_{\mathbf{X}}(\mathbf{x}; H_0)$ and $f_{\mathbf{X}}(\mathbf{x}; H_1)$. The probability of detection (P_D) can be calculated as

$$P_D = \int_{\mathbb{R}^N} \phi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}; H_1) d\mathbf{x}. \tag{1.4}$$

The probability of false alarm (P_{FA}) can be calculated as

$$P_{FA} = \int_{\mathbb{R}^N} \phi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}; H_0) d\mathbf{x}. \tag{1.5}$$

For BHTPs, the goal is to design $T(\mathbf{x})$ and determine η optimally. It is well known that the optimal detector is the likelihood ratio test (LRT) [1]–[3], which is

$$L(\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}; H_1)}{f_{\mathbf{X}}(\mathbf{x}; H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{1.6}$$

$L(\mathbf{x})$ is called the likelihood ratio function, which is the optimal test statistic. The optimal value of η depends on the optimality measure. For example, under Bayesian criterion, to have the minimum Bayesian cost, the optimal η is derived to be

$$\eta = \frac{(C_{10} - C_{00})\mathbb{P}(H_0)}{(C_{01} - C_{11})\mathbb{P}(H_1)}, \tag{1.7}$$

where $C_{ij}, i, j \in \{0, 1\}$ are the costs if we decide H_i when H_j is true and $\mathbb{P}(H_i)$ is the a-priori probability of H_i. Under Neyman-Pearson (NP) criterion, to have the maximum P_D subject to the constraint $P_{FA} \leq \alpha$, the optimal η can be calculated from

$$P_{FA} = \int_{L(\mathbf{x}) \geq \eta} \phi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}; H_0) d\mathbf{x} = \alpha. \tag{1.8}$$

NP criterion is more popular than Bayesian criterion as the costs C_{ij} and a-priori probabilities $\mathbb{P}(H_i)$ required for Bayesian criterion need not to be known.

In the next two sections, we will present some common BHTPs and their conventional detectors.

1.2 BHTP with Gaussian Noise

In this section, we consider BHTPs with i.i.d. Gaussian noise, i.e., the pdf of entries of $\mathbf{W}$, denoted as $f_W(w)$, is Gaussian. Gaussian noise can be observed very often. For example, in wireless and wired communication systems, the channel noise is usually modeled as Gaussian. The reason is that the noises in real world are often the sum of many independent random events. Based on the central limit theorem, they follow Gaussian distribution [4]. For Gaussian noise, the optimal detector is easy to design and implement. Two BHTPs with Gaussian noise and their conventionally used detectors are given below.

BHTP I Simple Gaussian-based BHTP, where the known signal $s[n]$ is in white Gaussian noise (WGN) $w[n]$ with variance σ^2. WGN is defined as zero mean Gaussian process with auto-correlation function $r_{ww}(k) = \mathbb{E}(w[n]w[n+k]) = \sigma^2\delta(k)$, where $\delta(k)$ is the discrete delta function. In another words, $f_W(w) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{w^2}{2\sigma^2}\right) \triangleq \mathcal{N}(w; 0, \sigma^2)$. Hence we have $f_{\mathbf{X}}(\mathbf{x}; H_0) = \prod_{n=0}^{N-1} f_W(x[n])$ and $f_{\mathbf{X}}(\mathbf{x}; H_1) = \prod_{n=0}^{N-1} f_W(x[n] - s[n])$. $L(\mathbf{x})$ can be calculated using (1.6), based on which the optimal test is derived as [1]

$$T(\mathbf{x}) = \sum_{n=0}^{N-1}(x[n]s[n]) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{1.9}$$

The optimal test statistic $T(\mathbf{x})$ is a linear function of $\mathbf{x}$. It is also linear in $\mathbf{s}$. This detector is typically termed as the matched filter (MF) or replica-correlator [1]. The schematic of this test is illustrated in Fig. 1.1.

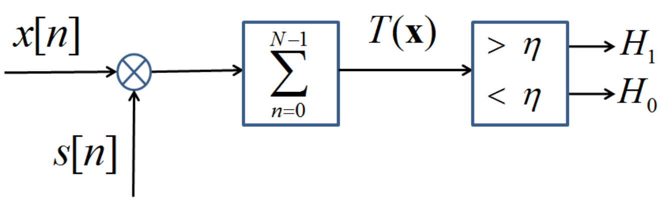


Fig. 1.1. Schematic of the matched filter (replica-correlator).

BHTP II Composite Gaussian-based BHTP, where signal is in WGN with some unknown parameters in the signal and/or the noise. In this case, $f_W(w), f_{\mathbf{X}}(\mathbf{x}; H_i), i = 0, 1$ are identical to the ones in BHTP I, but there are some parameters in $s[n]$ and $f_W(w)$ that are unknown. This composite BHTP is more realistic than the simple one. The generalized LRT (GLRT) is commonly used for the composite BHTP [1]. It first calculates the maximum likelihood (ML) estimations of the unknown parameters from the observations [5] and then use the estimated parameters to design the LRT. GLRT cannot be proved to be optimal but it works well in many applications [1], [2].

We now present a particular composite BHTP, which is the detection of a sinusoidal signal with known frequency f_0 but unknown amplitude and phase in WGN. This detection problem is of great interest in passive sonar system and radar system [1], [6]. The GLRT for this detection problem is a quadrature MF [1] and is shown as follows.

$$T(\mathbf{x}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]\exp(-j2\pi f_0 n)\right|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{1.10}$$

Its schematic is shown in Fig. 1.2.

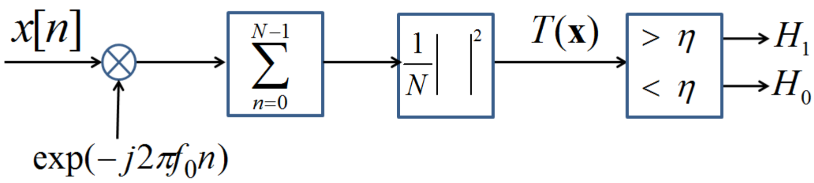


Fig. 1.2. Schematic of the quadrature matched filter.

From Figs. 1.1 and 1.2, we can see that the detectors for BHTPs with WGN do not depend on the parameters of the WGN. Also, the test statistics are linear or quadratic in the observations. The detectors are thus simple, easy to implement, and prone to be robust due to their independence of the noise pdf.

1.3 BHTP with Non-Gaussian Noise

In this section, we consider BHTPs with non-Gaussian noise. Non-Gaussian noises are revealed in many applications as well. For example, a type of Gaussian mixture pdf has been used to model the ocean acoustic noise [7], [8]. In nature, infrequent but powerful events, such as thunderstorms, iceberg breakup, tsunami, cause the “noise spikes", which also leads to non-Gaussian noise [1]. Two BHTPs with non-Gaussian noise and the corresponding detectors are given below.

BHTP III Simple non-Gaussian based BHTP, where the known signal $s[n]$ is in white non-Gaussian noise with known pdf $f_W(w)$. Since $f_{\mathbf{X}}(\mathbf{x}; H_0) = \prod_{n=0}^{N-1} f_W(x[n])$ and $f_{\mathbf{X}}(\mathbf{x}; H_1) = \prod_{n=0}^{N-1} f_W(x[n] - s[n])$ are known completely, using (1.6), the optimal detector can be obtained as [1]

$$T_O(\mathbf{x}) = \sum_{n=0}^{N-1} g_n(x[n]) \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{1.11}$$

where $g_n(x[n]) = \ln \frac{f_W(x[n]-s[n])}{f_W(x[n])}$.

In general, the $T_O(\mathbf{x})$ in (1.11) is nonlinear in $\mathbf{s}$ and in $\mathbf{x}$, which complicates the design and implementation of the detector. To simplify this detector, a suboptimal test statistic $T_L(\mathbf{x})$ that is linear in $\mathbf{s}$ with the following structure has been proposed [1]:

$$T_L(\mathbf{x}) = \sum_{n=0}^{N-1} g(x[n])s[n], \tag{1.12}$$

where $g(x)$ is generally nonlinear when the noise is non-Gaussian. Here we can observe that the design of an appropriate $g(x)$ is crucial to the detectability, complexity, and robustness of this detector. One possible $g(x)$ is obtained by calculating the first order Taylor expansion of $g_n(x)$ about the signal $s[n]$, which is

$$g_{LO}(x) = -\frac{1}{f_W(x)} \frac{df_W(x)}{dx}. \tag{1.13}$$

This $g_{LO}(x)$ design leads to the locally optimal (LO) detector [1], shown as

$$T_{LO}(\mathbf{x}) = \sum_{n=0}^{N-1} g_{LO}(x[n])s[n] \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{1.14}$$

which is illustrated in Fig. 1.3. When the signal is weak compared with the noise level, i.e., $|s[n]| \ll \sigma$, where σ is the standard deviation of the noise, the LO detector is expected to perform close to optimal. Hence, this LO detector is widely used in signal detection in non-Gaussian noise.

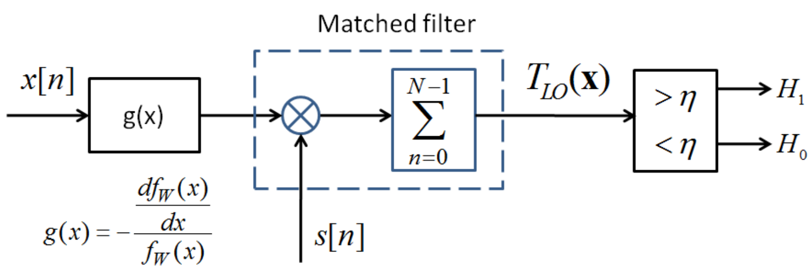


Fig. 1.3. Schematic of the LO detector for known signal in non-Gaussian noise.

BHTP IV Composite non-Gaussian based BHTP, where signal is in non-Gaussian noise with some unknown parameters in the signal and/or the noise. That is, we know the structure of the signal and the pdf form of the non-Gaussian noise. Therefore, $f_{\mathbf{X}}(x; H_i), i = 0, 1$ are identical to the ones in simple non-Gaussian BHTP, but some parameters in them are unknown.

As a special example of practical interests, we consider the detection of the sinusoidal signal with known frequency but unknown amplitude and phase in known non-Gaussian noise [6]. The GLRT is shown in Fig. 1.4, which is composed of a nonlinear component and a quadrature MF [1]. Compared with the schematic in Fig. 1.2, if the noise pdf changes from Gaussian to non-Gaussian, this detector is added with a nonlinear component $g(x)$ before the quadrature MF.

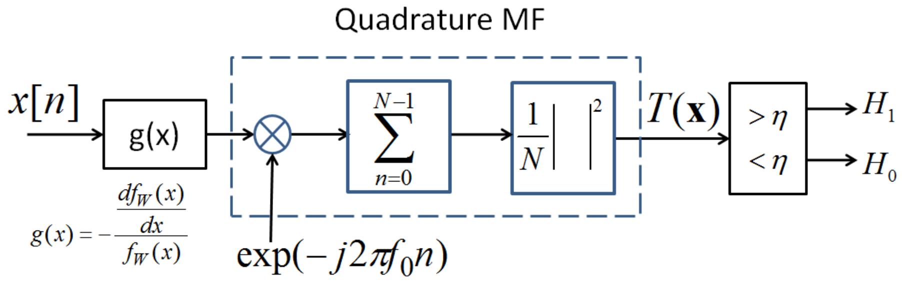


Fig. 1.4. Schematic of the LO detector for unknown sinusoidal signal detection in non-Gaussian noise.

For non-Gaussian noise, the detectors depend on the knowledge of the noise pdf via the component $g(x)$, and the detector design is more complicated compared with the one for Gaussian noise.

1.4 Research Goals and Methodology

This thesis focuses on signal detection in non-Gaussian noise. In this section, we present the challenges in this field, the goals of research, and the methodology to achieve the goals.

Conventionally, the following three strategies are used in signal detection in non-Gaussian noise, with known or unknown pdf.

Strategy I The noise is viewed as Gaussian, and the MF, which is optimal for Gaussian noise, is used to detect the signal. This is equivalent to using (1.12) where $g(x) = x$. This detector design, referred to as the MF, is unrelated to the noise pdf, but its performance is generally poor for non-Gaussian noise [1], [9].

Strategy II LRT and GLRT are employed. If signal and noise are known, LRT is employed and is the optimal test, as shown in the simple non-Gaussian based BHTP. More practical case is the composite non-Gaussian based BHTP in Sec. 1.3, where unknown parameters exist in the given $f_{\mathbf{X}}(\mathbf{x}; H_i), i = 0, 1$. For this composite BHTP, GLRT can be used [1], [2], [10]. Even though LRT or GLRT is commonly used to achieve optimal or near-optimal detection performance, for many applications, LRT or GLRT is unapplicable or impractical due to the following reasons.

1. As shown in (1.11), LRT needs to know $f_{\mathbf{X}}(\mathbf{x}; H_i)$,$i = 0, 1$, i.e., the full knowledge of the signal and the noise. In many real applications, there is no knowledge or only imprecise knowledge on the noise pdf and the signal. Thus LRT or GLRT is unavailable for these cases. If the form of the noise pdf and the signal is known but with unknown parameters, GLRT can be employed using estimations of the unknown parameters. However, accurate estimation of the noise pdf and the signal is difficult because the noise distribution may vary with time and the signal fades in traveling. Consequently, GLRT suffers the imprecise noise pdf and signal information, and thus risks poor performance.

2. LRT and GLRT are in general complex in implementation for detection problems with non-Gaussian noise. The test statistics of LRT or GLRT $T_O(\mathbf{x})$ in (1.11), $T_L(\mathbf{x})$ in (1.12), and $T_{LO}(\mathbf{x})$ in (1.14), are nonlinear in the observation and have high complexity in implementation. This further leads to the cost and delay issues in detector design. Many times, we would like to sacrifice a certain level of performance and seek for simple and cheap systems instead of the complicated optimal/local-optimal ones.

3. LRT and GLRT have low robustness. For applications where the noise keeps changing with time, the performance obtained from LRT and GLRT can degrade significantly because their performance is sensitive to the noise pdf and the estimated parameters.

Strategy III Specific nonlinear system based detector is employed. Comparing the LO detectors designed for non-Gaussian noise in Sec. 1.3 with the detectors for Gaussian noise in Sec. 1.2, we see that the LO detector has an extra nonlinear function $g(x)$ as shown in Figs. 1.3 and 1.4. Therefore, $g(x)$ is crucial for signal detection in non-Gaussian noise. Problems arise in LRT and GLRT because they calculate $g(x)$ from noise pdf that could be unknown or unprecise. In Strategy III, $g(x)$ is specified as a certain nonlinear function [6], [11], whose parameters can be optimized based on the knowledge of the signal and partial knowledge of the noise, if available. Compared with Strategy II, this strategy is expected to be more robust and less sensitive to errors in the noise pdf form and/or parameters. But when the noise pdf information is available, it is expected to perform worse than Strategy II, and hence it is more desirable for systems with unknown or constantly changing noise pdf. Another advantage of it over Strategy II is its complexity. For Strategy II, the complexity depends on the noise pdf, and non-Gaussian noises usually lead to highly complex detector structure. With Strategy III, we can control the implementation complexity via the design of the nonlinear function $g(x)$ and achieve the desired balance between complexity and performance. Compared with Strategy I, which uses a detector structure optimal to Gaussian noise only, Strategy III can achieve better performance for problems with non-Gaussian noise.

A summary of the comparison between the above three strategies is presented in Tab. 1.1, which

TABLE 1.1
COMPARISON OF DIFFERENT STRATEGIES IN DETECTOR DESIGN.

Character	I-MF	II-GLRT	III-specific $g(x)$
Performance	Generally poor	Excellent or poor, depending on accuracy level of estimated noise pdf	Good
Complexity	Low	High	Low
Robustness	High	Low	Could be high

shows that Strategy III has potential in detection problems with non-Gaussian noise. In this thesis, we choose some specific systems, including threshold system (TS) and bistable system (BS) as the $g(x)$ to obtain a simple and robust detector with a comparable performance to the optimal or near-optimal detectors. The detector based on TS is abbreviated as TD and the detector based on BS is abbreviated as BD. An introduction on TS and BS and the reasons why we choose TS and BS are presented in Sec. 2.1. A literature review on TD and BD will be given in Sec. 2.2.

Up to this point, we have introduced the first method in this thesis, which is to use TD or BD for BHTPs with non-Gaussian noise. However, when TD or BD is employed, there are two types of non-optimality. One non-optimality is induced by TS or BS because they are not the optimal nonlinearity in general. The other is that the TS or BS may not or, in some cases, cannot be designed optimally. In some circumstances, the detector already exists and cannot be adjusted. For example, in the human sensor systems like ears and eyes, the neurons are not optimal as far as the perception ability is concerned, but cannot be adjusted easily. In this case, we cannot conventionally design the detector, but need to consider other methods to improve the detectability.

If the optimal detector is unavailable or the detector design is difficult, an alternative method is to adjust the input. We can add an additional noise to the original input in the hope of improving the performance. This method can be more convenient than optimal detector design in some applications. For example, in some broadcast communication systems, it is easier to add an additional noise at the transmitter side than to adjust the distributed receiver at the receiver side.

Because noises are seen as destructive in general, techniques have been developed to filter noises. In some systems, however, the noise can play a constructive role. This nonintuitive physical phenomenon observed in some nonlinear systems that adding noise can improve the system performance was termed as stochastic resonance (SR) [12], [13]. In the context of signal detection, it has been shown that injecting additional noise to the input can improve the detectablity for some nonlinear detectors [14]–[20]. This effect was termed as “SR effect” or “noise enhanced effect” interchangeably used in literature [14]–[20]. The corresponding detector/detection are called “SR detector/detection” or “noise-enhanced detector/detection”. We use “noise enhanced effect” and “noise enhanced detector/detection” in this thesis because they are more appropriate. A clarification on the differences between “SR detection” and “noise enhanced detection” will be presented in Sec. 2.1.3.

In this thesis, we use the noise-enhanced effect as the second method, adding independent noise to the original observation to improve the performance. This effect is investigated in a binary TD for a general BHTP. A literature review on noise-enhanced detection will be given in Sec. 2.2.3.

To summarize, for signal detection in non-Gaussian noise, the challenges arise mainly from the cost/complexity consideration, the robustness to the ever-changing signal and noise parameters, and constraints in the detector adjustment. In this thesis, we aim at designing simple and robust detectors that still enjoy a detection performance comparable to LRT or GLRT. For this purpose, the methodology, including two techniques: TD or BD design (using TS or BS as $g(x)$), noise-enhanced effect (adding noise to the input), is demonstrated in Fig. 1.5.

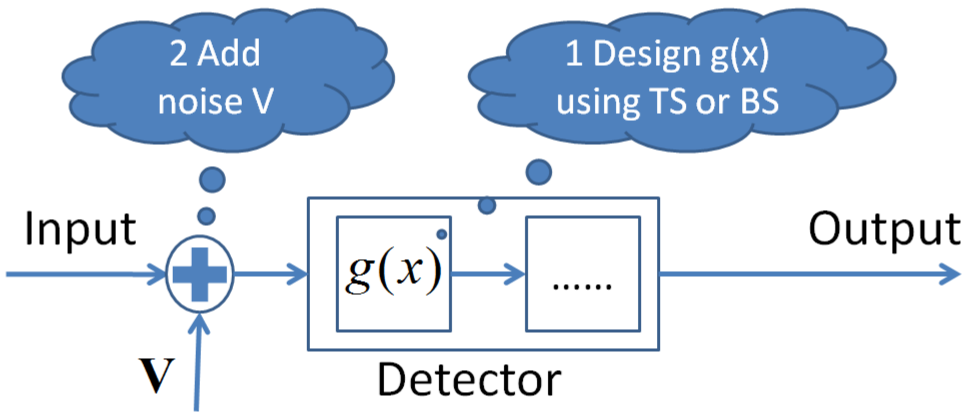


Fig. 1.5. Research methodology: design of TD and BD, and noise-enhanced effect.

Chapter 2

Background, Literature Review, and Summary of Contributions

In this chapter, we provide the background and the literature review of TD, BD, and the noise-enhanced detection in non-Gaussian noise. Then we summarize the major contributions of this thesis. This chapter is organized as follows. First, the background is introduced. The TS is briefly presented in Sec. 2.1.1. Followed is the introduction of the BS in Sec. 2.1.2. In Sec. 2.1.3, we explain the basic ideas of stochastic resonance (SR), SR based detection, and noise-enhanced detection. Second, we present the literature review. The TD design is reviewed in Sec. 2.2.1. Followed is the review of BD design in Sec. 2.2.2. We then review the noise-enhanced detection in Sec. 2.2.3. Third, in Sec. 2.3, to clarify the concept of the SR detection and elude the possible influence to this thesis due to the misuse of SR detection, we explain our understanding of the SR detection and clarify its differences to noise-enhanced detection and detector design. Fourth, the contributions of this thesis are summarized in Sec. 2.4. Finally, we conclude this chapter with a summary.

2.1 Background

2.1.1 Threshold System

In general, threshold system (TS) can be considered as a quantizer. It converts a continuous input to one of multiple discrete values.

TS is one class of system that is ubiquitous in nature and in man-made systems. For example, human sensory systems can be modeled as a TS because the stimulus becomes detectable only when the energy exceed a threshold [13]. Another TS example is the decision making in detectors, which can be find in any test in Secs. 1.2 and 1.3. In addition, TS is useful for detector design due to its clipping behavior. In many applications, the non-Gaussian noise commonly exhibits "spikes", which reveals heavy pdf tails. To reduce these "spikes", good detectors typically include nonlinearities like clipper [1]. Failure to do so leads to poor detectability. For some special cases, for example, if the noise is Laplacian, TS is the optimal nonlinear function for the detection. Finally, TS is simple in implementation, which is one major concern in our design.

We present the typically used TS's in detection in the following. We assume one-dimensional input $x \in \mathbb{R}$ and one-dimensional output $y \in \mathbb{R}$ when formulating these TS's. But these TS's can be extended to multiply-dimensional ones straightforwardly when needed.

1. Simple binary TS. It is defined as,

$$y = \begin{cases} 1 & x \geq \tau \\ 0 & x < \tau \end{cases}, \tag{2.1}$$

   where τ is the threshold of the TS. Here, "binary" refers to the two possible outputs 0 or 1; "simple" refers to the fact that the TS has one threshold τ. In other words, the input space is divided by τ into two continuous intervals corresponding to the two possible outputs.

2. Composite binary TS. It is defined as,

$$y = \begin{cases} 1 & x \in \mathcal{D} \\ 0 & x \notin \mathcal{D} \end{cases}, \tag{2.2}$$

   where $\mathcal{D}$ is a subset of $\mathbb{R}$. This TS is named in contrast to simple binary TS. It still outputs binary values 0 or 1 but there are more than one thresholds. That is, the input space may be divided by multiple thresholds, $\tau_1, ..., \tau_n, ...$, into multiple continuous intervals, where for the x in each interval, the TS outputs one of the two binary values. For example, $\mathcal{D} = \{x| - \infty < x \leq -1, 1 \leq x < \infty\}$. In this TS, there are two thresholds $\tau_1 = -1$ and $\tau_2 = 1$. The TS is equivalent to

$$y = \begin{cases} 1 & x \in (-\infty, -1] \cup [1, \infty) \\ 0 & otherwise \end{cases}.$$

   Note that the simple binary TS is one special case of the composite binary TS.

3. Three-level TS. A three-level TS has 3 possible output values, represented as

$$y = \begin{cases} -1 & x \leq -\tau \\ 0 & -\tau < x < \tau \\ 1 & x \geq \tau \end{cases}, \tag{2.3}$$

   where $\tau > 0$ is the threshold. This three-level TS has been used in Saha's detector in [6].

4. Multi-level quantizer. It is a staircase function with multi-level discrete outputs corresponding to multiple input regions. Conventional quantizer is a monotonically increasing function of the input. However, to be more general, this constraint does not apply to the TS mentioned here. The TS's listed in 1, 2 and 3 are all special cases of the multi-level quantizer.

2.1.2 Bistable System

A typical bistable system (BS) describes the overdamped motion of a ball that is in a bistable potential [21]. The BS is chosen as the nonlinear system in designing the detector because it is a good candidate as a "clipper" or "limiter", which is the important character in dealing with the "spikes" in non-Gaussian noise. In addition, the BS has been widely used in exploiting the noise enhanced effect in signal detection.

The BS details are presented below. The speed of the ball along y is governed by

$$\dot{y}(t) = -U'(y) + x(t), \tag{2.4}$$

where $U(y)$ denotes the quartic bistable potential

$$U(y) = -\left(\frac{a}{2}\right) y^2 + \left(\frac{b}{4}\right) y^4. \tag{2.5}$$

The bistable potential is shown in Fig. 2.1. Two parameters (a, b) decide the size of the bistable

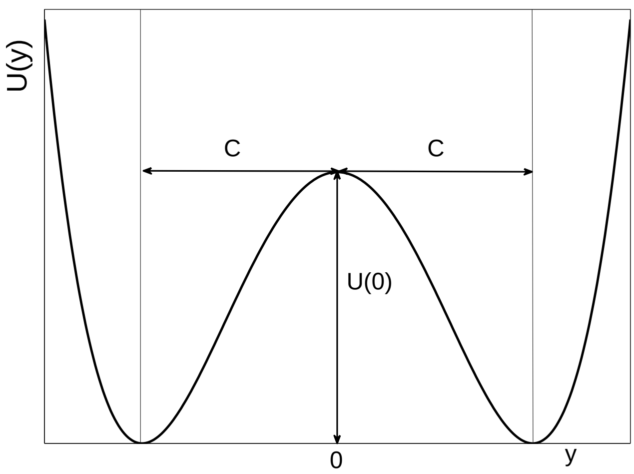


Fig. 2.1. Bistable potential.

potential, as shown in (2.5). The barrier height is $|U_0| = a^2/4b$. The potential minima are located at $y = \pm c = \pm\sqrt{a/b}$.

The input to the BS is $x(t)$, and the output is $y(t)$, the position of the ball in the bistable potential.

The motion of the ball in the bistable potential can be summarized as follows. If $x(t)$ is absent, because the ball is overdamped, the ball will go down slowly to one of the two equilibrium points at $y = \pm c$. If $x(t)$ is present, the ball moves with the speed given by (2.4), and may hop between the two wells.

When this BS is used in signal detection, the input is $x(t)$, which is $s(t) + w(t)$ (a signal $s(t)$ embedded in noise $w(t)$) for H_1 or only $w(t)$ (noise) for H_0. For both cases, the BS output $y(t)$ is a random process that is not wide sense stationary. Therefore, solving $y(t)$ from (2.4) is impossible in general. We can only obtain $y[n]$ for the input $x[n], n \in [0, N-1]$ using numerical method.

According to (2.4), the discrete time simulation model of the BS can be obtained. Given a starting position of the ball $y[0]$ and discrete time input $x[n]$, the discrete version of (2.4) can be obtained using Euler's method [22]:

$$y[n+1] = y[n] + \Delta t(ay[n] - by^3[n] + x[n]), \tag{2.6}$$

where Δt is the time interval (between $y[n]$ and $y[n+1]$) during which one sample applies to the system. If the parameters (a, b) and Δt are known and $y[0]$ is given, $y[n]$ can be calculated recursively using (2.6). It is worthy mentioning that the global truncation error caused by Euler's method is proportional to Δt, denoted as $\mathcal{O}(\Delta t)$. To improve the accuracy, we use a variation of the fourth-order Runge-Kutta's method [22], [23] in this thesis, which is

$$y[n+1] = y[n] + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \tag{2.7}$$

where

$$\begin{aligned}
k_1 &= \Delta t(ay[n] - by^3[n] + x[n]), \\
k_2 &= \Delta t\left(a\left(y[n] + \frac{k_1}{2}\right) - b\left(y[n] + \frac{k_1}{2}\right)^3 + x[n+1]\right), \\
k_3 &= \Delta t\left(a\left(y[n] + \frac{k_2}{2}\right) - b\left(y[n] + \frac{k_2}{2}\right)^3 + x[n+1]\right), \\
k_4 &= \Delta t\left(a\left(y[n] + \frac{k_3}{2}\right) - b\left(y[n] + \frac{k_3}{2}\right)^3 + x[n+2]\right).
\end{aligned}$$

This method can reduce the global error to $\mathcal{O}(\Delta t^4)$ [22]. Another consideration is the stability of the numerical solution of (2.4). The numerical solution is said to be unstable if the error grows exponentially since there is a bounded solution for (2.4). No explicit condition on the stability can be derived. However, reducing Δt is always helpful to have a stable solution.

2.1.3 Stochastic Resonance, Stochastic Resonance Detection, and Noise-enhanced Detection

As stated in Sec. 1.4, the second method of our research is the noise-enhanced detection. In this part, we focus on the basic idea behind the noise-enhanced detection. First, we present the conception of stochastic resonance (SR) because it invokes research in SR detection and noise-enhanced detection. Secondly, we present the concepts on SR detection. It is used with noise-enhanced detection interchangeably in literature. But the two have significant difference. At the same time, noise-enhanced signal detection is introduced, and its difference to SR detection is clarified. Here we only present the basics of these concepts. A more detailed note on SR detection will be given in Sec. 2.3.

2.1.3.1 Stochastic Resonance

Stochastic Resonance (SR) was originally proposed by Benzi, et al. to model the periodic recurrences of the earth's ice ages [24]–[26]. To explain the switching of the earth's climate between ice ages and periods of relative warmth with a periodic of about 100,000 years, the earth's orbit is assumed as the cause because it varies with this period. However, the variation is not strong enough to cause such a significant climate change. Therefore, they proposed a bistable "climatic potential" and the climate shall locate at one of the two stable states: "ice age" and "warm age". Only the earth's orbit variation cannot cause the climate jump from one stable state to another. However, with the help of other random fluctuations, strong climate changes may happen. There is a cooperative phenomenon between the weak periodic variation (the "signal") and the random fluctuations (the "noise"). The output (the strong climate change) is "resonate" with the "signal" with the help of the "noise", and hence this phenomenon is termed as "stochastic resonance".

In general, SR can be illustrated using Fig. 2.2. When a signal is applied to a nonlinear system, if the nonlinear system is in a form of threshold and the signal is subthreshold [1], all input information is blocked and the output has no information related to the input signal. In this case, by injecting additional noise to the subthreshold signal, the output can gain some information on the signal. The system gain, which can be signal-to-noise ratio (SNR), mutual information, or cross-correlation, reveals an increase with the increase of the additional noise intensity until to a level resulting in the maximum system gain. After that level, further increase of the noise intensity leads to decrease of the system gain. The non-monotonic relationship between system gain and noise intensity is considered as the signature of SR effect.

Fig. 2.2 shows that the SR effect requires three basic ingredients [21]: (1) a nonlinear system (in a form of threshold); (2) a subthreshold (weak) input; (3) a source of noise. Because the signal is subthreshold, not surprisingly, adding noise may provide a possible way to enhance the signal at the

[1] A subthreshold signal is a signal that is not strong enough to over-pass the threshold.

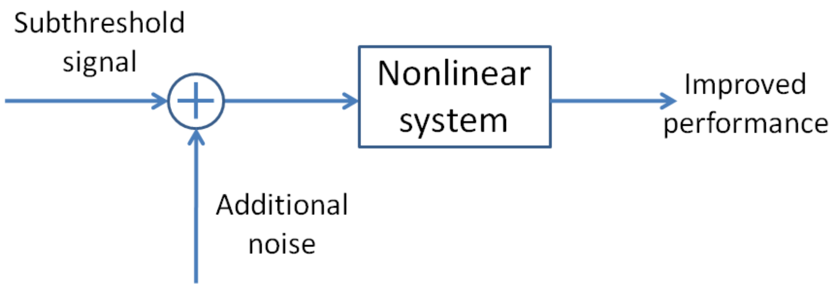


Fig. 2.2. Illustration of SR.

output side. SR mechanism can be used in detectors with TS or BS. SR has been studied intensively in 1980's and 1990's. Some review papers can be found in [12], [13], [21], [27].

2.1.3.2 Stochastic Resonance Detection and Noise-Enhanced Detection

Since SR effect can amplify weak signal in noise through a nonlinear system as shown in Fig. 2.2, it can bring potential improvement for signal detection in non-Gaussian noise. A typical SR detector is given in Fig. 2.3, which is composed of a nonlinear system and a followed detector [28]. The output of the nonlinear system is not a detection decision. Therefore, for the sake of detection, a detector uses the output of the nonlinear system as the input and make a decision. We term it as "inner detector" in this thesis because it is the detector inside the SR detector.

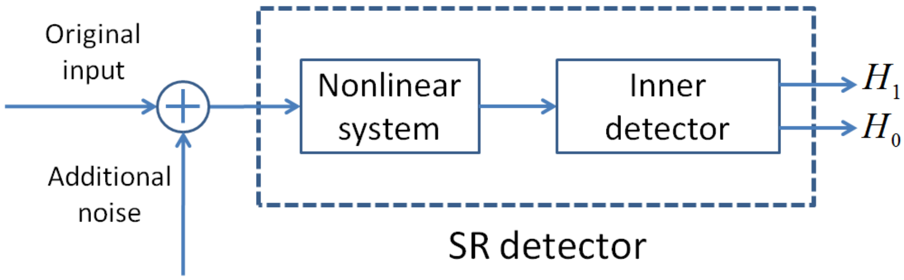


Fig. 2.3. General model of SR detector.

SR detector is termed because the SR effect occurs in the nonlinear system in the detector. However, the improvement in system gain, such as SNR, through the nonlinear system, does not guarantee an improvement in detection performance [29], [30]. In addition, the SR detection put more efforts on the SR effect than the detection performance. Hence, to have the SR effect occur, it suffers many constraints and is difficult to have a competitive detectability. In this thesis, we do not address much on SR detection. Instead, we focus on the noise-enhanced detection to achieve optimal detectability.

Unlike the SR detection, the noise-enhanced detection focuses on the optimal detectability via adding noise to the original input. Comparing with SR detection, in noise-enhanced detection, the major difference is that we do not need to consider the constraints in SR detection. For example, in SR detection, signal should be subthreshold, SR effect should occur. Noise-enhanced detection can occur for any signal (subthreshold or superthreshold), and does not require SR effect. Without these constraints, noise-enhanced detection provide more efficient and practical solution than SR detection in robust signal detection.

2.2 Literature Review

In this part, we review the literature on TD design, BD design, and noise-enhanced signal detection.

2.2.1 Detector Design Using Threshold System

As we stated in Sec. 2.1.1, TS is a promising candidate for the nonlinear component $g(x)$ in a detector designed for non-Gaussian noise. The objective of using TS is to achieve a detector with the following desired features: simple implementation, high robustness, good detectability, or a balance among these features. In this section, we will review the TDs proposed for different detection problems. At the end of this section, we represent a table to summarize the features of these TDs.

Thomas derived a nonparametric detector in [31], which is composed of a sign function and a replica-correlator. The sign function is one simple binary TS, described in (2.1). This TD offers advantages in implementation simplicity and robustness to imprecision in signal information. But its performance is poor compared with the optimal detector.

In [32], an optimum multi-level quantizer or TS was proposed for $g(x)$. As introduced in (1.13), when the noise pdf is known, $g_{LO}(\cdot)$ is the locally optimal nonlinear function. The optimum quantizer in [32] is designed to minimize the mean squared error (MSE) between the quantized output and the output of the $g_{LO}(\cdot)$ for a given number of the quantization levels. The detector based on this quantizer can achieve superior detectability to the detector based on the usual minimum distortion quantizer when the noise is non-Gaussian.

Miller et al. [33] investigated the detectability and robustness of the detectors using sign function, amplifier limiter, six-level and four-level TS as the $g(x)$. The detectability is measured by asymptotic relative efficiency (ARE) and the robustness is measured by the degradation of ARE. The parameters of these nonlinear systems are calculated to maximize the ARE.

In [14], Jung addressed the SNR behavior when inputting a known sinusoidal signal in Gaussian color noise with known pdf to the simple binary TS. The expression of the output SNR shows that there is an optimal threshold of the TS, which yields to the maximum SNR. The SNR behavior of the simple binary TS was also investigated in [15] when the input is a periodic train of pulse embedded in arbitrarily distributed white noise. By using the output SNR as the performance measure, it showed that there is an optimal threshold that leads to maximum SNR, and adding noise can improve the output SNR if the input signal is subthreshold to the given TS. However, as explained in [29], [30], high SNR at the TS output does not guarantee high detectability.

Saha et al. in [6] proposed to use a three-level TS as shown in (2.3) followed by a quadratic MF to detect sinusoidal signals with known frequency but unknown amplitude and phase in non-Gaussian noise with known pdf. The expression of the SNR at the output of the three-level TS was derived and used to calculate the optimal parameter of the three-level TS. It was shown that the proposed detector has superior detectabitlity to the quadratic MF for the Gaussian mixture noise and general Gaussian noise.

Chapeau-Blondeau [11] proposed a simple binary TS based maximum a-posteriori probability detector to detect DC signals in non-Gaussian noise. The detector was shown to have better performance than the linear MF.

A brief summary of the features of the TDs reviewed above is provided as Tab. 2.1. It shows that these TDs are usually designed in certain senses of optimality. But for all these works, the pdfs of the non-Gaussian noises are assumed to be known perfectly. This assumption leads to impractical issue because the noise pdf usually cannot be obtained perfectly. Furthermore, the design based on this assumption can result in complex implementation, low robustness, and poor detectability when applying the TDs. Even though the TS has simple implementation, the optimal parameters of the TS need to be calculated from the noise pdf, which adds complexity to the TD. Also, because the TD design depends on the noise pdf and the signal, the robustness of these detectors to the change of the noise and signal is expected to be low. When the noise pdf information is imprecise, the detection performance may degrade severely. These weakness limits the applications of these TDs. We will address these problems in this thesis.

2.2.2 Detector Design Using Bistable System

BD uses BS to replace the nonlinear component $g(x)$ in a detector designed for non-Gaussian noise (see Figs. 1.3 and 1.4). In other words, a BD is composed of a BS and a conventional detector, such as the linear MF in Fig. 1.3 or the quadrature MF in Fig. 1.4. Therefore, the BD design is to find the optimal or near optimal parameters of the BS to achieve the optimality in a certain sense.

BS has been used in exploiting the SR effect in signal detection [23], [29], [30], [34]. For this purpose, the BS design is subject to the constraints applied to SR. However, it has been shown that better performance can be obtained without these constraints [35]. Hence, we will consider the BS design for high detectability without considering any constraints related to SR detection.

Since analytical optimization of the BS parameters is difficult, the BS parameters are commonly determined by brute-force simulation or experiments [35]–[42].

In [36], the detection of a known signal in WGN was considered and Xu et al. proposed to

TABLE 2.1
COMPARISON OF REVIEWED TDS.

Reference	Signal	Noise	Optimality measure	Implemen tation	Robust ness	Detect ability
[31]	DC unknown	Gaussian unknown	NP	Simple	Good	Poor
[32]	Arbitrary known	Non-Gaussian known	MSE	Complex	Poor	Good
[33]	Arbitrary known	Known Gaussian-Laplace mixture	ARE	Depends	Good	Good
[14]	Sinusoidal known	Colored Gaussian known	SNR	Simple	N/A	N/A
[15]	Periodic pulse known	Arbitrary noise known	SNR	Simple	N/A	N/A
[6]	Sinusoidal known freq. unknown amp., phase	Gaussian mixture & Generalized Gaussian known	SNR	Complex	Poor	Good
[11]	Known DC	Non-Gaussian known	probability of error	Good	Good	Good

N/A: Not applicable

obtain the maximum output SNR via tuning the BS parameters by experiments. In [37], Duan et al. used a BS as a receiver to decode a binary pulse amplitude modulation (PAM) signal passing through a WGN. The BS parameters were chosen from the given guide rules for an acceptable (not minumum) bit error rate (BER) in decoding. For the same problem, Xu, Duan et al. [35] compared the performance of two methods, noise-enhanced effect and BS design, when using a BS as the receiver. It was pointed out that using noise-enhanced effect can be viewed as a special case of the BS design, thus the BS design achieves better performance than noise-enhanced effect. However, in this investigation, the performance of the proposed receivers cannot exceed linear MF, the optimal receiver for WGN channel.

BD can have superior detection performance to linear MF in signal detection with non-Gaussian noise. For example, in watermark detection in discrete cosine transform (DCT) domain, the signal is the watermark (or a signature), while the DCT coefficients of an image is the noise, whose pdf is non-Gaussian and unknown in general [43], [44]. Hence the watermark extraction can be considered as a detection problem of a known signal in non-Gaussian noise, and BD has been used for this problem [38]–[42]. Sun et al. first proposed to use BD in watermark extraction [38], [39], [42]. Wu et al. [40] employed the same strategy but extended the idea by using the DCT of 8×8 blocks and adding an effective permutation of DCT coefficients. Duan et al. [41] proposed to use an array of BS's to further improve the performance of BD. These BD's were called SR detector in [38]–[42]. But they are more suitable to be considered as nonlinear BS design [35] for BD because no SR effect occurs in these BDs.

In watermark extraction, the BDs reviewed above achieve better performance than the linear MF. However, no analytical and systematic parameter optimization for the BS is proposed. The BS parameters were given either arbitrarily in [38]–[40], [42] or via exhaustive search in [41], [45]. The former cannot guarantee performance. For the latter, finding the optimal parameters by numerical tuning is generally infeasible for watermark extraction due to the difficulties in estimating the extraction performance. We will address these problems and propose a practical technique on the BD design.

2.2.3 Noise-Enhanced Detection

Noise-enhanced detection attempts to improve the detection performance via adding noise to the original observation. Hence, the major design problem it to derive the optimal noise.

Kay initiated noise-enhanced signal detection in [16]. It was shown that a suboptimal TD can have an improved probability of detection via adding a WGN. For a fixed test, denoted as $T(\mathbf{x})$, and a fixed critical function $\phi(\mathbf{x})$ in (1.3), Kay et al. [17] showed that the pdf of the optimal noise for minimizing the probability of decision error under Beyesian criterion is a Dirac delta function, thus the optimal noise is a constant.

For a general binary detection problem with an arbitrarily given detector, under the NP criterion, Chen et al. [18], [46] investigated the noise-enhanced effect and derived the optimal pdf form of the noise. The results are summarized as follows.

Consider a BHTP with known $p_0(\mathbf{x}) \triangleq f_{\mathbf{X}}(\mathbf{x}; H_0)$ and $p_1(\mathbf{x}) \triangleq f_{\mathbf{X}}(\mathbf{x}; H_1)$. For a fixed test and critical function $\phi(\mathbf{x})$, P_D and P_{FA} can be calculated by (1.4) and (1.5). Let $\mathbf{U} = \mathbf{X} + \mathbf{V}$, where $\mathbf{V}$ is the noise added to the original input $\mathbf{X}$. The goal is to find the optimal pdf for $\mathbf{V}$ that can maximize P_D subject to $P_{FA} \leq \alpha$.

Because $f_{\mathbf{U}}(\mathbf{u}; H_i), i = 0, 1$ are the convolution of $p_i(\mathbf{x})$ and $f_{\mathbf{V}}(\mathbf{x})$, when applying $\mathbf{U}$ to the fixed detector, $P_D^{\mathbf{U}}$ and $P_{FA}^{\mathbf{U}}$ are

$$
\begin{aligned}
P_D^{\mathbf{U}} &= \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) \left(\int_{\mathbb{R}^N} \phi(\mathbf{u}) p_1(\mathbf{u} - \mathbf{x}) d\mathbf{u} \right) d\mathbf{x} = \int_{\mathbb{R}^N} F_1(\mathbf{x}) f_{\mathbf{V}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}_{\mathbf{V}}\{F_1(\mathbf{x})\}, \\
P_{FA}^{\mathbf{U}} &= \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) \left(\int_{\mathbb{R}^N} \phi(\mathbf{u}) p_0(\mathbf{u} - \mathbf{x}) d\mathbf{u} \right) d\mathbf{x} = \int_{\mathbb{R}^N} F_0(\mathbf{x}) f_{\mathbf{V}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}_{\mathbf{V}}\{F_0(\mathbf{x})\},
\end{aligned}
$$

where $\mathbb{E}_{\mathbf{V}}\{\}$ stands for the expectation over $\mathbf{V}$ and $F_i(\mathbf{v}) \triangleq \int_{\mathbb{R}^N} \phi(\mathbf{u}) p_i(\mathbf{u} - \mathbf{v}) d\mathbf{u}, i = 0, 1$. Note that $F_1(\mathbf{v})$ and $F_0(\mathbf{v})$ are the P_D and P_{FA} when a given $\mathbf{u} = \mathbf{x} + \mathbf{v}$ is applied to the detector. In particular, $F_1(\mathbf{0})$ and $F_0(\mathbf{0})$ are the P_D and P_{FA} when the original input $\mathbf{X}$ is applied to the detector. $F_1(\mathbf{v}), F_0(\mathbf{v})$ are functions in domain $\mathbf{V} \in \mathbb{R}^N$. Let $f_0 = F_0(\mathbf{v})$ and $f_1 = F_1(\mathbf{v})$. $f_0 \in [0, 1]$ and $f_1 \in [0, 1]$ are both one-dimensional scalars and there is a many-to-many mapping between f_0 and f_1 based on same $\mathbf{v}$.

The optimal $P_D^{\mathbf{U}}$ is obtained from two (f_0, f_1) pairs, denoted as (f_{01}, f_{11}) and (f_{02}, f_{12}), which are located in the top region of the convex hull of all the (f_0, f_1) pairs and should satisfy $P_{FA}^{\mathbf{U}} = \lambda f_{01} + (1-\lambda) f_{02} = \alpha$ for $\lambda \in [0,1]$. Correspondingly, the optimal noise pdf $f_{\mathbf{V}}(\mathbf{v}) = \lambda\delta(\mathbf{v} - \mathbf{v}_1) + (1-\lambda)\delta(\mathbf{v} - \mathbf{v}_2)$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are determined by

$$\begin{cases} F_0(\mathbf{v}_1) = f_{01} \\ F_1(\mathbf{v}_1) = f_{11} \\ F_0(\mathbf{v}_2) = f_{02} \\ F_1(\mathbf{v}_2) = f_{12} \end{cases}.$$

It shows that the optimal noise is randomization between two constants $\mathbf{v}_1$ and $\mathbf{v}_2$ with probabilities λ and $1-\lambda$, respectively. The maximum detection probability is: $P_D^{\mathbf{U}} = \lambda f_{11} + (1-\lambda) f_{12}$ where $\lambda = \frac{\alpha - f_{02}}{f_{01} - f_{02}}$.

In the same spirit, the theory of the noise-enhanced detection in [18] was extended to variable detectors [19]. For fixed detectors, all (f_0, f_1) pairs are obtained for all the possible $\mathbf{v}$. While for variable detectors, all (f_0, f_1) pairs are obtained for all possible $\mathbf{v}$ and all variable parameters, such as η and P_{FA}. With this new set of (f_0, f_1) pairs, the optimal noise $\mathbf{v}$ is obtained as did for the fixed detectors.

The theory in [18], [19] gives the optimal noise pdf under NP criterion. It evokes a lot interest in the study of noise-enhanced signal detection. Chen et al. [47] and Guerriero et al. [48] demonstrated that noise-enhanced effect can improve detection performance of sequential detectors. Kay in [49] pointed out that optimal noise-enhanced detection can be viewed as a randomized decision rule, which is a commonly used technique in signal detection [1], [2]. Bayram exploited noise-enhanced effect under NP criterion [50], in the minimax framework [51], and in Bayesian framework [52]. Patel et al. [20], [53] exploited the necessary and sufficient conditions for the existence of the optimal noise effect under NP criterion, and proposed a numerical tuning technique to find a near-optimal noise.

We point out that the major problem in existing results on noise-enhanced detection for a general detector is its practicability. First, in [18]–[20], the optimal noise pdf was derived. The noise pdf results however are in implicit form and numerical methods are required to use the results in real applications. Generally speaking, closed-form is hard to find. Secondly, under NP criterion, for different P_{FA}'s, the optimal noise pdf can be different. Thus, the optimal noise pdf can be sensitive to parameter values P_{FA} and the decision threshold η in the critical function $\phi(\cdot)$. It is difficult to use the noise-enhanced effect, especially for applications with imperfect information (on noise or signal parameters) or changing environment. Finally, for all existing results, numerical methods are required in finding the optimal noise, the computational complexity is in general high.

2.3 A Note on SR Detection

In this part, we clarify the concept of SR detection and its misuse in two perspectives: one is noise-enhanced detection; the other is detector design. First, a review on SR detection is presented. Secondly, we provide our understanding of SR detection. Finally, we clarify the relationship and the difference of SR detection, noise-enhanced detection, and detector design.

2.3.1 Review on SR Detection

The early SR detection aimed at detecting weak periodic signals in Gaussian noise using TS or BS. Inchiosa et al. [29] used SR detector for the detection problem BHTP II shown in Sec. 1.2. The SR detector uses a BS as the nonlinear system and the quadrature MF as the inner detector. Given that the signal is subthreshold to the BS, it was demonstrated that the P_D increases with the increase of the WGN variance until a critical level, then the P_D decreases with further increase of the WGN variance. Note that the original input (observation) includes signal and noise. This showed that adding WGN noise to the original input can increase P_D if the variance of the WGN noise in the original input is weaker than the critical level. Galdi et al. [30] addressed the sinusoidal signal detection in WGN. The proposed SR detector also used a BS as the nonlinear system but simple mean or sign counting was used as the inner detector. Instead of tuning the noise variance, the authors proposed to choose suitable parameters of the BS to match the noise level in the original input. In other words, the BS is adjusted for the purpose that the noise level in the original input is the critical level in the adjusted BS.

However, these attempts are generally unsuccessful in the view of signal detection. Because for the systems with WGN, the linear MF is optimal. Any additional nonlinearity, such as the nonlinearity induced by a TS or a BS, only degrades the detectability.

SR detectors were then used for signal detection in non-Gaussian noise. In [15], Chapeau-Blondeau derived the expression of the SNR at the output of a TS for detecting a rectangular pulse signal embedded in a known noise. It was pointed out that there is an optimal threshold of the TS to have the maximum SNR. Also, if the threshold is not set optimally, adding noise can increase the SNR. Saha et al. [6] proposed to use a SR detector, composed of a three-level TS and a quadrature MF, for the detection problem BHTP IV in Sec. 1.3. The expression of the SNR gain of the three-level TS is derived, which depends on the noise pdf and the TS threshold. The optimal threshold was solved to maximize the SNR gain. It was shown that this detector can have a superior performance to the quadratic MF. Also for non-Gaussian noise, Zozor et al. [54] demonstrated that SR effect can occur in LO detector. Rousseau et al. [55] investigated the SR effect in a TD to detect a known deterministic signal in a noise with known pdf. It showed that adding a WGN can decrease the probability of error.

The SR detectors reviewed above can be divided into two groups: in one group, SR effect is obtained by tuning the noise level [15], [29], [54], [55]; in the other group, the effect is obtained by tuning the parameters of the detector [6], [15], [30], [34]. However, the term SR detection was misused for some cases [36]–[42] because in these detector designs, no SR effect occurs. For a correct and clear presentation on these concepts, we should answer the following two questions: 1) what is the difference between the SR detection via tuning the noise level and the noise-enhanced detection? 2) what is the difference between the SR detection via tuning system parameters and detector design?

2.3.2 Understanding of SR Detection

In the following, taking BD as an example, we demonstrate the SR effect via adding noise and the SR effect via tuning parameters in SR detectors, respectively.

We first show the SR effect via adding noise in SR detectors. Assume that the noise in the original input W is WGN with variance σ_w^2, the additional noise V is also WGN with variance σ_v^2. Thus, the adjusted noise $U = W + V$ has a variance $\sigma_u^2 = \sigma_w^2 + \sigma_v^2$. As demonstrated in [29], we plot a diagram of the detection performance versus the noise variance σ^2 in Fig. 2.4. For a given SR detector, for example, a BS followed by a quadratic MF in [29], the noise variance in the original input is σ_w^2 and its performance is P_{ori}. We increase σ_v^2 until σ_u^2 reaches the critical level σ_c^2, and the optimal performance, denoted as $P_{opt}^{add\_n}$, is obtained as marked in Fig. 2.4. We can see that by adding noise, the performance is improved. Thus, SR effect happens.

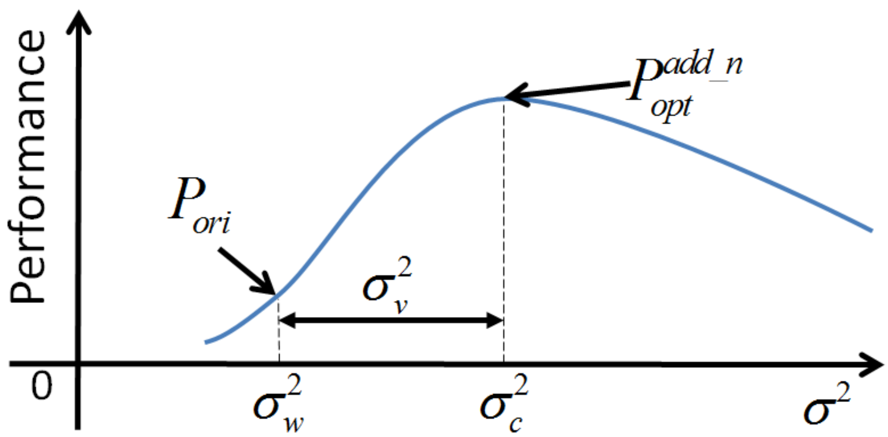


Fig. 2.4. SR effect via adding noise in SR detector.

Next we explain the SR effect induced by parameter tuning in SR detectors, as in [30]. We present the performance of the BDs with different BS parameters versus noise variance in Fig. 2.5. BS_1 is what we used in Fig. 2.4. We can change the BS parameters to have BS_2 such that for BS_2, its critical noise level σ_c^2 coincides with the original noise level σ_w^2. Under this situation, the SR effect is obtained by tuning the BS parameters, and the performance is denoted as $P_{opt}^{tun\_p\_SR}$.

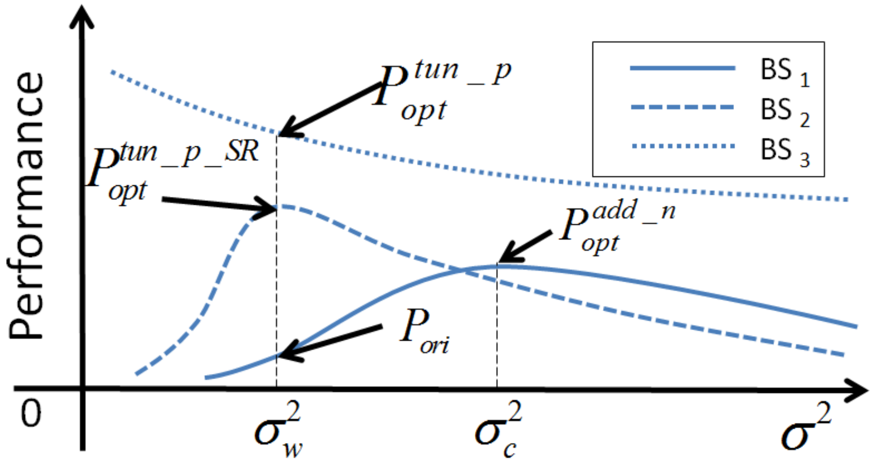


Fig. 2.5. SR effect via tuning parameters in BS based detector.

We should note here that for the given input, BS_1 and BS_2 both work at the SR regime because the constructive role of the noise can be observed. For BS_1, adding noise will increase its performance. For BS_2, the noise amount in the original input is just what the system needs for the optimal SR effect. BS_1 and BS_2 based detectors have SR effect and are no doubt to be called SR detector.

However, if we can adjust the BS parameters optimally/suboptimally as did in [36], [37], [42], we will have a BS_3 whose achievable performance is denoted as $P_{opt}^{tun\_p}$. This $P_{opt}^{tun\_p}$ is superior to $P_{opt}^{tun\_p\_SR}$ for any noise variance. For BS_3, noise does NOT play a constructive role, as shown in Fig. 2.5. Therefore, we would like to clarify that this tuning on BS parameters follows detector design rather than SR detector (although claimed to be SR detector) because no SR effect occurs in BS_3.

2.3.3 Relationship of SR Detector, Noise-Enhanced Detector, and Detector Design

SR detector has the signature that SR effect occurs in the nonlinear system. For the sake of SR effect, the input signal should be subthreshold to the given nonlinear system, which results in the constructive role of the noise. As we presented in Sec. 2.3.2, there are two types of SR detection based on two methods of achieving SR effects: via adding noise and via tuning parameters.

Noise-enhanced detector does not has the constraints related to the SR effect. It can be any type of detector, linear or nonlinear. If it is nonlinear, the nonlinear system can be arbitrary, and needs not to have the signal subthreshold to the system. The signature of the noise-enhanced detector is that adding an extra noise can improve the performance. Therefore, noise-enhanced detection can be viewed as an extension of SR detection based on adding noise.

Detector design is to determine the system or subsystem in a detector and hence is more general than SR detection based on tuning parameters, without considering the constraints of SR effect.

In summary, there is no overlap between the noise-enhanced detection and the detector design. The SR detection via adding noise is a special case of the noise-enhanced detection. The SR detection via tuning parameters is a special case of the detector design. This relationship can be illustrated in Fig. 2.6.

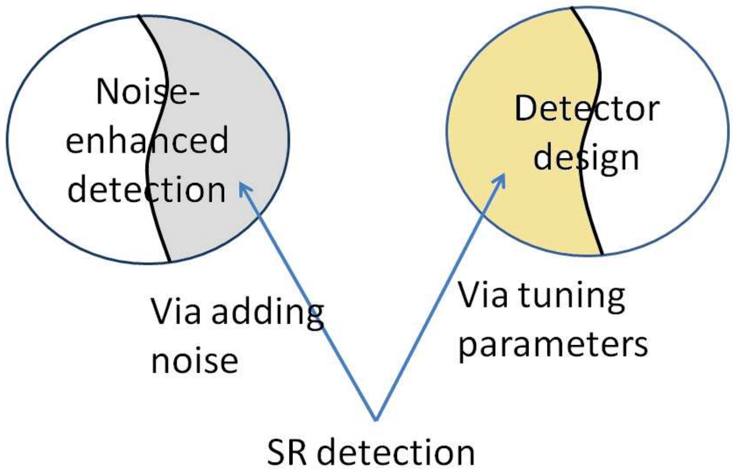


Fig. 2.6. Relationship of SR detection, noise-enhanced detection, and detector design.

According to the above clarification, we point out that [14], [15], [23], [29], [30], [34] should be classified as SR detectors. [16]–[18], [20], [46]–[53] should be classified as the noise-enhanced signal detection. [6], [36]–[42] should be classified as detector design.

2.4 Thesis Contributions

This thesis aims at designing robust and simple detectors with satisfying detection performance through TD design, BD design, and noise-enhanced effect. An overview of the contributions of this thesis is shown in Fig. 2.7. The four major contributions are listed as follows.

Contribution I: an optimal TD design under NP criterion for a known DC signal in known non-Gaussian noise. The TD is composed of a TS and a linear MF, and hence the major task is to design the TS. The optimal TS's, including simple binary TS and composite binary TS, are derived analytically. Experimental results show the validity of the derived optimal TS. For non-Gaussian noise with heavy pdf tails, the proposed TD significantly outperforms the linear MF, and has comparable performance to the LO detector. This work has been published as [56], [57] and is

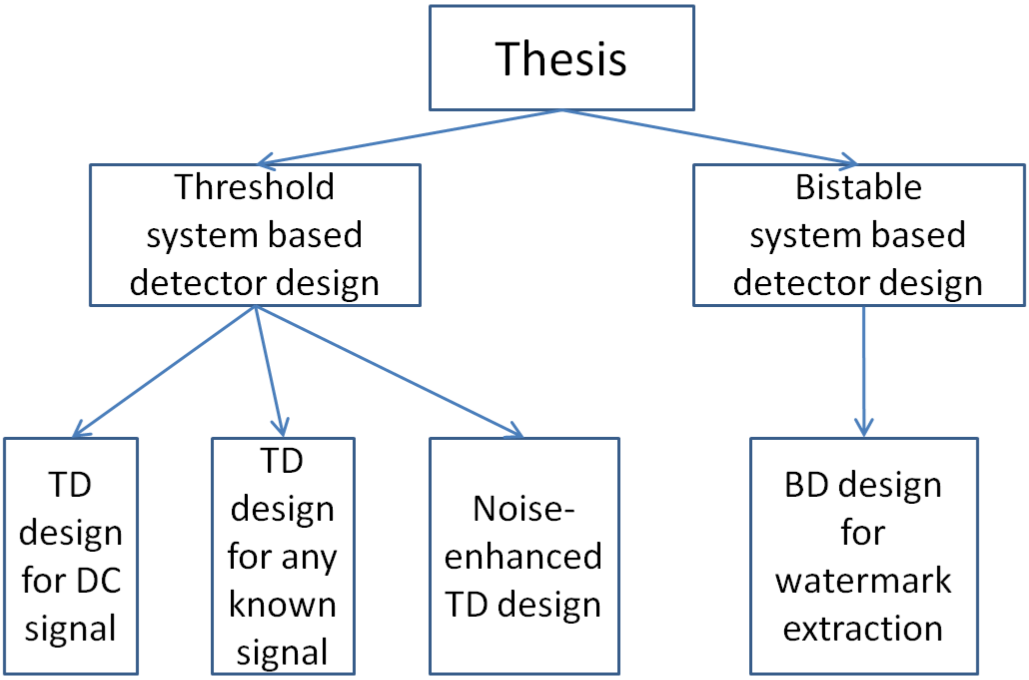


Fig. 2.7. Overview of the thesis contributions.

presented in Chap. 3.

Contribution II: a robust TD of any known signal in unknown non-Gaussian noise. We design a TD for detecting any known deterministic signal in independent non-Gaussian noise whose pdf is unknown but is symmetric and unimodal. Under the assumptions of white noise, small signal, and a large number of samples, the proposed TD is shown to maximize the AUC, which is the abbreviation for area under receiver operating characteristic (ROC) curve. While previous TD designs need accurate information of the noise pdf, the proposed TD is independent of the noise pdf. The detection probability and the ROC of the proposed TD are analyzed both theoretically and numerically. It is shown that even without knowing the noise pdf, the proposed TD has close performance to the optimal detector designed with precise noise pdf information. It also performs significantly better than the linear MF when the noise pdf has heavy tails. The practical implementation, robustness to both the noise pdf and the signal, and region of validity of the proposed TD are also investigated. This work has been published as [58] and is presented in Chap. 4.

Contribution III: optimal design of noise-enhanced TD under AUC measure. We investigate the noise-enhanced effect for a general BHTP using a binary TD. We adopt AUC as the performance measure for its implementation simplicity and robustness. First the optimal TS design that maximizes the AUC has been derived. Then we consider the noise-enhanced effect in this detector. The optimal noise pdf that maximizes the AUC is shown to be a delta function, indicating that the optimal noise is deterministic. Performance of the proposed design and comparison with other designs are shown via an example. This work has been submitted as [59]. The details are presented in Chap. 5.

Contribution IV: an adaptive BD for watermark extraction. In this part, we focus on BD design for watermark extraction. Since a BD is composed of a BS and a matched filter, we only need to consider the BS design, i.e., optimizing BS parameters. There is no existing efficient and systematic BS design method except exhaustive search. We propose to use the cross-correlation between the watermark signal and the BS output as the measure in determining the BS parameters. A key observation is that the optimal BS parameters depend on the noise and the watermark level but are not sensitive to the noise pdf form and the watermark sequence. This feature provides an easy and practical way to build an adaptive BS since the signal and the noise level can be estimated easily. Experimental results in watermark extraction show that the performance obtained from the proposed adaptive BD is satisfactory to various extraction scenarios, and performs better than existing BD's and the WGN-based ML detector for most cases. This work has been published as [28], [60]–[62] and is presented in Chap. 6.

Contribution I, II, and IV follow the technique of TD and BD design in the methodology stated in Sec. 1.4. Contribution III follows the technique of noise-enhanced effect.

2.5 Summary

In this chapter, we have presented a comprehensive review on TD design, BD design, and noise-enhanced detection. We first introduced the background on TS, BS, and noise-enhanced detection. Followed is the literature review on the topics addressed in this thesis, which are detector designs using TS or BS, and noise-enhanced detection. As SR detection always emerges when discussing these two topics and can bring confusion in concepts, we provided an extra note on the SR detection to clarify its relationship with the topics of this thesis. The contributions of this thesis were summarized at the end of this chapter.

Chapter 3

Optimal Design of Threshold Detector for a DC Signal in Non-Gaussian Noise

TS is widely used in nonlinear signal detection for a simple, robust, and suboptimal solution. The optimal design of the TS is crucial for good detection performance. In this chapter, we propose the optimal TS design in the detection of a DC signal in non-Gaussian noise under NP criterion. First, we propose a novel performance indicator to replace the probability of detection as the design criterion. Using this indicator, we derive the optimal design of simple binary TS and composite binary TS, respectively. Experiments show that the proposed TD can perform close to the LO detection with much simpler implementation and much less computational complexity. It performs significantly better than the MF for non-Gaussian noise with heavy pdf tails.

3.1 Introduction

The problem of detecting a known signal in additive noise with known distribution has well-known solutions, which is LRT or GLRT [1]. But LRT or GLRT can be highly complex in implementation, and has low robustness to parameter changes, thus has poor performance when there is imprecision or change in the noise or signal information. As stated in the research methodology in Sec. 1.4, for many applications, suboptimal but robust detectors provide more practical choices [31]. Threshold system (TS) based detector, or TD in short, is one of the suboptimal detectors widely used in many applications [6], [11], [31]–[33], [63]. TD has several advantages: high speed, low resource (computing capacity, memory, storage) requirement, simple implementation, high robustness, and good detection performance, as reviewed in Sec. 2.2.1.

Consider the BHTP described in (1.1). In this chapter, we assume that the signal is a known DC signal, i.e., $\mathbf{s} = A$ is a constant. The noises are i.i.d., whose pdf, $f_W(w)$, is known. It is also assumed that the signal is weak compared to the noise, i.e., $A \ll \sigma$, where σ is the standard deviation of the noise. For a good performance, large N ($N \gg 1$) is conventionally assumed in weak signal detection applications [1].

The schematic of the proposed TD is shown in Fig. 3.1. It is composed of a TS, followed by a simple mean calculator. Note that the MF for DC signal is equivalent to the simple mean. As shown in Fig. 1.3, the LO detector for simple non-Gaussian based BHTP is a nonlinear system $g_{LO}(x)$ defined in (1.13) followed by a MF. Here, we use TS to replace the complicated nonlinear system $g_{LO}(x)$.

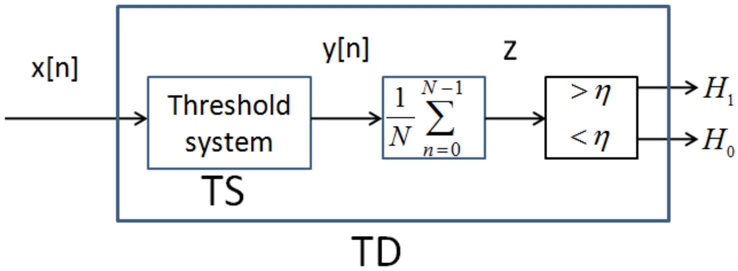


Fig. 3.1. Schematic of the proposed TD.

For the TS, we choose to use the simple binary TS shown in (2.1) and the composite binary TS shown in (2.2) since implementation simplicity is the major consideration in our design. Note that simple binary TS can be viewed as a special case of composite binary TS. Its design is represented separately since it has simpler implementation than the general composite binary TS and is widely used in practical applications.

If we use the simple binary TS in (2.1), the only parameter needs to be designed is the threshold τ. For later convenience, we present the complemental version of the simple binary TS, called complemental simple binary TS, as

$$y[n] = \begin{cases} 0 & x[n] \geq \tau \\ 1 & x[n] < \tau \end{cases} . \tag{3.1}$$

Naturally, when the DC signal is positive, i.e., $A > 0$, we use the simple binary TS in (2.1); when the DC signal is negative, we use the complemental simple binary TS in (3.1). Using this alternative, we can derive identical statistical feature of the TS output $y[n]$'s for either $A > 0$ or $A < 0$. As shown in Fig. 3.1, the test statistic $z \triangleq \frac{1}{N}\sum_{n=0}^{N-1} y[n]$ is the simple mean of $y[n]$'s. A decision H_1 or H_0 is made if $z > \eta$ or $z < \eta$.

Similar TD based on simple binary TS has been used in [11], [16]–[18]. In [16]–[18], the TD is used to show the occurrence of noise-enhanced effect. Chapeau-Blondeau in [11] demonstrated that the TD can achieve better detection performance than MF for non-Gaussian noises. However, the optimal design of the TS parameter has not been addressed. To the best of our knowledge, no optimal design of the composite binary TS was available in literature.

In this chapter, we use the NP criterion to design the TS parameters for both simple binary TS (the one in (2.1) or its complement form in (3.1)) and composite binary TS. That is, we design the threshold τ for simple binary TS and the set $\mathcal{D}$ for composite binary TS to have the maximum P_D subject to $P_{FA} \leq \alpha$. We first derive P_D and P_{FA} formulas for the proposed TD, based on which a simpler indicator of P_D is discovered. Using the proposed indicator as the design criterion, we proposed a low-complexity algorithm to calculate the optimal threshold τ_{opt} of the TS in (2.1) or (3.1) for simple TS. Also based on the indicator, we determine the optimal set $\mathcal{D}_{\text{opt}}$ if the composite TS in (2.2) is used. Experimental results demonstrate that the proposed optimal TD design can perform very close to the LO detector, and much better than the MF, for non-Gaussian noise with heavy pdf tails.

This chapter is organized as follows. In Sec. 3.2, we derive the P_D formula of the TD and propose a new indicator of P_D. We then present the optimal design of simple binary TS and composite binary TS in Sec. 3.3. In Sec. 3.4, we discuss the situation when the noise pdf is unknown. In Sec. 3.5, simulation results show the performance of the proposed optimal TD, and the comparison with LO detector and MF. Finally, we conclude this chapter in Sec. 3.6.

3.2 An Indicator of P_D

In this section, we derive the formula of P_D. Since it is too complicated to be used in the optimal TS Design, a simpler indicator is proposed, of which P_D is approximated as monotonically increasing function.

3.2.1 Calculations of P_D

We first calculate $f_Z(z; H_0)$ and $f_Z(z; H_1)$. Recall that $f_X(x; H_0)$ and $f_X(x; H_1)$ are the pdf's of each entry $x[n], n \in [0, N-1]$ under the hypotheses H_0 and H_1, respectively. Since the signal A and the noise pdf $f_W(w)$ are known, $f_X(x; H_i)$ are both known as well and $f_X(x; H_0) = f_W(w), f_X(x; H_1) = f_W(w - A)$. Given the threshold τ for the simple binary TS described in (2.1), or the set $\mathcal{D}$ for the composite binary TS in (2.2), $Y = y[n]$ is a random variable (RV) taking only two values, 0 or 1. Hence, Y can be considered as an output of a Bernoulli trial. When the simple TS is used, we have the following probabilities.

$$\mathbb{P}(Y = 0; H_0) = \int_{-\infty}^{\tau} f_X(x; H_0) dx \triangleq q_0, \tag{3.2}$$

$$\mathbb{P}(Y=1;H_0)=\int_{\tau}^{\infty} f_X(x;H_0)dx \triangleq p_0, \tag{3.3}$$

$$\mathbb{P}(Y=0;H_1)=\int_{-\infty}^{\tau} f_X(x;H_1)dx \triangleq q_1, \tag{3.4}$$

$$\mathbb{P}(Y=1;H_1)=\int_{\tau}^{\infty} f_X(x;H_1)dx \triangleq p_1, \tag{3.5}$$

where $p_0+q_0=1$ and $p_1+q_1=1$. Observe that p_0 and p_1 are the probability of $Y=1$ under the conditions of H_0 and H_1 respectively. If the DC signal $A>0$, $f_X(x;H_1)$ is a right shift of $f_X(x;H_0)$, and the TS in (2.1) is employed. It can be verified that $p_1>p_0$. For $A<0$, we use the complemental TS in (3.1) to have $p_1>p_0$. Without loss the generality, we only consider the case for $A>0$. While for $A<0$, the same results can be obtained.

If the composite binary TS is used, we have

$$\mathbb{P}(Y=0;H_0)=\int_{x\notin\mathcal{D}} f_X(x;H_0)dx \triangleq q_0, \tag{3.6}$$

$$\mathbb{P}(Y=1;H_0)=\int_{x\in\mathcal{D}} f_X(x;H_0)dx \triangleq p_0, \tag{3.7}$$

$$\mathbb{P}(Y=0;H_1)=\int_{x\notin\mathcal{D}} f_X(x;H_1)dx \triangleq q_1, \tag{3.8}$$

$$\mathbb{P}(Y=1;H_1)=\int_{x\in\mathcal{D}} f_X(x;H_1)dx \triangleq p_1. \tag{3.9}$$

Let $m\in[0,N]$ be the number of 1's in the TS output sequence. Under hypothesis H_0, the distribution of m is a binomial distribution, and its probability mass function (pmf) is $f_M(m)=\begin{pmatrix} N \\ m \end{pmatrix} p_0^m q_0^{N-m}$. Notice that $z=\frac{1}{N}\sum_{n=0}^{N-1} y[n]=\frac{m}{N}$, we then have the following pmf for $Z;H_0$

$$f_Z\left(z;H_0\right)=\begin{pmatrix} N \\ zN \end{pmatrix} p_0^{zN} q_0^{N-zN}, \qquad z=0,\frac{1}{N},\frac{2}{N},...,1. \tag{3.10}$$

Under H_1, similarly, we have

$$f_Z\left(z;H_1\right)=\begin{pmatrix} N \\ zN \end{pmatrix} p_1^{zN} q_1^{N-zN}, \qquad z=0,\frac{1}{N},\frac{2}{N},...,1. \tag{3.11}$$

Notice that $f_Z(z;H_i), i=0,1$ are discrete functions that are valid for $z=0,1/N,2/N,...,1$ only. However, when N approaches infinity, the functions approach continuous ones. Since $N\gg 1$, for the tractability of analysis, we use the continuous form of (3.10) as the pdf of Z under H_0 in the

P_{FA} calculation, which is shown as follows.

$$f_Z(z;H_0) = \frac{N \cdot \Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)} p_0^{zN} q_0^{N-zN}, \qquad z \in [0,1], \tag{3.12}$$

where $\Gamma(\cdot)$ is the *Gamma* function defined as,

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt.$$

Similarly, we obtain

$$f_Z(z;H_1) = \frac{N \cdot \Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)} p_1^{zN} q_1^{N-zN}, \qquad z \in [0,1]. \tag{3.13}$$

For a given $P_{FA} = \alpha$, η can be calculated using the following equation

$$P_{FA} = \int_{z>\eta} f_Z(z;H_0) dz = \alpha. \tag{3.14}$$

With the above η, P_D can be calculated by

$$P_D = \int_{z>\eta} f_Z(z;H_1) dz. \tag{3.15}$$

Since we consider continuous $f_Z(z;H_i), i = 0,1$, we use integrals in (3.14) and (3.15) to calculate P_{FA} and P_D. Note that for discrete $f_Z(z;H_i), i = 0,1$, the integrals in (3.14) and (3.15) should be replaced by summations.

3.2.2 A Detectability Indicator Δp

Define

$$\Delta p \triangleq p_1 - p_0 = \mathbb{P}(Y=1;H_1) - \mathbb{P}(Y=1;H_0), \tag{3.16}$$

where p_0, p_1 are defined in (3.3) and (3.5) for simple binary TS, in (3.7) and (3.9) for composite binary TS, respectively. In this subsection, we show that Δp is a good indicator of detectability.

Under NP criterion, the TS design problem for the TD is

$$\tau_{\text{opt}} = \arg\max_{\tau} P_D, \qquad \text{s.t.} \qquad P_{FA}(\tau) \le \alpha, \tag{3.17}$$

when simple binary TS is used, and

$$\mathcal{D}_{\text{opt}} = \arg\max_{\mathcal{D}} P_D, \qquad \text{s.t.} \qquad P_{FA}(\mathcal{D}) \le \alpha, \tag{3.18}$$

when composite binary TS is used.

We first consider the case of simple binary TS. By using the derived results in the previous section, (3.14), (3.15), (3.12) (3.13), (3.2)-(3.5), the τ_{opt} can be calculated by

$$\tau_{\text{opt}}=\arg\max_{\tau}\underbrace{\int_{z>\eta}\frac{N\cdot\Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)}\left(\int_{\tau}^{\infty}f_X(x;H_1)dx\right)^{zN}\left(\int_{-\infty}^{\tau}f_X(x;H_1)dx\right)^{N-zN}dz}_{P_D}, \tag{3.19}$$

$$\text{s.t.}\ \int_{z>\eta}\underbrace{\frac{N\cdot\Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)}\left(\int_{\tau}^{\infty}f_X(x;H_0)dx\right)^{zN}\left(\int_{-\infty}^{\tau}f_X(x;H_0)dx\right)^{N-zN}}_{f_Z(z;H_0)}dz=\alpha. \tag{3.20}$$

In general, it is difficult to find the close form solution for τ_{opt}. The complicated formulas even make advanced numerical methods difficult to find. One natural method is an exhaustive search of τ. Assume that α is given. First, for a given τ, we can calculate η numerically from (3.20). With this η, we can then calculate the P_D shown in (3.19) numerically. The two steps above are repeated for all the possible τ values in exhaustive search, and the τ that leads to the maximum P_D is τ_{opt}. However, exhaustive search has very high computational complexity and is impractical in real applications.

For the case of the composite binary TS, similarly, we have

$$\mathcal{D}_{\text{opt}}=\arg\max_{\mathcal{D}}\underbrace{\int_{z>\eta}\frac{N\cdot\Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)}\left(\int_{x\in\mathcal{D}}f_X(x;H_1)dx\right)^{zN}\left(\int_{x\notin\mathcal{D}}f_X(x;H_1)dx\right)^{N-zN}dz}_{P_D}, \tag{3.21}$$

$$\text{s.t.}\ \int_{z>\eta}\frac{N\cdot\Gamma(N+1)}{\Gamma(zN+1)\Gamma(N-zN+1)}\left(\int_{x\in\mathcal{D}}f_X(x;H_0)dx\right)^{zN}\left(\int_{x\notin\mathcal{D}}f_X(x;H_0)dx\right)^{N-zN}dz=\alpha. \tag{3.22}$$

For this optimization problem, there is even no way to conduct exhaustive search since the set of all possible $\mathcal{D}$ is the set of all subsets of $\mathbb{R}$, which has infinite dimensions.

For tractable TS design, in the following, we propose an indicator of P_D and use it as the design criterion.

Lemma 3.1. When $A \ll \sigma$ and $N \gg 1$, for any given P_{FA}, P_D is approximately a monotonically increasing function of Δp.

Proof. We first represent P_D as a monotonically decreasing function of Θ, where Θ is a function of Δp and p_1. We then show that $|\frac{\partial\Theta}{\partial p_1}| \ll 1$ and $\frac{\partial\Theta}{\partial\Delta p} < 0$, which shows that Θ is approximately a monotonically decreasing function of Δp, and hence P_D is approximately a monotonically increasing function of Δp.

When $N \gg 1$, using DeMoivre-Laplace theorem [4], $f_Z(z; H_0)$ and $f_Z(z; H_1)$ in (3.12) and (3.13) can be approximated by the normal distributions,

$$f_Z(z; H_0) \approx \mathcal{N}(z; p_0, p_0(1-p_0)/N), \tag{3.23}$$

$$f_Z(z; H_1) \approx \mathcal{N}(z; p_1, p_1(1-p_1)/N), \tag{3.24}$$

where $\mathcal{N}(z; \mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2. Using (3.23) in (3.14) gives

$$Q\left(\frac{\eta - p_0}{\sqrt{p_0(1-p_0)/N}}\right) = \alpha, \tag{3.25}$$

where $Q(x) = \int_x^\infty \mathcal{N}(z; 0, 1)dz$, which is the complementary cumulative distribution function (cdf) of standard normal pdf. We solve η from (3.25) and obtain

$$\eta = p_0 + Q^{-1}(\alpha)\sqrt{p_0(1-p_0)/N}, \tag{3.26}$$

where Q^{-1} is the inverse function of $Q(\cdot)$. Substituting (3.26) into (3.15), we obtain

$$P_D = Q\left(\frac{\eta - p_1}{\sqrt{p_1(1-p_1)/N}}\right) = Q\left(\frac{p_0 - p_1 + Q^{-1}(\alpha)\sqrt{p_0(1-p_0)/N}}{\sqrt{p_1(1-p_1)/N}}\right). \tag{3.27}$$

Substituting $p_0 = p_1 - \Delta p$ into (3.27) gives

$$P_D = Q\left(\frac{-\Delta p + Q^{-1}(\alpha)\sqrt{(p_1 - \Delta p)(1 - p_1 + \Delta p)/N}}{\sqrt{p_1(1-p_1)/N}}\right).$$

Because $Q(\cdot)$ is a monotonically decreasing function, we can consider only the argument inside the $Q(\cdot)$ function, which is denoted as

$$\Theta \triangleq \frac{-\Delta p + Q^{-1}(\alpha)\sqrt{(p_1 - \Delta p)(1 - p_1 + \Delta p)/N}}{\sqrt{p_1(1-p_1)/N}}. \tag{3.28}$$

The partial derivatives of Θ with respect to Δp and p_1 can then be obtained as follows:

$$\frac{\partial \Theta}{\partial \Delta p} = \frac{1}{t}\left(-\sqrt{N} + Q^{-1}(\alpha)\frac{-1 + 2p_1 - 2\Delta p}{2s}\right) \tag{3.29}$$

$$\frac{\partial \Theta}{\partial p_1} = \frac{Q^{-1}(\alpha)\left[(1-2p_1)\left(\frac{t}{2s} - \frac{s}{2t}\right) + \frac{t}{s}\Delta p\right] - t\Delta p\sqrt{N}}{t^2}, \tag{3.30}$$

where $s \triangleq \sqrt{(p_1 - \Delta p)(1 - p_1 + \Delta p)}$ and $t \triangleq \sqrt{p_1(1-p_1)}$.

In the following, we show that Claim 1. $|\frac{\partial \Theta}{\partial p_1}| \ll 1$, which means that the changing of p_1 has little influence to Θ and therefore Θ can be approximated as a function of only Δp; and Claim 2. $\frac{\partial\Theta}{\partial\Delta p} < 0$, which means that Θ monotonically decreases with Δp.

Claim 1. $|\frac{\partial\Theta}{\partial p_1}| \ll 1$. Because $A \ll \sigma$, and $f_X(x; H_1) = f_X(x - A; H_0) = f_W(x - A)$, for simple binary TS, we have, from (3.16) and using (3.3) and (3.5),

$$\begin{aligned} \Delta p &= \int_\tau^\infty f_X(x; H_1)dx - \int_\tau^\infty f_X(x; H_0)dx \\ &= \int_\tau^\infty f_W(x - A)dx - \int_\tau^\infty f_W(x)dx \\ &\approx \int_\tau^\infty (f_W(x) - Af'_W(x) - f_W(x))dx \\ &\approx Af_W(\tau) \ll p_1, p_0, 1. \end{aligned} \tag{3.31}$$

For composite binary TS, using (3.7) and (3.9) also gives $\Delta p \ll p_1, p_0, 1$. There we have

$$\left|\frac{t}{2s} - \frac{s}{2t}\right| = \frac{1}{2ts}\left|t^2 - s^2\right| = \frac{2p_1 - 1}{2ts} = \frac{2p_1 - 1}{2p_1(1 - p_1)}\Delta p + \mathcal{O}[(\Delta p)^2] \ll 1.$$

Here $\mathcal{O}[(\Delta p)^2]$ is a value proportional to the Δp raised to the second power. Thus from (3.30), $|\frac{\partial\Theta}{\partial p_1}| = \mathcal{O}(\Delta p) \ll 1$.

Claim 2. $\frac{\partial\Theta}{\partial\Delta p} < 0$. To show this, we recast (3.29) as,

$$\frac{\partial\Theta}{\partial\Delta p} = \frac{1}{\sqrt{p_1(1 - p_1)}}\left(\frac{Q^{-1}(\alpha)(p_0 - 1/2)}{\sqrt{p_0(1 - p_0)}} - \sqrt{N}\right). \tag{3.32}$$

Because $\sqrt{p_1(1 - p_1)}$ is positive, we only need to decide the sign of the following equation,

$$\Psi(p_0, \alpha, N) = \left(\frac{Q^{-1}(\alpha)(p_0 - 1/2)}{\sqrt{p_0(1 - p_0)}} - \sqrt{N}\right). \tag{3.33}$$

Since $p_0 \in (0, 1)$ for any given $P_{FA} = \alpha \in (0, 1)$, when N is large enough, $\Psi(p_0, \alpha, N)$ is always negative.

From Claim 1 and Claim 2, we conclude that for small signal and large N, Θ is approximately a monotonically decreasing function of Δp. Hence P_D is approximately a monotonically increasing function of Δp. $\square$

Simulation are conducted to justify the result in Lemma 3.1. One of the examples is shown in Fig. 3.2. In this example, we use $P_{FA} = 0.1$ and $N = 100$. We calculate P_D with respect to p_1 and p_0 using (3.12)-(3.15), and plot the contour of P_D in the (p_1, p_0) plane in Fig. 3.2. Since $P_{FA} = 0.1$, the contour line $P_D = 0.1$ is the line $p_0 = p_1$. It can be observed that the other contour lines, such as $P_D = 0.2$, are like straight lines parallel to the line $p_0 = p_1$. Because the contour lines is parallel to the line $p_0 = p_1$, the points in the same contour line have the same Δp. Also because the points in the same contour line have the same P_D, we can observe that the same Δp leads to the same P_D, regardless of the p_1 or p_0 values, which verifies Claim 1 in the proof of Lemma 3.1. We also observe that the P_D monotonically increases from top isoline $P_D = 0.1$ to the bottom isoline $P_D = 0.9$ with the increase of Δp, which verifies Claim 2 in the proof of Lemma 3.1. The result in Lemma 3.1 is proved for small signal and large N only. Simulation shows that the result works well in practical situations even for moderate N.

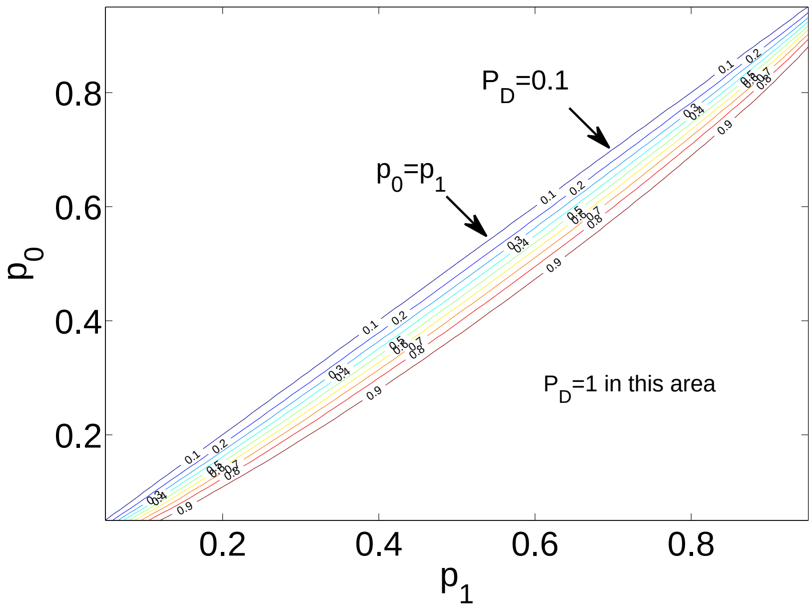


Fig. 3.2. Contour of P_D with respect to (p_1, p_0).

Lemma 3.1 shows that P_D is approximately a monotonically increasing function of Δp. Therefore in our TS optimization, instead of P_D, we can use Δp as the optimization criterion. Δp has a much simpler format than P_D and the optimization based on it is more tractable. The optimization of simple TS over P_D needs an exhaustive search of τ, and optimization of composite TS over P_D has no available solution. While with Δp, the design complexity can be largely reduced as will be demonstrated in the next section.

3.3 Optimal TS Design Using Δp

In this section, using Δp, we calculate τ_{opt} if simple binary TS is employed. We also determine $\mathcal{D}_{\text{opt}}$ if composite binary TS is used.

3.3.1 Optimal Design of Simple Binary TS

Under the NP criterion, if simple binary TS is used, τ_{opt} is the solution of the optimization in (3.19) and (3.20). From Lemma 3.1, we see that approximately, P_D monotonically increases with Δp for any given $P_{FA} = \alpha \neq 0$. Therefore, for tractable TS design, we replace P_D by Δp and study the following design.

$$\arg\max_{\tau} \Delta p(\tau). \tag{3.34}$$

Notice that unlike P_D, Δp is independent of α. Thus the constraint in the original problem (3.20) is not needed for the new problem in (3.34).

Lemma 3.2. Let $x_i, i = 1, ..., K$ be the K intersections between $f_X(x; H_0)$ and $f_X(x; H_1)$, i.e., $f_X(x_i; H_0) = f_X(x_i; H_1)$. The solution of (3.34) is $\tau_{\text{opt}} = \arg\max_{\tau \in \{x_i\}} \Delta p(\tau)$, which is independent of P_{FA}. When $A \ll \sigma$, $\tau_{\text{opt}} \approx \arg\max_{w} f_W(w)$.

Proof. The solution of (3.34) is the global maximum points of $\Delta p(\tau)$. Because $\Delta p(\tau) = p_1 - p_0 = \int_{\tau}^{\infty} [f_X(x; H_1) - f_X(x; H_0)]\, dx$, the local extreme points of Δp should satisfy $\frac{d\Delta p}{d\tau} = f_X(\tau; H_0) - f_X(\tau; H_1) = 0$. Therefore, τ_{opt} is the one of the $x_i, i = 1, ..., K$ that maximizes Δp, which is independent of P_{FA}.

If $A \ll \sigma$, according to (3.31), we have

$$\tau_{\text{opt}} \approx \arg\max_{\tau} A f_W(\tau) = \arg\max_{w} f_W(w). \tag{3.35}$$

In other words, τ_{opt} is the value that has the maximum probability in the noise pdf, i.e., the mode that has the largest pdf value. □

Based on Lemma 3.2, we can derive τ_{opt} for some specific cases, which is shown below.

Corollary 3.1. If the noise pdf is unimodal and symmetric, we have $\tau_{\text{opt}} = w_0 + A/2$, where w_0 is the only mode of noise pdf and A is the signal amplitude.

Proof. From Lemma 3.2, we know that τ_{opt} locates at one intersection of $f_X(x; H_0)$ and $f_X(x; H_1)$. Unimodal pdf means that the distribution has a single maximum. Thus there is only one intersection between $f_X(x; H_0)$ and $f_X(x; H_1)$. If the noise pdf is further symmetric, the only intersection has to be $x = w_0 + A/2$. □

Corollary 3.1 shows that τ_{opt} can be obtained in easy close-form if the noise pdf is unimodal and symmetric. For arbitrary noise pdf, we can find the $x_i, i = 1, ..., K$, the intersections between $f_X(x; H_0)$ and $f_X(x; H_1)$, by solving $\frac{d\Delta p}{d\tau} = f_X(\tau; H_0) - f_X(\tau; H_1) = 0$. This could be fulfilled using fast numerical methods like Newton-Raphson method [22].

In the following, we compare the complexity of the proposed TS design using Δp with that of using P_D. As we explained in Sec. 3.2, if P_D is used, to find τ_{opt}, one natural method is exhaustive search. To be able to compare the complexity of the design using P_D with that using Δp, we consider exhaustive search for both with the same range and resolution. If using Δp, $p_0(\tau)$ and $p_1(\tau)$ need to be calculated for every τ. If using P_D, beside the calculation of $p_0(\tau)$ and $p_1(\tau)$, it needs two numerical integrals for every τ to calculate the η for the given P_{FA} and the corresponding P_D as shown in (3.14) and (3.15). This shows that the computational complexity of τ_{opt} calculation using Δp is significantly reduced compared to that using P_D. For a specific example, the computer running times to find τ_{opt} using Δp and using P_D will be given in Sec. 3.5.1.

3.3.2 Optimal Design of Composite Binary TS

For composite binary TS, we similarly replace P_D by Δp in the TS design problem and study the following problem:

$$\arg\max_{\mathcal{D}} \Delta p(\mathcal{D}). \tag{3.36}$$

The result is presented in the following lemma.

<u>Lemma</u> 3.3. The optimal set that maximizes Δp is $\mathcal{D}_{\text{opt}} = \{x; f_X(x;H_1) > f_X(x;H_0)\}$.

Proof. If composite binary TS is employed, using p_0 and p_1 in (3.7) and (3.9), we have

$$\mathcal{D}_{\text{opt}} = \arg\max_{\mathcal{D}} \int_{x\in\mathcal{D}} (f_X(x;H_1) - f_X(x;H_0))dx = \{x; f_X(x;H_1) > f_X(x;H_0)\}.$$

□

For a DC signal, the result can be further simplified.

Corollary 3.2. Assume that the DC signal $A > 0$. Denote the ascending sorted intersections of $f_X(x;H_0)$ and $f_X(x;H_1)$ as $x_i, i = 1, 2, ..., K$. $\mathcal{D}_{opt} = \{x; x_j < x < x_{j+1}$ for odd $j, j \leq K\}$, where $x_{K+1} = \infty$. When $A \ll \sigma$, $x_i \approx w_i$, where $w_i, i = 1, 2, ..., K$ is the ascending sorted modes of $f_W(w)$.

Proof. Since $A > 0$, $f_X(x;H_1)$ is a right shift of $f_X(x;H_0)$. Hence we have $f_X(x;H_1) > f_X(x;H_0)$ if and only if $x_j < x < x_{j+1}$ for j odd.

The intersections $\{x_i\}$ satisfy $f_X(x;H_0) = f_X(x;H_1)$. If $A \ll \sigma$, we have $f_X(x;H_1) = f_X(x-A;H_0) \approx f_X(x;H_0) - Af'_X(x;H_0)$. Hence

$$\{x_i\} \approx \{m_i; f'_X(m_i;H_0) = 0\}, \tag{3.37}$$

which means that the intersections can be approximated as the modes of $f_W(w)$.

□

From Corollary 3.2, we can see that if there is only one intersection, the optimal composite binary TS will reduce to a simple binary TS. If there are multiple intersections, the optimal composite binary TS design will result in better performance than the optimal simple TS design, which will be demonstrated in Sec. 3.5.

3.4 Discussion on Detection with Unknown Noise pdf

In Sec. 3.2, we proposed an indicator of P_D and in Sec. 3.3, we derived the optimal designs of simple binary TS and composite binary TS. These results are obtained assuming that noise pdf is known. In practice, nevertheless, we often need to deal with detection with unknown noise pdf. In this section, we will discuss this practical issue. We classify the noises into two categories.

1. The noises have unimodal and symmetric pdf, which we call *Type I* noise. This type of noise covers a wide range of applications and the assumptions do not appear to be overly restrictive for practical applications [64]. For example, the generalized Gaussian (GG) and unimodal Gaussian mixture (GM) [1] belong to this type. Note that GG (including Gaussian, Laplacian, uniform) and unimodal GM are widely used noise types in a variety of applications [1], [6]–[8], [43], [44].

   For *Type I* noise, as we show in Corollaries 3.1 and 3.2, either using simple binary TS or composite one leads to identical optimal TS with threshold $\tau = w_0 + A/2$. Therefore, assuming that the noise pdf is symmetric and unimodal, the optimal TS design only depends on the mode of the noise pdf and the amplitude of the DC signal, which usually can be estimated with good precision. In particular, if $w_0 = 0$, the TS design does not depend on the noise pdf.

2. Other noises. For noises that do not belong to *Type I* noise, referred to as *Type II* noise, the optimal TS is not as straightforward as that for *Type I* noise. For example, in Sec. 3.5.2, we show an example with a bimodal Gaussian mixture (GM) noise (the pdf has two modes). For these noises, we need to calculate τ_{opt} for simple binary TS using Lemma 3.2, or to determine $\mathcal{D}_{\text{opt}}$ for composite binary TS using Corollary 3.2. For these designs, instead of the full knowledge of the pdf, only the intersections of $f_X(x; H_0)$ and $f_X(x; H_1)$, or approximately, only the modes of $f_W(w)$, are needed as shown in Corollary 3.1 and Corollary 3.2. Intuitively, the less the noise pdf knowledge is needed in design, the higher the robustness of the design to the noise pdf. The modes can be obtained via estimation theory [5], which is easier than obtaining the full knowledge of the pdf needed for the LO detector and Saha's detector in [6].

3.5 Simulation Results

Two examples are presented in this section to illustrate the performance of the proposed TD based on simple binary TS and composite binary TS, respectively. One example is a DC signal detection in unimodal GM noise, and the other one is a DC signal detection in bimodal GM noise. The unimodal GM noise is widely used in signal detection to model underwater noise [7], [8] and DCT coefficients of images [43], [44]. The bimodal GM noise is also widely used in research, especially when investigating the noise-enhanced effect [18], [19]. We use a bimodal GM noise in second example to demonstrate the potential performance improvement of using composite TS over that of simple TS.

3.5.1 Mean-Shift Unimodal Gaussian Mixture Detection

We consider a DC signal detection in unimodal GM noise [1], [6], which is defined as,

$$f_W(w) = \frac{c}{\sigma\sqrt{2\pi}}\left[\alpha \exp\left(-\frac{c^2w^2}{2\sigma^2}\right) + \frac{1-\alpha}{\beta}\exp\left(-\frac{c^2w^2}{2\beta^2\sigma^2}\right)\right], \tag{3.38}$$

where $c = [\alpha + (1-\alpha)\beta^2]^{1/2}, 0 < \alpha < 1, \beta > 0$. A DC signal $A = 0.1$ is embedded in GM noise with $\alpha = 0.9$, $\beta = 5$, and $\sigma = 1$. The number of the observations is $N = 100$.

Since the GM is unimodel and symmetric, there is only one intersection between $f_X(x; H_0)$ and $f_X(x; H_1)$, located at 0.05. For this case, according to Corollary 3.2, the optimal composite TS reduces to the optimal simple TS with $\tau_{\text{opt}} = 0.05$. Using the simple binary TS, we show the theoretical and simulated P_D with respect to τ for different P_{FA} in Fig. 3.3. Theoretical results are calculated based on (3.12) and (3.13). For a given τ and P_{FA}, simulated P_D is calculated from $f_Z(z; H_0)$ and $f_Z(z; H_1)$, which are approximated based on the histograms of the values of Z under H_0 and H_1 generated from the multiple simulations. We can see that when $\tau = 0.05$, the maximum P_D is obtained and is independent of P_{FA}. This provides a justification of Lemma 3.2.

Next, in Fig. 3.4, we show the ROC of the TD with the proposed TS from Monte-Carlo simulations and compare it with the theoretical one calculated from (3.12) and (3.13). Consistency between the simulated ROC and the theoretical one is revealed. The ROC of the LO detector is also shown as the upper bounder since it is the optimal detector for small signals, along with the ROC of the MF used as the lower bounder. To quantify the performance difference, we calculate the average uniform loss (AUL) between two ROCs (detector ι and detector κ) in dB, defined as

$$\text{AUL}_{\kappa}^{\iota} \triangleq \int_0^1 10\log_{10}\left(\frac{P_D^{\iota}}{P_D^{\kappa}}\right) dP_{FA}. \tag{3.39}$$

A positive and large AUL means that the performance of detector ι is significantly better than that of

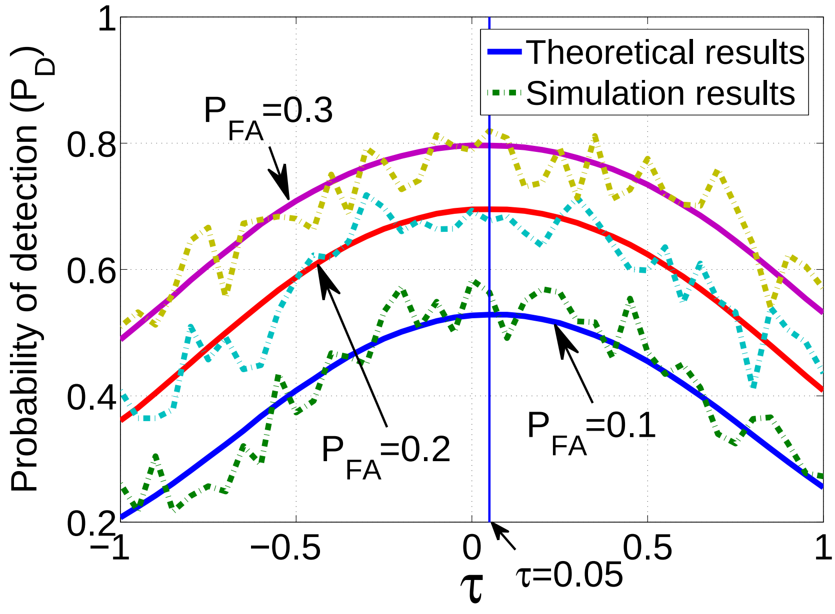


Fig. 3.3. P_D with respect to τ and P_{FA} for detection of a DC signal $A = 0.1$ in GM noise with $\alpha = 0.9$, $\beta = 5$, and $\sigma = 1$.

detector κ. It is shown that the TD with optimal TS (simple binary TS with $\tau = 0.05$) has an AUL of only 0.17dB (4% degradation) compared with the LO detector. Note that the LO detector design needs full knowledge of the noise pdf, while the proposed TD does not require any information of the noise. In addition, the proposed TD has a much lower complexity in implementation. Compared to the proposed TD, the MF has an AUL of 0.595dB (15% degradation).

We note that the P_D curve obtained from simulations has a staircase shape. This is because the pmf's $f_Z(z; H_i), i = 0, 1$ obtained from simulated histograms are discrete. We used the continuous pdf's (3.12) and (3.13) instead of pmf's for the theoretical ROC calculation.

We now compare the computational complexity of deriving the optimal TS from P_D with that of the proposed TS design. Because the GM (unimodal) is *Type I* noise (zero mean), we can obtain the $\tau_{\text{opt}} = A/2 = 0.05$ without any calculation. However, to show the simplicity of Δp with respect to P_D, we use exhaustive search even in the proposed design, as represented in Sec. 3.3. In the exhaustive search of τ, the range [-5,5] is considered with a resolution of 0.01. The average running time (Matlab) is about 0.189 second if P_D is used as the criterion in TS design and 0.034 second if the proposed design is used in a computer with AMD 8450 triple-core processor (2.1GHz) and 4GB memory.

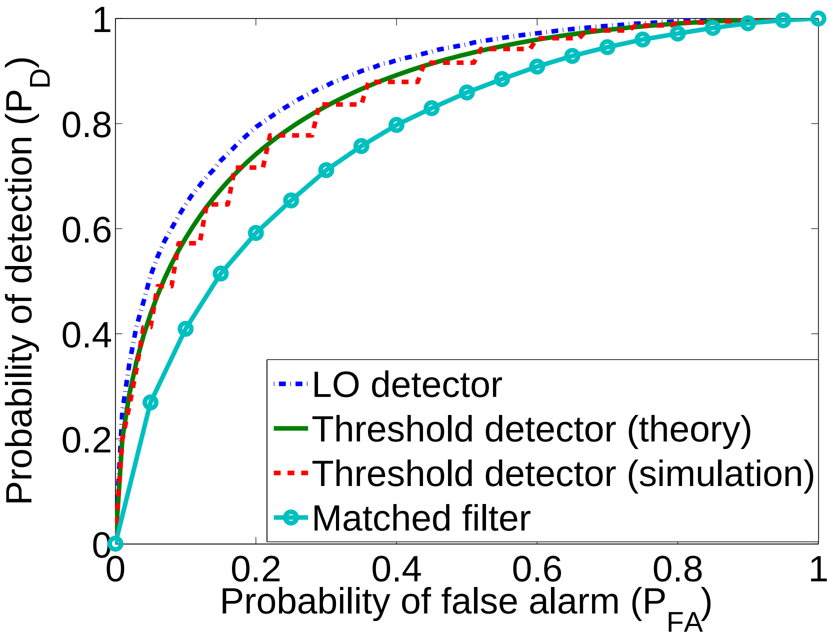


Fig. 3.4. ROCs from LO detector, MF and the proposed TD for detection of a DC signal $A = 0.1$ in GM noise with $\alpha = 0.9$, $\beta = 5$, and $\sigma = 1$.

3.5.2 Mean-Shift Bimodal Gaussian Mixture Detection

In this section, we show the optimal design and ROC of the proposed TD for the detection of a DC signal in a bimodal GM noise, which has the following pdf:

$$f_W(w) = \frac{1}{2}\mathcal{N}(w; \mu, \sigma^2) + \frac{1}{2}\mathcal{N}(w; -\mu, \sigma^2). \tag{3.40}$$

Thus

$$f_X(x; H_0) = f_W(x) = \frac{1}{2}\mathcal{N}(x; \mu, \sigma^2) + \frac{1}{2}\mathcal{N}(x; -\mu, \sigma^2), \tag{3.41}$$

and

$$f_X(x; H_1) = \frac{1}{2}\mathcal{N}(x; \mu + A, \sigma^2) + \frac{1}{2}\mathcal{N}(x; -\mu + A, \sigma^2), \tag{3.42}$$

where A is the DC signal. For $\mu = 3, \sigma^2 = 1, A = 0.1$, two modes of this GM noise can be calculated as $w_1 = -3$ and $w_2 = 3$, and the three intersections are $x_1 = -2.95, x_2 = 0.05, x_3 = 3.05$.

We first use simple binary TS and calculate its τ_{opt}. For a given τ, we have

$$\Delta p(\tau) = \frac{1}{2}\left[Q\left(\frac{\tau - \mu - A}{\sigma}\right) - Q\left(\frac{\tau - \mu}{\sigma}\right) + Q\left(\frac{\tau + \mu - A}{\sigma}\right) - Q\left(\frac{\tau + \mu}{\sigma}\right)\right],$$

and $\Delta p(\tau = x_1) = \Delta p(\tau = x_3) = 0.0199$, $\Delta p(\tau = x_2) = 0.0004$. Therefore $\tau_{\text{opt}} = x_1$ or

$\tau_{\rm opt} = x_3$.

We then use composite binary TS. According to Corollary 3.2, we have $\mathcal{D}_{\rm opt} = \{x; -2.95 \leq x \leq 0.05 \text{ and } 3.05 \leq x\}$. Thus, the optimal composite TS can be represented as

$$\mathrm{TS}_{\rm opt}: \quad Y = \begin{cases} 1 & x \in [-2.95, 0.05] \cup [3.05, \infty) \\ 0 & \text{elsewhere} \end{cases}. \tag{3.43}$$

The Δp of this $\mathrm{TS}_{\rm opt}$ is 0.0394, which is calculated by

$$\Delta p = \frac{1}{2}\sum_{i=1}^{3}(-1)^{i-1}\left[Q\left(\frac{x_i-\mu-A}{\sigma}\right) - Q\left(\frac{x_i-\mu}{\sigma}\right) + Q\left(\frac{x_i+\mu-A}{\sigma}\right) - Q\left(\frac{x_i+\mu}{\sigma}\right)\right].$$

In Fig. 3.5, we show the ROCs obtained from simulations for the TD with the optimal simple TS, the TD with the optimal composite TS, the LO detector, and the MF. It reveals that the TD with the optimal composite TS performs very close to the LO detection, only has an AUL of 0.1553 (3.64% degradation). The TD with the optimal simple TS has simpler implementation than the TD with the optimal composite TS, but the performance degradation is obvious, with an AUL of 0.7304dB (18.32% degradation). Compared to the MF, both TDs have superior performance. Compared to the TD with optimal composite TS, the MF has an AUL of 1.4972dB (41.16% degradation). Compared to the TD with the optimal simple TS, the MF has an AUL of 0.7668dB (19.31% degradation).

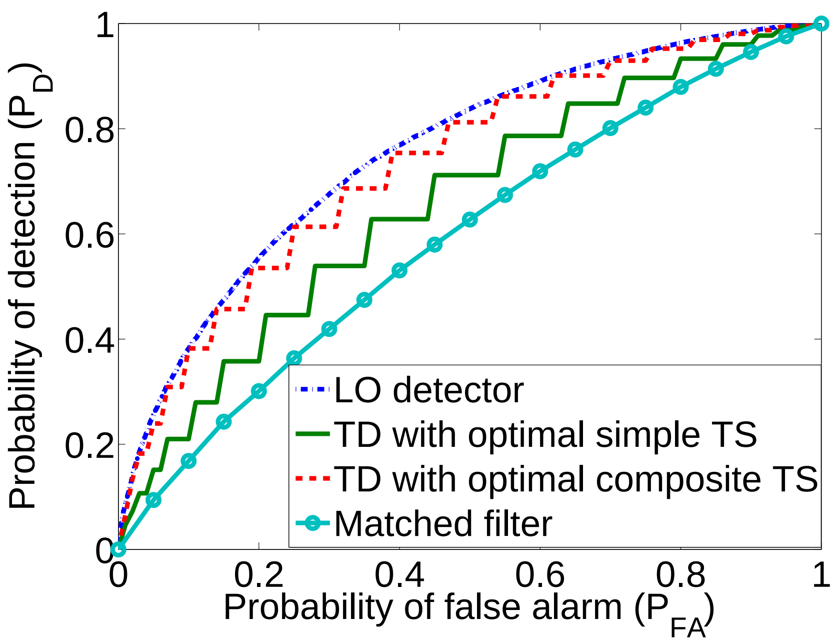


Fig. 3.5. ROCs from LO detector, MF and the proposed TD in the detection of a DC signal $A = 0.1$ in the bimodal GM noise with $\mu = 3$, $\sigma = 1$, and $N = 100$.

In the above two examples, the proposed designs are compared with the LO detector to show that its performance is close to optimal. They are not compared with the optimal LRT. But note that for weak signal detection, the LRT and the LO detector have almost the same performance [1], [2].

3.6 Conclusions

In this chapter, we considered the optimal TD design for detecting a known DC signal in non-Gaussian noise with known pdf. Under NP criterion, we showed that P_D can be represented by an indicator Δp in the optimal design of the TS. We investigated two types of TS's: simple binary TS and composite binary TS. For simple binary TS, using the indicator Δp, we derived an easy and fast way to calculate the optimal TS threshold. For composite binary TS, the otpimal TS structure was derived, also with the help of Δp. Experimental results show the validity of the proposed TS designs. The performance of the proposed TDs are shown to be superior to the MF, for non-Gaussian noises with heavy pdf tails, and can perform very close to the LO detector, with a much simpler implementation than the LO detector.

Chapter 4

A Robust Detector of Known Signal in Non-Gaussian Noise Using Threshold Systems

In this chapter, we propose a TD for detecting a known deterministic signal in independent non-Gaussian noise whose pdf is unknown but is symmetric and unimodal. The optimality of the proposed TD is proved under the assumptions of white noise, small signal, and a large number of samples. While previous TD designs need accurate information of the noise pdf, the proposed TD is independent of the noise pdf, and thus is robust to the noise pdf. The detection probability and the ROC of the proposed TD are analyzed both theoretically and numerically. It is shown that even without knowing the noise pdf, the proposed TD has close performance to the optimal detector designed with the noise pdf information. It also performs significantly better than the MF when the noise pdf has heavy tails. The practical implementation, robustness to both the noise pdf and the signal, and region of validity of the proposed TD are also investigated.

4.1 Introduction

The detection of a known deterministic signal in unknown non-Gaussian noise is a problem of great interest in many fields, such as communications and image processing [1], [2]. For example, in watermark detection in discrete cosine transform (DCT) domain, the signal is the watermark (or a signature), which is usually known [1], while the DCT coefficients of an image is the noise, whose pdf is non-Gaussian and unknown in general [43], [44], [62]. Other applications include the

[1]The watermark sequence is unknown for blind watermark. However, when extracting one watermark bit in the watermark sequence, this watermark bit is usually known signal. For example, for binary watermark, the watermark bit is either 0 or 1, which is known.

feature extraction in images [65] and underwater communications, where we can have precise signal information or obtain a reliable estimation of the signal, while the noise pdf is non-Gaussian and is difficult to estimate [6]–[8].

Consider the BHTP described in (1.1). As stated in Sec. 1.4, since LRT or GLRT results in complex detector with low robustness and poor performance when the knowledge of the signal and noise is imprecise, we resort to TD for a suboptimal solution. We reviewed the existing TD's in Sec. 2.2.1 and proposed optimal TD designs for DC signal detection in known non-Gaussian noise in Chap. 3.

However, the aforementioned TD and the LO detector need full knowledge of the noise pdf. This makes its application limited. They cannot be used for applications where the noise pdf cannot be obtained perfectly. Furthermore, the assumption on perfect noise pdf information causes problems in simplicity, robustness, and detectability when applying the TDs in real systems. Even though the TS can be implemented simply, the optimal parameters of the TS need to be calculated from the noise pdf, which adds complexity to the TD. The robustness has also been influenced because the TD design depends on the noise pdf and the signal. As to the detectablity, degradation cannot be avoid because practically speaking we only have imprecise noise pdf. These weakness limits the applications of these TDs. We will address these problems in this chapter.

To the best of our knowledge, there is no previous work on TD for detecting an arbitrary signal in non-Gaussian noise with unknown pdf, which is the focus of this chapter. Our goal is to find a robust and low complexity detector that also enjoys near-optimal performance. We consider detection problems where the noise pdf is unimodal and symmetric. This covers a wide range of noise pdfs, for example, the Gaussian mixture (GM) and the generalized Gaussian (GG). We propose a TD composed of a binary TS array and a linear correlator that is independent of the noise pdf. The optimality of the design is analyzed for the case of white noise, small signal, and a large number of samples. Simulation results show that the proposed TD performs very close to the LO detector and is much better than the MF for noises with heavy pdf tails. Properties of the proposed detector such as robustness, complexity, and region of validity are also investigated.

The remainder of the chapter is organized as follows. In Sec. 4.2, we present the detection problem and the proposed TD structure. The optimality of the TD design including the binary TS and the linear correlator is also proved. In Sec. 4.3, we derive the detection probability of the proposed TD and present simulation results. In Sec. 4.4, we discuss the robustness, implementation complexity, and region of validity of the proposed TD. Finally, we draw conclusions in Sec. 4.5. Involved proofs are included in the appendices.

4.2 Problem Statement and Proposed Detector Structure

In this section, we explain the detection problem and the proposed detector.

4.2.1 Problem Statement

Consider the detection problem described in (1.1). The signal $s[n]$'s are assumed to be known. The noise is assumed unknown but subject to the following constraints: 1) $f_W(w)$ is symmetric about $w = 0$, i.e., $f_W(-w) = f_W(w)$; 2) $f_W(w)$ is unimodal; and 3) $f_W(w)$ is continuous. Thus, $f_W(w)$ has a unique maximum at $w = 0$; and is non-decreasing when $w < 0$, non-increasing when $w > 0$. The above assumptions are not overly restrictive for practical applications [64]. For example, for underwater communication and DCT-domain watermarking mentioned in Sec. 4.1, the noise pdfs, although unknown, are shown to satisfy these assumptions [7], [8], [43], [44]. Furthermore, we also assume that compared with the noise standard deviation σ, the signal is weak [1], i.e., $|s[n]| \ll \sigma$ for $n = 0, 1, ..., N-1$.

Because $f_W(w)$ is unknown and sometimes ever-changing, the optimal and LO detectors, which require the noise pdf information, cannot be realized. We aim at designing a detector whose parameters are independent of the noise pdf, thus robust to the noise pdf; but at the same time, good detection probability is desired.

4.2.2 Proposed TD Structure

The proposed TD is shown in Fig. 4.1, in which each data point $x[n]$ is separately quantized at the threshold $s[n]/2$ (by a TS with binarization threshold $s[n]/2$) to yield a binary $y[n]$. Then linear correlation is performed between the sequence $y[n]$ and the absolute value of the signal to be detected, i.e., $|s[n]|$. A decision is made via comparing the correlation result Z with the threshold η.

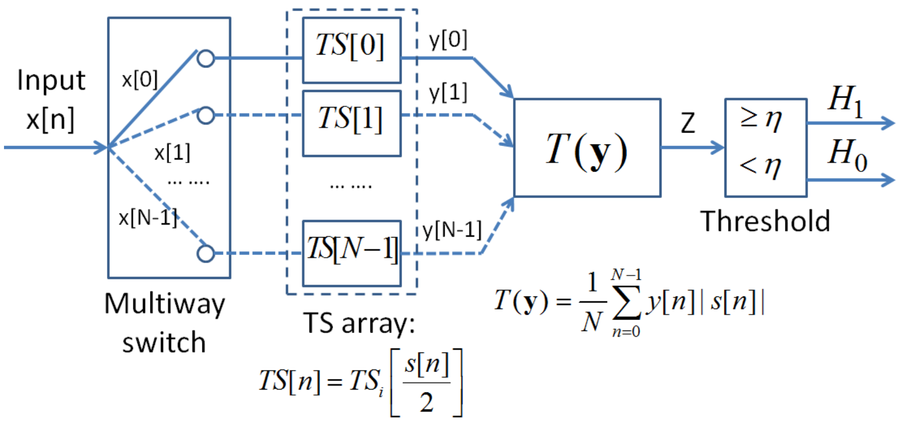


Fig. 4.1. Proposed TD structure.

Here are more detailed explanations on different parts of the proposed TD. The multiway switch directs the observation $x[n]$ to its corresponding TS, denoted as $TS[n]$, and outputs $y[n]$. For simplicity in implementation, we use the simple binary TS shown in (2.1), denoted as $\text{TS}_1(\tau)$, and the complemental binary TS shown in (3.1), denoted as $\text{TS}_2(\tau)$, where τ is the threshold in TS. We use $\text{TS}_1(\tau)$ when $s[n] \geq 0$, and use $\text{TS}_2(\tau)$ otherwise. The binarization threshold is set to be $s[n]/2$ for both cases, i.e., $\tau = s[n]/2$. The linear correlator produces

$$Z = T^{TD}(\mathbf{y}) = \frac{1}{N} \sum_{n=0}^{N-1} y[n]|s[n]|. \tag{4.1}$$

The threshold η of the test statistic is calculated from the desired false alarm level using $\int_{\eta}^{\infty} f_Z(z; H_0)dz = P_{FA}$.

4.2.3 Optimality of the Proposed TD

In this subsection, we prove the optimality of the binarization threshold of the TS and the correlator design in (4.1).

Our TS design problem can be stated as follows: for the one-dimensional binary detection problem where $x[n]$ is the observation and $y[n]$ (output of the $TS[n]$) is the detection result, find the optimal binarization threshold. For the optimality measure, NP criterion is widely used, i.e., to find the maximum detection probability P_D for a given P_{FA}. However, with the NP criterion, the threshold optimization requires the noise pdf, which is unknown in our model. Furthermore, for different values of P_{FA}, the optimal threshold will be different, which complicates the implementation of the TD. Therefore, in this chapter, we use the area under the receiver operating characteristic (ROC) curve, denoted as AUC, as the optimality measure [66]. It fully characterizes the detectability of a detector and is shown to be a good detectability measure according to Area Theorem [67]. Most importantly, it results in an optimal binarization threshold independent of P_{FA}.

We first prove the optimality of the optimal binarization threshold $s[n]/2$.

Theorem 4.1. For each observation $x[n]$, consider the binary detection problem with the binary TS output $y[n]$ as the test statistic. The binarization threshold that leads to the maximum AUC is $\tau = s[n]/2$.

Proof. See Appendix 4.A. □

It is noteworthy that the proposed binarization threshold is optimal only for the single observation $x[n]$ and may not be globally optimal for the overall detection problem. As will be shown later, however, this design leads to a robust TD and close-to-optimal performance.

Next, we prove the optimality of the test statistic in (4.1).

Theorem 4.2. Consider the detection problem with observation $\mathbf{y}$, the outputs of the TS array in the proposed TD. When $|s[n]| \ll \sigma$ and $N \gg 1$, the test statistic $T^{TD}(\mathbf{y})$ in (4.1) is the optimal LRT.

Proof. See Appendix 4.B. □

The proposed TD has simple structure, which is easy in implementation. Theorems 4.1 and 4.2 in addition show its advantage in performance.

4.3 Investigation on Detection Probability

In this section, the detection probability of the proposed TD is analyzed and simulation results are presented. The performance of previously proposed detectors is also investigated for comparison.

4.3.1 Detection Probability of the Proposed TD

To obtain the P_D of the proposed TD for a given P_{FA}, we need to derive the pdfs of $T^{TD}(\mathbf{y})$ under H_1 and H_0. $T^{TD}(\mathbf{y})$ is the summation of N independent random variables $y[n], n = 0, 1, ...N-1$ with a weight of $|s[n]|$. Because $y[n]$'s are independent, according to the central limit theorem, when N is large, the distribution of $T^{TD}(\mathbf{y})$ can be approximated as Gaussian. Therefore, we only need to calculate the mean and the variance of $T^{TD}(\mathbf{y})$.

Denote the probabilities of $y[n] = 1$ under Hypothesis H_0 and H_1, respectively, as

$$p_0[n] \triangleq \mathbb{P}(y[n] = 1; H_0), \quad p_1[n] \triangleq \mathbb{P}(y[n] = 1; H_1). \tag{4.2}$$

It can be derived that (see Appendix 4.B for the calculations)

$$p_0[n] \approx \tfrac{1}{2} - \tfrac{|s[n]|}{2} f_W(0), \quad p_1[n] \approx \tfrac{1}{2} + \tfrac{|s[n]|}{2} f_W(0). \tag{4.3}$$

The probability of $y[n] = 0$ under Hypothesis H_i is $1 - p_i[n]$ for $i = 0, 1$. The mean and the variance of $y[n]$ under Hypothesis H_i are

$$\mathbb{E}\{y[n]; H_i\} = 1 \cdot p_i[n] + 0 \cdot (1 - p_i[n]) = p_i[n], \tag{4.4}$$

$$\mathrm{Var}\{y[n]; H_i\} = 1^2 \cdot p_i[n] + 0^2 \cdot (1 - p_i[n]) - \mathbb{E}^2\{y[n]; H_i\} = p_i[n](1 - p_i[n]), \tag{4.5}$$

where $\mathbb{E}$ stands for the expectation. With (4.4), (4.5), and the independence of $y[n]$'s, the mean and variance of $T^{TD}(\mathbf{y})$ are

$$\mu_i = \mathbb{E}\{T^{TD}(\mathbf{y}); H_i\} = \mathbb{E}\left\{\frac{1}{N}\sum_{n=0}^{N-1} y[n]|s[n]|\right\} = \frac{1}{N}\sum_{n=0}^{N-1} p_i[n]|s[n]|, \tag{4.6}$$

$$\sigma_i^2 = \mathrm{Var}\{T^{TD}(\mathbf{y}); H_i\} = \mathrm{Var}\left\{\frac{1}{N}\sum_{n=0}^{N-1} y[n]|s[n]|\right\} = \frac{1}{N^2}\sum_{n=0}^{N-1} s^2[n]p_i[n](1-p_i[n]). \quad (4.7)$$

Thus,

$$f_Z(z; H_0) = \mathcal{N}(z; \mu_0, \sigma_0^2), \quad f_Z(z; H_1) = \mathcal{N}(z; \mu_1, \sigma_1^2), \quad (4.8)$$

where $\mathcal{N}(z; \mu, \sigma^2)$ is the Gaussian distribution with mean μ and variance σ^2. The pdfs in (4.8) are the same as those for a standard Gaussian binary detection problem. It can be derived that for a given P_{FA}, the P_D of the proposed TD can be expressed as

$$P_D^{TD} = Q\left(\frac{\sigma_0 Q^{-1}(P_{FA}) + \mu_0 - \mu_1}{\sigma_1}\right), \quad (4.9)$$

where $Q(\cdot)$ is the complementary cumulative distribution function (cdf) of the standard normal pdf and $Q^{-1}(\cdot)$ is the inverse of $Q(\cdot)$.

For a DC signal, i.e., $s[n] = A, n = 0, 1, ..., N-1$, $p_i[n]$'s are identical, denoted as p_i. We have $\mu_i = A \cdot p_i$ and $\sigma_i^2 = A^2 \cdot p_i \cdot (1-p_i)/N$. The result in (4.9) reduces to the one in (3.27).

4.3.2 Detection Probability of the MF and the LO Detector

If the MF is used, which is optimal for Gaussian noise, the test statistic is given by

$$T^{MF}(\mathbf{x}) = \frac{1}{N}\sum_{n=0}^{N-1} s[n]x[n].$$

Again, according to the central limit theorem, $T^{MF}(\mathbf{x})$ is Gaussian. We have

$$\mathbb{E}\left\{T^{MF}(\mathbf{x}); H_0\right\} = \frac{1}{N}\mathbb{E}\left\{\sum_{n=0}^{N-1} s[n]w[n]\right\} = 0.$$
$$\mathrm{Var}\left\{T^{MF}(\mathbf{x}); H_0\right\} = \frac{1}{N^2}\mathrm{Var}\left\{\sum_{n=0}^{N-1} s[n]w[n]\right\} = \frac{\sigma^2}{N^2}\sum_{n=0}^{N-1} s^2[n].$$
$$\mathbb{E}\left\{T^{MF}(\mathbf{x}); H_1\right\} = \frac{1}{N}\mathbb{E}\left\{\sum_{n=0}^{N-1} s[n](s[n]+w[n])\right\} = \frac{1}{N}\sum_{n=0}^{N-1} s^2[n].$$
$$\mathrm{Var}\left\{T^{MF}(\mathbf{x}); H_1\right\} = \frac{1}{N^2}\mathrm{Var}\left\{\sum_{n=0}^{N-1} s[n](s[n]+w[n])\right\} = \frac{\sigma^2}{N^2}\sum_{n=0}^{N-1} s^2[n].$$

Therefore, for a given P_{FA}, the P_D of the MF can be expressed as

$$P_D^{MF} = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{\sum_{n=0}^{N-1} s^2[n]}{\sigma^2}}\right). \quad (4.10)$$

For comparison, we also present the detection probability of the LO detector [1]:

$$P_D^{LO} = Q\left(Q^{-1}(P_{FA}) - \sqrt{I\sum_{n=0}^{N-1} s^2[n]}\right), \tag{4.11}$$

where

$$I = \int_{-\infty}^{\infty} \frac{\left(\frac{df_W(w)}{dw}\right)^2}{f_W(w)} dw. \tag{4.12}$$

4.3.3 Simulation Results

In this section, we present and compare the P_D's and the ROCs of the proposed TD, the MF, the LO detector, and the detector in [6] via simulation. Note that the performance of the MF is utilized as the lower bound benchmark and the one of the LO detector as the upper bound benchmark.

We consider two signals:

1) sinusoidal signal

$$s[n] = A\sin(0.02\pi n + \phi), \tag{4.13}$$

where $A = 0.1$ and $\phi = 0$ unless otherwise stated;

2) a randomly generated signal

$$s[n] = B \cdot l[n], \tag{4.14}$$

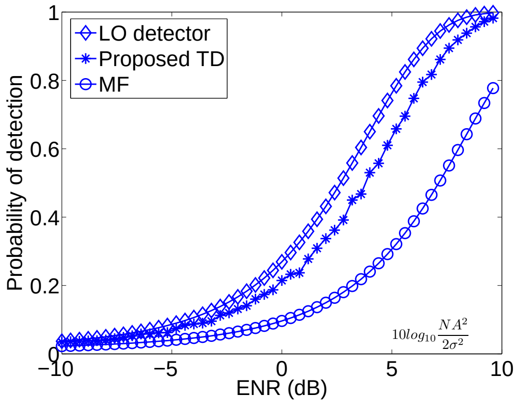
where $l[n]$'s are randomly generated by the uniform distribution on $[0, 1]$ and $B = 0.1$ unless otherwise stated. The sinusoidal signal is largely used in signal detection. The random signal is used to model a known signal with an arbitrary structure. We set $N = 100$.

As to the noise, GM noise and GG noise are considered because they are widely used in practical applications such as underwater noise [6]–[8] and DCT coefficients [43], [44]. The GM pdf has three parameters α, β, and σ, and is defined in (3.38). The GG pdf has two parameters β and σ, and is defined as

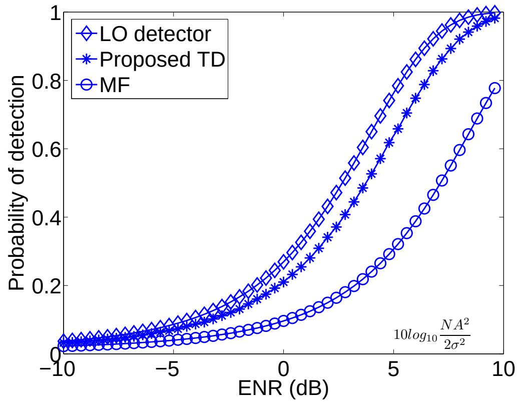
$$f_W(w) = \frac{c_1(\beta)}{\sigma}\exp\left(-c_2(\beta)\left|\frac{w}{\sigma}\right|^{\frac{2}{1+\beta}}\right), \tag{4.15}$$

where $c_1(\beta) = \frac{\Gamma^{1/2}(\frac{3}{2}(1+\beta))}{(1+\beta)\Gamma^{3/2}(\frac{1}{2}(1+\beta))}$, $c_2(\beta) = \left[\frac{\Gamma(\frac{3}{2}(1+\beta))}{\Gamma(\frac{1}{2}(1+\beta))}\right]^{\frac{1}{1+\beta}}$ and $\Gamma(\cdot)$ is the Gamma function. We set $(\alpha, \beta, \sigma) = (0.3, 5, 1)$ for GM noise and $(\beta, \sigma) = (0.9, 1)$ for GG noise unless otherwise stated.

First we compare the theoretical results on P_D for both the sinusoid signal (4.13) and the random signal (4.14) in the GM noise (3.38). P_{FA} is set to be 0.01. We show the detection probabilities of the proposed TD and other schemes for different energy-to-noise ratios (ENRs) defined as $10\log_{10}\left(\sum_{n=0}^{N-1} s^2[n]/\sigma^2\right)$ dB. By having the magnitude A of the sinusoidal signal range from 0.045 to 0.45, and the magnitude B of the random signal range from 0.055 to 0.55, the ENR ranges from -10dB to 10dB. The P_D's of the proposed TD, the MF, and the LO detector are calculated using (4.9), (4.10), and (4.11), respectively. Note that for the proposed TD, when calculating P_D^{TD} using (4.9), $\mu_i, \sigma_i, i = 0, 1$ are calculated by (4.6) and (4.7), in which $p_i[n], i = 0, 1$ are calculated according to (4.3). For the LO detector, the value of I given in (4.12) is calculated numerically. From Fig. 4.2 we can see that the proposed TD has close performance to the LO detector and is significant better than the MF for both signals. At $P_D = 0.3$, the proposed TD is about 4dB better than the MF, and is only 1dB worse than the LO detector. The advantage of the proposed TD over the MF is even bigger at higher P_D levels.



(a) Sinusoidal signal



(b) Random signal

Fig. 4.2. P_D versus ENR of the proposed TD, the MF, and the LO detector in GM noise with $\alpha = 0.3$, $\beta = 5$, and $\sigma = 1$ for $N = 100$ and $P_{FA} = 0.01$.

We now compare the theoretical ROCs calculated by (4.9) with the ROCs obtained from Monte Carlo simulations for the proposed TD. To obtain the ROC from Monte Carlo simulations, observations are generated under H_0 and H_1 (20000 times for each) and the test statistic is calculated as $Z = T^{TD}(\mathbf{y})$ for all observations. $f_Z(z; H_i), i = 0, 1$ are approximated as the normalized histograms of the 20000 outputs for both hypothesis, from which the ROC is generated. We use the sinusoid signal (4.13) and the random signal (4.14) in GM noise and GG noise, respectively. Fig. 4.3 shows that the ROCs calculated from (4.9) are consistent with the ROCs obtained from simulations for all cases.

Next, we compare the ROCs of the proposed TD, the MF, the LO detector, and the detector in [6] called Saha's detector, obtained from Monte Carlo simulations in Fig. 4.4. Saha's detector has the detection structure in Fig. 1.3, in which $g(x)$ is designed as a three-level TS and a quadrature MF is used instead of the replica-correlator due to unknown parameters in signal. The optimal design of the three-level TS requires noise pdf information, thus this detector can only be used when the noise pdf is available. Here, we change the quadrature MF to linear MF, which is a better design for the detection of known signals, and to make it comparable with the proposed TD. The optimal thresholds of the 3-level TS in [6] are numerically calculated to be 0.08 for the GM noise and 0.01

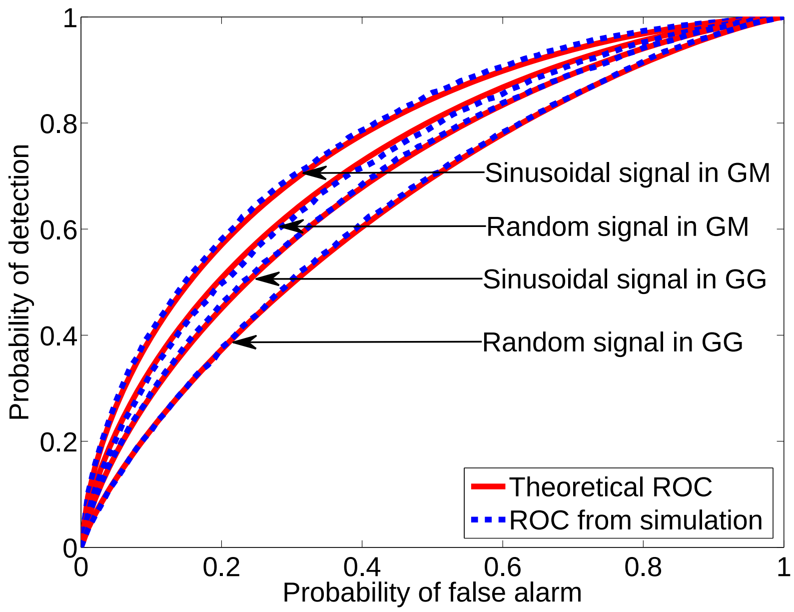
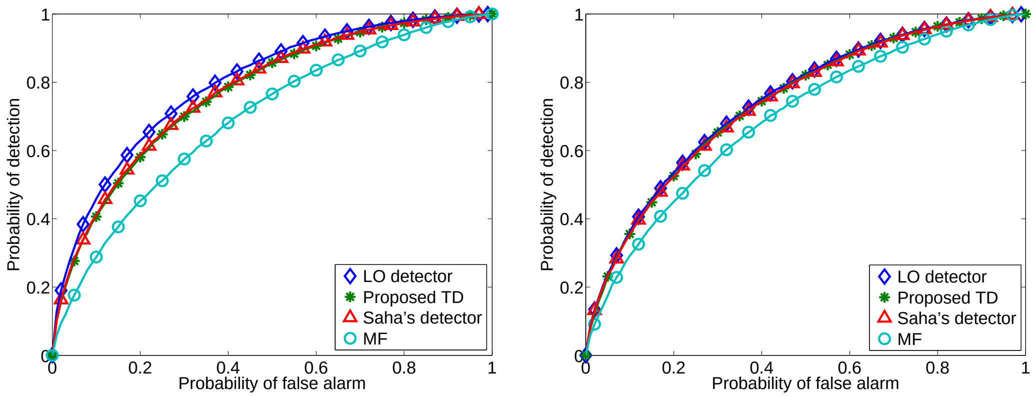


Fig. 4.3. Comparison between theoretical and simulated ROCs for the proposed TD. GM is with $\alpha = 0.3, \beta = 5, \sigma = 1$; GG is with $\beta = 0.9, \sigma = 1$; and $N = 100$.

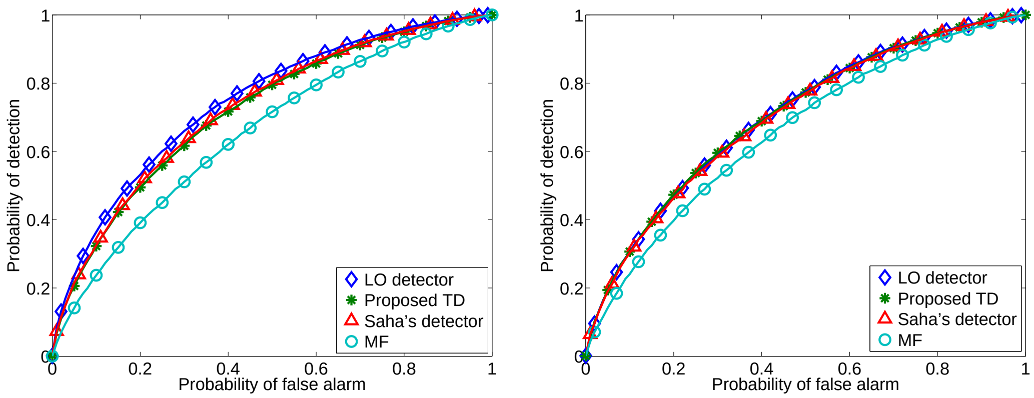
for the GG noise. We show the ROCs for the detection of the sinusoidal signal (4.13) in GM noise in Fig. 5(a) and GG noise in Fig. 5(b), respectively. It can be seen that the ROCs of the proposed TD are close to those of the LO detector and Saha's detector, especially for the GG noise, although it requires no noise pdf information. The proposed TD has a significant improvement compared to the MF. For the sinusoidal detection in GM noise, compared with the LO detector, the proposed TD has an AUL of 0.195dB (4.59% degradation). Compared with the proposed TD, the MF has an AUL of 0.642dB (15.93% degradation). Compared with the detector in [6], the proposed TD has an AUL of 0.0265dB (0.61% degradation). For the sinusoidal detection in GG noise, compared with the LO detector, the proposed TD has an AUL of 0.0493dB (1.14% degradation) only. Compared with the proposed TD, the MF has an AUL of 0.386dB (9.29% degradation). Compared with Saha's detector, the proposed TD has an AUL of 0.0059dB (0.14% degradation). It should be mentioned that the closeness of the proposed TD to the optimal detector varies with the noise pdf parameters. In general, the proposed TD performs closer to the optimal detector when the noise pdf has heavier tails.

In the next experiment, we compare the ROCs for the detection of the random signal (4.14) in GM noise and GG noise. The ROCs are shown in Fig. 4.5. Again we can see that the proposed TD has close performance to the LO detector and Saha's detector, and is a lot better than the MF. For the detection in GM noise, compared with the LO detector, the proposed TD has an AUL of 0.199dB (4.69% degradation). Compared with Saha's detector, the proposed TD has an AUL of



(a) GM noise with $\alpha = 0.3, \beta = 5, \sigma = 1$ (b) GG noise with $\beta = 0.9, \sigma = 1$

Fig. 4.4. Comparison of the ROCs obtained from Monte Carlo simulations in detecting the sinusoid signal (4.13) in GM and GG noise, $N = 100$.



(a) GM noise with $\alpha = 0.3, \beta = 5, \sigma = 1$ (b) GG noise with $\beta = 0.9, \sigma = 1$

Fig. 4.5. Comparison of the ROCs in detecting the random signal (4.14) in GM and GG noise, $N = 100$.

0.04dB (0.93% degradation). Compared with the proposed TD, the MF has an AUL of 0.597dB (14.74% degradation). For the detection in GG noise, compared with the LO detector, the proposed TD has an AUL of 0.016dB (0.37% degradation) only. Compared with Saha's detector, the proposed TD has an AUL of -0.031dB (0.72% increase). Compared with the proposed TD, the MF has an AUL of 0.41dB (9.9% degradation).

4.4 Discussions

In this section, we discuss properties of the proposed TD, including the robustness, implemental complexity, and region of validity.

4.4.1 Robustness

The robustness in signal detection refers to the stability of the detection performance to changes in parameters in the system. For example, we can observe from (4.11) and (4.12) that the detection probability of the LO detector depends on P_{FA}, $s[n]$, the form of $f_W(w)$ and its parameters. Let $\rho_m, m = 1, ..., M$ be the set of parameters involved in the robustness evaluation. In this paper, we use a quantitative measure proposed in [68] to evaluate the robustness, which is defined as

$$\Phi \triangleq \left(1 + \sum_{m=1}^{M} \left(\frac{\partial P_D}{\partial \rho_m}\right)^2\right)^{-\frac{1}{2}}. \tag{4.16}$$

It reflects how P_D fluctuates with changes/inaccuracy in ρ_m's. It is normalized to be between 0 and 1. A lower Φ means lower robustness. Note that we assume unknown noise pdf in the design of the proposed TD. But as shown in (4.16), we can only evaluate the quantitative robustness with respect to a specific noise pdf form. Thus, in what follows, we consider GM noise and calculate the robustness of the detectors with respect to the pdf form of GM noise.

We derive the expressions of Φ for the proposed TD, the MF, and the LO detector in the case of the sinusoidal signal (4.13) in GM noise. We consider the robustness with respect to inaccuracy in the noise parameters (α, β, σ) and the signal parameters (A, ϕ), respectively. Analysis on the robustness with respect to other factors, such as the noise pdf form, is more involved and left for future work. We define

$$\Phi_n \triangleq \left(1 + \sum_{\rho_n = \{\alpha,\beta,\sigma\}} \left(\frac{\partial P_D}{\partial \rho_n}\right)^2\right)^{-\frac{1}{2}}, \tag{4.17}$$

which is the robustness to the noise parameters $\rho_n = \{\alpha, \beta, \sigma\}$; and

$$\Phi_s \triangleq \left(1 + \sum_{\rho_s = \{A,\phi\}} \left(\frac{\partial P_D}{\partial \rho_s}\right)^2\right)^{-\frac{1}{2}}, \tag{4.18}$$

which is the robustness to the signal parameters $\rho_s = \{A, \phi\}$.

For the proposed TD, P_D^{TD} depends on $p_i[n]$ via μ_i, σ_i^2 for $i = 0, 1$ as shown in (4.6) and (4.7). From (4.3), we can see that $p_i[n]$ only depends on the signal $s[n]$ and the value of the noise pdf at 0. Using (4.3), (4.6), and (4.7), we obtain

$$\mu_0 \approx \frac{1}{2N} \sum_{n=0}^{N-1} |s[n]|[1 - |s[n]| f_W(0)], \quad \sigma_0^2 \approx \frac{1}{4N^2} \sum_{n=0}^{N-1} s^2[n][1 - s^2[n] f_W^2(0)],$$
$$\mu_1 \approx \frac{1}{2N} \sum_{n=0}^{N-1} |s[n]|[1 + |s[n]| f_W(0)], \quad \sigma_1^2 \approx \frac{1}{4N^2} \sum_{n=0}^{N-1} s^2[n][1 - s^2[n] f_W^2(0)] \approx \sigma_0^2.$$

Using these approximation in (4.9), we have

$$P_D^{TD} \approx Q\left(Q^{-1}(P_{FA}) - \frac{2vf_W(0)}{\sqrt{v-uf_W^2(0)}}\right), \tag{4.19}$$

where we have defined

$$v \triangleq \sum_{n=0}^{N-1} s^2[n] \quad u \triangleq \sum_{n=0}^{N-1} s^4[n]. \tag{4.20}$$

For any parameter $\rho_n \in \{\alpha, \beta, \sigma\}$, we have

$$\frac{\partial P_D^{TD}}{\partial \rho_n} = \sqrt{\frac{2}{\pi}}\left[v-uf_W^2(0)\right]^{-\frac{3}{2}} v^2 \exp\left[-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \frac{2vf_W(0)}{\sqrt{v-uf_W^2(0)}}\right)^2\right]\frac{\partial f_W(0)}{\partial \rho_n}. \tag{4.21}$$

From (3.38), we obtain $f_W(0) = \frac{c}{\sigma\sqrt{2\pi}}\left[\alpha + \frac{1-\alpha}{\beta}\right]$, where $c = \sqrt{\alpha + (1-\alpha)\beta^2}$. Thus

$$\frac{\partial f_W(0)}{\partial \alpha} = \frac{c}{\sigma\sqrt{2\pi}}\left(1 - \frac{1}{\beta}\right) + \left(\frac{1-\beta^2}{2\big(\alpha + (1-\alpha)\beta^2\big)}\right) f_W(0), \tag{4.22}$$

$$\frac{\partial f_W(0)}{\partial \beta} = \frac{c}{\sigma\sqrt{2\pi}}\left(-\frac{1-\alpha}{\beta^2}\right) + \left(\frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta^2}\right) f_W(0), \tag{4.23}$$

$$\frac{\partial f_W(0)}{\partial \sigma} = -\frac{1}{\sigma} f_W(0). \tag{4.24}$$

Using (4.21)-(4.24) in (4.17), we can obtain Φ_n^{TD}.

For the robustness to the signal, for $\rho_s = \{A, \phi\}$, we have

$$\frac{\partial P_D^{TD}}{\partial \rho_s} = -\frac{f_W(0)}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \frac{2vf_W(0)}{\sqrt{v-uf_W^2(0)}}\right)^2\right]\frac{[v-2uf_W^2(0)]\frac{\partial v}{\partial \rho_s} + vf_W^2(0)\frac{\partial u}{\partial \rho_s}}{(v-uf_W^2(0))^{\frac{3}{2}}}, \tag{4.25}$$

where

$$\frac{\partial v}{\partial A} = 2A\sum_{n=0}^{N-1}\sin^2(0.02\pi n + \phi), \quad \frac{\partial u}{\partial A} = 4A^3\sum_{n=0}^{N-1}\sin^4(0.02\pi n + \phi), \tag{4.26}$$

$$\frac{\partial v}{\partial \phi} = A^2\sum_{n=0}^{N-1}\sin(0.04\pi n + 2\phi), \quad \frac{\partial u}{\partial \phi} = 4A^4\sum_{n=0}^{N-1}\sin^3(0.02\pi n + \phi)\cos(0.02\pi n + \phi). \tag{4.27}$$

Using (4.26) and (4.27) in (4.25) and using (4.25) in (4.18), we can obtain Φ_s^{TD}.

For the MF, the detection probability is shown in (4.10), which depends on the noise parameter σ^2 and the signal parameters $\rho_s = \{A, \phi\}$. It can be shown that

$$\frac{\partial P_D^{MF}}{\partial \sigma} = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \frac{\sqrt{v}}{\sigma}\right)^2\right)\left(-\frac{\sqrt{v}}{\sigma^2}\right). \tag{4.28}$$

$$\frac{\partial P_D^{MF}}{\partial \rho_s} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \frac{\sqrt{v}}{\sigma}\right)^2\right)\left(-\frac{1}{2\sigma\sqrt{v}}\right)\left(\frac{\partial v}{\partial \rho_s}\right). \quad (4.29)$$

Φ_n^{MF} can be calculated by using (4.28) in (4.17). Φ_s^{MF} can be calculated by using (4.26) and (4.27) in (4.29) and using (4.29) in (4.18).

For the LO detector, P_D relies on **s** and I defined in (4.12), and I is influenced by all parameters α, β, σ in (3.38). For $\rho_n = \{\alpha, \beta, \sigma\}$, we have

$$\frac{\partial P_D^{LO}}{\partial \rho_n} = -\frac{1}{2}\sqrt{\frac{v}{2\pi I}} \exp\left(-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \sqrt{vI}\right)^2\right)\left(\frac{\partial I}{\partial \rho_n}\right), \quad (4.30)$$

where

$$\frac{\partial I}{\partial \rho_n} = \int_{-\infty}^{\infty}\left(\frac{2f'_W(w)\frac{\partial f'_W(w)}{\partial \rho_m} f_W(w) - (f'_W(w))^2\frac{\partial f_W(w)}{\partial \rho_m}}{f_W^2(w)}\right) dw. \quad (4.31)$$

Note that we resort to numerical calculation because the close form result of (4.31) is unavailable. For the signal parameters $\rho_s = \{A, \phi\}$, we have

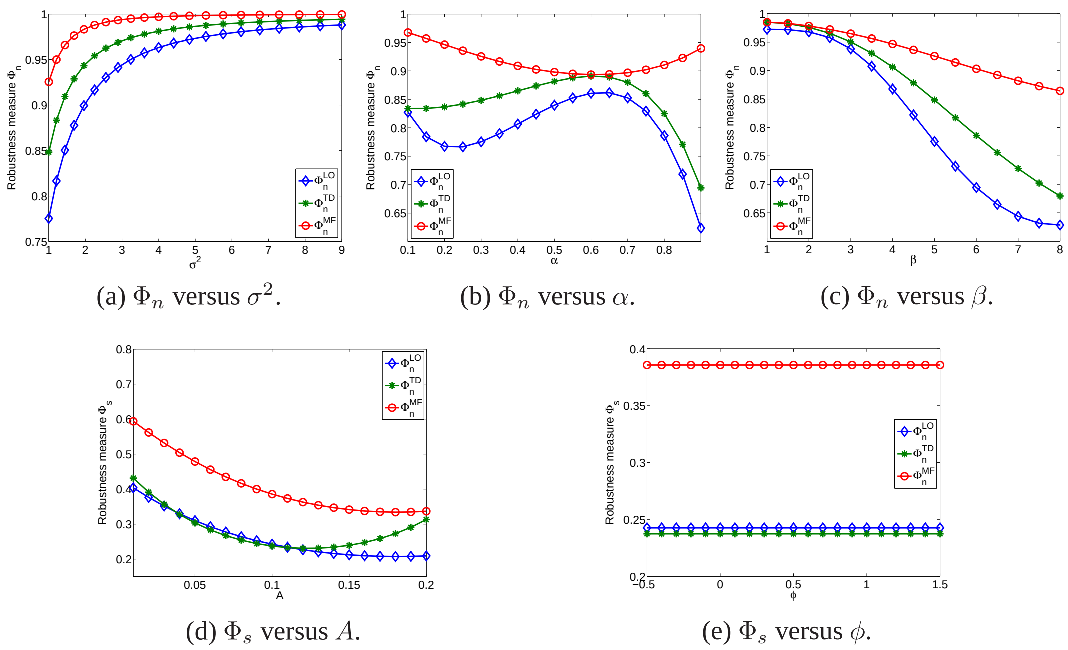
$$\frac{\partial P_D^{LO}}{\partial \rho_s} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(Q^{-1}(P_{FA}) - \sqrt{vI}\right)^2\right)\left(-\frac{\sqrt{I}}{2\sqrt{v}}\right)\left(\frac{\partial v}{\partial \rho_s}\right). \quad (4.32)$$

By using (4.30) in (4.17) and (4.32) in (4.18), Φ_n^{LO} and Φ_s^{LO} can be calculated.

For Saha's detector, since its close form P_D is not available, its robustness measure is not calculated here. But simulation results are provided for comparison in Fig. 4.7.

In Fig. 4.6, we plot Φ_n^{TD}, Φ_n^{MF}, Φ_n^{LO} and Φ_s^{TD}, Φ_s^{MF}, Φ_s^{LO} as functions of σ^2, α, β, A, and ϕ. We set $P_{FA} = 0.1$, $N = 100$. In each subfigure of Fig. 4.6, the robustness measures are shown as functions of one of the parameters while the other parameters are fixed. The fixed parameters values are set to be $\sigma^2 = 1$, $\alpha = 0.3$, $\beta = 5$, $A = 0.1$, $\phi = 0$. For example, in Fig. 4.6 (a), the values of the robustness measure are shown as σ^2 ranges from 1 to 9 while $\alpha = 0.3$, $\beta = 5$, $A = 0.1$, $\phi = 0$. It can be observed that among the three detectors, the MF always has the highest robustness to the noise parameters. The robustness of the proposed TD is superior to that of the LO detector, and is close to Φ_n^{MF} for some parameter values. In term of the signal, not surprisingly, due to its dependency on the signal, the robustness of the proposed TD is a lot worse than the MF. It is also worse than the LO detector for all phase values and small amplitude values, but the difference is small. For large amplitude, the proposed TD is more robust to signal than the LO detector.

In Fig. 4.7, we show the ROCs obtained via Monte Carlo simulation under inaccurate parameters of the noise pdf and the signal for the sinusoidal signal (4.13) with $A = 0.1, \phi = 0$ in the GM noise with $\alpha = 0.3, \beta = 5, \sigma = 1$. In Fig. 4.7 (a), the ROCs based on accurate signal information but an inaccurate estimation of the noise pdf, where $\alpha = 0.4, \beta = 2.5, \sigma = 1$, are shown. We can see that the performance of the LO detector and the performance of Saha's detector is worse than the



(a) Φ_n versus σ^2. (b) Φ_n versus α. (c) Φ_n versus β.

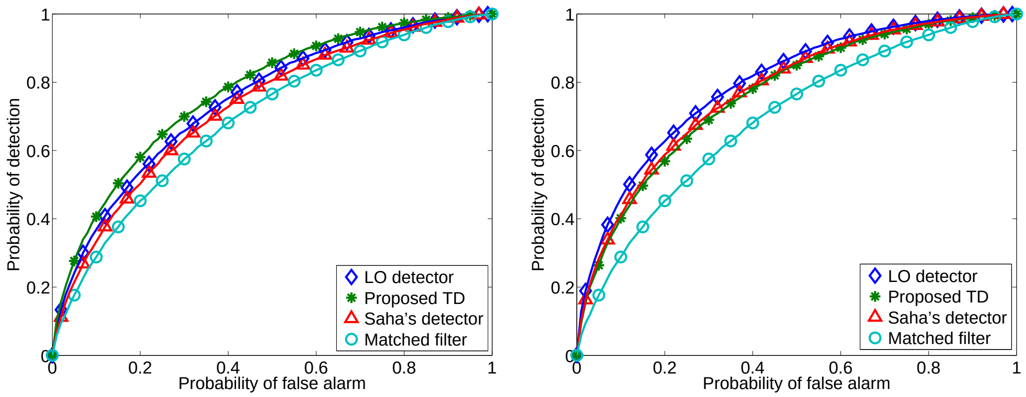
(d) Φ_s versus A. (e) Φ_s versus ϕ.

Fig. 4.6. Comparison of robustness measure.

proposed TD due to inaccurate noise estimation. Compared with the ROCs shown in Fig. 4.4, the AUL of the LO detector is 0.392dB (9.45% degradation) and the AUL of Saha's detector is 0.4dB (9.65% degradation). The AUL of the proposed TD and the MF is 0, which shows that they are immune to inaccuracy in noise pdf. In Fig. 4.7 (b), the ROCs based on accurate noise information but an inaccurate estimation of the signal amplitude and phase, where $A = 0.12, \phi = 0.1\pi$, are shown. We can observe that the proposed TD still performs slightly worse than the LO detector, comparable to Saha's detector, and significantly better than the MF. Compared with the ROCs shown in Fig. 4.4, the AULs of the LO detector, the proposed TD, Saha's detector, and the MF are 0.022dB (0.5% degradation), 0.034dB (0.7% degradation), 0.022dB (0.5% degradation), and 0.018dB (0.4% degradation), respectively. This reveals that the proposed TD, although is based on perfect signal information, can bear inaccuracy in signal to some extent. In particular, we compare the robustness of the proposed TD with Saha's detector, which depends on both the noise and the signal. The above simulation shows that the proposed TD is significantly more robust to the noise pdf. As to the robustness to the signal, the proposed TD is inferior, as expected, but still comparable to Saha's detector.

4.4.2 Implementation Complexity

In this part, the implementation complexity of the proposed TD is discussed. The multiway switch and the binary TS array can both be easily implemented in circuit design. The correlator is also easy



(a) ROCs under inaccurate noise pdf

(b) ROCs under inaccurate signal

Fig. 4.7. ROCs under inaccurate estimations of the noise pdf and the signal.

in implementation, since it is linear in both the absolute value of the signal $|s[n]|$ and the TS output. In addition, the structure of the proposed TD can be further simplified. In the proposed TD, an array of TS's instead of one TS is used because $s[n]$ can take different values, which requires different binarization thresholds. Thus, for different samples in the signal sequence with the same value, the same TS can be used. If the signal is DC, only one TS is needed and the multiway switch is not necessary. In this case, the proposed TD structure reduces to the one we proposed in [56]. If the signal is periodic and the sampling time is appropriate, we only need to consider the samples in one period and the required number of TSs can be reduced. In general, based on the given $s[n]$, we can divide the range of $s[n]$'s values into groups. For example, the $s[n]$'s whose value is between $C - \epsilon$ and $C + \epsilon$ can be put into one group and the TS with binarization threshold $C/2$ is applied, where C is a constant and ϵ is a positive scalar. This reduces the implementation complexity with some penalty on the performance. One can balance performance and complexity by adjusting the number of groups.

For the LO detector, in general, the $g(x)$ component in (1.13) is nonlinear in the observations, and its implementation can be highly complicated. Further, since it requires perfect noise pdf information, its complexity even increases if the noise pdf estimation component is taken into account. Saha's detector in [6] is simple in implementation, but it also requires noise pdf estimation and, in addition, a numerical optimization of the three-level TS threshold [69]. Hence it may be even more complicated in implementation than the LO detector.

4.4.3 Region of Validity

The proposed TD is applicable for noises with zero-mean, unimodal, and symmetric pdf. However, it may not perform efficiently for all noises in this category. Therefore, it is helpful to address the validity of the proposed TD corresponding to the form and the parameters of the noise pdf. Using the MF as the benchmark, we define the validity of the proposed TD based on whether it has a higher detection probability than the MF. Accordingly, for a given type of noise pdf, the validity region is defined as the set of the parameters of the noise pdf with which the proposed TD is superior to the MF.

In general, theoretical derivation of the validity region is difficult because for a certain noise pdf, the close-form of the ROC is usually unavailable. We often need to resort to numerical method and conduct simulation for a large number of noise pdf parameters and signal forms. Nevertheless, we show in the following theorem that when the signal is weak, and the number of samples is large, the validity region reduces to a simple form, which is independent of the signal $s[n]$, P_{FA}, and for some cases, even the noise variance σ^2.

Theorem 4.3. If $|s[n]| \ll \sigma$ and $N \gg 1$, the validity region of the proposed TD with respect to the MF is independent of the signal sequence $s[n]$'s and P_{FA}, and can be approximated as the set of noise pdf parameters that satisfy $f_W(0) > \frac{1}{2\sigma}$.

Proof. To find the validity region, we need to solve

$$P_D^{TD} > P_D^{MF} \Leftrightarrow Q\left(\frac{\sigma_0 Q^{-1}(P_{FA}) + \mu_0 - \mu_1}{\sigma_1}\right) > Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{\sum_{n=0}^{N-1} s^2[n]}{\sigma^2}}\right). \tag{4.33}$$

Using (4.19), we can recast (4.33) as

$$Q\left(Q^{-1}(P_{FA}) - \frac{2v f_W(0)}{\sqrt{v - u f_W^2(0))}}\right) > Q\left(Q^{-1}(P_{FA}) - \frac{\sqrt{v}}{\sigma}\right), \tag{4.34}$$

where v, u are defined in (4.20). Since $|s[n]| \ll \sigma$, ignoring the second order term $u f_W^2(0)$ in (4.34), we have

$$Q\left(Q^{-1}(P_{FA}) - 2 f_W(0)\sqrt{v}\right) > Q\left(Q^{-1}(P_{FA}) - \frac{\sqrt{v}}{\sigma}\right) \quad \Leftrightarrow \quad f_W(0) > \frac{1}{2\sigma}.$$

$\square$

Theorem 4.3 shows that the validity region only depends on $f_W(0)$ and σ, but is independent of P_{FA} and $s[n]$. For the GM noise and the GG noise, we further investigate the validity region in the following corollary and show that it is also independent of σ.

Corollary 4.1. For the GM noise, the validity region of the proposed TD is

$$\left\{(\alpha, \beta) \,\middle|\, \sqrt{\alpha + (1-\alpha)\beta^2}\left(\alpha + \frac{1-\alpha}{\beta}\right) > \sqrt{\frac{\pi}{2}}\right\}. \tag{4.35}$$

For the GG noise, the validity region of the proposed TD is

$$\{\beta \,|\beta \in [0.42, 1]\} \, . \tag{4.36}$$

For both cases, the validity region of the proposed TD is independent of the noise variance.

Proof. For the GM noise in (3.38), according to Theorem 4.3, we have

$$\begin{aligned}
& f_W(0) > \frac{1}{2\sigma} \quad \Leftrightarrow \quad f_W(0) = \frac{c}{\sigma\sqrt{2\pi}}\left(\alpha + \frac{1-\alpha}{\beta}\right) > \frac{1}{2\sigma} \\
\Leftrightarrow \quad & \frac{c}{\sqrt{2\pi}}\left(\alpha + \frac{1-\alpha}{\beta}\right) > \frac{1}{2} \quad \Leftrightarrow \quad \sqrt{\alpha + (1-\alpha)\beta^2}\left(\alpha + \frac{1-\alpha}{\beta}\right) > \sqrt{\frac{\pi}{2}}.
\end{aligned}$$

Similarly, for the GG noise in (4.15), we have

$$\begin{aligned}
& f_W(0) > \frac{1}{2\sigma} \Leftrightarrow \quad f_W(0) = \frac{c_1(\beta)}{\sigma} > \frac{1}{2\sigma} \\
\Leftrightarrow \quad & c_1(\beta) = \frac{\Gamma^{1/2}\left(\frac{3}{2}(1+\beta)\right)}{(1+\beta)\Gamma^{3/2}\left(\frac{1}{2}(1+\beta)\right)} > \frac{1}{2} \quad \Leftrightarrow \quad \{\beta|\beta \in [0.42, 1]\}.
\end{aligned}$$

□

Corollary 4.1 provides a simply way to calculate the validity regions for the two noise pdf's. In the remaining of this subsection, we show simulation results on the region of validity.

First, for the GM noise, we justify the analytical results in Theorem 4.3 and Corollary 4.1. We set $P_{FA} = 0.1$ and $N = 100$, and simulate the validity regions for three detection problems. Problem I: sinusoidal signal (4.13) in GM noise with $\sigma^2 = 1$; Problem II: sinusoidal signal (4.13) in GM noise with $\sigma^2 = 4$; and Problem III: random signal (4.14) in GM noise with $\sigma^2 = 1$. For all problems, we obtain P_D^{TD}'s from simulation for different (α, β)'s, and determine the validity regions by $P_D^{TD} > P_D^{MF} = 0.2828$. These validity regions are compared with the analytic result, which is calculated using (4.35) in Corollary 4.1. The validity regions are shown in Fig. 4.8, where the ones for Problems I, II, III, and Corollary 4.1 are marked as red, green, blue, and gray, respectively. For better demonstration, we shape the regions by the order of analytical region (gray), Problem III (blue), Problem II (green), and Problem I (red) due to the size of the regions. It is observed that the borders of the validity regions of all three problems are almost the same and match the analytical one obtained from (4.35) precisely. Comparing the validity regions of Problem I and Problem II, we see that the validity region is independent of noise variance σ^2. Comparing the validity regions of Problem I and Problem III, we see that the validity region is independent of signal $s[n]$. Comparing the validity regions obtained from simulation and the analytical one, we see that the (4.35) is a sound approximation in determining the validity region of the proposed TD.

We then turn to the GG noise. We consider the detection of the sinusoidal signal (4.13) in GG

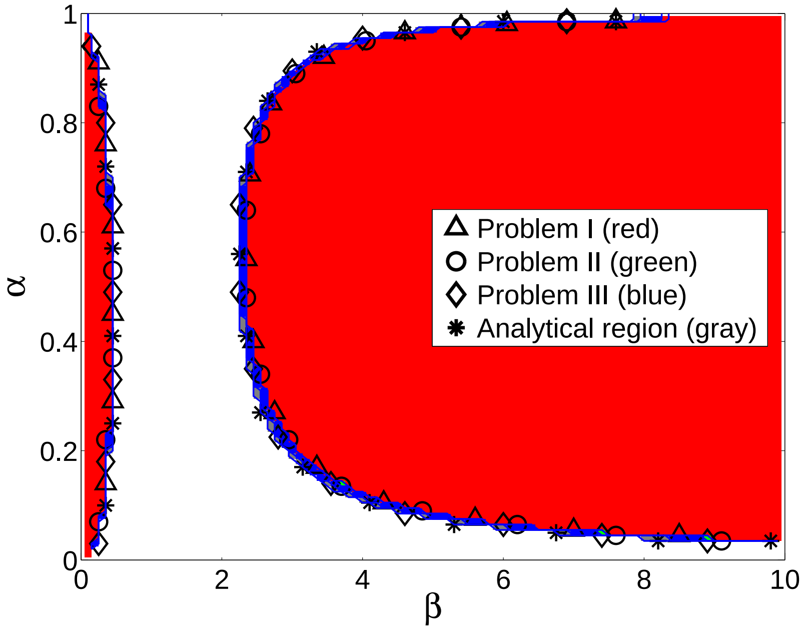


Fig. 4.8. The validity regions for different detection problem with GM noise.

noise with $\sigma^2 = 1$, obtain P_D^{TD} for different β by simulation, and show them in Fig. 4.9. It can be observed that the validity region is about $\{\beta|\beta \in [0.44, 1]\}$ obtained from$\{\beta|P_D^{TD} > P_D^{MF} = 0.2828\}$, where P_D^{MF} is calculated from (4.10). This is close to the result (4.36) in Corollary 4.1, which is $\{\beta|\beta \in [0.42, 1]\}$.

4.5 Conclusions

In this chapter, we proposed a low-complexity threshold system based detector to detect any known deterministic signal embedded in independent unknown non-Gaussian noise. We assumed that the noise pdf is unimodal and symmetric, the signal is small compared to the noise variance, and there are a large number of samples. The optimality of the two parts of the proposed detector, the binary threshold system and the correlator, was proved. The detection probability and the ROC of the proposed TD were investigated both analytically and numerically. For noises with heavy pdf tails, simulation showed that the performance of the proposed TD approaches that of the LO detector and Saha's detector, the design of which need exact noise pdf information, and is much better than the MF. Through a robustness measure, we showed that the proposed TD is highly robust to the noise pdf. On the other hand, its robustness to the signal is inferior but comparable to the LO detector and Saha's detector. The implementation complexity of the proposed detector was discussed and compared with other detection designs. The validity region of the proposed detector was defined

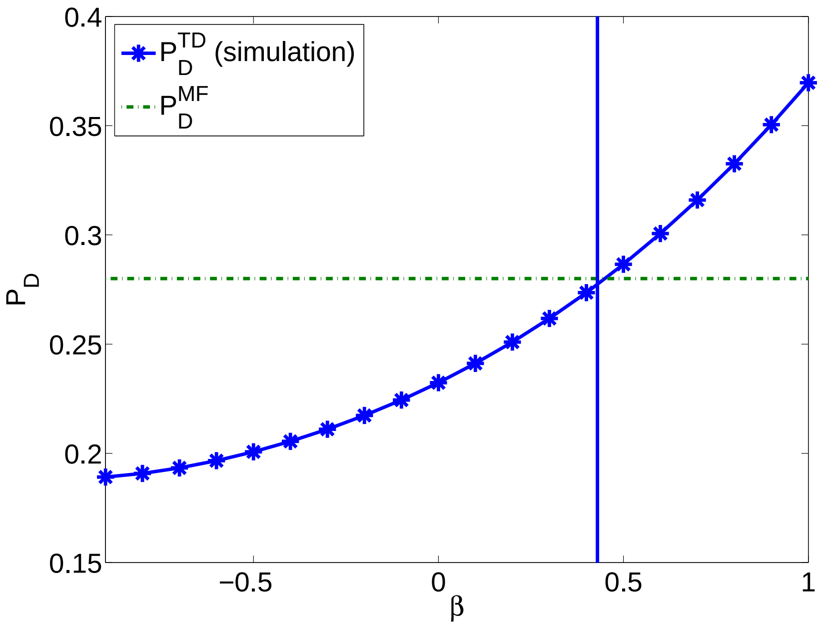


Fig. 4.9. P_D^{TD} versus β for the sinusoidal signal detection in GG noise.

and analyzed using the MF as the benchmark.

Appendix 4.A Proof of Theorem 4.1

Without loss of generality, we assume that $x[n] \geq 0$ and show that the binarization threshold $\tau = s[n]/2$ maximizes the AUC. The case of $x[n] < 0$ can be proved similarly.

Since $x[n] \geq 0$, $TS_1(\tau)$ is applied. With the binarization threshold τ, we define $p_0(\tau)$ and $p_1(\tau)$ as follows:

$$p_0(\tau) \triangleq \mathbb{P}(Y[n] = 1; H_0) = \int_{\tau}^{\infty} f_X(x; H_0)dx, \tag{4.37}$$

$$p_1(\tau) \triangleq \mathbb{P}(Y[n] = 1; H_1) = \int_{\tau}^{\infty} f_X(x; H_1)dx. \tag{4.38}$$

Since $f_X(x; H_0) = f_W(w)$, $f_X(x; H_1) = f_W(w - s[n])$ and $s[n] \geq 0$, $f_X(x; H_1)$ is a right shift of $f_X(x; H_0)$ by $s[n]$. Thus, $p_0(\tau) \leq p_1(\tau)$ for all τ's.

The TS output $y[n]$ has two possibilities $y[n] = 0$ or $y[n] = 1$, based on which we will decide on H_0 or H_1. Therefore the likelihood ratio values of the binary detection problem are as follows:

$$L(y[n] = 0) = \frac{\mathbb{P}(y[n] = 0; H_1)}{\mathbb{P}(y[n] = 0; H_0)} = \frac{1 - p_1(\tau)}{1 - p_0(\tau)},$$
$$L(y[n] = 1) = \frac{\mathbb{P}(y[n] = 1; H_1)}{\mathbb{P}(y[n] = 1; H_0)} = \frac{p_1(\tau)}{p_0(\tau)}.$$

Since $0 < p_0 \leq p_1 < 1$, we have $0 < L(y[n]=0) \leq 1 \leq L(y[n]=1)$, and the decision rule of the binary detection problem is

$$\delta(y[n]) = \begin{cases} H_1 & \text{if } L(y[n]) > \gamma \\ H_0 & \text{if } L(y[n]) < \gamma \\ H_0 \text{ or } H_1 & \text{if } L(y[n]) = \gamma \end{cases} . \tag{4.39}$$

If $\gamma < L(y[n]=0)$, we have $P_{FA} = 1$ and $P_D = 1$. If $L(y[n]=0) < \gamma < L(y[n]=1)$, we have $P_{FA} = p_0(\tau)$ and $P_D = p_1(\tau)$. If $\gamma > L(y[n]=1)$, we have $P_{FA} = 0$ and $P_D = 0$. With the help of randomization decision functions, the ROC of the detection problem is the combination of the segment from $(0,0)$ to $(p_0(\tau), p_1(\tau))$ and the segment from $(p_0(\tau), p_1(\tau))$ to $(1,1)$.

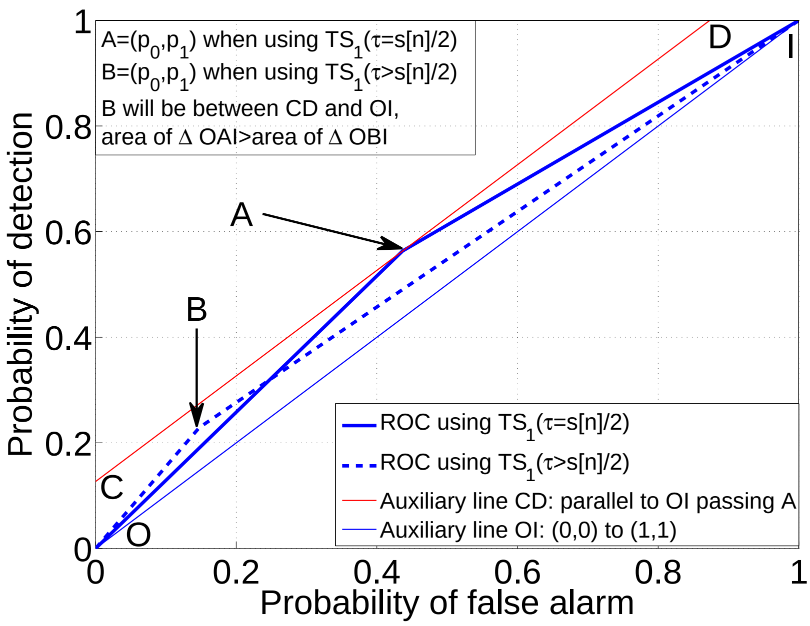


Fig. 4.10. ROCs of the TS with different τ.

Now we show that when $\tau = s[n]/2$, the AUC is the largest. Denote the $(0,0)$ point in the $P_{FA} - P_D$ square as O, the $(1,1)$ point as I, and the $\left(p_0\left(\frac{s[n]}{2}\right), p_1\left(\frac{s[n]}{2}\right)\right)$ point as A, which are shown in Fig. 4.10. For another $\tau \neq s[n]/2$, denote the $(p_0(\tau), p_1(\tau))$ point as B. It is thus sufficient to show that the area of $\triangle OAI$ is no smaller than that of $\triangle OBI$.

Without loss of generality, assume that $\tau > s[n]/2$. The other case can be proved similarly. Note that $p_0(\tau)$ and $p_1(\tau)$ are decreasing functions of τ. Define $\Delta p_0 \triangleq p_0\left(\frac{s[n]}{2}\right) - p_0(\tau) = \int_{s[n]/2}^{\tau} f_W(w)dw$ and $\Delta p_1 \triangleq p_1\left(\frac{s[n]}{2}\right) - p_1(\tau) = \int_{s[n]/2}^{\tau} f_W(w - s[n])dw$. Since $f_W(w)$ is symmetric at $w = 0$ and unimodal, we have $f_W(w - s[n]) \geq f_W(w)$ for $\frac{s[n]}{2} < w < \tau$, and hence $\Delta p_1 \geq \Delta p_0$. This means that the point B is on or under the segment CD in Fig. 4.10, where segment

CD includes point A and is parallel to the segment OI. Thus, the area of $\triangle OAI$ is no smaller than that of $\triangle OBI$.

Appendix 4.B Proof of Theorem 4.2

For the binary detection problem with observation **y**, the optimal test is the LRT, defined as

$$L(\mathbf{y}) = \frac{\mathbb{P}(\mathbf{y}; H_1)}{\mathbb{P}(\mathbf{y}; H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta'. \tag{4.40}$$

Since entries of $\mathbf{y}$ are independent, we have

$$L(\mathbf{y}) = \prod_{n=0}^{N-1} \frac{\mathbb{P}(y[n]; H_1)}{\mathbb{P}(y[n]; H_0)} = \prod_{n=0}^{N-1} L(y[n]), \tag{4.41}$$

where $L(y[n]) \triangleq \frac{\mathbb{P}(y[n];H_1)}{\mathbb{P}(y[n];H_0)}$.

Notice that $y[n]$ only takes 0 or 1, we calculate the values of $L(y[n])$ for $y[n] = 1$ and $y[n] = 0$, respectively. First, we consider the case of $s[n] \geq 0$, for which the optimal TS is $TS_1\left(\frac{s[n]}{2}\right)$. Because $f_X(x; H_1) = f_X(x - s[n]; H_0)$, $f_X(x; H_0) = f_W(w)$, and $f_W(w)$ is unimodal and symmetric at $w = 0$, we have

$$\begin{aligned}
\mathbb{P}(y[n] = 1; H_1) &= \int_{\frac{s[n]}{2}}^{\infty} f_X(x; H_1)dx = \int_{-\frac{s[n]}{2}}^{\infty} f_X(x; H_0)dx = \int_{-\frac{s[n]}{2}}^{\infty} f_W(w)dx \\
&= \int_0^{\infty} f_W(w)dw + \int_{-\frac{s[n]}{2}}^{0} f_W(w)dw = \frac{1}{2} + \int_0^{\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n] = 1; H_0) &= \int_{\frac{s[n]}{2}}^{\infty} f_X(x; H_0)dx = \int_0^{\infty} f_W(w)dw - \int_0^{\frac{s[n]}{2}} f_W(w)dw = \frac{1}{2} - \int_0^{\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n] = 0; H_1) &= 1 - \mathbb{P}(y[n] = 1; H_1) = \frac{1}{2} - \int_0^{\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n] = 0; H_0) &= 1 - \mathbb{P}(y[n] = 1; H_0) = \frac{1}{2} + \int_0^{\frac{s[n]}{2}} f_W(w)dw.
\end{aligned}$$

For small signal, i.e., $|s[n]| \ll \sigma$, we have

$$\begin{aligned}
&\mathbb{P}(y[n] = 1; H_1) \approx \tfrac{1}{2} + \tfrac{s[n]}{2} f_W(0), \quad \mathbb{P}(y[n] = 1; H_0) \approx \tfrac{1}{2} - \tfrac{s[n]}{2} f_W(0), \\
&\mathbb{P}(y[n] = 0; H_1) \approx \tfrac{1}{2} - \tfrac{s[n]}{2} f_W(0), \quad \mathbb{P}(y[n] = 0; H_0) \approx \tfrac{1}{2} + \tfrac{s[n]}{2} f_W(0),
\end{aligned} \tag{4.42}$$

and thus

$$\begin{aligned}
\ln L(y[n]=1) &\approx \ln\left(\frac{\frac{1}{2}+\frac{s[n]}{2}f_W(0)}{\frac{1}{2}-\frac{s[n]}{2}f_W(0)}\right) = \ln\left(1+s[n]f_W(0)\right) - \ln\left(1-s[n]f_W(0)\right) \\
&= s[n]f_W(0) - (-s[n]f_W(0)) + \mathcal{O}\Big(\big(s[n]f_W(0)\big)^2\Big) \approx 2s[n]f_W(0), \\
\ln L(y[n]=0) &\approx \ln\left(\frac{\frac{1}{2}-\frac{s[n]}{2}f_W(0)}{\frac{1}{2}+\frac{s[n]}{2}f_W(0)}\right) \approx -2s[n]f_W(0).
\end{aligned} \tag{4.43}$$

Now, for the case of $s[n]<0$, the optimal TS is $TS_2\left(\frac{s[n]}{2}\right)$. Similarly, we can show that

$$\begin{aligned}
\mathbb{P}(y[n]=1;H_1) &= \frac{1}{2} + \int_0^{-\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n]=1;H_0) &= \frac{1}{2} - \int_0^{-\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n]=0;H_1) &= 1-\mathbb{P}(y[n]=1;H_1) = \frac{1}{2} - \int_0^{-\frac{s[n]}{2}} f_W(w)dw, \\
\mathbb{P}(y[n]=0;H_0) &= 1-\mathbb{P}(y[n]=1;H_0) = \frac{1}{2} + \int_0^{-\frac{s[n]}{2}} f_W(w)dw.
\end{aligned}$$

For small signal, i.e., $|s[n]| \ll \sigma$, we have

$$\begin{aligned}
&\mathbb{P}(y[n]=1;H_1) \approx \tfrac{1}{2} - \tfrac{s[n]}{2}f_W(0), \quad \mathbb{P}(y[n]=1;H_0) \approx \tfrac{1}{2} + \tfrac{s[n]}{2}f_W(0) \\
&\mathbb{P}(y[n]=0;H_1) \approx \tfrac{1}{2} + \tfrac{s[n]}{2}f_W(0), \quad \mathbb{P}(y[n]=0;H_0) \approx \tfrac{1}{2} - \tfrac{s[n]}{2}f_W(0),
\end{aligned} \tag{4.44}$$

and thus

$$\begin{aligned}
\ln L(y[n]=1) &\approx \ln\left(\frac{\frac{1}{2}+\frac{s[n]}{2}f_W(0)}{\frac{1}{2}-\frac{s[n]}{2}f_W(0)}\right) \approx -2s[n]f_W(0), \\
\ln L(y[n]=0) &\approx \ln\left(\frac{\frac{1}{2}-\frac{s[n]}{2}f_W(0)}{\frac{1}{2}+\frac{s[n]}{2}f_W(0)}\right) \approx 2s[n]f_W(0).
\end{aligned} \tag{4.45}$$

Combining (4.43) and (4.45), for any $s[n]$, we have

$$\ln L(y[n]=1) \approx 2|s[n]|f_W(0), \quad \ln L(y[n]=0) \approx -2|s[n]|f_W(0). \tag{4.46}$$

Note that the values of $\ln L(y[n]=1)$ and $\ln L(y[n]=0)$ are independent of $x[n]$ but only depend on the signal $|s[n]|$ and the value of the noise pdf at 0.

For a given **y** vector, let D be the number of 1's in **y**. The number of 0's is thus $N-D$. From (4.41) and (4.46), we have

$$\ln L(\mathbf{y}) \approx 2\left(\sum_{\{n|y[n]=1\}} |s[n]| - \sum_{\{n|y[n]=0\}} |s[n]|\right) f_W(0)$$

$$
\begin{aligned}
&= 2\left[\sum_{\{n|y[n]=1\}} |s[n]| - \left(\sum_{n=0}^{N-1} |s[n]| - \sum_{\{n|y[n]=1\}} |s[n]|\right)\right] f_W(0) \\
&= 4 \sum_{\{n|y[n]=1\}} |s[n]| f_W(0) - 2\sum_{n=0}^{N-1} |s[n]| f_W(0) \\
&= 4\left(\sum_{n=0}^{N-1} |s[n]| y[n]\right) f_W(0) - 2\sum_{n=0}^{N-1} |s[n]| f_W(0). \qquad (4.47)
\end{aligned}
$$

Note that the second term in (4.47) and $f_W(0)$ are constants, independent of the hypotheses and the observation, Thus from (4.40), the optimal test rule becomes

$$
T^{TD}(\mathbf{y}) \triangleq \frac{1}{N} \sum_{n=0}^{N-1} y[n]|s[n]| \underset{H_0}{\overset{H_1}{\gtrless}} \eta,
$$

which shows that the proposed $T^{TD}(\mathbf{y})$ is the optimal test statistic.

Chapter 5

Optimal Design and Noise-Enhanced Effect for Binary Threshold Detector under AUC Measure

This chapter considers the binary threshold system (TS) based detector for a general binary testing problem. First, the optimal binary TS that maximizes AUC is derived. Then the noise-enhanced effect is investigated. The optimal noise that can achieve the maximum AUC is derived and shown to be deterministic. An example is shown to justify the derived results.

5.1 Introduction

As explained in Sec. 1.4, threshold system (TS) based detector, or TD in short, is one widely used suboptimal detector for detection problems with non-Gaussian noise [11], [56], [58]. For a DC signal detection with known noise, Chapeau-Blondeau in [11] proposed a maximum a-posteriori probability detector for a given binary TS, but did not address the optimal design of the TS. We filled this gap in Chap. 3 to design the optimal TS using NP criterion. In Chap. 4, for an arbitrary known signal detection in non-Gaussian noise with unknown pdf, we proposed an optimal TD and analyzed its properties. In this chapter, we consider a general multiple dimensional binary detection problem, and derive the optimal composite binary TS that maximizes the AUC, where AUC stands for the area under the receiver operating characteristic (ROC) curve.

We also investigates the noise-enhanced effect in this binary TD. The idea of noised-enhanced effect arises from the phenomena that the noise sometimes can play a constructive role. For example, when paddlefish try to catch a bug, its perceptive abilities are improved if an external noise is added in his ambient environment [13]. Many other interesting noise-enhanced examples can be found in

[12], [13]. In the context of signal detection, noised-enhanced effect refers to improved performance obtained from injecting additional independent “noise" into the observation [16]–[20]. Kay in [16] showed that when the threshold of the binary TS in a TD is not optimal, for a given P_{FA}, P_D of the TD can be increased by adding WGN. Kay et al. also calculated the optimal pdf for minimizing the probability of decision error under Bayesian criterion [17]. For a general binary detection problem with an arbitrarily given detector, under the NP criterion, Chen et al. [18], [19] and Patel [20] investigated the noise-enhanced effect and derived the optimal pdf form of the noise. However, the results suffer the following disadvantages. The derived optimal noise pdf is implicitly represented as a function of the conditional probabilities, which in general are difficult to be obtained in closed-form. An numerical method for finding the optimal noise pdf was proposed in [20]. But it is computationally expensive. With the NP criterion, the optimal noise pdf depends on the desired level of P_{FA}. This adds on more computational cost, especially for applications with changing environment or requirements.

In this chapter, for a general binary detection problem with binary TD, we derive the optimal TD and the noise pdf for the best noise-enhanced effect under the AUC measure. A simple closed-form of the optimal noise pdf is derived. The computational cost in finding the optimal noise is very low. The properties of the noise-enhanced effect are also discussed. An illustrative example is presented as well.

5.2 Problem Statement and AUC Measure

In this section, we introduce the detection problem, the binary TD structure, and the AUC measure.

5.2.1 Detection Problem and Detector Structure

We recast the BHTP in (1.1) as

$$\begin{cases} H_0 : \mathbf{X} \sim f_{\mathbf{X}}(\mathbf{x}; H_0) \triangleq f_0(\mathbf{x}) \\ H_1 : \mathbf{X} \sim f_{\mathbf{X}}(\mathbf{x}; H_1) \triangleq f_1(\mathbf{x}) \end{cases} , \tag{5.1}$$

where $\mathbf{X}$ represents the N-dimensional observation, $f_{\mathbf{X}}(\mathbf{x}; H_i), i = 0, 1$ are the pdf’s of $\mathbf{X}$ under hypotheses $H_i, i = 0, 1$, abbreviated as $f_i(\mathbf{x}), i = 0, 1$ for later convenience.

For this detection problem, we will use a binary TD with possible noise-enhanced effect. The structure is shown in Fig. 5.1. First an N-dimensional noise $\mathbf{V}$ is added to the original observation $\mathbf{X}$ to produce the new observation $\mathbf{U}$. Then $\mathbf{U}$ is applied to the binary TS and converted into a binary signal $Y \in \{0, 1\}$. Finally a decision H_1 or H_0 will be made based on the rule $y \underset{H_0}{\overset{H_1}{\gtrless}} \eta$.

In this detector design, we use a multiple-dimensional composite binary TS as the test statistic for its low complexity. For a binary testing problem, the optimal test statistic is the LR, which requires $f_0(\mathbf{x})$ and $f_1(\mathbf{x})$ to be perfectly known. When there are unknown parameters, the generalized LR can be used. These detectors however generally have high complexity and low robustness [1]. Binary TD, other than its low complexity, has been shown to achieve good detectability and high robustness, especially for noises with a heavy-tail pdf [58].

The noise $\mathbf{V}$ is introduced to improve the detector performance, when possible, called the noised-enhanced effect. Noise-enhanced effect in signal detection was first proposed in [16] and further developed in [17], [18], [20].

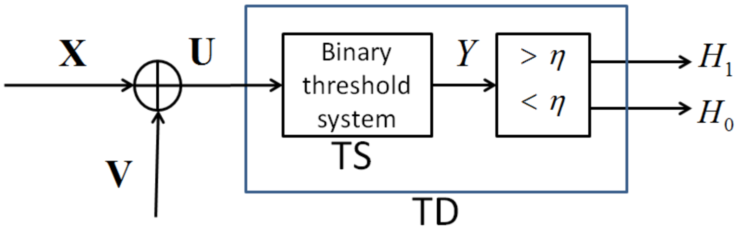


Fig. 5.1. Noise-enhanced binary TS based detector.

Here are the detailed formulation of different parts of the proposed noise-enhanced TD. Let $f_{\mathbf{V}}(\mathbf{x})$ be the pdf of the additional noise $\mathbf{V}$. Since $\mathbf{U} = \mathbf{X} + \mathbf{V}$, we have

$$f_0^{\mathbf{U}}(\mathbf{u}) \triangleq f_{\mathbf{U}}(\mathbf{u}; H_0) = f_0(\mathbf{u}) * f_{\mathbf{V}}(\mathbf{u}) = \int_{\mathbb{R}^N} f_0(\mathbf{u}-\mathbf{x}) f_{\mathbf{V}}(\mathbf{x}) d\mathbf{x},$$
$$f_1^{\mathbf{U}}(\mathbf{u}) \triangleq f_{\mathbf{U}}(\mathbf{u}; H_1) = f_1(\mathbf{u}) * f_{\mathbf{V}}(\mathbf{u}) = \int_{\mathbb{R}^N} f_1(\mathbf{u}-\mathbf{x}) f_{\mathbf{V}}(\mathbf{x}) d\mathbf{x},$$

where $f_i^{\mathbf{U}}(\mathbf{u})$ is the pdf of $\mathbf{U}$ under H_i for $i = 0, 1$ and $*$ stands for convolution. The function of the binary TS can be expressed as

$$Y = T(\mathbf{u}) = \begin{cases} 1 & \mathbf{u} \in \mathcal{D} \\ 0 & \text{elsewhere} \end{cases}, \tag{5.2}$$

where $\mathcal{D}$ is a subset in $\mathbb{R}^N$. This binary TS is a high-dimensional generalization of the composite binary TS in (2.2). The decision of the TD can be represented by the critical function:

$$\phi(\mathbf{u}, \eta) = \begin{cases} 1 : T(\mathbf{u}) > \eta \\ \nu : T(\mathbf{u}) = \eta \\ 0 : T(\mathbf{u}) < \eta \end{cases}, \tag{5.3}$$

where $\nu \in [0, 1]$.

The detection probability and false alarm probability can be calculated as

$$
\begin{aligned}
P_D(\eta) &= \int_{\mathbb{R}} \phi(\mathbf{u},\eta) f_1^{\mathbf{U}}(\mathbf{u}) d\mathbf{u} = \int_{\mathbb{R}^N} \phi(\mathbf{u},\eta) \left(\int_{\mathbb{R}^N} f_1(\mathbf{u}-\mathbf{x}) f_{\mathbf{V}}(\mathbf{x}) d\mathbf{x} \right) d\mathbf{u} \\
&= \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) \left(\int_{\mathbb{R}^N} \phi(\mathbf{u},\eta) f_1(\mathbf{u}-\mathbf{x}) d\mathbf{u} \right) d\mathbf{x} = \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) F_1(\mathbf{x},\eta) d\mathbf{x}, \qquad (5.4)
\end{aligned}
$$

$$
P_{FA}(\eta) = \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) \left(\int_{\mathbb{R}^N} \phi(\mathbf{u},\eta) f_0(\mathbf{u}-\mathbf{x}) d\mathbf{u} \right) d\mathbf{x} = \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x}) F_0(\mathbf{x},\eta) d\mathbf{x}, \qquad (5.5)
$$

where we define

$$
F_i(\mathbf{x},\eta) \triangleq \int_{\mathbb{R}^N} \phi(\mathbf{u},\eta) f_i(\mathbf{u}-\mathbf{x}) d\mathbf{u}, \qquad (5.6)
$$

for $i = 0, 1$. Note that if $\mathbf{V} = \mathbf{0}$ (no added noise), we have $\mathbf{U} = \mathbf{X}$, and thus

$$
P_D(\eta) = F_1(\mathbf{0},\eta) \quad P_{FA}(\eta) = F_0(\mathbf{0},\eta). \qquad (5.7)
$$

The detection model is a general one since it applies to problems of any dimension with any conditional probabilities. The binary TS model is also a general one, which includes the TS's used in [16]–[18], [56], [58] as 1-dimensional special cases.

5.2.2 AUC Measure and Design Problems

One widely used criterion in signal detection is NP criterion. With NP criterion, the optimal TS design depends on the given level of P_{FA}. In general, the optimization of the test statistic $T(\mathbf{u})$ is computationally costly and sensitive to the value of P_{FA} [1], [56]. Similar problems exists in finding the best noise pdf for the noise-enhanced effect.

In this chapter, we use AUC as the performance measure to avoid the excessive computational load and produce practical and robust designs. The AUC of a detector is defined as the area enclosed by the curve (P_{FA}, P_D) together with the lines $P_D = 0$ and $P_{FA} = 1$. When both $P_D(\eta)$ and $P_{FA}(\eta)$ are continuous piecewise differentiable functions, this area can be calculated through either

$$
\text{AUC} = \int_0^1 P_D dP_{FA} = -\int_{-\infty}^{\infty} P_D(\eta) \frac{\partial P_{FA}(\eta)}{\partial \eta} d\eta \qquad (5.8)
$$

or

$$
\begin{aligned}
\text{AUC} &= \frac{1}{2} + \frac{1}{2} \left(\int_0^1 P_D dP_{FA} - \int_0^1 P_{FA} dP_D \right) \\
&= \frac{1}{2} + \frac{1}{2} \int_{-\infty}^{\infty} \left(P_{FA}(\eta) \frac{\partial P_D(\eta)}{\partial \eta} - P_D(\eta) \frac{\partial P_{FA}(\eta)}{\partial \eta} \right) d\eta.
\end{aligned} \qquad (5.9)
$$

The latter formula follows from Green's formula. When the functions involved are not continuous, however, the two formulas are not equivalent anymore. In such cases, (5.8) often leads to incorrect answers.[1] Nevertheless, one can verify through smoothing that (5.9) still gives the correct answer even for discontinuous $P_D(\eta), P_{FA}(\eta)$, due to the total cancellation of the ill-defined terms. In the following we will use (5.9) whenever analytic calculation of AUC is needed.

AUC represents the average performance of the detector over all possible P_{FA}'s and is independent of a particular one. The AUC is shown to be a valid measure of detection capacity [66], [67]. We use AUC as the performance measure for tractable design results and performance analysis. With the AUC measure, the optimal TS and the optimal noise pdf are independent of the P_{FA} level.

In this chapter, for the signal detection problem in (5.1), using the TD shown in Fig. 5.1, we investigate the following two design problems:

1. When there is no added noise, design the optimal binary TS in the TD that maximizes the AUC.

2. For a given TS, derive the optimal noise pdf for the TD that maximizes the AUC.

5.3 Optimal Binary TS Dedsign

We first consider the optimal binary TS design that maximizes the AUC when there is no added noise. Since the TS is binary, this is equivalent to the design of $\mathcal{D}$ in (5.2).

Theorem 5.1. The AUC maximizing $\mathcal{D}$ for the binary TS is $\mathcal{D}_{\text{opt}} = \{\mathbf{x} : f_1(\mathbf{x}) \geq f_0(\mathbf{x})\}$. Equivalently, the optimal binary TS that maximizes the AUC is

$$Y = T(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \mathcal{D}_{\text{opt}} = \{\mathbf{x} : f_1(\mathbf{x}) \geq f_0(\mathbf{x})\} \\ 0 & \text{elsewhere} \end{cases} . \tag{5.10}$$

Proof. For a TS specified in (5.2) and the critical function in (5.3), we have

$$\phi(\mathbf{u}, \eta) = \begin{cases} 0 & \eta > 1 \text{ or } (0 < \eta \leq 1, \mathbf{u} \notin \mathcal{D}) \\ \nu & (\eta = 1, \mathbf{u} \in \mathcal{D}) \text{ or } (\eta = 0, \mathbf{u} \notin \mathcal{D}) \\ 1 & \eta < 0 \text{ or } (0 \leq \eta < 1, \mathbf{u} \in \mathcal{D}) \end{cases} . \tag{5.11}$$

[1] The underlying reason is that in such cases, the integrals involved become products of the Dirac delta function and discontinuous functions. Such pairings are in general ill-defined.

Define $p_i \triangleq \int_{\mathbf{u}\in\mathcal{A}} f_i(\mathbf{u})d\mathbf{u}, i = 0, 1$. Using (5.11) in (5.7) gives

$$(P_{FA}, P_D) = \begin{cases} (0,0) & \eta > 1 \\ (p_0, p_1) & 0 < \eta < 1 \\ (1,1) & \eta < 0 \\ (\nu p_0, \nu p_1) & \eta = 1 \\ (\nu + p_0(1-\nu), \nu + p_1(1-\nu)) & \eta = 0 \end{cases}. \tag{5.12}$$

Since $\nu \in [0, 1]$, the ROC obtained from this TD is the combination of the two segments: the segment from $(0, 0)$ to (p_0, p_1), and the segment from (p_0, p_1) to $(1, 1)$. With this ROC, we can calculate that

$$\text{AUC} = \frac{1}{2} + \text{Area of triangle of } \triangle\,(0,0)(p_0,p_1)(1,1) = \frac{1}{2} + \frac{1}{2}(p_1 - p_0).$$

Therefore, maximizing AUC is equivalent to maximizing $p_1 - p_0$. We have

$$\mathcal{D}_{opt} = \arg\max_{\mathcal{S}\subset\mathbb{R}^N} (p_1 - p_0) = \arg\max_{\mathcal{S}\subset\mathbb{R}^N} \int_{\mathcal{S}} [f_1(\mathbf{x}) - f_0(\mathbf{x})]d\mathbf{x} = \{\mathbf{x} : f_1(\mathbf{x}) \geq f_0(\mathbf{x})\}.$$

$\square$

Remark 5.1. Straightforward calculation shows that in the above simple case, Formula (5.8) gives $\frac{1}{2} + \frac{p_1 - p_0}{2} + (\nu - \frac{1}{2})[(p_1 p_0 + (1-p_1)(1-p_0)]$ which is wrong unless $\nu = \frac{1}{2}$, while (5.9) gives the correct answer as the extra terms cancel.

5.4 Optimal Noise-Enhanced Effect

In this section, we investigate the noise-enhanced effect and derive the optimal noise pdf that maximizes the AUC. With the additive noise $\mathbf{V}$, using (5.4) and (5.5) in (5.9), we have the following AUC calculation.

$$\begin{aligned}
\text{AUC} =& \frac{1}{2} + \frac{1}{2}\int_{\mathbb{R}} \left(\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})F_0(\mathbf{x},\eta)d\mathbf{x}\right)\left(\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{y})\frac{\partial F_1(\mathbf{y},\eta)}{\partial\eta}d\mathbf{y}\right)d\eta \\
& - \frac{1}{2}\int_{\mathbb{R}} \left(\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})F_1(\mathbf{x},\eta)d\mathbf{x}\right)\left(\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{y})\frac{\partial F_0(\mathbf{y},\eta)}{\partial\eta}d\mathbf{y}\right)d\eta \\
=& \frac{1}{2} + \frac{1}{2}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})f_{\mathbf{V}}(\mathbf{y})\left(\int_{\mathbb{R}} F_0(\mathbf{x},\eta)\frac{\partial F_1(\mathbf{y},\eta)}{\partial\eta}d\eta - \int_{\mathbb{R}} F_1(\mathbf{x},\eta)\frac{\partial F_0(\mathbf{y},\eta)}{\partial\eta}d\eta\right)d\mathbf{x}d\mathbf{y} \\
=& \frac{1}{2}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})f_{\mathbf{V}}(\mathbf{y})H(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y},
\end{aligned} \tag{5.13}$$

where we define

$$H(\mathbf{x},\mathbf{y}) \triangleq 1 + \int_{\mathbb{R}} F_0(\mathbf{x},\eta)\frac{\partial F_1(\mathbf{y},\eta)}{\partial \eta}d\eta - \int_{\mathbb{R}} F_1(\mathbf{x},\eta)\frac{\partial F_0(\mathbf{y},\eta)}{\partial \eta}d\eta. \tag{5.14}$$

The noise pdf optimization problem is thus

$$\arg\max_{f_{\mathbf{V}}(\mathbf{x})} \frac{1}{2}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})f_{\mathbf{V}}(\mathbf{y})H(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} \tag{5.15}$$

$$\text{s.t.} \qquad \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})d\mathbf{x} = 1, \qquad f_{\mathbf{V}}(\mathbf{x}) \geq 0. \tag{5.16}$$

The conditions in (5.16) are because that $f_{\mathbf{V}}(\mathbf{x})$ is a pdf.

Theorem 5.2. For the binary TS given in (5.2), define

$$G(\mathbf{x}) \triangleq \int_{\mathcal{D}} [f_1(\mathbf{u}-\mathbf{x}) - f_0(\mathbf{u}-\mathbf{x})]d\mathbf{u}. \tag{5.17}$$

Let $\mathbf{x}_{\mathrm{opt}}$ be the maximum point of $G(\mathbf{x})$, i.e.,

$$\mathbf{x}_{\mathrm{opt}} = \arg\max_{\mathbf{x}} G(\mathbf{x}). \tag{5.18}$$

The optimal noise pdf that maximizes the AUC of the TD in Fig. 5.1 is

$$f_{\mathbf{V}_{\mathrm{opt}}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_{\mathrm{opt}}), \tag{5.19}$$

where $\delta(\cdot)$ is the Kronecker delta function.

Proof. Using (5.11) in (5.6) gives

$$F_i(\mathbf{x},\eta) = \begin{cases} 0 & \eta > 1 \\ \int_{\mathbf{u}\in\mathcal{D}} \nu f_i(\mathbf{u}-\mathbf{x})d\mathbf{u} & \eta = 1 \\ \int_{\mathbf{u}\in\mathcal{D}} f_i(\mathbf{u}-\mathbf{x})d\mathbf{u} & 0 < \eta < 1 \\ \int_{\mathbf{u}\in\mathcal{D}} f_i(\mathbf{u}-\mathbf{x})d\mathbf{u} + \int_{\mathbf{u}\notin\mathcal{D}} \nu f_i(\mathbf{u}-\mathbf{x})d\mathbf{u} & \eta = 0 \\ 1 & \eta < 0 \end{cases}.$$

We then have

$$-\frac{\partial F_i(\mathbf{x},\eta)}{\partial \eta} = \left[\int_{\mathbf{u}\in\mathcal{D}} f_i(\mathbf{u}-\mathbf{x})d\mathbf{u}\right]\delta(\eta-1) + \left[\int_{\mathbf{u}\notin\mathcal{D}} f_i(\mathbf{u}-\mathbf{x})d\mathbf{u}\right]\delta(\eta).$$

Thus

$$
\begin{aligned}
&-\int_{\mathbb{R}} F_1(\mathbf{x},\eta)\frac{\partial F_0(\mathbf{y},\eta)}{\partial\eta}d\eta \\
=&F_1(\mathbf{x},1)\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} + F_1(\mathbf{x},0)\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
=&\nu\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} + \int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
&+\nu\int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
=&\nu\left[\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} + \int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u}\right] \\
&+\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u} - \int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
=&\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u} + \nu\int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
&-(1-\nu)\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u}.
\end{aligned}
\tag{5.20}
$$

Similarly, we can show that,

$$
\begin{aligned}
\int_{\mathbb{R}} F_0(\mathbf{x},\eta)\frac{\partial F_1(\mathbf{y},\eta)}{\partial\eta}d\eta = &-\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{x})d\mathbf{u} - \nu\int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{y})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{x})d\mathbf{u} \\
&+(1-\nu)\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{y})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{x})d\mathbf{u}.
\end{aligned}
\tag{5.21}
$$

Now using (5.20) and (5.21) in (5.14), we have

$$
H(\mathbf{x},\mathbf{y}) = \frac{1}{2} + \int_{\mathbf{u}\in\mathcal{D}} \left[p_1(\mathbf{u}-\mathbf{x}) - p_0(\mathbf{u}-\mathbf{x})\right] d\mathbf{u} + K(\mathbf{x},\mathbf{y}), \tag{5.22}
$$

where

$$
\begin{aligned}
K(\mathbf{x},\mathbf{y}) \triangleq\ &\nu\int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} - (1-\nu)\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{x})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{y})d\mathbf{u} \\
&-\nu\int_{\mathbf{u}\notin\mathcal{D}} f_1(\mathbf{u}-\mathbf{y})d\mathbf{u}\int_{\mathbf{u}\notin\mathcal{D}} f_0(\mathbf{u}-\mathbf{x})d\mathbf{u} + (1-\nu)\int_{\mathbf{u}\in\mathcal{D}} f_1(\mathbf{u}-\mathbf{y})d\mathbf{u}\int_{\mathbf{u}\in\mathcal{D}} f_0(\mathbf{u}-\mathbf{x})d\mathbf{u}
\end{aligned}
$$

It can be shown straightforwardly that $K(\mathbf{x},\mathbf{y})$ is skew-symmetric, i.e., $K(\mathbf{x},\mathbf{y}) = -K(\mathbf{y},\mathbf{x})$. Now we calculate the object function (5.15) using (5.22). Because the integral of the skew-symmetric terms is zero, we have

$$
\text{AUC} = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})f_{\mathbf{V}}(\mathbf{y})H(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} = \frac{1}{2} + \int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})G(\mathbf{x})d\mathbf{x}, \tag{5.23}
$$

where $G(\mathbf{x})$ is defined in (5.17). By using Holder's inequality,

$$\int_{\mathbb{R}^N} f_{\mathbf{V}}(\mathbf{x})G(\mathbf{x})d\mathbf{x} \leq ||f_{\mathbf{V}}(\mathbf{x})||_1||G(\mathbf{x})||_\infty = \max G(\mathbf{x})$$

with equality when $f_{\mathbf{V}_{\text{opt}}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_{\text{opt}})$ where $\mathbf{x}_{\text{opt}}$ is defined in (5.18).

□

This theorem says that the optimal added $\mathbf{V}$ is deterministic, whose value (an N-dimensional vector) is the $\mathbf{x}_{\text{opt}}$ defined in (5.18). This is equivalent to conduct an optimal mean shift on the observation for the largest AUC. This is due to the binary TS in the TD, which functions as the test statistic of the problem. Once the structure of the delta function of the optimal pdf is found, the determination of the optimal point is straightforward. Thus, the main contribution of the theorem is to discover that the optimal $\mathbf{V}$ is deterministic. Note that $\mathbf{x}_{\text{opt}}$ may not be unique. Any $\mathbf{x}_{\text{opt}}$ that leads to the same maximum $G(\mathbf{x})$ will have identically maximum AUC.

From the proof of Theorem 5.2, for a deterministic added noise $\mathbf{v}$, we have, from (5.23),

$$\text{AUC} = \frac{1}{2}[1 + G(\mathbf{v})]. \tag{5.24}$$

We can thus determine whether noise-enhanced effect exists by comparing $G(\mathbf{v})$ with $G(\mathbf{0})$.

Corollary 5.1. (Existence of noise-enhanced effect)

1. For $\mathbf{v} \in \mathbb{R}^{\mathbf{n}}$, if $G(\mathbf{v}) > G(\mathbf{0})$, the AUC of the TD will be improved by adding the constant $\mathbf{v}$.
2. If the TS is designed to be optimal as in Theorem 5.1, the AUC of the TD cannot be increased via adding noise.

Proof. The first part of the corollary can be seen directly from the AUC formula in (5.24). Now we prove the second part. It has been shown in Theorem 5.2 that the best noise is deterministic. With the optimal TS design in Theorem 5.1, $G(\mathbf{0})$ is the maximum of $G(\mathbf{v})$, thus the AUC can no long be improved by any $\mathbf{v}$. □

Corollary 5.1 shows that the AUC can be increased via adding noise if $G(\mathbf{0})$ is not the global maximum. It also shows that if the TS is designed to be optimal, adding noise will not improve the AUC. For some non-optimal TS, if its $G(\mathbf{0})$ is the global maximum point, we also cannot increase its AUC.

Now we discuss how to find $\mathbf{x}_{\text{opt}}$ defined in (5.18), the global maximum point of $G(\mathbf{x})$. First, candidate $\mathbf{x}_c$'s should satisfy $G'(\mathbf{x}_c) = 0$ and $G''(\mathbf{x}_c) < 0$. Thus we first solve $\mathbf{x}_c$'s from $G'(\mathbf{x}_c) = 0$ and $G''(\mathbf{x}_c) < 0$, then $\mathbf{x}_{\text{opt}}$ is one of the $\mathbf{x}_c$'s resulting in the largest value of $G(\mathbf{x})$. This can be done efficiently using standard numerical algorithms such as Newton's method.

In [18], [19], the noised-enhance effect was investigated under NP criterion. In this letter, we investigate the same problem under the AUC measure. In [18], [19], for a general given test statistic, the optimal noise pdf under NP criterion was proved to be a combination of maximum two delta functions. Although the noise pdf structure was found, closed-form is rarely available. In this work, we derived the optimal noise pdf in a semi-closed form, the calculation of which is significantly simpler. Under the NP criterion, the optimal noise pdf changes with η and P_{FA}, and hence the design is sensitive to η and P_{FA}. AUC leads to an optimal noise that is independent of η and P_{FA}, and intuitively can be more sustainable to errors in the system. We would also like to note that for a given P_{FA} and under perfect design, the proposed scheme may be inferior in P_D to those in [18], [19] since the goal of the proposed noise-enhanced TD is to maximize the AUC not the P_D for a particular P_{FA}.

5.5 An Example

In this section, we use an example to illustrate the results in Secs. 5.3 and 5.4. We consider the detection of a DC signal in bimodal GM noise, as in Sec. 3.5.2, where $f_X(x; H_i), i = 0, 1$ are shown in (3.41) and (3.42) with $A = 0.5$, $\mu = 3$ and $\sigma = 1$. The optimal TS, denoted as TS_{opt}, can be derived using Theorem 5.1 to be

$$\text{TS}_{\text{opt}}: \quad Y = \begin{cases} 1 & x \in [\tau_1, \tau_2] \cup [\tau_3, \infty) \\ 0 & \text{elsewhere} \end{cases}, \tag{5.25}$$

where $\tau_1 = -2.75$, $\tau_2 = 0.25$, and $\tau_3 = 3.25$. That is $\mathcal{D}_{\text{opt}} = [-2.75, 0.25] \cup [3.25, \infty)$.

First we consider the TD with TS_{opt}. We can calculate the $G(x)$ according to (5.17) to be

$$\begin{aligned} G(x) = \frac{1}{2} \sum_{i=1}^{3} (-1)^{i-1} \Bigg[& Q\left(\frac{\tau_i - 3.5 - x}{\sigma}\right) - Q\left(\frac{\tau_i - 3 - x}{\sigma}\right) \\ & + Q\left(\frac{\tau_i + 2.5 - x}{\sigma}\right) - Q\left(\frac{\tau_i + 3 - x}{\sigma}\right) \Bigg]. \end{aligned} \tag{5.26}$$

Solving $G'(x_c) = 0$ subject to $G''(x_c) < 0$ numerically, we have $x_c \in \{\pm 6.04, 0\}$. Since $G(6.04) = G(-6.04) = 0.0975$, $G(0) = 0.1945$, we have $x_{\text{opt}} = 0$. This justifies the second part of Corollary 5.1 that noise-enhanced effect cannot occur if the TD is optimal.

Next, we consider the TD with a non-optimal TS, denoted as TS_{nopt}, given by

$$Y = \begin{cases} 1 & x \geq 0 \\ 0 & \text{elsewhere} \end{cases}. \tag{5.27}$$

According to (5.17), for this case, we have

$$G(x) = \frac{1}{2}\left[Q\left(\frac{-3.5-x}{\sigma}\right) - Q\left(\frac{-3-x}{\sigma}\right) + Q\left(\frac{2.5-x}{\sigma}\right) - Q\left(\frac{3-x}{\sigma}\right)\right]. \qquad (5.28)$$

Solving $G'(x_c) = 0$ subject to $G''(x_c) < 0$, we obtain two candidate points for the maximum: $\{-3.25, 2.75\}$. Since $G(-3.25) = G(2.75) = 0.0987$, We have two optimal solutions: $x_{\mathrm{opt}} = -3.25$ or $x_{\mathrm{opt}} = 2.75$. They both result in a higher G value than $G(0) = 0.0024$. Based on the first part of Corollary 5.1, the noise-enhanced effect appears.

We can analytically find the ROCs of the TD for the six cases: 1) using $\mathrm{TS}_{\mathrm{opt}}$, 2) using TS_{nopt}, 3) using TS_{nopt} and adding $\mathbf{V}_{\mathrm{opt}} = -3.25$, 4) using TS_{nopt} and adding the optimal WGN with pdf $f_{\mathbf{V}}(x) = \mathcal{N}(x; 0, \sigma^2)$, 5) using TS_{nopt} and adding the optimal noises under NP criterion for $P_{FA} = 0.1$, and 6) using TS_{nopt} and adding the optimal noises under NP criterion for $P_{FA} = 0.5$, respectively. For Case 4, by adding the Gaussian noise, we have

$$f_0(x) = \frac{1}{2}\mathcal{N}(x; 3, 1+\sigma^2) + \frac{1}{2}\mathcal{N}(x; -3, 1+\sigma^2),$$
$$f_1(x) = \frac{1}{2}\mathcal{N}(x; 3.5, 1+\sigma^2) + \frac{1}{2}\mathcal{N}(x; -2.5, 1+\sigma^2).$$

We can calculate the the optimal σ to be,

$$\sigma_{\mathrm{opt}} = \arg\max_{\sigma} \mathrm{AUC} = \arg\max_{\sigma} \int_0^{\infty} (f_1(x) - f_0(x))dx = 2.81.$$

For Case 5, we have $f_{V_{opt}}(v) = \delta(v + 3.86)$; For Case 6, $f_{V_{opt}}(v) = 0.4012\delta(v + 3.25) + 0.5988\delta(v - 2.75)$ using results in [18].

The ROCs of the 6 detectors are shown in Fig. 5.2. We can see that when using $\mathrm{TS}_{\mathrm{opt}}$, the AUC is 0.5975 which is larger than the AUC of using TS_{nopt}, which is 0.5015. This justifies Theorem 5.1. When TS_{nopt} is used, noise-enhanced effect happens by adding V_{opt}. The AUC of using TS_{nopt} and V_{opt} is 0.5494, which is lower than that of $\mathrm{TS}_{\mathrm{opt}}$. This is because the structure of TS_{nopt} is not optimal. With TS_{nopt}, adding the best deterministic V_{opt} is better than adding the best WGN, whose AUC is 0.5202. This justifies Theorem 5.2. Finally, under NP criterion [18], different V_{opt}^{NP}'s are needed for different P_{FA}'s, which is shown in the previous paragraph. At the specific P_{FA} value, the achieved P_D using V_{opt}^{NP} is higher than that using V_{opt} (see the P_D when $P_{FA} = 0.1$ and V_{opt}^{NP} calculated for $P_{FA} = 0.1$ is added, and the one when $P_{FA} = 0.5$ and V_{opt}^{NP} calculated for $P_{FA} = 0.5$ is added). However, adding V_{opt} achieves a AUC no smaller than adding V_{opt}^{NP}. Also, at other P_{FA} values, the P_D obtained via adding V_{opt}^{NP} may be lower than that one via adding V_{opt} (e.g., $P_{FA} = 0.2$).

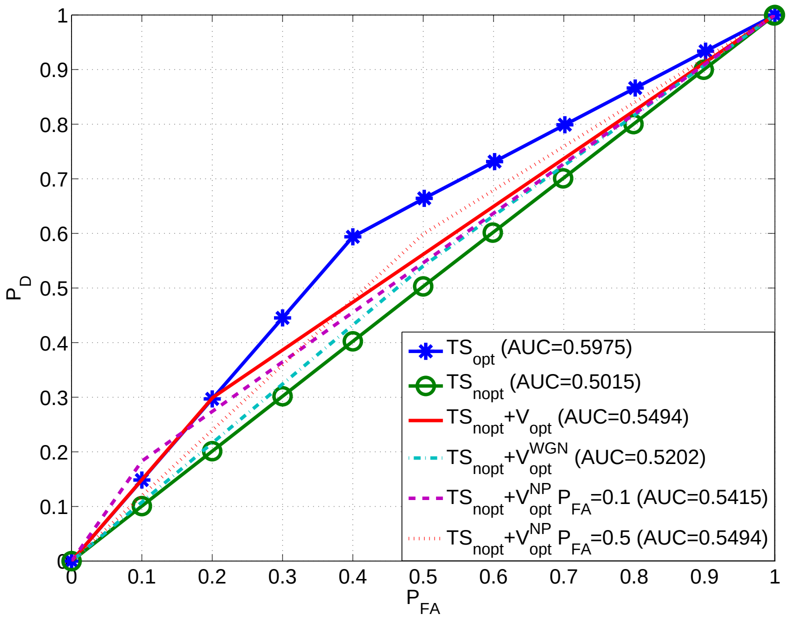


Fig. 5.2. ROC's obtained from the optimal TS, the non-optimal TS, adding optimal noise V_{opt} under AUC criterion, adding optimal WGN, adding V_{opt}^{NP} under NP criterion for $P_{FA} = 0.1$ and $P_{FA} = 0.5$, respectively.

5.6 Conclusions

In this chapter, we investigated the general BHTP using a binary TD. We adopted the AUC as the performance measure for its implementation simplicity and robustness. First the optimal multi-dimensional binary TS that maximizes the AUC was derived. Then we considered noise-enhanced effect of the detector. The optimal noise pdf that maximizes the AUC was shown to be a delta function, indicating that the optimal noise is deterministic. Performance of the proposed design was shown via an example and comparisons with other designs were made.

Chapter 6

An Adaptive Bistable System Based Detector for Watermark Extraction

In this chapter, we explore a bistable system (BS) based detector (BD) to detect a binary pulse amplitude modulation (PAM) signal embedded in unknown non-Gaussian noise. This BD has been used in discrete cosine transform (DCT) domain watermark extraction, where the watermark is the signal embedded in the DCT coefficients, and the DCT coefficients are the noises. In existing BD designs for watermark extraction, the BS parameters are determined using two methods: one is to find a set of the parameter values from extraction experiments, and this fixed set is used for different watermarks, images, and watermark extraction scenarios; the other is to obtain the best parameter values from exhaustive search of the parameters. However, we demonstrate that one specific set of BS parameter values may not provide good performance for different extraction cases. Also, exhaustive search has prohibitively high computational complexity and is inapplicable for unknown watermark sequence, i.e., blind watermark. To discover a tractable method for the BD design, we propose to use the cross-correlation of the watermark signal and the output of the BS as the performance measure, and the BS parameters are optimized for the maximum cross-correlation. Via experiments, we observe that the 3-dimensional optimization of the BS parameters can be reduced to a 1-dimensional optimization problem, which has reduced complexity. Further, when the noise pdf is unimodal and symmetric and with heavy tails, another key observation based on experiments is that the optimal BS parameters are sensitive to the variance of the noise and the amplitude of the watermark only, but not other noise statistics, such as the pdf form and the watermark sequence. Based on this observation, it is possible to generate a look-up table of the BS parameters for different watermark amplitudes and noise variances. With the help of this look-up table, an adaptive BD design is constructed, whose BS parameters are adaptive to the estimated amplitude of the watermark and the variance of the DCT coefficients. Experimental results show that the performance of the proposed adaptive BD is superior to those of the existing BDs and the white Gaussian noise (WGN)-based maximum likelihood (ML) detector.

6.1 Introduction

DCT-domain watermark extraction [38]–[42] can be considered as a BHTP, where the watermark sequence is the signal and the DCT coefficients of the host image are the i.i.d. noise samples. Consider a binary watermark sequence. When detecting a specific bit in the watermark, the signal is a binary DC signal (0 or 1), or in general, a binary PAM signal with unknown amplitude. The noise, which are the DCT coefficients of the host image, is shown to be non-Gaussian with heavy pdf tails [44]. It is usually modeled as some pdf forms, such as generalized Gaussian or Cauchy in [43], [44]. This kind of non-Gaussian noise exhibits spikes and a good detector typically includes a nonlinear limiter to reduce the noise spikes [1]. The BS can be a suitable limiter [28], [60]. Hence, the BD composed of a BS and a summation has been employed in watermark extraction [38]–[42]. The main results have been reviewed in Sec. 2.2.2.

It has been shown that BD can provide good performance in watermark extraction [38]–[42]. The BD design is to determine the BS parameters which is critical in obtaining a good extraction performance. There are two methods in BS parameter design in the literature. In the designs in [38]–[40], [42], the BS parameters are fixed for different watermarked images and for watermarked images suffered from different attacks such as JPEG compression and adding Gaussian noise. The BS parameters are determined by experimental experiences. For example, one parameter set of the BS is chosen if by some experiments it leads to a satisfied bit error rate (BER) in extraction [38], [40]. Different works [38]–[40], [42] use different BS parameter values. The other was proposed by Duan et al. in [41], [45] to obtain the appropriate BS parameters for the minimum BER using exhaustive search. However, both of the aforementioned methods require the watermark sequence to be known for the calculation of the BER in experiments or in exhaustive search. But the watermark sequence is often not available in applications and the watermark is called a blind watermark. In addition, for the first method, one fixed BS parameter set is not suitable for all scenarios in watermark extraction. For the second method, exhaustive search suffers from high computational complexity. To summarize, there is no systematic and practical method to design the BS parameters for blind watermark extraction in literature.

In this chapter, we first propose to use the cross-correlation of the watermark signal and the output of the BS as the performance measure in designing the BS. The main advantage of using cross-correlation measure to replace the BER measure in [38]–[42] is that the BS design is isolated out of the overall BD. Although this may lead to some performance penalty, it largely simplifies the design complexity so that the design can be used in real applications. It also leads to efficient extraction for blind watermark as will be shown later in this chapter. The BER measure in [38]–[42] only applies to applications with known watermark, and the BER-minimizing design is computationally prohibitive for real applications. We investigate how the cross-correlation is affected by the BS parameters via experiments, based on which, we propose a tuning one parameter (TOP) technique to determine the BS parameter values with manageable computational complexity. Furthermore, under the cross-correlation measure, we observe that the BS parameter design is sensitive to the variance of the noise and the amplitude of the signal only, not other statistics of the noise and the signal, such as the noise pdf form and the specific (watermark) signal sequence. Hence, via off-line experiments, we can generate a look-up table containing the desired BS parameters in term of different noise variances and signal amplitude values. We then design a BD whose BS parameters are adaptive to the input (watermark image) based on the look-up table. It is illustrated that applying one BS parameter set for various watermark extraction cases cannot guarantee performance and the proposed BD with adaptive BS parameters performs better.

The organization of this chapter is as follows. In Sec. 6.2, we briefly review watermark embedding and extraction using BD. In Sec. 6.3, we present the detection problem in watermark extraction, and demonstrate the limitation of keeping the BS parameters unchanged for different scenarios. In Sec. 6.4, we propose new BS design, including BS parameter optimization and the design of the adaptive BS. We then present the simulation results in Sec. 6.5, followed by conclusions in Sec. 6.6.

6.2 Review of Watermark Embedding and Extraction Using BD

A watermarking scheme includes two stages: embedding and extraction. Note that the BD is used only in the extraction. In this section, we present a brief review of the watermarking scheme and the use of BD in watermark extraction [40].

6.2.1 Embedding Algorithm

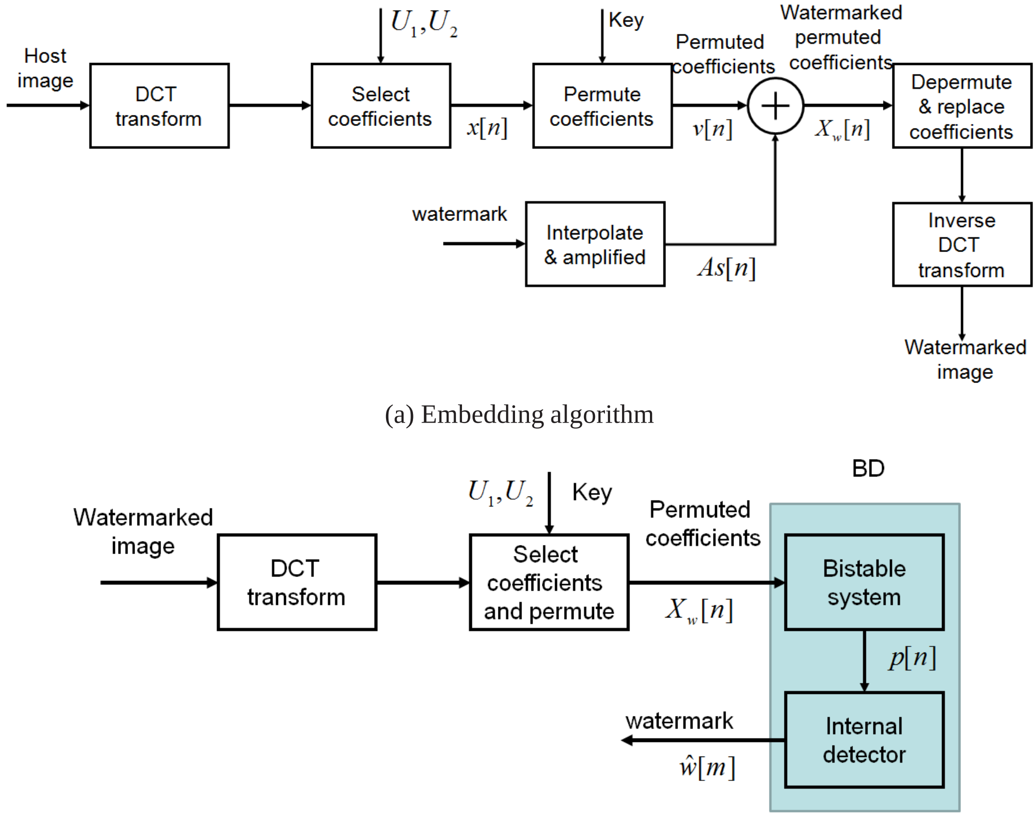
The schematic of the watermark embedding algorithm [40] is shown in Fig. 6.1 (a). The host image I is divided into K blocks, each with 8×8 pixels. The DCT coefficients of the k^{th} block also have size 8×8, which are zigzag scanned from low frequency to high frequency to obtain a one-dimensional sequence denoted as $X^k, k = 1, 2, ..., K$. The middle r DCT coefficients in X^k are denoted by $X_u^k(U_1 \leq u \leq U_2)$, where U_1 is the starting index, U_2 is the ending index, and $r = U_2 - U_1 + 1$. X_u^k's are then cascaded to generate a sequence X with a length of rK.

X is then permuted to V as follows. First, generate a random sequence R with the same length of X using a specific key. Second, generate R' by sorting elements of R in ascending order. Let L contain the index of the ascendent ordering. Finally, V is obtained via permuting X by L. For example, $X = \{10, 65, 43, 20\}$, $R = \{0.3, -0.1, -0.2, 3\}$, then $R' = \{-0.2, -0.1, 0.3, 3\}$, $L =$

$\{3, 2, 1, 4\}$, and $V = \{43, 65, 10, 20\}$.

Assume that $w[m], m \in [1, M]$ is a binary watermark sequence consisting of -1 and 1. Note that a binary watermark should be a sequence of 0 and 1. Here 0 is converted to -1, such that the least energy of watermark can be available for a required SNR. Every bit in $w[m]$ is repeated S times to generate the sequence $s[n]$, i.e., $s[n]$ is $\{\underbrace{w[1], ..., w[1]}_{S \text{ times}}, \underbrace{w[2], ..., w[2]}_{S \text{ times}}, ..., \underbrace{w[M], ..., w[M]}_{S \text{ times}}\}$, $n = 1, 2, ..., SM$. Note that $SM \leq rK$ should be satisfied to have enough DCT coefficients for $s[n]$ to be embedded.

The watermark embedding algorithm is given by $X_w[n] = v[n] + As[n]$, where A is the watermark amplitude, X_w is the watermarked DCT coefficients. X_w contains M segments, each of which contains one watermark bit. The watermarked image is generated by replacing the original DCT coefficients with $X_w[n]$, then conducting an inverse DCT transform.



(a) Embedding algorithm

(b) Extraction algorithm

Fig. 6.1. Schematic of the BD based watermarking scheme.

As an example, we consider the embedding of a 16×16 binary "W" watermark shown in Fig. 6.2 (a) into the 512×512 Lena image. $w[m], m = 1, 2, ..., 256$ is binary sequence obtained by zigzag scanning the "W" watermark. Let $A = 3$, $U_1 = 9$, $U_2 = 44$, $M = 256$ and $S = 500$. Using the embedding scheme in Fig. 6.1 (a), we have the watermarked Lena image shown in Fig. 6.2 (b).

(a) 16×16 binary watermark image (b) 512×512 watermarked Lena image

Fig. 6.2. Embedded watermark “W” and watermarked Lena image.

6.2.2 Extraction Algorithm

The schematic of watermark extraction is shown in Fig. 6.1 (b). The first two steps are exactly the inverse of the last two steps of the embedding scheme. After them, $X_w[n]$, the sequence containing watermarked DCT coefficients, is obtained. A BD is followed to extract the watermark. The $X_w[n]$ is applied to a BS to obtain the output $y[n]$. $y[n]$ is the input to the internal detector, where $y[n]$ is first partitioned into M segments, each of which includes S samples corresponding to one bit of the watermark, then a value is calculated for one segment as

$$\mu[m] = \sum_{n=(m-1)S+1}^{mS} y[n]. \tag{6.1}$$

The watermark bit $w[m]$ is recovered by,

$$\hat{w}[m] = \begin{cases} 1 & \mu[m] \geq 0 \\ -1 & \mu[m] < 0 \end{cases}. \tag{6.2}$$

One evaluation of the extraction performance is the BER, defined as

$$\text{BER} \triangleq \frac{\sum_{m=1}^{M} |w[m] - \hat{w}[m]|}{2M}. \tag{6.3}$$

BER$\in [0, 1]$. Smaller BER means better extraction performance.

6.3 Design Problems of BD in Watermark Extraction

In this section, we first model the watermark extraction as a BHTP in Sec. 6.3.1 and derive the optimal ML detector that minimizes the probability of error. In Sec. 6.3.2, we demonstrate the non-Gaussian nature of DCT coefficients, which motivates the use of BD. In Sec. 6.3.3, we show that the means and variances of the segments in one watermarked image after certain attacks vary significantly, which motivates the use of adaptive BS. Finally, in Sec. 6.3.4, we present the design problems of the BS.

6.3.1 Detection Problem in Watermark Extraction and Optimal ML Detector for WGN

In this subsection, we formulate the binary watermark extraction into a BHTP and represent the optimal ML detector if the noise is WGN. In Sec. 6.2, we reviewed the watermark embedding and extraction algorithms in [40]. The watermark bearing DCT coefficients sequence $X_w[n]$ has M segments. Each segment has S samples, and these S samples in the same segment can be considered as adding A if watermark is 1 or $-A$ if watermark is 0 to S permuted DCT coefficients. Therefore, the watermark extraction can be represented by a BHTP, shown as follows.

$$\begin{cases} H_0: & x[n] = -A + v[n] & n = 1, 2, ..., S \\ H_1: & x[n] = A + v[n] & n = 1, 2, ..., S \end{cases}. \tag{6.4}$$

For this detection problem, we want to find a detector that minimizes the probability of error, which is

$$P_e = \mathbb{P}(H_0|H_1)\mathbb{P}(H_1) + \mathbb{P}(H_1|H_0)\mathbb{P}(H_0), \tag{6.5}$$

where $\mathbb{P}(H_i|H_j), i, j = 0, 1$ is the conditional probability of deciding H_i when H_j is true. In blind watermark extraction, the prior probability, $\mathbb{P}(H_0)$ and $\mathbb{P}(H_1)$, is usually considered equal, i.e., $\mathbb{P}(H_0) = \mathbb{P}(H_1) = 1/2$. Therefore, the optimal detector reduces to the maximum likelihood (ML) detector, which is

$$L(\mathbf{x}) = \frac{f_{\mathbf{V}}(\mathbf{x} - A)}{f_{\mathbf{V}}(\mathbf{x} + A)} \underset{H_0}{\overset{H_1}{\gtrless}} 1. \tag{6.6}$$

To implement the ML detector in real applications, we need to know $f_V(v)$. Here, the noise V is DCT coefficient of a host image. Its mean is 0, which can be shown from the definition of DCT [70]. However, it is difficult to model $f_V(v)$ as a specific pdf. The simplest way is to assume that $f_V(v) = \mathcal{N}(v; 0, \sigma^2)$ [1], [9], based on which the ML detector reduces to

$$T(\mathbf{x}) = \sum_{n=1}^{S} x[n] \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{6.7}$$

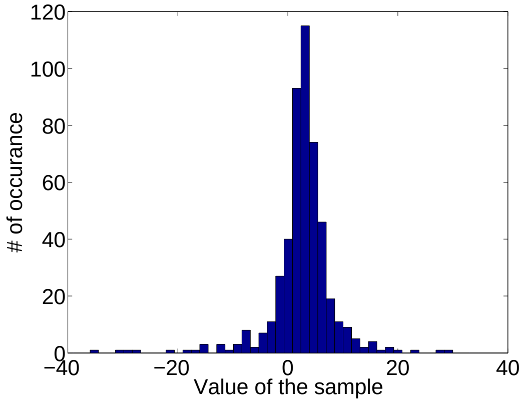
Since this detector is based on white Gaussian noise (WGN), we call it WGN-based ML detector. Because $x[n]$'s are Gaussian under either H_0 or H_1, the pdf of the output of $T(\mathbf{x})$ is also Gaussian. Therefore, via straightforward calculations, we have

$$P_e = \frac{1}{2}\Big[\mathbb{P}\big(T(\mathbf{x}) > 0|H_0\big) + \mathbb{P}\big(T(\mathbf{x}) < 0|H_1\big)\Big] = Q\left(\frac{A}{\sqrt{\sigma^2/S}}\right). \tag{6.8}$$

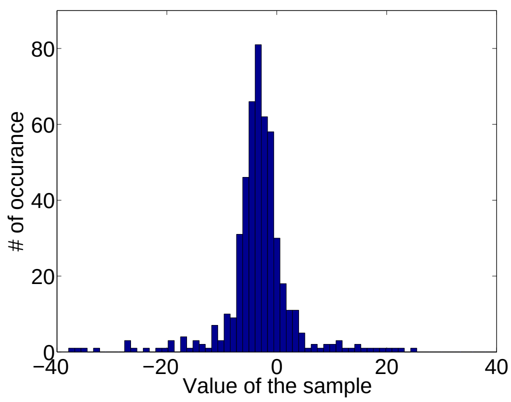
For the example considered in Sec. 6.2, where $A = 3$, $\sigma^2 = 100$, and $S = 500$, we have $P_e = 9.85 \times 10^{-12}$ from (6.8), which shows that errors occur very rarely.

The WGN-based ML detector is simple and easy to implemented. However, it is not optimal when $V[n]$ is not Gaussian. It is used as a benchmark because other detectors have no use if they cannot have superior performance to the WGN-based ML detector. Next, we will show that the distribution of DCT coefficients is far from Gaussian, which motivates the use of BD instead of WGN-based ML detector in watermark extraction.

6.3.2 Investigation on Noise Distribution and Motivation of Using BD



(a) First segment

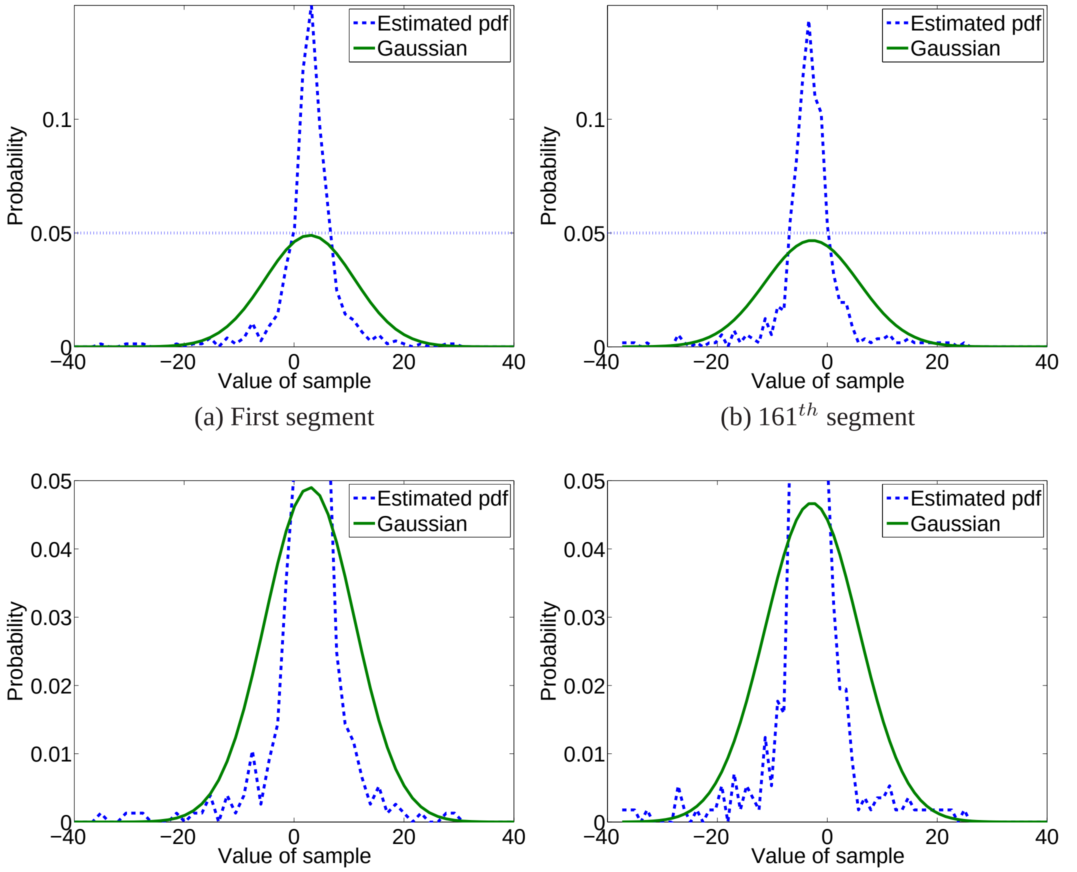


(b) 161^{th} segment

Fig. 6.3. Histograms of two segments in $X_w[n]$ of the watermarked image.

If the noise is white Gaussian, the ML detector is linear as shown in (6.7), and is optimal in the sense of BER. But it has been shown in [43], [44] that $f_V(v)$ is not Gaussian. It is bell-shaped with heavy pdf tails. In this thesis, a bell-shape pdf is unimodal and symmetric about zero. Hence, if the WGN-based ML detector is used, the performance is expected to be suboptimal and many time, unsatisfactory. The non-Gaussian $f_V(v)$ is modeled by generalized Gaussian distribution in [44] and Cauchy distribution in [43]. Based on these models with estimated parameters, the corresponding LO detectors were employed in watermark extraction, which are actually GLRTs. However, these detectors are complex in implementation and have low robustness to changes in the noise pdf.

In this subsection, we first show by experiments that $f_V(v)$ is not Gaussian but has heavy pdf tails. We then show that the distribution of $V[n]$ changes significantly after the watermarked Lena image [1] in Fig. 6.2 (b) is attacked by JPEG compression and Gaussian noise, which causes difficulties for GLRT. We also show that BD can have a better performance than the WGN-based ML detector. These motivate the use of BD in watermark extraction.



(a) First segment

(b) 161^{th} segment

(c) Enlargement of the boxed area in (a)

(d) Enlargement of the boxed area in (b)

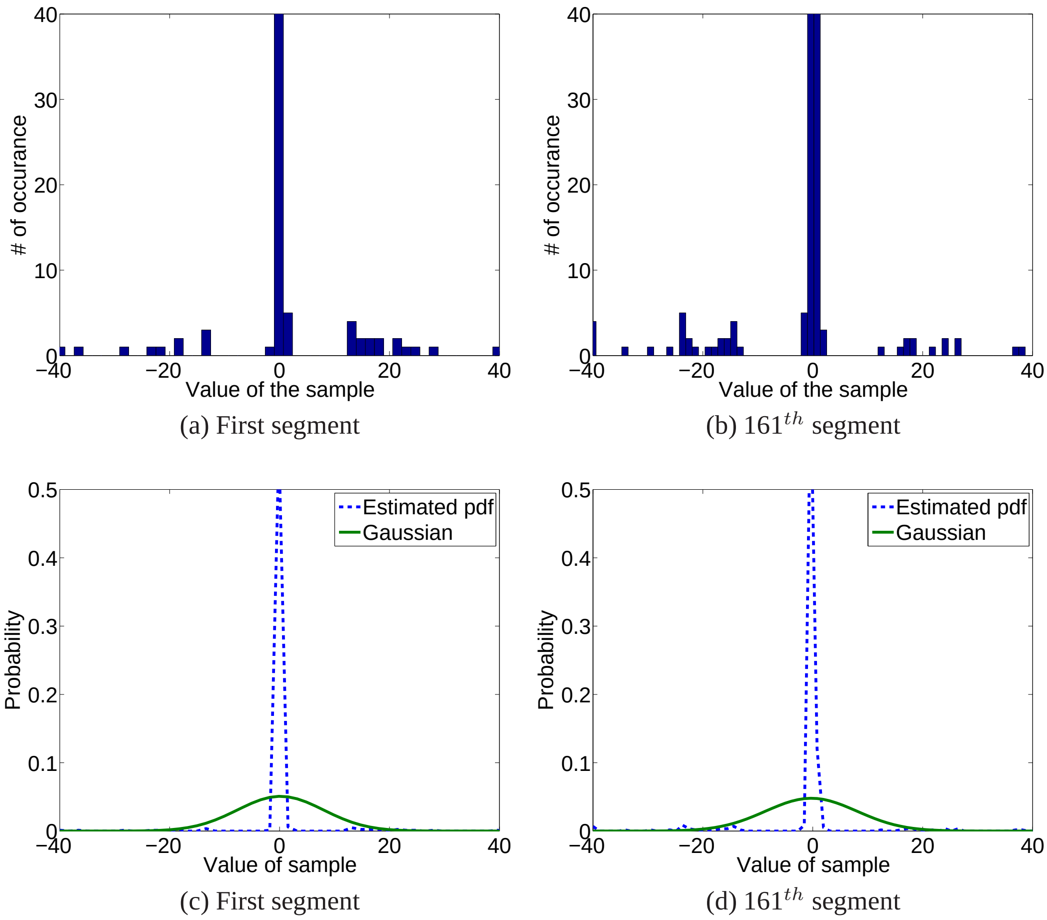
Fig. 6.4. pdfs of two segments in $X_w[n]$ and Gaussian pdfs.

We again use the example illustrated in Fig. 6.2, where watermark "W" is embedded in Lena image. Fig. 6.3 shows the histograms of the samples of two segments (the 1st and 161th segments) in $X_w[n]$ corresponding to two embedded watermark bits with values A and $-A$ respectively, where $A = 3$. To show that the pdf of the DCT coefficients is not Gaussian, we plot the pdfs estimated [2] from the two segments and the Gaussian pdfs with identical mean and variance in Fig. 6.4. We can

[1]Only the results from Lena image are presented in this thesis. We note that this claim applies for other watermarked images.

[2]This is performed by Matlab function "[vBin,xOut]=hist(data,100)" to have the occurrence times xOut in the 100 bins whose central values are in vBin. Then xOut is divided by the integral of (vBin,xOut) sequence and normalized xOut, denoted as $\mathrm{xOut}^{\mathrm{nor}}$, is obtained. The estimated pdf is the lines connecting the points ($\mathrm{vBin},\mathrm{xOut}^{\mathrm{nor}}$).

see that the pdfs of the DCT coefficients are largely different to Gaussian pdfs, but is approximately unimodal and symmetic with heavy pdf tails, as stated in [43], [44].



(a) First segment (b) 161^{th} segment

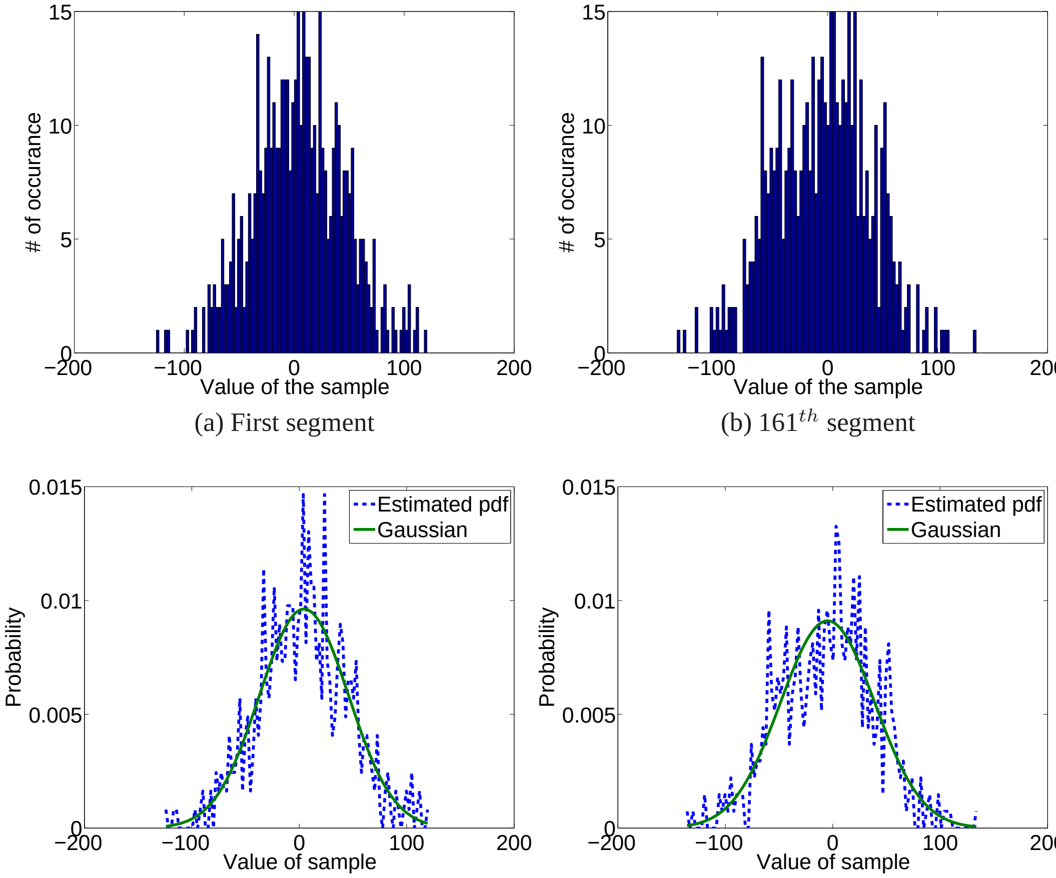
(c) First segment (d) 161^{th} segment

Fig. 6.5. Histograms and pdfs of two segments in $X_w[n]$ of the JPEG compressed watermarked image with quality of 50.

A good watermark scheme should be robust to attacks, such as JPEG compression, clipping, adding Gaussian noise, adding pepper salt noise, etc. Now, we study the distribution of the DCT coefficients when the watermarked Lena image is attached by JPEG compression and adding Gaussian noise, respectively. The histograms and pdfs along with the Gaussian pdfs with the same means and variances of the samples in the same two segments after the original watermarked image is attacked by JPEG compression [3] and Gaussian noise [4], are shown in Figs. 6.5 and 6.6, respectively.

[3]The JPEG compression is performed by the Matlab function “imwrite(imageDate, ’savedFile’, ’JPEG’, ’Quality’, qualityValue)”, which compresses the imageData into a JPEG image file named as “savedFile.jpg” with “Quality”=qualityValue. The “Quality” is a number between 0 and 100, where a higher number induces higher quality (less image degradation due to compression), but the size of the resulted “savedFile.jpg” is larger [71].

[4]The Gaussian noise is added into the watermarked image by using the Matlab function “imnoise (imageDate, ’gaussian’, mean, variance)”. We choose zero mean and variance=0.03 in the illustrated example. The mean and variance parameters for ’Gaussian’ noise are always specified as if the image were of class double in the range [0, 1]. Therefore it can be shown if the noise variance is 0.03, a quite strong noise is added to the image when the noisy image is converted back to the same class of the input [71].



(a) First segment (b) 161^{th} segment

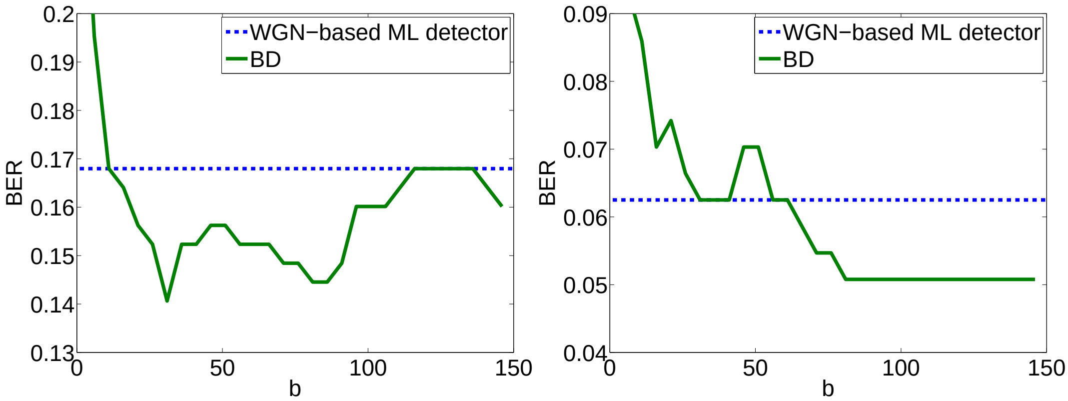
(c) First segment (d) 161^{th} segment

Fig. 6.6. Histograms and pdfs of two segments in $X_w[n]$ of the watermarked image added Gaussian noise with variance of 0.03.

We can observe from Fig. 6.5 (c) and (d) that after JPEG compression, the pdfs less resemble Gaussian pdfs. That is because when the watermarked image is compressed, many DCT coefficients are truncated to zero. As a consequence, both the watermark information embedded and the original DCT coefficients degrade. Therefore, we observe that many samples are zeros and non-zero samples are scattered. For the case of adding Gaussian noise, we can see from Fig. 6.6 (c) and (d) that the distributions of the samples in these two segments are close to Gaussian. This is because the added Gaussian noise is much stronger than the original DCT coefficients, thus dominates the noise.

Through the above experiments, we first show that $f_V(v)$ is not Gaussian. Thus, the WGN-based ML detector may have poor performance. We also observe that $f_V(v)$ is approximately unimodal and symmetric with heavy tails in all cases. For non-Gaussian noise with heavy pdf tails, a good detector should include a limiter to elute the spikes. Since the BS can be a nonlinear limiter illustrated in Sec. 2.1.2, the BD with appropriate BS parameters is expected to have better performance than the ML detector. In the following, we show an example to justify this.

We use the BD structure shown in Fig. 6.1 (b) to extract watermark bits from watermarked Lena image and calculate the BERs. The discrete BS is presented in (2.6) with parameters $(a, b, \Delta t)$. We set $a = 1$, $\Delta t = 0.01$. For different b's, the BER obtained from the BD is shown in Fig. 6.7 for the watermarked Lena image after JPEG compression with quality of 30 and 50, respectively. The BER of the WGN-based ML detector is also shown for reference. It reveals that the BD with suitable parameters, i.e., $b \in [15, 110]$ for quality 30 and $b > 70$ for quality 50, can have better performance than WGN-based ML detector.



(a) Quality 30 (b) Quality 50

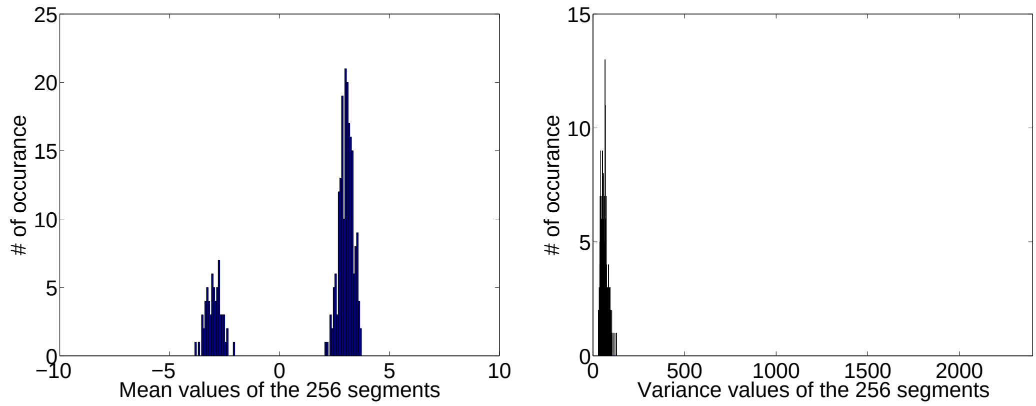
Fig. 6.7. BERs of the extracted watermark from watermarked Lena image after JPEG compression with quality 30 and 50 using BD with fixed a, Δt, and different b's.

6.3.3 Investigation on Mean and Variance of Watermark Bit Bearing Segments and Motivation of Using Adaptive BS

In this subsection, we investigate via experiments the means and variances of the segments (each will be used to extract one watermark bit) in $X_w[n]$, and motivate the use of adaptive BS based on the observations.

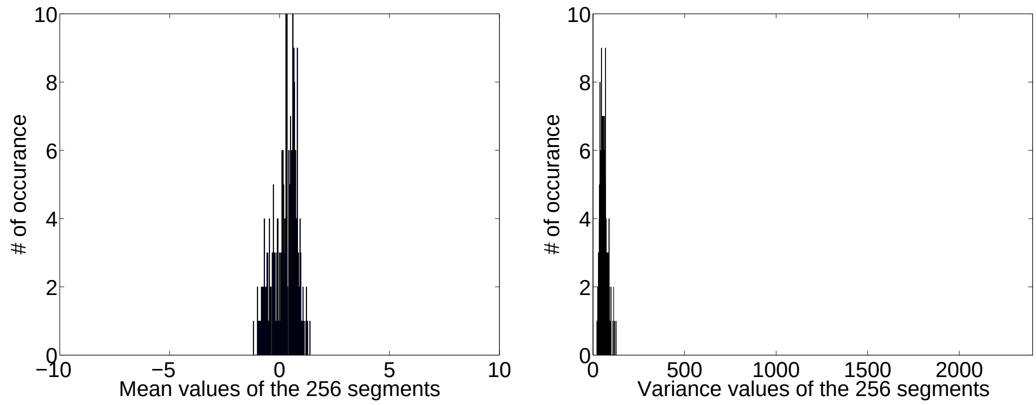
Consider the watermarked Lena image in Fig. 6.2 (b). In watermark bearing sequence $X_w[n]$, there are 256 segments, each having 500 samples corresponding to one watermark bit. Now we present histograms of the means and variances of the 256 segments in Fig. 6.8, Fig. 6.9 and Fig. 6.10, for the watermarked Lena image, the watermarked Lena image after JPEG compression, and watermarked Lena image after adding Gaussian noise, respectively.

For the original watermarked Lena image, Fig. 6.8 shows that the means of the samples in the 256 segments gather around 3 and -3 and the variances vary from 30 to 120. For this case, every segment bearing one watermarking bit contains 500 samples that are the summation of 500 DCT coefficients and 3 if the watermark bit is 1 and the summation of 500 DCT coefficients and -3 if the



(a) Mean (b) Variance

Fig. 6.8. Histograms of the means and variances of the 256 segments in $X_w[n]$ of the original watermarked Lena image.

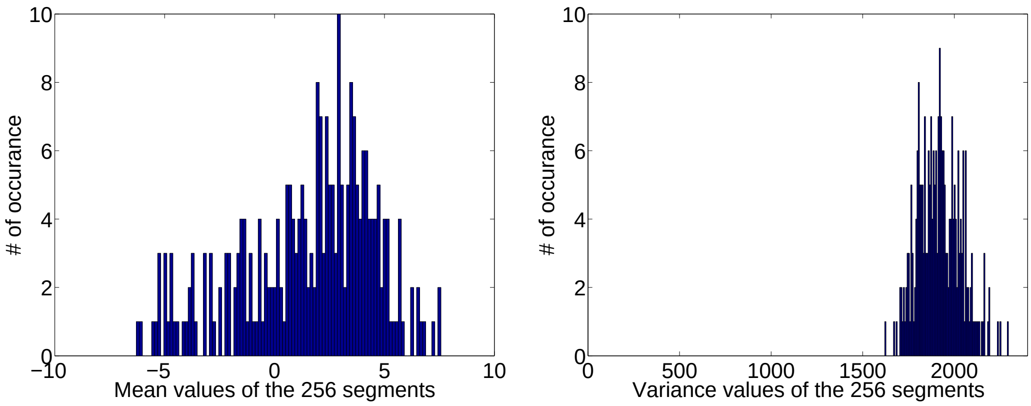


(a) Mean (b) Variance

Fig. 6.9. Histograms of the means and variances of the 256 segments in $X_w[n]$ of the watermarked Lena image after JPEG compression with quality of 50.

watermark bit is -1. Naturally, the segments whose means are near -3 correspond to the watermark -1 in watermark embedding, and the segments whose means are near 3 correspond to 1 in watermark embedding. From Fig. 6.8, watermark bits can be extracted with good performance because there is a big gap between the two categories.

When the watermarked image is attacked by JPEG compression or by adding Gaussian noise, it reveals in Figs. 6.9 and 6.10 that there is no obvious division for the mean values. For JPEG compression, the watermark information is filtered heavily, showing that the means gather around zero, but the variances has little change compared with those of the original watermarked image. With Gaussian noise attack, similarly, there is no obvious division by the mean values. This is because the added noise, which is much stronger than the original noise, masks the watermark



(a) Mean (b) Variance

Fig. 6.10. Histograms of the means and variances of the 256 segments in $X_w[n]$ of the watermarked Lena image after adding zero-mean Gaussian noise with variance 0.03.

information. This can be seen from the noise variance, which is about 2000, while for original watermarked Lena image, it is about 50.

Since the mean of the DCT coefficients is zero, the mean of the samples in a segment corresponding to one bit of the watermark can be used to estimate the watermark amplitude A, and the variance σ^2 of the samples can be used to estimate the noise strength. It can be observed from Figs. 6.8, 6.9 and 6.10, that A varies from very weak (such as 0.1) to very strong (such as 3), σ^2 also varies from very weak (such as 30) to very strong (such as 2300), for different situations. Although the variety of the mean and the variance is represented here for the watermarked Lena image attacked by JPEG compression and Gaussian noise only, further experiments show that this variety exists for other watermarked images and watermarked images subject to other attacks, such as clipping. It is thus expected that for different segments in watermark bearing DCT coefficients with different means and variances, the optimal BS design should be different. This motivates the use of adaptive BS, where the BS parameters $(a, b, \Delta t)$ are adaptive to the different segments with different means and variances.

If for simplicity consideration, it is desirable to used only one set of BS parameters for all different segments in one extraction case. It can be observed that for different extraction cases, the sample means and variances change dramatically. Thus, we expect to choose different BS parameter values for different watermarked images. This motivates the used of adaptive BS. A justification is shown in Fig. 6.11. We show the BERs obtained from the BD using BS with $a = 1$, $\Delta t = 0.01$, and different b for the watermark bearing Lena image attacked by JPEG compression and Gaussian noise. For the watermarked Lena after JPEG compression, the best b should be larger than 80; While for the watermarked Lena after adding Gaussian noise, the best b should be less than 10. This justifies that for different situations, the BS parameters should be chosen differently for good performance.

Using only one set $(a, b, \Delta t)$ for different situations will lead to degraded performance.

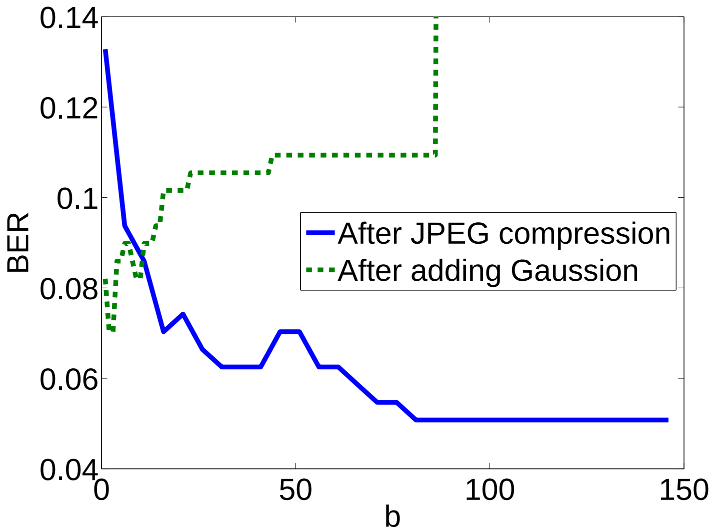


Fig. 6.11. BERs of the extracted watermark from watermarked Lena image after JPEG compression with quality 50 and after adding Gaussian noise with variance of 0.03 using BD with fixed a, Δt, and different b's.

Although not shown in Fig. 6.11, experiments for different images has been done to show similar results . The above investigations motive adaptive BD, i.e., BD whose BS parameters are adaptive to different segments, different watermarked images, or watermarked images with different attacks.

6.3.4 BS Design Problems

It has been shown that the pdf of the DCT coefficients of an image is unimodal and symmetric with heavy tails at the both sides [43], [44]. We also demonstrated the pdfs, means, and variances of the samples in the segments of the watermark bearing DCT coefficients in Secs. 6.3.2 and 6.3.3. With these observations, our detection problem is the one shown in (6.4) with the following assumptions.

1. A is a unknown DC signal.

2. $f_V(v)$ is unknown, but is unimodal and symmetric about zero with heavy pdf tails.

For this detection problem, we present the reason of using BD and the reason of having the BS parameters adaptive to the different situations in Secs. 6.3.2 and 6.3.3 respectively. Since the BD is a BS followed by a summation, the BD design is basically the design of the adaptive BS. We will address this design problem in the next section.

6.4 Proposed BS Design

The values of the BS parameters $(a, b, \Delta t)$ are crucial for the BD to have a good performance. However there is no systematic method for the BS parameter design in literature. In this section, we first propose to use the cross-correlation of the watermark signal and the BS output as the measure in finding the optimal values of the BS parameters. Secondly, we observe that the optimization of the 3-tuple $(a, b, \Delta t)$ can be reduced to the optimization of only one parameter b with fixed a and Δt. We also observe that the optimal value of parameter b depends on the signal amplitude and the noise variance only, but does not depend on signal sequence and the noise pdf form, assuming that the noise pdf is unimodal, symmetric, and with heavy pdf tails. Hence, a look-up table of optimal values of b in term of the signal amplitude A and the noise variance σ^2 can be generated via off-line simulations. This look-up table is then used to choose the BS parameters adaptively according to various segments of one watermarked image or various watermarked images according to the estimated A and σ^2.

6.4.1 Cross-Correlation as the Design Criterion

As shown in (6.3), BER measures the difference between the watermark and the extracted watermark in the BD output. BER is the direct performance measure in watermark extraction, and was used in the BD design [38]–[42]. However, the calculation of BER requires the watermark information, which is not available for blind watermark. More importantly, due to the complexity of BER calculation and analysis, no systematic method has been found in designing the BER-minimizing BS, other than exhaustive search. But the computational complexity of exhaustive search is prohibitive. Thus, in this thesis, for a tractable BS design, we consider the function of the BS itself, not the overall BER performance of the BD. The observation $x[n]$ is composed of the watermark signal $s[n]$ and the noise $v[n]$. For a good BS, the output $y[n]$ should be correlated with the watermark information $s[n]$ as much as possible. Therefore, we use cross-correlation between the watermark signal $s[n]$ and BS output $y[n]$ as a measure to determine the BS parameters, which is defined as follows,

$$C_m = \mathbb{E}\left(\frac{\sum_{n=0}^{N-m-1} s[n]y[n+m]}{\sqrt{\sum_{n=0}^{N-m-1} s^2[n]}\sqrt{\sum_{n=0}^{N-m-1} y^2[n+m]}}\right), \tag{6.9}$$

where $m > 0$ is the time lag, $\mathbb{E}(\cdot)$ denotes the average over the noises, realized via averaging over multiple experiments. The BS introduces system lag. The system lag of the BS, denoted ς, is the m that results in the maximum cross-correlation, which is

$$\varsigma = \arg\max_m C_m. \tag{6.10}$$

We attempt to find the $(a, b, \Delta t)$ to maximize C_ς, which is

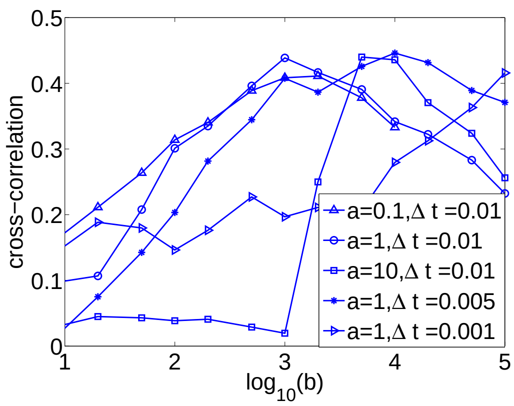
$$(a, b, \Delta t)_{\mathrm{opt}} = \arg \max_{(a,b,\Delta t)} C_\varsigma = \arg \max_{(a,b,\Delta t)} \max_m C_m. \tag{6.11}$$

It is noteworthy that in this BS design, watermark information is required. However, our experimental results, which will be shown later, illustrate that the BS design using cross-correlation measure depends on the amplitude of the watermark only, but not the specific sequence of the watermark. Thus, for blind watermark extraction, one can first obtain the BS parameters off-line via a known training watermark, whose amplitude is the same as the estimation of the watermark to be extracted, then use this BS parameter values in the watermark extraction.

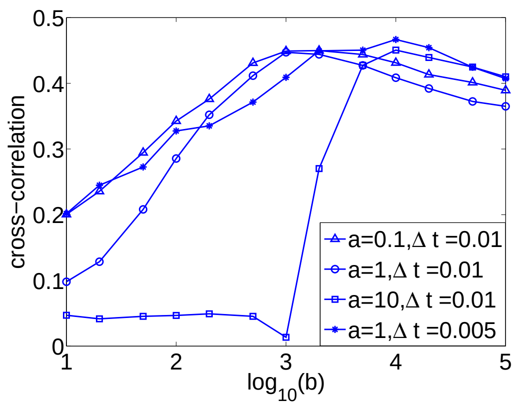
6.4.2 Observations on the Connection between Cross-Correlation and $(a, b, \Delta t)$

The task in (6.11) in general is difficult because of the complex nature of the BS. We cannot have analytical solution of (6.11). Exhaustive search is a natural method for the optimization, but it has high computational complexity. It is also impractical for blind watermark extraction. We propose an empirical method in designing the BS parameters based on the observations on the relationship between the cross-correlation and BS parameters, which are given below.

We first conduct experiments on known PAM signals with WGN and unimodal Gaussian mixture (GM) noise whose pdf is shown in (3.38). In the simulation, the signal is a 20 bit random binary PAM with $A = 0.1$. Every bit in the PAM signal is interpolated 50 times to generate a sequence $s[n]$ with length of 1000 samples. $s[n]$ is embedded in (a) WGN samples, and (b) GM noise ($(\alpha, \beta, \sigma) = (0.9, 5, 1)$) samples, to generate the input to the BS. Cross-correlation is calculated from 1000 simulations using (6.9) and (6.10), which is shown in Fig. 6.12 (a) for WGN and (b) for GM noise. We try on numerous fixed values of $(a, \Delta t)$ and calculate the cross-correlation as b



(a) White Gaussian noise



(b) Gaussian mixture noise

Fig. 6.12. Cross-correlation versus b and numerous fixed $(a, \Delta t)$ for WGN and GM noise.

changes. The optimal b that maximizes the cross-correlation with respect to each $(a, \Delta t)$ value can be found. We observe that for both cases, regardless of the $(a, \Delta t)$ values used in the experiments, the same (or very close) maximum cross-correlation is obtained. The optimal b values for different $(a, \Delta t)$'s are nevertheless different.

Based on the above observation, we propose a simplified optimization of the BS parameters by reducing the 3-dimensional optimization to a 1-dimensional optimization. First, we fix a, e.g., $a = 1$. Second, we determine a suitable Δt mainly in consideration of the stability of the BS. For the pre-determined $(a, \Delta t)$, we can conduct a 1-dimensional optimization to find the b that maximizes the cross-correlation. This can be done by exhaustive search. Limited simulation results shown in Fig. 6.12 indicate that this reduction in optimization dimension induces little cross-correlation degradation. We only need to tune one parameter (TOP) in this method, thus it is called TOP method.

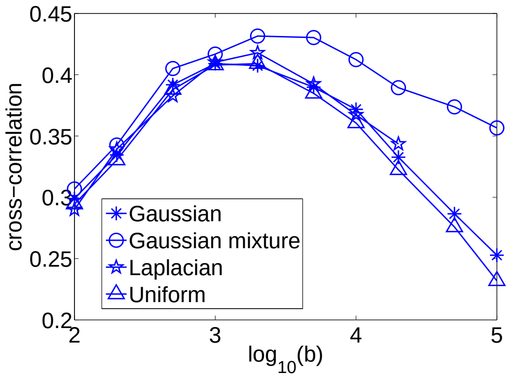
6.4.3 Observations on the Connection between Cross-Correlation and Watermark Signal and Noise

Even though we can use 1-dimensional optimization to determine the BS parameter, the computational complexity can still be too high in watermark extraction because exhaustive search is needed for every extraction task. Also, to calculate the cross-correlation, watermark signal must be known, which makes it impossible for blind watermark extraction. We obtain another important observation that the BS parameters are only sensitive to the signal amplitude A and noise variance σ^2, but not sensitive to other parameters, such as the signal sequence and the pdf form of the noise if the noise pdf is symmetric and unimodal with heavy tails.

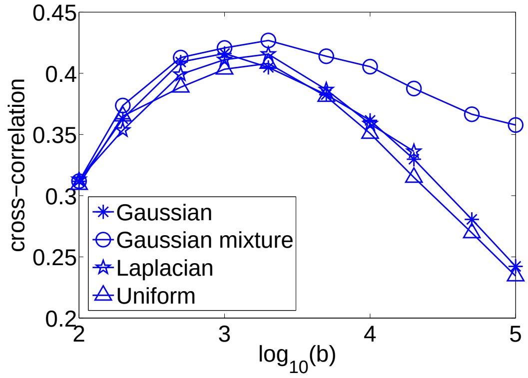
To see this, in Fig. 6.13, we show the cross-correlation versus the BS parameter b for several PAM signals with different amplitudes and several noises with different variances. Simulation setting is the same as previous example, whose results are shown in Fig. 6.12. In (a), the PAM signal $\#1$ is [1 1 -1 1 1 -1 -1 1 1 1 -1 1 1 -1 1 -1 -1 1 1 1] and $A = 0.1$. We choose 4 noises: Gaussian, GM, Laplacian, and uniform noises with $\sigma^2 = 1$. The pdf of GM is defined in (3.38) with $\alpha = 0.9$ and $\beta = 5$. GM, Laplacian, and uniform noises are unimodal and symmetric with heavy pdf tails. In (b), the PAM signal $\#2$ is [1 -1 1 1 1 1 1 -1 1 -1 1 -1 -1 -1 -1 1 1 -1 1 -1] again with $A = 0.1$. The same set of noises are used. Comparing (a) with (b), according to cross-correlation measure, we can observe that the optimal BS parameter b is not sensitive to the PAM signal sequence.

Since the optimal BS parameters are not sensitive to the PAM signal sequence, the only parameter related to PAM signal is the amplitude A. We illustrate the cross-correlation for randomly generated PAM signal with $A = 0.4$ in Fig. 6.13 (c). Compared (c) with (a) and (b), the influence of A to the BS parameter b can be observed.

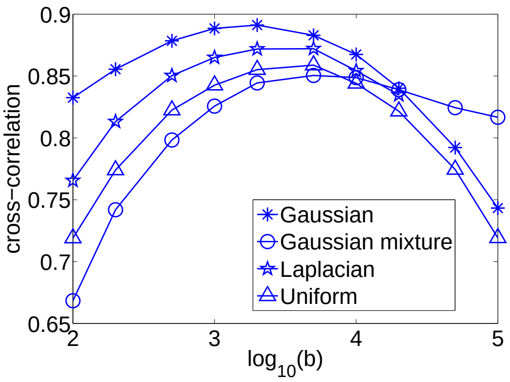
For the noise, from (a), (b), and (c), we can observe that the optimal BS parameter b is not



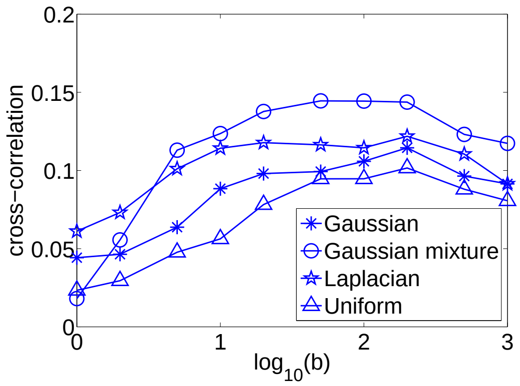
(a) For PAM signal #1.



(b) For PAM signal #2.



(c) Random PAM signal with $A = 0.4$.



(d) Noises with $\sigma^2 = 16$.

Fig. 6.13. Cross-correlation versus b for different PAM signals with different amplitudes and different noises with different variances.

sensitive to the noise pdfs if they are unimodal and symmetric with heavy tails and have the same variance, because for GM, Laplacian, and uniform noises, whose variances are the same, the optimal b are the same. Compared with the optimal b for PAM signal in Gaussian noise, the optimal b for GM, Laplacian, and uniform noises is larger. That is expected because GM, Laplacian, and uniform noises have heavy pdf tails and need strong "limiter", which is fulfilled by increasing b value. The optimal BS parameter b is sensitive to the noise variance σ^2. We demonstrate the influence of σ^2 to the cross-correlation in (d), where a randomly generated PAM signal with $A = 0.1$ is used, and for the noises, $\sigma^2 = 16$. The optimal b is changed by comparing (d) with (a) and (b).

This observation, the optimal BS parameters are not sensitive to the watermark sequence (PAM signal) and the pdf form of the noise, but to the amplitude of the PAM signal and the noise variance only, provides an easy way to determine the BS parameters off-line. We can generate a random PAM signal embedded in any noise whose pdf is unimodal, symmetric, and has heavy tails, and apply this to the BS to find the optimal BS parameters. Combining with the TOP technique in Sec. 6.4.2, for a fixed $(a, \Delta t)$ set, we can construct a look-up table containing the optimal parameter b in term of

A (amplitude of watermark) and σ^2 (power of the noise). In this thesis, we use the GM noise with $\alpha = 0.9$ and $\beta = 5$ in the simulation to generate the look-up table.

Next we will show how to use the look-up table to construct an adaptive BD and use it in watermark extraction.

6.4.4 Adaptive BS

Based on the previous observations and designs, instead of using one set of $(a, b, \Delta t)$ for all segments of a watermarked image, various watermarked images, and images under different attacks [39], [40], we propose a BD whose BS parameters are adaptive to different segments or different watermarked images.

The proposed watermark extraction algorithm is similar to that shown in Fig. 6.1 (b), except the BS module. The BS module in Fig. 6.1 (b) is changed to an adaptive BS and its schematic is shown in Fig. 6.14. There are 4 sub-modules in the adaptive BS design. These sub-modules are explained below.

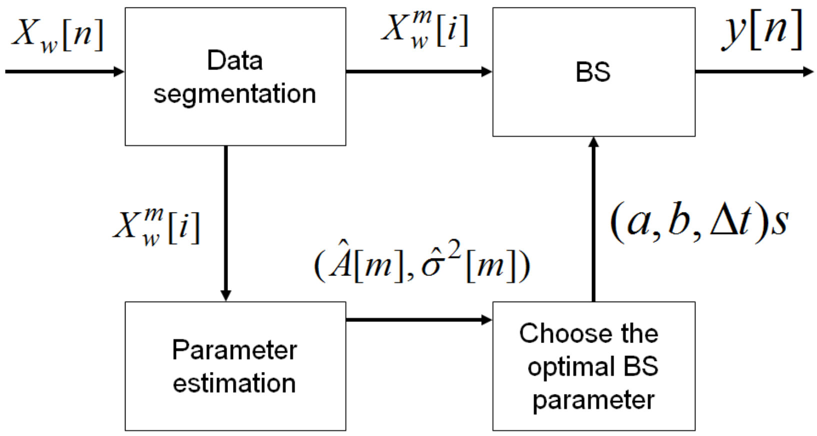


Fig. 6.14. Schematic of adaptive bistable system.

1. Data segmentation: $X_w[n]$ is partitioned into M segments, denoted as $X_w^m[i], m \in [1, M], i \in [1, S]$. $X_w^m[i] = X_w[(m-1)S + i]$ contains the samples that the m^{th} bit of the watermark is embedded in.

2. Parameter estimation of the mean $\hat{A}[m]$ and variance $\hat{\sigma}^2[m]$. We use the maximum likelihood estimation (MLE) of the embedded watermark amplitude A and the noise variance σ^2 in the m^{th} segment, which are shown as follows.

$$\hat{A}[m] = \frac{1}{S} \left| \sum_{i=1}^{S} X_w^m[i] \right|, \tag{6.12}$$

$$\hat{\sigma}^2[m] = \frac{1}{S}\sum_{j=1}^{S}\left(X_w^m[j] - \frac{1}{S}\sum_{i=1}^{S} X_w^m[i]\right)^2. \tag{6.13}$$

3. Choose the optimal BS parameter: We choose an appropriate b from the look-up table according to the estimated $(\hat{A}, \hat{\sigma}^2)$. Note that a and Δt are previously fixed.

4. BS: The watermark bearing sequence is processed by the BS with the chosen parameters. The output $y[n]$ is generated, which is used to extract watermark information according to (6.1) and (6.2).

The proposed BD has a BS whose parameters are adaptive to every segment $X_w^m[i]$. For every segments, the $(\hat{A}[m], \hat{\sigma}^2[m])$ are estimated and the suitable parameters are chosen from the look-up table. For systems with stringent processing limitation and delay tolerance, for simplicity, we can only estimate one $(\hat{A}, \hat{\sigma}^2)$ for one image (not for every segment) and use only one BS parameter set for the whole image. The MLE of the (A, σ^2) of one image is shown below:

$$\hat{A} = \frac{1}{M}\sum_{m=1}^{M} \hat{A}[m], \tag{6.14}$$

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{m=1}^{M} \hat{\sigma}^2[m], \tag{6.15}$$

where $\hat{A}[m]$ and $\hat{\sigma}^2[m]$ are the estimated watermark amplitude and variance of the m^{th} segment as shown in (6.12) and (6.13).

We name BS-I for the BD that chooses one set of BS parameters only for one image (using (6.14) and (6.15)). We name BS-II for the BD that chooses a set of BS parameters for every segment. Compared to existing BD designs in [38]–[42], our scheme chooses BS parameters adaptively based on the $(\hat{A}, \hat{\sigma}^2)$ estimated from different images or different segments of one image. The performance is expected to be better than the BD with fixed parameters. In addition, the design of the BS parameters is very fast and practical.

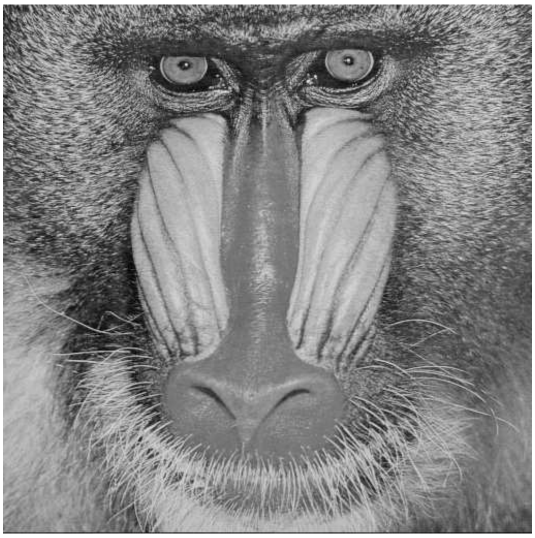
6.5 Experimental Results

In this section, experimental results are given to illustrate the BERs when extracting watermark from the watermarked images attacked by JPEG compression and Gaussian noise. The images, Lena, Peppers, Goldhill, and Baboon are chosen as the host images, which are shown in Fig. 6.15. We embed watermark “W” shown in Fig. 6.2 (a) into host images by the embedding algorithm presented in Sec. 6.2 with $A = 3$, $U_1 = 9$, $U_2 = 44$, $M = 256$, $S = 500$. We present the performance results from the WGN-based ML detector (abbreviated as ML detector below), Wu’s method [40], Sun's method [38], BS-I detector, and BS-II detector. Wu's method [40] and Sun's method [38] used different and fixed BS parameters in extraction. The peak SNRs (PSNRs) of the

(a) Lena.

(b) Peppers.

(c) Goldhill.



(d) Baboon.

Fig. 6.15. Host images used in watermark experiments.

watermarked images are all 41.07 dB. PSNR is defined as

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{Max}_I^2}{\mathrm{MSE}}\right)$$

Here, Max_I is the maximum possible pixel value of the original image. When the pixels are represented using 8 bits per sample, Max_I is 255. MSE is the mean square error between the pixels of the original image I and the pixels of watermarked image J, i.e.,

$$\text{MSE} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I(m,n) - J(m,n))^2,$$

where M, N are the numbers of pixels in every column and row, respectively. Comparing the original Lena image in Fig. 6.15 (a) and the watermarked Lena image in Fig. 6.2 (b), we can hardly distinguish any difference between them. That means that we can have satisfactory transparency of the watermark for the watermark embedding with PSNR of 41.07 dB.

First, we show the extraction performance for the watermarked images attacked by JPEG compression. The BERs for the 4 images and means are shown in Tab. 6.1 with respect to the five techniques and the compression qualities from 80 to 30. The BERs are all 0 when the compression qualities are 100 and 90 for all five techniques. Note that if the compression quality is 100, no compression attack is applied to the watermarked image. The BERs in bold font are the minimum values for a specific case. The parameters used in Wu's method are $(a, b, \Delta t) = (4000, 3 \times 10^{11}, 10^{-6})$. The parameters used in Sun's method are $(a, b, \Delta t) = (500, 3 \times 10^{10}, 10^{-5})$. For BS-I, the parameters used are $\Delta t = 0.01$, $a = 1$, and $b = [40, 40, 45, 45, 50, 60]$ for Lena image corresponding to the six compression qualities, $b = [30, 30, 40, 40, 30, 45]$ for Peppers image, $b = [15, 15, 15, 15, 30, 30]$ for Goldhill image, and $b = [3, 3, 3, 3, 3, 3]$ for Baboon image. For BS-II, different b is chosen for every segment with fixed $\Delta t = 0.01$ and $a = 1$. From Tab. 6.1, it is observed that the performance of WGN-based ML detector is always better than Wu's method. Sun's method can provide a better performance than the ML detector. The performance of the proposed BS-I and BS-II detectors are shown in the two right-most columns. It is observed that the performance of the proposed BDs are generally better than the ML detector Wu's method, and Sun's method. Especially, we can see from the mean values of BER that BS-II is better than other designs.

Second, we conduct experiments on the watermarked images attacked by Gaussian noise with variances from 0.01 to 0.03. The parameters used in Wu's method and in Sun's method are identical to the parameters used previously. For BS-I, the parameters used are $\Delta t = 0.01$, $a = 1$, and $b = [40, 20, 15, 10, 5]$ for Lena image corresponding to the five variances of the additive WGN, $b = [40, 40, 20, 15, 10]$ for Peppers image, $b = [40, 40, 15, 10, 10]$ for Goldhill image, and $b = [20, 15, 10, 5, 5]$ for Baboon image. For BS-II, different b is chosen for every segment with fixed $\Delta t = 0.01$ and $a = 1$. The results are shown in Tab. 6.2. The WGN-based ML detector has better performance than other designs. We can see from the mean values of BER, the proposed BDs are better than Wu's method and Sun's method, and closely reach the BER of the ML detector.

We show that the BD has an improved performance compared to the ML detector when the watermarked images are attacked by JPEG compression. However, the BD cannot have a better performance than the ML detector when the image is attacked by Gaussian noise. This can be

TABLE 6.1
BERS (IN %) OF THE EXTRACTED WATERMARK FOR THE FOUR WATERMARKED IMAGES ATTACKED BY JPEG COMPRESSION WITH DIFFERENT QUALITIES.

Image	Quality	Methods				
		ML method	Wu's method	Sun's method	BS-I	BS-II
Lena	80	**0**	0.39	**0**	**0**	**0**
	70	**0**	0.43	**0**	**0**	**0**
	60	**1.56**	9.38	**1.56**	1.95	1.95
	50	**6.25**	14.45	6.64	**6.25**	6.64
	40	10.55	18.75	9.77	9.77	**8.98**
	30	16.8	26.56	**14.45**	16.41	16.41
Peppers	80	**0**	1.56	**0**	**0**	**0**
	70	2.73	4.69	1.56	1.56	**1.17**
	60	7.03	10.16	2.34	**1.56**	1.95
	50	12.5	18.36	**8.98**	9.38	**8.98**
	40	18.75	25.39	16.41	16.41	**15.23**
	30	30.86	35.55	25	**24.61**	25
Goldhill	80	**0**	0.39	**0**	0.39	0.39
	70	0.39	1.56	**0**	**0**	**0**
	60	1.17	3.52	**0.78**	**0.78**	**0.78**
	50	**1.95**	5.47	**1.95**	2.34	**1.95**
	40	7.42	7.42	**4.69**	**4.69**	**4.69**
	30	13.67	14.06	10.16	10.55	**9.77**
Baboon	80	1.17	1.56	1.17	1.17	**0.78**
	70	1.95	1.95	**1.56**	1.95	**1.56**
	60	**2.73**	3.52	**2.73**	3.13	**2.73**
	50	4.3	6.25	4.3	**3.91**	**3.91**
	40	5.86	7.81	5.47	**5.08**	**5.08**
	30	10.94	11.33	11.72	10.16	**9.77**
Mean	80	**0.29**	0.98	**0.29**	0.39	**0.29**
	70	1.27	2.16	0.78	0.88	**0.69**
	60	3.12	6.65	**1.85**	1.86	**1.85**
	50	6.25	11.13	5.47	5.47	**5.37**
	40	10.65	14.84	9.09	8.99	**8.5**
	30	18.07	21.88	15.33	15.43	**15.24**

seen from the mean values of BER. The reasons are depicted as follows. If the noise is WGN, the WGN-based ML detector is the optimal detector. When the image is compressed by JPEG, many coefficients are converted to 0 that leads to the loss of the signal (watermark) as well as the noise.

TABLE 6.2
BERS (IN %) OF THE EXTRACTED WATERMARK FOR THE FOUR WATERMARKED IMAGES ATTACKED BY GAUSSIAN NOISE WITH DIFFERENT VARIANCES.

Image	Variance	Methods				
		ML method	Wu's method	Sun's method	BS-I	BS-II
Lena	0.01	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**
	0.015	**2.73**	**2.73**	3.52	**2.73**	3.13
	0.02	4.69	5.47	**3.91**	3.52	3.52
	0.025	3.91	3.13	3.13	**2.34**	2.73
	0.03	8.2	8.59	**7.42**	7.81	7.81
Peppers	0.01	**0.78**	1.95	1.17	1.17	1.56
	0.015	**2.34**	3.91	**2.34**	**2.34**	3.13
	0.02	**3.91**	7.03	5.08	4.3	5.08
	0.025	8.2	8.2	**6.25**	7.42	7.03
	0.03	**7.81**	9.77	8.59	8.98	9.38
Goldhill	0.01	**1.56**	**1.56**	**1.56**	**1.56**	**1.56**
	0.015	**2.34**	3.52	**2.34**	**2.34**	2.73
	0.02	4.69	**4.3**	5.47	5.47	5.47
	0.025	**5.47**	**5.47**	7.81	6.64	6.64
	0.03	**8.59**	11.33	9.77	10.55	10.94
Baboon	0.01	**3.52**	**3.52**	**3.52**	**3.52**	3.91
	0.015	**5.08**	6.25	6.64	6.25	6.25
	0.02	**5.86**	7.03	7.42	7.03	6.64
	0.025	**5.86**	7.42	6.25	**5.86**	7.03
	0.03	10.55	10.94	**9.77**	10.55	10.94
Mean	0.01	**1.56**	1.86	1.66	1.66	1.86
	0.015	**3.12**	4.35	3.71	3.66	3.81
	0.02	**4.79**	5.96	5.47	5.07	5.18
	0.025	5.86	6.06	5.86	**5.52**	5.86
	0.03	**8.79**	10.16	8.89	9.47	9.77

The coefficients in each segment is not WGN but with heavy tails, which is shown in Fig. 6.5. The WGN-based ML detector is not optimal detector, and the BD provides an improved performance due to the clipping feature of BS. When adding Gaussian noise to the watermarked image, the noise is dominated by the added Gaussian. Hence, the noise pdf is close to Gaussian (see Fig. 6.6). In this case, the WGN-based ML detector tends to be the optimal detector. The BD cannot have an improved performance. However, the BD has comparable performance to the ML detector.

6.6 Conclusions

The motivation for the research in this chapter is to find a simple and practical method to determine the BS parameters when it is used in a BD in watermark extraction. We first proposed to use the cross correlation of the watermark signal and the BS output as the design measure. We then observed that the same maximum cross-correlation can be obtained by choosing the best b for different $(a, \Delta t)$'s. Hence, the optimization of the BS parameters can be reduced from the 3-dimensional optimization of $(a, b, \Delta t)$ to the 1-dimensional optimization of b while fixing a and Δt. We also observed that the optimal b depends only on the signal amplitude and the noise variance. Hence, a look-up table of the optimal b in term of the signal amplitude and noise variance can be generated via off-line simulations. Using this look-up table, we proposed a BD design, whose BS parameter values are adaptive to the various watermarked images or even the various segments in one watermarked image. It showed that the proposed BD performs better than the WGN-based ML detector and the BD with a fixed BS when the watermarked images are attacked by JPEG compression. When adding Gaussian noise, the proposed BD has BERs close to the ones of using ML detector. Note that in this scenario ML detector is optimal .

Chapter 7

Conclusions and Future Work

7.1 Conclusions

For signal detection in non-Gaussian noise, the challenges arise mainly from the cost/complexity consideration and the robustness in the ever-changing environment. In this thesis, we aim at designing simple and robust detectors that still enjoy a detection performance comparable to LRT or GLRT. For this purpose, two techniques: (1) TD or BD design and (2) noise-enhanced effect, are investigated.

In Chap. 3, we considered the optimal TD design for detecting a known DC signal in known non-Gaussian noise. Under NP criterion, we showed that the detection probability monotonically increases with an alternative indicator. Based on this much more tractable indicator, we derived the optimal designs when using simple binary TS and composite binary TS. Experimental results show the validation of the optimal TS design. The performance of the proposed TDs were shown to be superior to the MF for non-Gaussian noise with heavy pdf tails, and can perform very close to the LO detector with a much simpler implementation.

In Chap. 4, we proposed a low-complexity TD to detect any known deterministic signal embedded in independent unknown non-Gaussian noise. The optimality of the two parts of the proposed detector, the binary TS array and the correlator, was proved. The detection probability and the ROC of the proposed TD were investigated both analytically and numerically. For noises with heavy pdf tails, simulation showed that the performance of the proposed TD approaches that of the LO detector and Saha's detector in [6], in which the two designs need exact noise pdf information, and is much better than that of the MF. Through a robustness measure, we showed that the proposed TD is highly robust to the noise pdf. On the other hand, its robustness to the signal is inferior but comparable to the LO detector and Saha's detector. The implementation complexity of the proposed detector was discussed and compared with other detecter designs. The validity region of the proposed detector was defined and analyzed using the MF as the benchmark.

In Chap. 5, we investigated the noise-enhanced effect for the general BHTP using a binary TD. We adopted the AUC as the performance measure for its simplicity and robustness potential. First the optimal TS that maximizes the AUC was derived. Then the optimal noise pdf that maximizes the AUC was shown to be a delta function, indicating that the optimal noise is deterministic. Experiments showed that the noise-enhanced effect can be employed for some fixed non-optimal detectors.

In Chap. 6, we considered the use of BD in watermark extraction. We first proposed to maximize the cross-correlation of the watermark signal and the BS output in the BS design for the BD. We observed that the same maximum cross-correlation can be obtained when we reduce the 3-dimensional optimization of the BS parameters to a 1-dimensional optimization. We also observed that the optimal BS parameter values depend only on the watermark amplitude and the noise variance, but not on other parameters. These motivated the use of a look-up table of optimal BS parameters in term of the watermark amplitude and noise variance. Based on this look-up table, we proposed a BS design that is adaptive to various watermarked images or even various segments in one watermarked image. Experiments showed that the adaptive BD can achieve a better performance than WGN-based ML detector and existing BDs in DCT domain watermark extraction.

7.2 Future Work

We have proposed several possible robust detectors with simple implementation, including TD, BD, and noise-enhanced TD. One future work is how to further improve the proposed detectors and their practical use in real applications.

The proposed robust TD for known signal in unknown noise in Chap. 4 provides potentials in real applications, such as communication and image processing, where the signal information is more reliable than the noise information. However, the performance of the proposed TD depends on the noise pdf even the design does not. In some applications, the noise pdf or its form can be obtained from estimation. Thus, GLRT can have very good performance. For some noise pdf with certain parameters, the proposed TD can have worse performance than the MF, as shown in Sec. 4.4.3. This leads to the following detection design, in which GLRT, proposed TD, and MF are used alternatively for different situations, i.e., a hybrid strategy to choose among MF, GLRT, and proposed TD, based on the accuracy level of the pdf estimation. GLRT has ideal performance if the noise pdf information is accurate. But with imperfect noise pdf information, GLRT detector does not work well. For this case, we resort to a more robust detector, proposed TD or MF, whichever results in better performance.

For a specific application, we may have a certain pdf form to model the noise. For example, the GM can be used to model the ocean acoustic noise [7], [8]. With the known noise pdf form, we can estimate the parameters of the noise pdf based on observations. We then evaluate the quality of the estimation by comparing the observations with the estimated pdf. If the quality level is larger than a threshold, the pdf is considered as well fitted and GLRT will be chosen [1], [2]. If the quality level is smaller than the threshold, we consider the noise pdf information to be inaccurate and resort to proposed TD or MF. To choose between proposed TD and MF, if the estimated noise pdf parameters are within the validity region of proposed TD, proposed TD is used; otherwise, MF is used.

Using this hybrid strategy, we can take advantage of the three detectors: GLRT, proposed TD, and MF. We expect an improved overall performance from the hybrid strategy compared with using any one detector only. The practical implementation of the hybrid strategy, for example, how to evaluate the accuracy level of the estimated pdf, simple way to choose between proposed TD and MF without knowing exact pdf, need to be studied further.

Another future work is to design a simple, robust detector for the BHTP IV presented in Sec. 1.3. BHTP IV is a more realistic model for radar and sonar applications. We are seeking and designing suitable simple systems that lead to robust detector, which can also have comparable performance to the optimal detector.

References

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. New Jersey: Prentice Hall PTR, 2008.

[2] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. New York: Springer, 2008.

[3] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Springer, 2005.

[4] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. New Delhi: Tata McGraw-Hill, 2002.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, K. Gettman, Ed. New Jersey: Prentice Hall PTR, 1993.

[6] A. A. Saha and G. V. Anand, "Design of detectors based on stochastic resonance," *Signal Processing*, vol. 83, no. 6, pp. 1193–1212, Jun. 2003.

[7] W. Machell, C. S. Penrod, and G. E. Ellis, "Statistical characteristics of ocean acoustic noise processes," in *Topics in Non-Gaussian Signal Processing*, E. J. Wegman, S. C. Schwartz, and J. B. Thomas, Eds. New York: Springer, 1989, ch. 3, pp. 29–57.

[8] D. R. Powell and G. R. Wilson, "Class a modeling of ocean acoustic noise processes," in *Topics in Non-Gaussian Signal Processing*, E. J. Wegman, S. C. Schwartz, and J. B. Thomas, Eds. New York: Springer, 1989, ch. 2, pp. 17–28.

[9] S. M. Kay and D. Sengupta, "Detection in incompletely characterized colored non-Gaussian noise via parametric modeling," *IEEE Trans. Signal Process.*, vol. 41, no. 10, pp. 3066–3070, Oct. 1993.

[10] S. M. Saberali and H. Amindavar, "Blind mem pdf estimation for detection in the presence of unknown non-Gaussian noise," *IEEE Signal Process. Lett.*, vol. 16, no. 12, pp. 1087 –1090, Dec. 2009.

[11] F. Chapeau-Blondeau, “Nonlinear test statistic to improve signal detection in non-Gaussian noise,” *IEEE Signal Process. Lett.*, vol. 7, no. 7, pp. 205–207, Jul. 2000.

[12] K. Wiesenfeld and F. Moss, “Stochastic resonance and the benefits of noise: from ice ages to crayfish and squids,” *Nature*, vol. 373, pp. 33–36, Jan. 1995.

[13] F. Moss, L. M. Ward, and W. G. Sannita, “Stochastic resonance and sensory information processing: A tutorial and review of application,” *Clinical Neurophysiology*, vol. 115, no. 2, pp. 267–281, Feb. 2004.

[14] P. Jung, “Stochastic resonance and optimal design of threshold detectors,” *Physics Letters A*, vol. 207, no. 1-2, pp. 93–104, Oct. 1995.

[15] F. Chapeau-Blondeau, “Stochastic resonance and optimal detection of pulse trains by threshold devices,” *Digital Signal Processing*, vol. 9, no. 3, pp. 162–177, Jul. 1999.

[16] S. Kay, “Can detectability be improved by adding noise?” *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.

[17] S. Kay, J. H. Michels, H. Chen, and P. Varshney, “Reducing probability of decision error using stochastic resonance,” *IEEE Signal Process. Lett.*, vol. 13, no. 11, pp. 695–698, Nov. 2006.

[18] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, “Theory of the stochastic resonance effect in signal detection: part I-fixed detectors,” *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3171–3184, July 2007.

[19] H. Chen and P. K. Varshney, “Theory of the stochastic resonance effect in signal detection: part II-variable detectors,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2008.

[20] A. Patel and B. Kosko, “Optimal noise benefits in neyman-pearson and inequality-constrained statistical signal detection,” *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1655 –1669, May 2009.

[21] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, “Stochastic resonance,” *Review Modern Physics*, vol. 70, no. 1, pp. 223–287, Jan. 1998.

[22] S. C. Chapra, *Applied Numerical Methods with MATLAB for Engineers and Scientists*. New York: McGraw-Hill, 2006.

[23] A. S. Asdi and A. H. Tewfik, “Detection of weak signals using adaptive stochastic resonance,” in *Proceedings of ICASSP 1995*, Washington DC, USA, 1995, pp. 1332–1335.

[24] R. Benzi, A. Sutera, and A. Vulpiani, “The mechanism of stochastic resonance,” *Journal of Physics A: Mathematical and General*, vol. 14, no. 11, pp. L453–L457, Nov. 1981.

[25] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, “Stochastic resonance in climatic change,” *Tellus*, vol. 24, 1982.

[26] ——, “A theory of stochastic resonance in climatic change,” *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 565–578, June 1983.

[27] K. Wiesenfeld, T. Wellens, and A. Buchleitner, *Stochastic Resonance*, book series lecture notes in physics ed., ser. Prentice-Hall signal processing series, S. Haykin, Ed. Berlin/Heidelberg: Springer, 2002, vol. 611.

[28] G. Guo and M. Mandal, “Design of stochastic-resonator-based detector using bistable system,” in *Proceeding of IEEE SPCOM 2010*, july 2010.

[29] M. E. Inchiosa and A. R. Bulsara, “Signal detection statistics of stochastic resonators,” *Phys. Rev. E*, vol. 53, no. 3, pp. R2021–R2024, Mar. 1996.

[30] V. Galdi, V. Pierro, and I. M. Pinto, “Evaluation of stochastic-resonance-based detectors of weak harmonic signals in additive white Gaussian noise,” *Phys. Rev. E*, vol. 57, no. 6, pp. 6470–6479, June 1998.

[31] J. B. Thomas, “Nonparametric detection,” *Proc. IEEE*, vol. 58, no. 5, pp. 623–631, May 1970.

[32] S. Kassam, “Optimum quantization for signal detection,” *IEEE Trans. Commun.*, vol. 25, no. 5, pp. 479–484, May 1977.

[33] J. Miller and J. Thomas, “Robust detectors for signals in non-Gaussian noise,” *IEEE Trans. Commun.*, vol. 25, no. 7, pp. 686–690, July 1977.

[34] P. Hänggi, M. E. Inchiosa, D. Fogliatti, and A. R. Bulsara, “Nonlinear stochastic resonance: The saga of anomalous output-input gain,” *Phys. Rev. E*, vol. 62, no. 5, pp. 6155–6163, Nov. 2000.

[35] B. Xu, F. Duan, and F. Chapeau-Blondeau, “Comparison of aperiodic stochastic resonance in a bistable system realized by adding noiseand by tuning system parameters,” *Phys. Rev. E*, vol. 69, no. 6, p. 061110, June 2004.

[36] B. Xu, F. Duan, R. Bao, and J. Li, “Stochastic resonance with tuning system parameters: the application of bistable systems in signal processing,” *Chaos, Solitons and Fractals*, vol. 13, no. 4, pp. 633 – 644, 2002.

[37] F. Duan and B. Xu, “Parameter-induced stochastic resonance and baseband binary pam signals transmission over an awgn channel,” *International Journal of Bifurcation and Chaos*, vol. 13, no. 2, pp. 411–425, 2003.

[38] S. Sun and P. Qiu, “Algorithm of digital watermarking based on parameter-induced stochastic resonance,” *Journal on Communications*, vol. 26, no. 12, pp. 48–55, Dec. 2005.

[39] S. Sun and S. Kwong, “Digital watermarking based on stochastic resonance signal processor,” *PCM 2007, LNCS4810*, pp. 367–376, 2007.

[40] G. Wu and Z. Qiu, “A novel watermarking scheme based on stochastic resonance,” in *Proceedings of the 8th International Conference on Signal Processing*, 2006.

[41] F. Duan and D. Abbott, “The application of nonlinear bistable detectors to dct-domain watermarking schemes,” in *Procedings Of SPIE*, vol. 6802, no. 15, 2008.

[42] S. Sun and B. Lei, “On an aperiodic stochastic resonance signal processor and its application in digital watermarking,” *Signal Processing*, vol. 88, no. 8, pp. 2085 – 2094, 2008.

[43] A. Briassouli and M. G. Strintzis, “Locally optimum nonlinearities for DCT watermark detection,” *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1604–1617, Dec. 2004.

[44] J. R. Hernandez, M. Amado, and F. Perez-Gonzalez, “DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure,” *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 55–68, Jan. 2000.

[45] D. Rousseau, J. Rojas, F. Duan, and F. Chapeau-Blondeau, “Evaluation of a nonlinear bistable filter for binary signal detection,” *International Journal of Bifurcation and Chaos*, vol. 15, no. 2, pp. 667–679, 2005.

[46] H. Chen, P. K. Varshney, J. H. Michels, and S. Kay, “Approaching near optimal detection performance via stochastic resonance,” in *Proceedings of ICASSP 2006* , vol. 3, May 2006.

[47] H. Chen, P. Varshney, and J. Michels, “Improving sequential detection performance via stochastic resonance,” *IEEE Signal Process. Lett.*, vol. 15, pp. 685–688, 2008.

[48] M. Guerriero, S. Marano, V. Matta, and P. Willett, “Stochastic resonance in sequential detectors,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 2–15, Jan. 2009.

[49] S. Kay, “Noise enhanced detection as a special case of randomization,” *IEEE Signal Process. Lett.*, vol. 15, pp. 709 –712, 2008.

[50] S. Bayram and S. Gezici, “On the improvability and nonimprovability of detection via additional independent noise,” *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 1001 –1004, Nov. 2009.

[51] ——, “Noise-enhanced m-ary hypothesis-testing in the minimax framework,” in *Proceedings of ICSPCS 2009*, Sept. 2009, pp. 1–6.

[52] S. Bayram, S. Gezici, and H. V. Poor, “Noise enhanced detection in the restricted bayesian framework,” in *Proceedings of ICASSP 2010* , Mar. 2010, pp. 3778 –3781.

[53] A. Patel and B. Kosko, “Optimal noise benefits in neyman-pearson signal detection,” in *Proceedings of ICASSP 2008*, Apr. 2008, pp. 3889 –3892.

[54] S. Zozor and P.-O. Amblard, “Stochastic resonance in locally optimal detectors,” *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177 – 3181, Dec. 2003.

[55] D. Rousseau, G. Anand, and F. Chapeau-Blondeau, “Noise-enhanced nonlinear detector to improve signal detection in non-Gaussian noise,” *Signal Processing*, vol. 86, no. 11, pp. 3456 – 3465, 2006.

[56] G. Guo and M. Mandal, “On optimal threshold and structure in threshold system based detector,” *Signal Processing*, vol. 92, no. 1, pp. 170–178, Jan. 2012.

[57] ——, “Optimal design of a threshold system based detector,” in *Proceedings of IEEE MWSCAS 2011*, Aug. 2011.

[58] G. Guo, M. Mandal, and Y. Jing, “A robust detector of known signal in non-Gaussian noise using threshold systems,” *Signal Processing*, vol. 92, no. 11, pp. 2676–2688, Nov. 2012.

[59] G. Guo, X. Yu, Y. Jing, and M. Mandal, “Optimal design and noise-enhanced effect for binary threshold detector under AUC measure,” *(Submitted to) IEEE Signal Process. Lett.*, May 2012.

[60] G. Guo and M. Mandal, “Design of bistable system based nonlinear detector,” in *Proceedings of IEEE MWSCAS 2010*, 2010.

[61] ——, “An adaptive bistable system based detector and its application in watermark extraction,” in *Proceedings of IEEE MWSCAS 2011*, Aug. 2011.

[62] ——, “An adaptive stochastic-resonator-based detector and its application in watermark extraction,” *WSEAS Transaction on Signal Processing*, vol. 7, no. 2, pp. 65–81, Apr. 2011.

[63] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, “Low-complexity digital encoding strategies for wireless sensor networks,” in *Proceedings of ICASSP*, vol. 6, Seattle, Washington, USA, May 1998, pp. 3273–3276.

[64] S. Kay and D. Sengupta, “Recent advances in the theory and application of non-Gaussian autoregressive processes,” in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Ed. Englewood Cliffs: Prentice-Hall, Jan. 1991.

[65] M. S. Nixon and A. S. Aguado, *Feature Extraction and Image Processing*, 2nd ed. London: Elsevier, 2008.

[66] S. Atapattu, C. Tellambura, and H. Jiang, “Analysis of area under the ROC curve of energy detection,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1216–1225, Mar. 2010.

[67] T. D. Wickens, *Elementary Signal Detection Theory*. New York: Oxford University Press, 2002.

[68] M. Thompson, “Applications of robustness measures in signal detection,” in *Advances in Communications and Signal Processing*, Berlin: Springer, 1989, vol. 129, pp. 83–90.

[69] A. A. Saha and V. G. Guha, “Detectors based on stochastic resonance, Part 2: Convergence analysis and perturbative corrections,” *Signal Processing*, vol. 87, no. 1, pp. 134–147, Jan. 2007.

[70] M. Mandal, *Multimedia Signals and Systems*, 2nd ed. Norwell: Kluwer Academic Publishers, Nov. 2002.

[71] “Matlab Help Document,” 2008.