NOISE ENHANCED DETECTION

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCES

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

By

Suat Bayram

June 2009

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Sinan Gezici (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Orhan Arıkan

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. İbrahim Körpeoğlu

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

NOISE ENHANCED DETECTION

Suat Bayram
M.S. in Electrical and Electronics Engineering
Supervisor: Asst. Prof. Dr. Sinan Gezici
June 2009

Performance of some suboptimal detectors can be improved by adding independent noise to their measurements. Improving the performance of a detector by adding a stochastic signal to the measurement can be considered in the framework of stochastic resonance (SR), which can be regarded as the observation of "noise benefits" related to signal transmission in nonlinear systems. Such noise benefits can be in various forms, such as a decrease in probability of error, or an increase in probability of detection under a false-alarm rate constraint. The main focus of this thesis is to investigate noise benefits in the Bayesian, minimax and Neyman-Pearson frameworks, and characterize optimal additional noise components, and quantify their effects.

In the first part of the thesis, a Bayesian framework is considered, and the previous results on optimal additional noise components for simple binary hypothesis-testing problems are extended to M-ary composite hypothesis-testing problems. In addition, a practical detection problem is considered in the Bayesian framework. Namely, binary hypothesis-testing via a sign detector is studied for antipodal signals under symmetric Gaussian mixture noise, and the effects of shifting the measurements (observations) used by the sign detector are investigated. First, a sufficient condition is obtained to specify when the sign detector

based on the modified measurements (called the "modified" sign detector) can have smaller probability of error than the original sign detector. Also, two sufficient conditions under which the original sign detector cannot be improved by measurement modification are derived in terms of desired signal and Gaussian mixture noise parameters. Then, for equal variances of the Gaussian components in the mixture noise, it is shown that the probability of error for the modified detector is a monotone increasing function of the variance parameter, which is not always true for the original detector. In addition, the maximum improvement, specified as the ratio between the probabilities of error for the original and the modified detectors, is specified as 2 for infinitesimally small variances of the Gaussian components in the mixture noise. Finally, numerical examples are presented to support the theoretical results, and some extensions to the case of asymmetric Gaussian mixture noise are explained.

In the second part of the thesis, the effects of adding independent noise to measurements are studied for M-ary hypothesis-testing problems according to the minimax criterion. It is shown that the optimal additional noise can be represented by a randomization of at most M signal values. In addition, a convex relaxation approach is proposed to obtain an accurate approximation to the noise probability distribution in polynomial time. Furthermore, sufficient conditions are presented to determine when additional noise can or cannot improve the performance of a given detector. Finally, a numerical example is presented.

Finally, the effects of additional independent noise are investigated in the Neyman-Pearson framework, and various sufficient conditions on the improvability and the non-improvability of a suboptimal detector are derived. First, a sufficient condition under which the performance of a suboptimal detector cannot be enhanced by additional independent noise is obtained according to the Neyman-Pearson criterion. Then, sufficient conditions are obtained to specify

when the detector performance can be improved. In addition to a generic condition, various explicit sufficient conditions are proposed for easy evaluation of improvability. Finally, a numerical example is presented and the practicality of the proposed conditions is discussed.

Keywords: Hypothesis testing, noise enhanced detection, Bayes decision rule, minimax, Neyman-Pearson, stochastic resonance (SR), sign detector.

ÖZET

GÜRÜLTÜ İLE GELİŞTİRİLMİŞ SEZİM

Suat Bayram
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Asst. Prof. Dr. Sinan Gezici
Haziran 2009

Optimal olmayan bazı detektörlerin girdisine bağımsız gürültü eklenerek, detektörün performansı artırılabilir. Bu olay, doğrusal olmayan sistemlerde sinyal iletimi sırasında gürültü yararının gözlemlenmesi şeklinde de tanımlanabilen stokastik rezonans (SR) kavramı ile ilintilidir. Gürültü yararı, hata ihtimalinin azalması ya da belirli yanlış tespit seviyesi altında doğru tespit ihtimalinin artması gibi birçok farklı şekilde gözlemlenebilir. Bu tezin temel olarak yoğunlaştığı konu, gürültü yararının *Bayesian*, *minimax* ve *Neyman-Pearson* kriterlerinde çalışılması ve optimal gürültünün formunun ve etkilerinin incelenmesidir.

Tezin ilk kısmında, ikili basit hipotez testleri için optimal gürültü formu ile ilgili literatürdeki önceki sonuçlar, çoklu bileşik hipotez testlerine genişletilmektedir. Buna ek olarak, iki kutuplu sinyallerin Gauss karışımı (*Gaussian mixture*) gürültüsü altında işaret detektörü ile tespit edilmesi problemi, Bayesian kriterine göre analiz edilmektedir. Bu analizde, işaret detektörü tarafından kullanılan gözlemleri kaydırmanın sonuçları araştırılmaktadır. İlk olarak, kaydırılmış gözlemler kullanan işaret detektörünün, orijinal işaret detektöründen daha düşük hata olasılığına sahip olması için bir yeterli koşul sunulmaktadır. Bunun yanında, orijinal detektörün geliştirilemediği iki yeterli koşul elde edilmektedir. Bu yeterli koşullar, sinyal ve Gauss karışımı gürültüsünün

parametreleri cinsinden bulunmaktadır. Gauss karışımı gürültüsündeki Gauss bileşenlerinin standard sapmaları eşit olduğu zaman, hata ihtimali değistirilmiş detektör için monoton artan bir fonksiyondur. Bu durum, orijinal detektör için her zaman geçerli değildir. Buna ek olarak, Gauss karışımı gürültüsündeki Gauss bileşenlerin standard sapmaları sıfıra gittiği zaman, orijinal hata ihtimalinin değistirilmiş detektörün hata ihtimaline oranının, yani gelişim oranının, en fazla ikiye eşit olduğu gösterilmektedir. Son olarak, sayısal örneklerle teorik sonuçlar desteklenmekte ve teorik sonuçların simetrik olmayan Gauss karışımı gürültüsüne nasıl genişletilebileceğiyle ilgili yorumlar yapılmaktadır.

Tezin ikinci kısmında, çoklu (M'li) hipotez testlerinde, detektörlerin kullandığı gözlemlere bağımsız gürültü eklemenin, minimax kriteri altındaki etkileri analiz edilmektedir. Optimal gürültünün ihtimal yoğunluk fonksiyonunun en fazla M farklı değer icin sıfırdan farklı olabileceği ispatlanmaktadır. Buna ek olarak, polinom zamanda optimal gürültünün ihtimal yoğunluk fonksiyonunun yaklaşık olarak "convex relaxation" yöntemiyle elde edilebileceği gösterilmektedir. Ayrıca, detektör performansının gürültüyle hangi durumlarda geliştirilip geliştirilemiyeceğiyle ilgili yeterli koşullar sunulmaktadır. Son bölümde ise, sayısal bir örnek üzerine çalışılmaktadır.

Son olarak, Neyman-Pearson kriteri altında, gözleme bağımsız gürültü eklemenin detektör performansı üzerindeki etkileri incelenmektedir. Bu bağlamda, optimal olmayan bir detektörün performansının geliştirilip geliştirilemiyeceği durumlarla ilgili yeterli koşullar çıkarılmaktadır. İlk olarak, Neyman-Pearson kriterine göre, optimal olmayan bir detektörün performansının hangi durumda gözleme bağımsız gürültü ekleme yoluyla geliştirilemiyeceğiyle ilgili yeterli koşul sunulmaktadır. Daha sonra, detektörün geliştirilebilmesiyle ilgili yeterli koşullar elde edilmektedir. Genel koşulların yanında, geliştirilebilirliliğin kolay test edilebilmesine imkan sağlayan çeşitli yeterli koşullar da önerilmektedir. Son olarak, sayısal örnekler sunulmakta ve önerilen koşulların pratik değerleri tartışılmaktadır.

Anahtar Kelimeler: Hipotez testi, gürültüyle geliştirilmiş sezim, Bayes kuralı, minimax, Neyman-Pearson, stokastik rezonans (SR), işaret detektörü.

ACKNOWLEDGMENTS

I was so lucky to have Asst. Prof. Dr. Sinan Gezici as my advisor. He has been one of the few people who had vital influence on my life. His endless energy, perfectionist personality and inspirational nature have been a great admiration for me. It was a real privilege and honor for me to work with such a visionary and inspirational advisor. I would like to, especially, thank him for providing me great research opportunities and environment.

I also extend my special thanks to Sara Bergene, Zakir Sözduyar, Mehmet Barış Tabakcıoğlu, Mustafa Ürel, Abdülkadir Eryıldırım, Saba Öz, Mahmut Yavuzer, Hamza Soğancı, M. Emin Tutay, Yasir Karışan, Sevinç Figen Öktem, Aslı Ünlügedik, and Başak Demirbağ for being wonderful friends and sharing unforgettable moments together. Also I would like to thank Professor Orhan Arıkan and Asst. Prof. Dr. İbrahim Körpeoğlu for agreeing to serve on my thesis committee. I would also like to thank TÜBİTAK for its financial support which was vital for me.

I would also like to thank Elif Eda Demir, who has been a true friend for my life, for the spectacular moments we had.

Finally, I would like to give a special thank to my mother Zehra, and my brothers Yakup, Cavit and Acar Alp for their unconditional love and support throughout my studies. They mean everything to me. I have much gratitude towards my mother for helping me believe that there is nothing one cannot accomplish.

Contents

List of Figures

In memory of my father . . .

Chapter 1

Introduction

1.1 Objectives and Contributions of the Thesis

Performance of some suboptimal detectors can be improved by adding independent noise to their measurements. Improving the performance of a detector by adding a stochastic signal to the measurement is referred to as noise enhanced detection [1], [2]. Noise enhanced detection can also be considered in the framework of stochastic resonance (SR), which can be regarded as the observation of "noise benefits" related to signal transmission in nonlinear systems [3]-[17]. Such noise benefits can be in various forms, such as an increase in output signal-to-noise ratio (SNR) [5], [3], a decrease in probability of error [18], or an increase in probability of detection under a false-alarm rate constraint [1], [17].

Although adding noise to a system commonly degrades its output, SR presents an exception to that intuition, which is observed under special circumstances. The SR was first studied in [3] to explain the periodic recurrence of ice gases. In that work, presence of noise was taken into account to explain a natural phenomenon. Since then, the SR concept has been employed in numerous nonlinear systems, such as optical, electronic, magnetic, and neuronal systems [8].

The first experimental verification of the SR phenomenon was the investigation of the behavior of the Schmitt trigger in an electronic bistable system [19].

Considering the probability of error as the performance criterion, it is shown in [18] that the optimal additional signal that minimizes the probability of error of a suboptimal detector has a constant value. In other words, addition of a stochastic signal to the measurement corresponds to a shift of the measurements in that scenario. Hence, when the aim is to minimize the probability of error, improvement of detector performance by utilizing SR can be regarded as threshold adaptation, which has been applied in various fields, such as in radar problems [20]. Although the formulation of the optimal signal value is provided in [18], no studies have investigated sufficient conditions for improvability and non-improvability of specific suboptimal detectors according to the minimum probability of error criterion, and quantified performance improvements that can be achieved by measurement modifications. In addition, the effects of additional noise have not been investigated for composite hypothesis-testing problems in the Bayesian framework.

An important application of the results in [18] includes the investigation of the effects of additional noise (effectively, measurement shifts) on sign detectors that operate under Gaussian mixture noise. Motivated by the fact that, under zero-mean Gaussian noise, signals with opposite polarities minimize the error probability of a sign detector based on correlation outputs [21], antipodal signaling with sign detection has been extensively used in communications systems [22]. In fact, sign detectors can be employed as suboptimal detectors in symmetric non-Gaussian noise environments as well due to their low complexity [23]. Therefore, it is of interest to investigate techniques that preserve the low complexity structure of the sign detector but improve the overall receiver performance by modifying the measurements (observations) used by the detector. In this thesis, the effects of adding a stochastic signal to measurements are investigated for sign detection of antipodal signals under symmetric Gaussian mixture noise. The Gaussian mixture model is encountered in many practical scenarios, such as characterization of multiple-access-interference (MAI) [24], ultra-wideband (UWB) communications systems [25], localization [26] and acquisition [27] problems.

In Chapter 2 of the thesis, optimal additional noise is shown to have a constant value for M-ary composite hypothesis-testing problems in the Bayesian framework, which extends the results in [18]. In other words, the optimal additional noise corresponds to a shift of the measurements for M-ary composite hypothesis-testing problems as well. Then, the effects of measurement shifts are investigated for sign detection of antipodal signals under symmetric Gaussian mixture noise according to the minimum probability of error criterion. First, a sufficient condition is obtained for measurement shifts to reduce the probability error of a sign detector in terms of desired signal and Gaussian mixture noise parameters. Then, two conditions under which the performance of the original detector cannot be improved are derived. Also, for equal variances of the Gaussian components in the mixture noise, the probability of error for the modified detector is characterized as a monotone increasing function of the variance. It is also shown via numerical examples that the original detector does not have this property in general. In addition, a theoretical performance comparison is made between the original and the modified detectors for small variances of the Gaussian components in the mixture noise, and it is shown that the maximum ratio between the probabilities of error for the original and the modified detectors is equal to two. As a byproduct of this result, sufficient conditions for improvability and non-improvability of the sign detector are obtained for infinitesimally small variance values. Finally, numerical examples are presented to support the theoretical results, and some concluding remarks are made.

In addition to the Bayesian criterion, performance of some detectors can be evaluated according to the minimax criterion in the absence of prior information about the hypotheses [21], [28]. The study in [29] utilizes the results in [18] and [1] in order to investigate optimal additional noise for suboptimal variable detectors in the Bayesian and minimax frameworks. Although the formulation of optimal additional noise is studied for a binary hypothesis-testing problem in [29], no studies have investigated M-ary hypothesis problems under the minimax framework, and provided the structure of the optimal noise probability density functions (PDFs) and sufficient conditions for the improvability and the non-improvability of a given detector.

In Chapter 3 of this thesis, noise enhanced detection is studied for M-ary hypothesis-testing problems in the minimax framework. First, the formulation of optimal additional noise is provided for an M-ary hypothesis-testing problem according to the minimax criterion. Then, it is shown that the optimal additional noise can be represented by a randomization of no more than M signal levels. In addition, a convex relaxation approach is proposed to obtain an accurate approximation to the noise PDF in polynomial time. Also, sufficient conditions are provided regarding the improvability and non-improvability of a given detector via additional noise.

In the absence of prior information about the hypotheses, the Neyman-Pearson criterion considers the maximization of detection probability under a constraint on the probability of false alarm [21]. In the framework of noise enhanced detection, the aim is to obtain the optimal additional noise that maximizes the probability of detection under a constraint on the probability of false alarm [1], [17]. In [1], a theoretical framework is developed for this problem, and the PDF of optimal additional noise is specified. Specifically, it is proven that optimal noise can be characterized by a randomization of at most two discrete signals, which is an important result as it greatly simplifies the calculation of optimal noise PDFs. Moreover, [1] provides sufficient conditions under which the performance of a suboptimal detector can or cannot be improved via additional independent noise. The study in [17] focuses on the same problem and obtains the optimal additional noise PDF via an optimization theoretic approach. In addition, it derives alternative improvability conditions for the case of *scalar* observations.

In Chapter 4 of this thesis, new improvability and non-improvability conditions are proposed for detectors in the Neyman-Pearson framework, and the improvability conditions in [17] are extended. The results also provide alternative sufficient conditions to those in [1]. In other words, new sufficient conditions are derived, under which the detection probability of a suboptimal detector can or cannot be improved by additional independent noise, under a constraint on the probability of false alarm. All the proposed conditions are defined in terms of the probabilities of detection and false alarm for specific additional noise values without the need for any other auxiliary functions employed in [1]. In addition to deriving generic conditions, simpler but less generic improvability conditions are provided for practical purposes. The results are compared to those in [1], and the advantages and disadvantages are specified for both approaches. In other words, comments are provided regarding specific detection problems, for which one approach can be more suitable than the other. Moreover, the improvability conditions in [17] for *scalar* observations are extended to both more generic conditions and to the case of vector observations.

1.2 Organization of the Thesis

The organization of the thesis is as follows. In Chapter 2, optimal additional noise is characterized for M-ary composite hypothesis-testing problems in the Bayesian framework, and the effects of additional noise are investigated for conventional sign detectors under symmetric Gaussian mixture noise.

In Chapter 3, noise benefits are investigated for M-ary hypothesis-testing problems under the minimax framework. Both the optimal additional noise characterization is provided, and a technique for obtaining the optimal additional noise components is proposed.

In Chapter 4, new improvability and non-improvability conditions are proposed for suboptimal detectors in the Neyman-Pearson framework, and the improvability conditions in [17] are extended. The results also provide alternative sufficient conditions to those in [1].

Chapter 2

Noise Enhanced Detection in the Bayesian Framework and Its Application to Sign Detection under Gaussian Mixture Noise

This chapter is organized as follows. In Section 2.1, optimal additional noise is characterized for M-ary composite hypothesis-testing problems according to the Bayesian criterion. Then, based on the results in Section 2.1, noise enhanced detection is studied for sign detectors under Gaussian mixture noise in the remaining sections. In Section 2.2.1, the system model is introduced, and the Gaussian mixture measurement noise is described. Then, Section 2.2.2 studies the optimal additional independent noise for minimizing the probability of decision error for a sign detector under symmetric Gaussian mixture noise. In Section 2.2.3, conditions on desired signal amplitude and/or the parameters of Gaussian mixture noise are derived in order to specify whether the performance of the detector can be improved. After that, the probability of error performance of the noise enhanced detector is investigated, and a monotonicity property of the probability of error and the maximum improvement ratio are derived in Section 2.2.4. Finally, numerical examples are studied in Section 2.2.5, and concluding remarks and extensions are presented in Section 2.3.

2.1 Noise Enhanced M-ary Composite Hypothesis-Testing in the Bayesian Framework

2.1.1 Generic Solution

Consider the following M-ary composite hypothesis-testing problem:

$$\mathcal{H}_i \ : \ p_\theta^X(\boldsymbol{x}) \ , \ \theta \in \Lambda_i \ , \quad i = 0, 1, \ldots, M-1 \ , \tag{2.1}$$

where $\mathcal{H}_i$ denotes the ith hypothesis and $p_\theta^X(\boldsymbol{x})$ represents the probability density function (PDF) of observation $\boldsymbol{X}$ for a given value of $\Theta = \theta$. Each observation (measurement) $\boldsymbol{x}$ is a vector with K components; i.e., $\boldsymbol{x} \in \mathbb{R}^K$, and $\Lambda_0, \Lambda_1, \ldots, \Lambda_{M-1}$ form a partition of the parameter space Λ. The prior distribution of the unknown parameter Θ, denoted by $w(\theta)$, is assumed to be known, considering a Bayesian framework.

A generic decision rule can be defined as

$$\phi(\boldsymbol{x}) = i \ , \quad \text{if} \ \ \boldsymbol{x} \in \Gamma_i \ , \tag{2.2}$$

for $i = 0, 1, \ldots, M-1$, where $\Gamma_0, \Gamma_1, \ldots, \Gamma_{M-1}$ form a partition of the observation space Γ.

As shown in Fig. 2.1, the aim is to add noise to the original observation $\boldsymbol{x}$ in order to improve the performance of the detector according to the Bayesian criterion. By adding noise $\boldsymbol{c}$ to the original observation $\boldsymbol{x}$, the modified observation is formed as $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{c}$, where $\boldsymbol{c}$ has a PDF denoted by $p_{\boldsymbol{C}}(\cdot)$, and is

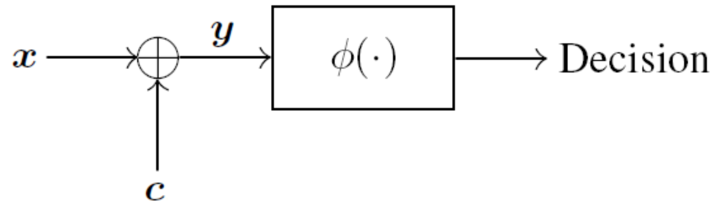

Figure 2.1: Additional independent noise $\boldsymbol{c}$ is added to observation $\boldsymbol{x}$ in order to improve the performance of the detector $\phi(\cdot)$.

independent of $\boldsymbol{x}$. It is assumed that the detector ϕ, described by (2.2), is fixed, and the only means for improving the performance of the detector is to optimize the additional noise $\boldsymbol{c}$. In other words, the aim is to find $p_{\boldsymbol{C}}(\cdot)$ that minimizes the Bayes risk $r(\phi)$; that is,

$$p_{C}^{\mathrm{opt}}(\boldsymbol{c}) = \arg\min_{p_{\boldsymbol{C}}(\boldsymbol{c})} r(\phi) \ , \tag{2.3}$$

where the Bayes risk is given by [21]

$$r(\phi) = \mathrm{E}\{R_{\Theta}(\phi)\} = \int_{\Lambda} R_{\theta}(\phi) w(\theta)\, d\theta \ , \tag{2.4}$$

with $R_{\theta}(\phi)$ denoting the conditional risk that is defined as the average cost of decision rule ϕ for a given $\theta \in \Lambda$. The conditional risk can be calculated from [21]

$$R_{\theta}(\phi) = \mathrm{E}\{\mathsf{C}[\phi(\boldsymbol{Y}), \Theta] \mid \Theta = \theta\} = \int_{\Gamma} \mathsf{C}[\phi(\boldsymbol{y}), \theta]\, p_{\theta}^{Y}(\boldsymbol{y})\, d\boldsymbol{y} \ , \tag{2.5}$$

where $p_{\theta}^{Y}(\boldsymbol{y})$ is the PDF of the modified observation for a given value of $\Theta = \theta$, and $\mathsf{C}[i, \theta]$ is the cost of selecting $\mathcal{H}_i$ when $\Theta = \theta$, for $\theta \in \Lambda$. Thus, $r(\phi)$ can be expressed as

$$r(\phi) = \int_{\Lambda} \int_{\Gamma} \mathsf{C}[\phi(\boldsymbol{y}), \theta]\, p_{\theta}^{Y}(\boldsymbol{y})\, w(\theta)\, d\boldsymbol{y}\, d\theta. \tag{2.6}$$

Due to the addition of independent noise, the modified observation has the following PDF:

$$p_{\theta}^{Y}(\boldsymbol{y}) = \int_{\mathbb{R}^K} p_{\theta}^{X}(\boldsymbol{y} - \boldsymbol{c}) p_{\boldsymbol{C}}(\boldsymbol{c})\, d\boldsymbol{c} \ . \tag{2.7}$$

Then, from (2.6) and (2.7), the following expressions are obtained:

$$\begin{aligned} r(\phi) &= \int_{\Lambda}\int_{\Gamma}\int_{\mathbb{R}^K} \mathsf{C}[\phi(\boldsymbol{y}),\theta]\, p_{\theta}^{X}(\boldsymbol{y}-\boldsymbol{c})\, p_{\boldsymbol{C}}(\boldsymbol{c})\, w(\theta)\, d\boldsymbol{c}\, d\boldsymbol{y}\, d\theta && (2.8)\\ &= \int_{\mathbb{R}^K} p_{\boldsymbol{C}}(\boldsymbol{c}) \left[\int_{\Lambda}\int_{\Gamma} \mathsf{C}[\phi(\boldsymbol{y}),\theta] p_{\theta}^{X}(\boldsymbol{y}-\boldsymbol{c})\, w(\theta)\, d\boldsymbol{y}\, d\theta\right] d\boldsymbol{c} && (2.9)\\ &= \int_{\mathbb{R}^K} p_{\boldsymbol{C}}(\boldsymbol{c})\, f(\boldsymbol{c})\, d\boldsymbol{c} && (2.10)\\ &= \mathrm{E}\{f(\boldsymbol{C})\} && (2.11) \end{aligned}$$

where

$$f(\boldsymbol{c}) \doteq \int_{\Lambda}\int_{\Gamma} \mathsf{C}[\phi(\boldsymbol{y}),\theta]\, p_{\theta}^{X}(\boldsymbol{y}-\boldsymbol{c})\, w(\theta)\, d\boldsymbol{y}\, d\theta\ . \tag{2.12}$$

From (2.11), it is observed that the solution of (2.3) can be obtained by assigning all the probability to the minimizer of $f(\boldsymbol{c})$; i.e.,

$$p_{\boldsymbol{C}}^{\text{opt}}(\boldsymbol{c}) = \delta(\boldsymbol{c}-\boldsymbol{c}_0)\ , \tag{2.13}$$

where

$$\boldsymbol{c}_0 = \arg\min_{\boldsymbol{c}}\ f(\boldsymbol{c}). \tag{2.14}$$

In other words, the optimal additional noise that minimizes the Bayes risk can be expressed as a constant corresponding to the minimum value of $f(\boldsymbol{c})$. Of course, when $f(\boldsymbol{c})$ has multiple minima, then the optimal noise PDF can be represented as $p_{\boldsymbol{C}}(\boldsymbol{c}) = \sum_{i=1}^{N} \lambda_i \delta(\boldsymbol{c}-\boldsymbol{c}_{0i})$, for any $\lambda_i \geq 0$ such that $\sum_{i=1}^{N} \lambda_i = 1$, where $\boldsymbol{c}_{01}, \ldots, \boldsymbol{c}_{0N}$ represent the values corresponding to the minimum values of $f(\boldsymbol{c})$.

The main implication of the result in (2.13) is that among all PDFs for the additional independent noise $\boldsymbol{c}$, the ones that assign all the probability to a single noise value can be used as the optimal additional signal components in Fig. 2.1. In other words, in the Bayesian framework, addition of independent noise to observations corresponds to shifting the decision region of the detector.

2.1.2 Special Cases

The analysis in the previous section considers a Bayes risk based on a very generic cost function $\mathsf{C}[j,\theta]$, which can assign different costs even to the same decision j for a given true hypothesis $\theta \in \Lambda_i$ when different values of θ in set Λ_i are considered. In this section, various special cases are studied for some specific structures of the cost function. In addition, the binary hypothesis-testing problem ($M = 2$) is analyzed in more detail.

If it is assumed, for all i, j, that the cost of deciding $\mathcal{H}_j$ when $\mathcal{H}_i$ is true is the same for all $\theta \in \Lambda_i$ (i.e., if a uniform cost is assumed in each Λ_i for $i = 0, 1, \dots, M-1$), the cost function satisfies

$$\mathsf{C}[\phi(\boldsymbol{y}) = j \,,\, \theta] = \mathsf{C}_{ji} \ , \quad \forall \theta \in \Lambda_i \ , \ \ \forall i, j \in \{0, 1, \dots, M-1\} \ , \tag{2.15}$$

where C_{ji} is a non-negative constant that is independent of θ [21]. Then, $f(\boldsymbol{c})$ in (2.12) becomes

$$\begin{aligned} f(\boldsymbol{c}) &= \int_\Lambda \int_\Gamma \mathsf{C}[\phi(\boldsymbol{y}), \theta] \; p_\theta^X(\boldsymbol{y} - \boldsymbol{c}) \; w(\theta) \, d\boldsymbol{y} \, d\theta \\ &= \sum_{i=0}^{M-1} \int_{\Lambda_i} \left(\sum_{j=0}^{M-1} \int_{\Gamma_j} \mathsf{C}_{ji} \; p_\theta^X(\boldsymbol{y} - \boldsymbol{c}) \, d\boldsymbol{y} \right) w(\theta) \, d\theta \\ &= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \mathsf{C}_{ji} \int_{\Gamma_j} \int_{\Lambda_i} \; p_\theta^X(\boldsymbol{y} - \boldsymbol{c}) w(\theta) \, d\theta \, d\boldsymbol{y} \\ &= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \mathsf{C}_{ji} f_{ji}(\boldsymbol{c}) \end{aligned} \tag{2.16}$$

where

$$f_{ji}(\boldsymbol{c}) \doteq \int_{\Gamma_j} \int_{\Lambda_i} \; p_\theta^X(\boldsymbol{y} - \boldsymbol{c}) w(\theta) \, d\theta \, d\boldsymbol{y} \ . \tag{2.17}$$

In addition to (2.15), if uniform cost assignment (UCA) is considered, the costs are specified as $\mathsf{C}_{ji} = 1$ for $j \neq i$ and $\mathsf{C}_{ji} = 0$ for $j = i$. In other words, the correct decisions are assigned zero cost, whereas the wrong ones are assigned unit cost. In this case, $f(\boldsymbol{c})$ in (2.16) becomes

$$f(\boldsymbol{c}) = \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} f_{ji}(\boldsymbol{c}) = 1 - \sum_{i=0}^{M-1} f_{ii}(\boldsymbol{c}) \ . \tag{2.18}$$

Next, let $M = 2$ (i.e., binary hypothesis-testing) and assume uniform costs in Λ_i for $i = 0, 1$. Then, $f(\boldsymbol{c})$ can be calculated as follows:

$$\begin{aligned}
f(\boldsymbol{c}) &= \sum_{i=0}^{1}\sum_{j=0}^{1} \mathsf{C}_{ji} f_{ji}(\boldsymbol{c}) = \sum_{i=0}^{1}\sum_{j=0}^{1} \mathsf{C}_{ji} \int_{\Gamma_j}\int_{\Lambda_i} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} \\
&= \mathsf{C}_{10} \int_{\Gamma_1}\int_{\Lambda_0} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} + \mathsf{C}_{00} \int_{\Gamma_0}\int_{\Lambda_0} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} \\
&+ \mathsf{C}_{01} \int_{\Gamma_0}\int_{\Lambda_1} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} + \mathsf{C}_{11} \int_{\Gamma_1}\int_{\Lambda_1} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} \\
&= \pi_1 \mathsf{C}_{01} + \pi_0 \mathsf{C}_{00} + \int_{\Gamma_1} \Bigg[(\mathsf{C}_{10} - \mathsf{C}_{00}) \int_{\Lambda_0} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta \\
&- (\mathsf{C}_{01} - \mathsf{C}_{11}) \int_{\Lambda_1} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta \Bigg] d\boldsymbol{y} \ ,
\end{aligned} \tag{2.19}$$

where the following relation is employed in obtaining the final expression:

$$\int_{\Gamma_0}\int_{\Lambda_i} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} = \pi_i - \int_{\Gamma_1}\int_{\Lambda_i} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta\, d\boldsymbol{y} \ , \tag{2.20}$$

for $i = 0, 1$, with $\pi_i = P(\mathcal{H}_i) = \int_{\Lambda_i} w(\theta)\, d\theta$.

Then, the Bayes risk in (2.10) can be expressed from (2.19) as

$$\begin{aligned}
r(\phi) &= \int_{\mathbb{R}^K} p_{\boldsymbol{C}}(\boldsymbol{c})\ f(\boldsymbol{c})\, d\boldsymbol{c} = \mathrm{E}\{f(\boldsymbol{C})\} \\
&= \pi_1 \mathsf{C}_{01} + \pi_0 \mathsf{C}_{00} - \mathrm{E}\{g(\boldsymbol{C})\} \ ,
\end{aligned} \tag{2.21}$$

where

$$\begin{aligned}
g(\boldsymbol{c}) = \int_{\Gamma_1} \Bigg[&- (\mathsf{C}_{10} - \mathsf{C}_{00}) \int_{\Lambda_0} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta \\
&+ (\mathsf{C}_{01} - \mathsf{C}_{11}) \int_{\Lambda_1} p_\theta^X(\boldsymbol{y}-\boldsymbol{c}) w(\theta)\, d\theta \Bigg] d\boldsymbol{y} \ .
\end{aligned} \tag{2.22}$$

From (2.21), it is observed that $r(\phi)$ is minimized for $p_{\boldsymbol{C}}(\boldsymbol{c}) = \delta(\boldsymbol{c} - \boldsymbol{c}_0)$, where

$$\boldsymbol{c}_0 = \arg\max_{\boldsymbol{c}} \ g(\boldsymbol{c}) \ . \tag{2.23}$$

Therefore, the optimal additional noise that minimizes the Bayes risk can be expressed as a constant corresponding to the maximum value of $g(\boldsymbol{c})$.

In order to obtain a more explicit expression for $g(\boldsymbol{c})$, the following result is employed [21].

$$p^X(\boldsymbol{y}-\boldsymbol{c}|\Theta\in\Lambda_i)=\frac{1}{\pi_i}\int_{\Lambda_i}p_\theta^X(\boldsymbol{y}-\boldsymbol{c})w(\theta)\,d\theta\ ,\quad i=0,1. \tag{2.24}$$

Then, $g(\boldsymbol{c})$ in (2.22) can be expressed as

$$\begin{aligned} g(\boldsymbol{c}) &= \int_{\mathbb{R}^K}\phi(\boldsymbol{y})\Bigg[(\mathsf{C}_{01}-\mathsf{C}_{11})\int_{\Lambda_1}p_\theta^X(\boldsymbol{y}-\boldsymbol{c})w(\theta)\,d\theta \\ &\quad -(\mathsf{C}_{10}-\mathsf{C}_{00})\int_{\Lambda_0}p_\theta^X(\boldsymbol{y}-\boldsymbol{c})w(\theta)\,d\theta\Bigg]d\boldsymbol{y} && (2.25) \\ &= \int_{\mathbb{R}^K}\phi(\boldsymbol{x}+\boldsymbol{c})\Big[(\mathsf{C}_{01}-\mathsf{C}_{11})\pi_1 p^X(\boldsymbol{x}|\theta\in\Lambda_1) \\ &\quad -(\mathsf{C}_{10}-\mathsf{C}_{00})\pi_0 p^X(\boldsymbol{x}|\theta\in\Lambda_0)\Big]d\boldsymbol{x}\ , && (2.26) \end{aligned}$$

where the result in (2.24), as well as a change of variables ($\boldsymbol{x}=\boldsymbol{y}-\boldsymbol{c}$) are used in obtaining the final result. If we define a new function $h(\boldsymbol{x})$ as

$$h(\boldsymbol{x})=(\mathsf{C}_{01}-\mathsf{C}_{11})\pi_1 p^X(\boldsymbol{x}|\theta\in\Lambda_1)-(\mathsf{C}_{10}-\mathsf{C}_{00})\pi_0 p^X(\boldsymbol{x}|\theta\in\Lambda_0)\ , \tag{2.27}$$

we then have

$$g(\boldsymbol{c})=\int_{\mathbb{R}^K}\phi(\boldsymbol{x}+\boldsymbol{c})h(\boldsymbol{x})\,d\boldsymbol{x}\ . \tag{2.28}$$

As can be seen from (2.28), $g(\boldsymbol{c})$ is the correlation between the decision function and $h(\boldsymbol{x})$.

If we also assume that correct decisions have zero cost, and wrong ones have unit cost, we have $r(\phi)=\pi_1-\mathrm{E}\{g(\boldsymbol{C})\}$ and

$$h(\boldsymbol{x})=\pi_1 p^X(\boldsymbol{x}|\theta\in\Lambda_1)-\pi_0 p^X(\boldsymbol{x}|\theta\in\Lambda_0)\ . \tag{2.29}$$

For simple hypotheses (i.e., when Λ_0 and Λ_1 contain single elements), (2.29) yields $h(\boldsymbol{x})=\pi_1 p_1^X(\boldsymbol{x})-\pi_0 p_0^X(\boldsymbol{x})$, which is the result obtained in [18]. In Section 2.2, the result for simple hypotheses is used to investigate noise enhanced sign detectors under Gaussian mixture noise.

2.1.3 A Detection Example

In this section, the following composite hypothesis-testing problem is studied in order to present an example of the theoretical results obtained in the previous sections.

$$\begin{aligned} \mathcal{H}_0 &: \theta \in \Lambda_0 = [-\alpha, 0] \,, \\ \mathcal{H}_1 &: \theta \in \Lambda_1 = (0, 2\alpha] \,, \end{aligned} \tag{2.30}$$

where α is a known positive real number. For a given value of $\Theta = \theta$, the observation X has the following PDF:

$$p_\theta^X(x) = \frac{1}{3}\left[\gamma(x;\theta - A, \sigma^2) + \gamma(x;\theta,\sigma^2) + \gamma(x;\theta + A,\sigma^2)\right] \,, \tag{2.31}$$

where $\gamma(x;\mu,\sigma^2) \doteq \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. In other words, the observation is distributed as the mixture of three Gaussian distributions with the same variance σ^2 and means $\theta - A$, θ, and $\theta + A$. In addition, the prior distribution of Θ is modeled by a uniform random variable between $-\alpha$ and 2α, which is denoted as $\Theta \sim \mathcal{U}[-\alpha\,, 2\alpha]$. Therefore, the prior probabilities of the hypotheses can be obtained as $\pi_0 = P(\mathcal{H}_0) = \int_{-\alpha}^{0} \frac{1}{3\alpha}\, d\theta = 1/3$ and $\pi_1 = P(\mathcal{H}_1) = \int_0^{2\alpha} \frac{1}{3\alpha}\, d\theta = 2/3\,$.

The sign detector is considered as the decision rule in this example, which is expressed as

$$\phi(y) = \begin{cases} 1 & \text{if } \ y \geq 0 \\ 0 & \text{if } \ y < 0 \end{cases} \,. \tag{2.32}$$

The aim is to obtain the optimal value of additional signal c such that $y = x + c$ results in the minimum Bayes risk for this composite hypothesis-testing problem (cf. (2.3)).

Assuming uniform costs in Λ_0 and Λ_1 and for UCA, the optimal value of c can be obtained from (2.14), (2.17) and (2.18), or from (2.23), (2.28) and (2.29). When the solution based on (2.14), (2.17) and (2.18) is considered, $f(c)$ can be obtained as $f(c) = 1 - f_{00}(c) - f_{11}(c)$. From (2.17), $f_{00}(c)$ and $f_{11}(c)$ can be expressed for $\Theta \sim \mathcal{U}[-\alpha\,, 2\alpha]$ as

$$f_{00}(c) = \frac{1}{3\alpha}\int_{-\alpha}^{0}\int_{-\infty}^{0} p_\theta^X(y-c)\,dy\,d\theta\;, \tag{2.33}$$

$$f_{11}(c) = \frac{1}{3\alpha}\int_{0}^{2\alpha}\int_{0}^{\infty} p_\theta^X(y-c)\,dy\,d\theta\;. \tag{2.34}$$

Then, after some manipulation, $f(c) = 1 - f_{00}(c) - f_{11}(c)$ can be obtained, from (2.31), (2.33) and (2.34), as

$$f(c) = 1 - f_{00}(c) - f_{11}(c) = \frac{2}{3} - \frac{1}{9\alpha}\left(\int_0^{2\alpha} v_\theta(c)\,d\theta - \int_{-\alpha}^{0} v_\theta(c)\,d\theta\right)\;, \tag{2.35}$$

where

$$v_\theta(c) \doteq Q\left(\frac{-\theta + A - c}{\sigma}\right) + Q\left(\frac{-\theta - c}{\sigma}\right) + Q\left(\frac{-\theta - A - c}{\sigma}\right)\;. \tag{2.36}$$

Therefore, the optimal value of c can be calculated from (2.14) and (2.35) as

$$c_0 = \arg\max_c \left\{\int_0^{2\alpha} v_\theta(c)\,d\theta - \int_{-\alpha}^{0} v_\theta(c)\,d\theta\right\}\;. \tag{2.37}$$

In Fig. 2.2, the Bayes risks are plotted against α for the original sign detector (i.e., without additional signal c) and for the noise enhanced sign detector (i.e., with optimal additional signal c_0) when $A = 2$ and $\sigma = 1$. Note from (2.11) that the Bayes risks are given by $r(\phi) = f(0)$ and $r(\phi) = f(c_0)$ for the original and the noise enhanced sign detectors, respectively. It is observed from the figure that there is significant improvement for small values of α and the amount of improvement decreases as α increases. For example, for $\alpha = 0.5$, the Bayes risks are 0.429 and 0.323, respectively, for the conventional and the noise enhanced sign detectors, whereas they are 0.298 and 0.262 for $\alpha = 1.5$.

In Fig. 2.3, the Bayes risk is plotted versus additional signal c when $A = 2$ and $\sigma = 1$ for various values α. For each α, there is a unique minimizer of the Bayes risk for a positive value of c. In addition, it is observed that the optimal additional signal value c_0 in (2.37) decreases as α increases. In fact, for large

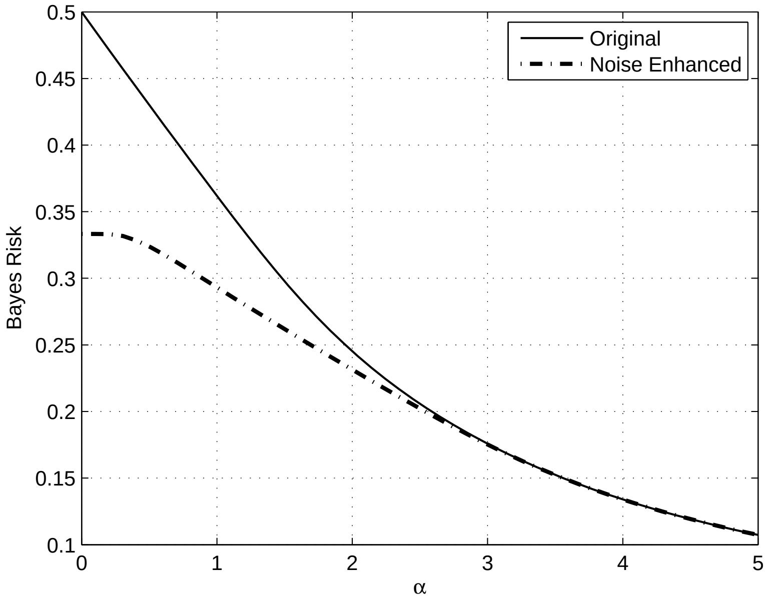

Figure 2.2: Bayes risk versus α for $A = 2$ and $\sigma = 1$.

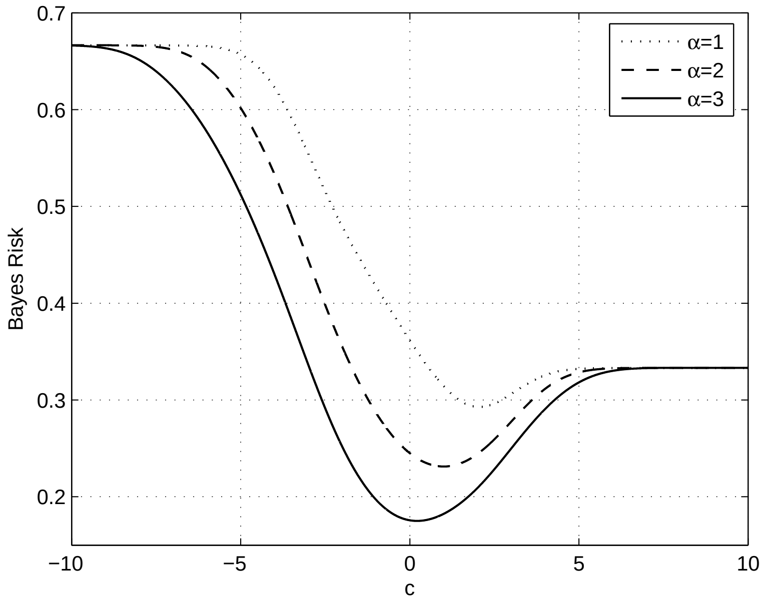

Figure 2.3: Bayes risk versus c for $A = 2$ and $\sigma = 1$.

values of α, c_0 goes to zero, which implies that the detector cannot be improved by additional noise; that is, the original detector is non-improvable, which is in compliance with the result in Fig. 2.2.

2.2 Noise Enhanced Sign Detection under Gaussian Mixture Noise

After showing, in the previous section, that an optimal additional noise corresponds to a shift of the measurements used by the detector, this section investigates the effects of measurement shifts for sign detection of antipodal signals under symmetric Gaussian mixture noise.

2.2.1 Signal Model

Consider the following measurement (observation) model

$$x = A\,b + n \ , \tag{2.38}$$

where $b \in \{-1, +1\}$ represents the equiprobable binary symbol to be detected, $A > 0$ is the known amplitude coefficient,[1] and n is the measurement noise, which is modeled as symmetric Gaussian mixture noise. The PDF of the noise is given by

$$p_N(x) = \sum_{i=1}^{M} w_i\,\psi_i(x - x_i) \ , \tag{2.39}$$

where $w_i \geq 0$ for $i = 1, \ldots, M$, $\sum_{i=1}^{M} w_i = 1$, and

$$\psi_i(x) = \frac{1}{\sqrt{2\pi}\,\sigma_i}\,\exp\left(\frac{-x^2}{2\,\sigma_i^2}\right) \ , \tag{2.40}$$

[1]The results in the thesis can be extended to $A < 0$ cases as well, by switching the decision regions of the detector in (2.42).

for $i = 1, \ldots, M$. Due to the symmetry assumption, $x_i = -x_{M-i+1}$, $w_i = w_{M-i+1}$ and $\sigma_i = \sigma_{M-i+1}$ for $i = 1, \ldots, \lfloor M/2 \rfloor$.

The symmetric Gaussian mixture model specified above is observed in many practical scenarios [25]-[30]. One important scenario is multiuser wireless communications, in which the desired signal is corrupted by interference from other users as well as zero-mean Gaussian background noise. In that case, the overall noise has a symmetric Gaussian mixture model when the user symbols are symmetric and equiprobable (e.g., ± 1 with equal probability) [23].

The problem can be stated as the following binary hypothesis test

$$\begin{aligned} \mathcal{H}_0 &: X \sim p_N(x + A)\ , \\ \mathcal{H}_1 &: X \sim p_N(x - A)\ , \end{aligned} \tag{2.41}$$

where hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond to $b = -1$ and $b = +1$ cases, respectively. The following conventional sign detector is considered to determine the index of the true hypothesis, which is expressed as

$$\phi(x) = \begin{cases} 0\ , & x < 0 \\ 1\ , & x > 0 \end{cases}\ . \tag{2.42}$$

In the case of $x = 0$, the detector decides $\mathcal{H}_0$ or $\mathcal{H}_1$ randomly (i.e., with equal probabilities). It is well-known that the conventional detector in (2.42) is not optimal in general for Gaussian mixture noise [18], [31]. However, its main advantage is that it has very low complexity, which makes it very practical for low cost applications. Therefore, the main aim in this work is to keep the low complexity of the detector but to modify the measurement in (2.38) in order to improve detection performance.

2.2.2 Formulation of Optimal Measurement Shifts

Instead of the original measurement x, consider a noise modified version of the measurement as

$$y = x + c \; , \tag{2.43}$$

where c represents additional independent noise term.

As studied in [18] and in Section 2.1, the optimal additional noise c that minimizes the probability of decision error[2] is a constant that solves the following maximization problem:

$$c_{\text{opt}} = \arg\max_{c} \int_{-\infty}^{\infty} \phi(y+c) \left[p_N(y-A) - p_N(y+A) \right] dy \; , \tag{2.44}$$

where $p_N(\cdot)$ represents the PDF of the measurement noise in (2.38).

For the detector in (2.42), the optimal additional noise in (2.44) is given by

$$c_{\text{opt}} = \arg\max_{c} \int_{-c}^{\infty} \left[p_N(y-A) - p_N(y+A) \right] dy \; , \tag{2.45}$$

which, after some manipulation, can be expressed, from (2.39), as

$$c_{\text{opt}} = \arg\max_{c} \sum_{i=1}^{M} w_i \left[Q\left(\frac{-c - A - x_i}{\sigma_i} \right) - Q\left(\frac{-c + A - x_i}{\sigma_i} \right) \right] \; , \tag{2.46}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \mathrm{e}^{-t^2/2} dt$ represents the Q-function. Note that the optimization in (2.46) can be performed over $c \geq 0$ only, since it can be shown that the term in the square brackets is an even function of c for the symmetric Gaussian mixture noise model.

The probability of decision error when a constant noise c is added in (2.43) is given by [18]

$$\mathrm{P}(c) = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{M} w_i \left[Q\left(\frac{-c - A - x_i}{\sigma_i} \right) - Q\left(\frac{-c + A - x_i}{\sigma_i} \right) \right] \; . \tag{2.47}$$

[2]This criterion is equivalent to the minimization of the Bayes risk for uniform cost assignment and equal priors [21].

When the optimal value of c is calculated as in (2.46), $\mathrm{P_{SR}} = \mathrm{P}(c_{\mathrm{opt}})$ specifies the error probability obtained via measurement modification. The conventional case corresponds to using no additional noise, i.e., $\mathrm{P_{conv}} = \mathrm{P}(0)$. Note that $c_{\mathrm{opt}} = 0$ corresponds to the *non-improvability* case, in which it is not possible to improve the detector performance by adding noise to the measurement; that is, $\mathrm{P_{SR}} = \mathrm{P_{conv}}$. On the other hand, if $\mathrm{P_{SR}} < \mathrm{P_{conv}}$, the detector is *improvable* [1].

One justification for using additional noise (measurement shifts) to improve performance of suboptimal detectors as in (2.42) instead of employing an optimal detector based on the likelihood ratio test is reduced implementation complexity [1]. Instead of calculating the likelihood ratio for each observation, the detector in (2.42) just checks the sign of the observation shifted by c_{opt}. Note that the calculation of c_{opt} requires the solution of the optimization problem in (2.46), but that problem needs to be solved only when the noise statistics and/or the signal amplitude change.

2.2.3 Conditions for Improvability and Non-improvability of Detection

In this section, sufficient conditions are derived in order to determine whether additional noise can enhance the performance of the conventional detector in (2.42) in the presence of symmetric Gaussian mixture measurement noise. Such improvability and non-improvability conditions carry practical importance, since determination of whether additional noise is useful or not based on desired signal and measurement noise parameters helps specify when to solve the optimization problem in (2.46) for the optimal additional noise.

First, a sufficient condition on the signal amplitude and the measurement noise statistics is obtained in order for additional noise to improve detection performance.

Proposition 1: *The detector in* (2.42) *is improvable if the signal amplitude* A *in* (2.38) *and the measurement noise specified by* (2.39) *and* (2.40) *satisfy*

$$\sum_{i=1}^{M} \frac{w_i}{\sigma_i^3} (A + x_i)\, \mathrm{e}^{-\frac{(A+x_i)^2}{2\sigma_i^2}} < 0 \ . \tag{2.48}$$

Proof: From (2.46), a first-order necessary condition for optimal additional noise value can be obtained by equating the first derivative with respect to c to zero.

$$\sum_{i=1}^{M} \frac{w_i}{\sqrt{2\pi}\,\sigma_i} \left(\mathrm{e}^{-\frac{(-c-A-x_i)^2}{2\sigma_i^2}} - \mathrm{e}^{-\frac{(-c+A-x_i)^2}{2\sigma_i^2}} \right) = 0 \ . \tag{2.49}$$

Note that the condition in (2.49) is satisfied by the conventional solution, i.e., for $c = 0$. In addition, the second derivative at $c = 0$ can be calculated from (2.49) as

$$\sum_{i=1}^{M} \frac{w_i}{\sqrt{2\pi}\,\sigma_i^3} \left(-(A + x_i)\, \mathrm{e}^{-\frac{(A+x_i)^2}{2\sigma_i^2}} - (A - x_i)\, \mathrm{e}^{-\frac{(A-x_i)^2}{2\sigma_i^2}} \right) \ . \tag{2.50}$$

Due to the symmetry of the Gaussian mixture PDF, the expression in (2.50) is always positive when the condition in the proposition is satisfied. Since the first derivative is zero and the second derivative is positive at $c = 0$, it is a minimum point of the objective function in (2.46). Therefore, (2.47) implies that there exists $c \neq 0$ such that $\mathrm{P}_{\mathrm{SR}}(c) < \mathrm{P}_{\mathrm{conv}}$, which proves the improvability of the detector. $\square$

Proposition 1 provides a simple sufficient condition to determine if the use of additional noise can improve the performance of the detector in (2.42). When the condition in (2.48) is satisfied, the optimal additional noise can be calculated from (2.46) (which is non-zero since the system is improvable), and the updated measurement in (2.43) can be used for improved error performance.

Similar to determining the improvability of the system, it is also important to know when the system cannot be improved via additional noise. Such a knowledge prevents efforts for solving (2.46) to find the additional noise, which yields

$c_{\mathrm{opt}} = 0$ when the system is non-improvable. In the following, two conditions are provided to classify the system as non-improvable.

Proposition 2: *Assume that the variances of the Gaussian components in the mixture noise, specified by* (2.39) *and* (2.40)*, converge to infinity; that is,* $\sigma_i^2 \to \infty$ *for* $i = 1, \ldots, M$*. Then, the detector in* (2.42) *is non-improvable.*

Proof: This result can be proven by showing that $\lim_{\sigma_1^2,\ldots,\sigma_M^2 \to \infty} \frac{P(c)}{P(0)} = 1$, where $P(c)$ is as in (2.47). In other words, no improvement can be obtained for any value of c. Hence, the detector is non-improvable. $\square$

The main implication of Proposition 2 is that when the variance of each Gaussian component in the Gaussian mixture noise is very large, the conventional decision rule, which decides $\mathcal{H}_0$ for negative measurements and $\mathcal{H}_1$ otherwise, has lower probability of error than any other decision rule that applies the sign rule in (2.42) on shifted measurements as in (2.43) for $c \neq 0$. In other words, for large variances, $c_{\mathrm{opt}} = 0$ and $\mathrm{P}_{\mathrm{SR}} = \mathrm{P}_{\mathrm{conv}}$.

Another non-improvability condition can be obtained when the signal amplitude A in (2.38) is larger than or equal to all the mass points in the Gaussian mixture noise.

Proposition 3: *Assume that the signal amplitude* A *in* (2.38) *is larger than or equal to the maximum of the mean values of the Gaussian components in the Gaussian mixture in* (2.39)*; that is,*

$$A \geq \max_{i=1,\ldots,M} \{x_i\} \; . \tag{2.51}$$

Then, the detector in (2.42) *is non-improvable.*

Proof: The first-order necessary optimality condition in (2.49) is given by

$$\sum_{i=1}^{M} \frac{w_i}{\sigma_i} \, \mathrm{e}^{-\frac{(c+A+x_i)^2}{2\sigma_i^2}} = \sum_{i=1}^{M} \frac{w_i}{\sigma_i} \, \mathrm{e}^{-\frac{(c-A+x_i)^2}{2\sigma_i^2}} \; . \tag{2.52}$$

Due to the symmetry of the Gaussian mixture noise, (2.52) can be expressed as

$$\sum_{i=1}^{\lfloor M/2 \rfloor} \frac{w_i}{\sigma_i}\left(\mathrm{e}^{-\frac{(c+A+x_i)^2}{2\sigma_i^2}} + \mathrm{e}^{-\frac{(c+A-x_i)^2}{2\sigma_i^2}}\right) = \sum_{i=1}^{\lfloor M/2 \rfloor} \frac{w_i}{\sigma_i}\left(\mathrm{e}^{-\frac{(-c+A+x_i)^2}{2\sigma_i^2}} + \mathrm{e}^{-\frac{(-c+A-x_i)^2}{2\sigma_i^2}}\right) . \tag{2.53}$$

Since $A \geq \max_{i=1,\ldots,M}\{x_i\}$, $A+x_i \geq 0$ and $-A+x_i \leq 0$ for $i = 1, \ldots, M$. Then, for $c > 0$, it is observed that $\mathrm{e}^{-\frac{(c+A+x_i)^2}{2\sigma_i^2}} < \mathrm{e}^{-\frac{(-c+A+x_i)^2}{2\sigma_i^2}}$ and $\mathrm{e}^{-\frac{(c+A-x_i)^2}{2\sigma_i^2}} < \mathrm{e}^{-\frac{(-c+A-x_i)^2}{2\sigma_i^2}}$ for $i = 1, \ldots, M$. Therefore, the term on the right-hand-side (RHS) of (2.53) is always larger than that on the left-hand-side (LHS) for $c > 0$. Similarly, it can be shown that the term on the LHS of (2.53) is always larger than that on the RHS for $c < 0$. The equality is satisfied only when $c = 0$. In addition, the second derivative at $c = 0$, given in (2.50), is always negative since $A \pm x_i \geq 0$ for $i = 1, \ldots, M$. Hence, $c = 0$ is the unique maximum of the problem in (2.46).
$\square$

Proposition 3 states that if the signal amplitude A is larger than or equal to all the mean values of the Gaussian components in the mixture noise, then there is no need to search for optimal additional noise as $c_{\text{opt}} = 0$ in that case, which implies that the conventional algorithm cannot be improved. In fact, if $A > \max_{i=1,\ldots,M}\{x_i\}$ and if σ_i's in (2.40) are very small, then the conventional system can have very small probability of error, hence, may not need performance improvement in some cases.

2.2.4 Performance Analysis of Noise Enhanced Detection

After the investigation of improvability and non-improvability conditions in the previous section, this section focuses on some properties of noise enhanced detection, and theoretical limits on performance improvements that can be obtained by adding noise to measurements.

First, the effects of additional noise are investigated as a function of the standard deviations of the Gaussian noise components in the Gaussian mixture noise specified by (2.39) and (2.40). Let $\boldsymbol{\sigma} = [\sigma_1 \cdots \sigma_M]$ represent the standard deviation terms in (2.40). Then, the probability of decision error of the noise enhanced detector can be expressed, from (2.46) and (2.47), as

$$\mathrm{P}_{\mathrm{SR}}(\boldsymbol{\sigma}) = \frac{1}{2} - \frac{1}{2} \max_c \sum_{i=1}^{M} w_i \left[Q\left(\frac{-c - A - x_i}{\sigma_i}\right) - Q\left(\frac{-c + A - x_i}{\sigma_i}\right) \right] . \tag{2.54}$$

In the conventional case, no additional noise is used; hence, the probability of decision error is given by

$$\mathrm{P}_{\mathrm{conv}}(\boldsymbol{\sigma}) = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{M} w_i \left[Q\left(\frac{-A - x_i}{\sigma_i}\right) - Q\left(\frac{A - x_i}{\sigma_i}\right) \right] . \tag{2.55}$$

For certain parameters of the Gaussian mixture noise, the probabilities of decision error in (2.54) and (2.55) may not be monotonically decreasing as the standard deviations, $\sigma_1, \ldots, \sigma_M$, decrease. Although this might seem counter-intuitive at first, it mainly due to the multi-modal nature of the Gaussian mixture distribution. In Section 2.2.5, numerical examples are provided to illustrate that behavior. Although the probabilities of error can exhibit non-monotonic behaviors in general, the following proposition states that for equal standard deviations, a decrease in the standard deviation value can never result in an increase in the probability of decision error for the noise enhanced detector.

Proposition 4: *Assume $\sigma_i = \sigma$ for $i = 1, \ldots, M$. Then, $\mathrm{P}_{\mathrm{SR}}(\boldsymbol{\sigma})$ in (2.54) is a monotone increasing function of σ.*

Proof: When $\sigma_i = \sigma$ for $i = 1, \ldots, M$, $\mathrm{P}_{\mathrm{SR}}(\boldsymbol{\sigma})$ in (2.54) is expressed as

$$\mathrm{P}_{\mathrm{SR}}(\sigma) = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{M} w_i \left[Q\left(\frac{-c_{\mathrm{opt}}(\sigma) - A - x_i}{\sigma}\right) - Q\left(\frac{-c_{\mathrm{opt}}(\sigma) + A - x_i}{\sigma}\right) \right] , \tag{2.56}$$

where $c_{\text{opt}}(\sigma)$ represents the maximizer of the summation term in (2.54), which satisfies the following first and second derivative conditions[3]

$$\sum_{i=1}^{M} \frac{w_i}{\sigma} \left(\mathrm{e}^{-\frac{(-c_{\text{opt}}(\sigma)-A-x_i)^2}{2\sigma^2}} - \mathrm{e}^{-\frac{(-c_{\text{opt}}(\sigma)+A-x_i)^2}{2\sigma^2}} \right) = 0 \ , \tag{2.57}$$

$$\sum_{i=1}^{M} \frac{w_i}{\sigma^3} \Big[(-c_{\text{opt}}(\sigma) - A - x_i)\, \mathrm{e}^{-\frac{(-c_{\text{opt}}(\sigma)-A-x_i)^2}{2\sigma^2}} - (-c_{\text{opt}}(\sigma) + A - x_i)\, \mathrm{e}^{-\frac{(-c_{\text{opt}}(\sigma)+A-x_i)^2}{2\sigma^2}} \Big] < 0 \ . \tag{2.58}$$

In order to prove the monotonicity of $\mathrm{P}_{\text{SR}}(\sigma)$ in (2.56) with respect to σ, the first derivative of $\mathrm{P}_{\text{SR}}(\sigma)$ is calculated as follows:

$$\frac{d\mathrm{P}_{\text{SR}}(\sigma)}{d\sigma} = \frac{1}{2} \sum_{i=1}^{M} \frac{w_i}{\sqrt{2\pi}\,\sigma^2} \Bigg\{ \left[-\frac{dc_{\text{opt}}(\sigma)}{d\sigma}\sigma + c_{\text{opt}}(\sigma) + A + x_i \right] \mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)+A+x_i)^2}{2\sigma^2}} - \left[-\frac{dc_{\text{opt}}(\sigma)}{d\sigma}\sigma + c_{\text{opt}}(\sigma) - A + x_i \right] \mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)-A+x_i)^2}{2\sigma^2}} \Bigg\} \ , \tag{2.59}$$

which can be manipulated to obtain

$$\frac{d\mathrm{P}_{\text{SR}}(\sigma)}{d\sigma} = -\frac{1}{2\sqrt{2\pi}} \frac{dc_{\text{opt}}(\sigma)}{d\sigma} \sum_{i=1}^{M} \frac{w_i}{\sigma} \left[\mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)+A+x_i)^2}{2\sigma^2}} - \mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)-A+x_i)^2}{2\sigma^2}} \right] + \frac{1}{2\sqrt{2\pi}\,\sigma^2} \sum_{i=1}^{M} w_i \Bigg[(c_{\text{opt}}(\sigma) + A + x_i)\, \mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)+A+x_i)^2}{2\sigma^2}} - (c_{\text{opt}}(\sigma) - A + x_i)\, \mathrm{e}^{-\frac{(c_{\text{opt}}(\sigma)-A+x_i)^2}{2\sigma^2}} \Bigg] \ . \tag{2.60}$$

Since $c_{\text{opt}}(\sigma)$ satisfies (2.57), the first term in (2.60) becomes zero. In addition, (2.58) implies that the second term in (2.60) is always positive. Therefore, $d\mathrm{P}_{\text{SR}}(\sigma)/d\sigma > 0$ is satisfied; hence, $\mathrm{P}_{\text{SR}}(\sigma)$ is a monotone increasing function of σ. $\square$

It is noted from the proof of Proposition 4 that the result is valid also for asymmetric Gaussian mixture noise. In other words, as long as $\sigma_i = \sigma$ for $i = 1, \ldots, M$, $\mathrm{P}_{\text{SR}}(\sigma)$ in (2.54) is a monotone increasing function of σ.

[3] The inequalities in (2.57) and (2.58) can be obtained similar to those in (2.49) and (2.50) by taking the derivatives of the summation term in (2.54), which is equal to that in (2.46), with respect to c.

One implication of Proposition 4 is that for equal $\sigma_1, \ldots, \sigma_M$ in (2.40), the noise enhanced detector utilizes any decrease in the standard deviations for decreasing the probability of decision error. In other words, a decrease in the standard deviation can never increase error probability. This statement is not true in general for the conventional algorithm, which does not employ any additional noise. Addition of noise provides such a desirable monotonicity property since it effectively provides an adaptive detector structure depending on the characteristics of the noise. Note that addition of a constant to the decision variable, as in (2.43), for the detector in (2.42) is equivalent to using the original observation but adjusting the threshold of the detector.

The condition in Proposition 4 about equal $\sigma_1, \ldots, \sigma_M$ values may not hold in all scenarios. However, one important scenario in which such Gaussian mixture noise components are observed includes measurement noise that is composed of zero-mean Gaussian noise and discrete noise components. An important example of such a scenario is binary detection in the presence of multiple-access interference (MAI) [23], where the measurement is modeled as

$$x = A_1 b_1 + \sum_{k=2}^{K} A_k b_k + n \ , \tag{2.61}$$

with $b_i \in \{\pm 1\}$ and n representing a zero-mean Gaussian noise component. The aim is to detect b_1 in the presence of MAI, $\sum_{k=2}^{K} A_k b_k$, and background noise, n. Therefore, the total noise, $\sum_{k=2}^{K} A_k b_k + n$, can be modeled as Gaussian mixture noise with mean values at $\sum_{k=2}^{K} A_k b_k$ for all possible $b_2, \ldots, b_K$ values (that is, for $[b_2, \ldots, b_K] \in \{\pm 1\}^{K-1}$) and standard deviation terms being all equal to that of the background noise term n. Therefore, the result in Proposition 4 applies in this practical scenario.

As studied in Proposition 2, additional noise cannot improve detector performance for very large variances of the Gaussian mixture noise. Another important case is to investigate the behavior of the noise enhanced detector for very small variances. As $\sigma_i \to 0$ for $i = 1, \ldots, M$, the probability of decision error in (2.55)

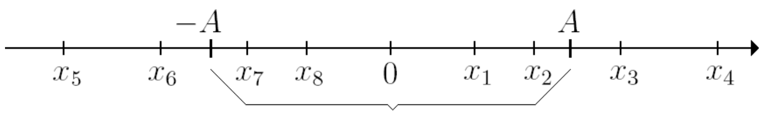

Figure 2.4: Mean values (x_j's) in a symmetric Gaussian mixture noise for $M = 8$, and signal amplitude A.

for the conventional algorithm can be expressed as[4]

$$\mathrm{P}_{\text{conv}} = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{M} w_i \, u(A - |x_i|) \ , \tag{2.62}$$

where $u(\cdot)$ is the unit step function defined as

$$u(x) \doteq \begin{cases} 1 & , \ x > 0 \\ 0.5 & , \ x = 0 \\ 0 & , \ x < 0 \end{cases} \ . \tag{2.63}$$

Similarly, as $\sigma_i \to 0$ for $i = 1, \ldots, M$, the probability of decision error in (2.54) for the noise enhanced algorithm is given by

$$\mathrm{P}_{\text{SR}} = \frac{1}{2} - \frac{1}{2} \max_{c} \sum_{i=1}^{M} w_i \, u(A - |x_i + c|) \ . \tag{2.64}$$

The expressions in (2.62) and (2.64) provide a simple interpretation of the probability of decision error. For example, consider the values of $x_1, \ldots, x_M$ and A as in Fig. 2.4. Since the probability of error expression in (2.62) states that the x_i values that are between $-A$ and A contribute to the summation term, only the weights w_1, w_2, w_7 and w_8 are employed in the calculation of the probability of error for the settings in Fig. 2.4. For the noise enhanced scenario, various values of c in (2.64) correspond to various shifts of the interval in Fig. 2.4 as shown in Fig. 2.5. Then, the value of c that results in the minimum probability of error is selected as the additional noise component.

[4] $x_1, \ldots, x_M$ are assumed to be distinct such that $|x_j - x_k| \gg \sigma_i$ as $\sigma_i \to 0$, $\forall j \neq k$, $\forall i$.

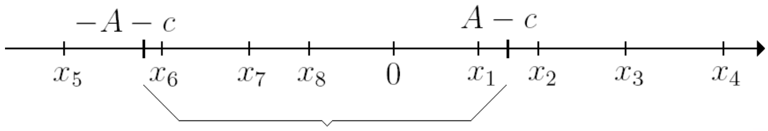

Figure 2.5: Mean values (x_j's) in a symmetric Gaussian mixture noise for $M = 8$, signal amplitude A, and additional noise c.

The previous interpretation of noise enhanced detection for very small variance values facilitates calculation of theoretical limits on performance improvements that can be obtained via additional noise.

Proposition 5: *Let M be an even number*[5] *and $0 < x_1 < \cdots < x_{M/2}$ without loss of generality. As $\sigma_i \to 0$ for $i = 1, \ldots, M$, the maximum improvement of the sign detector in* (2.42) *under symmetric Gaussian mixture noise given by* (2.39) *and* (2.40) *is specified as*

$$\max_{A, x_1, \ldots, x_M, w_1, \ldots, w_M} \frac{\mathrm{P_{conv}}}{\mathrm{P_{SR}}} = 2 \,, \tag{2.65}$$

which is achieved when there exists $i \in \{1, \ldots, M/2 - 1\}$ such that $x_{i+1} > A > (x_i + x_{M/2})/2$.

Proof: Let $x_i < A < x_{i+1}$ for any $i \in \{1, \ldots, M/2 - 1\}$. Note that there is no need to consider $i = M/2$ since there can be no improvement by adding noise to the measurement for $A > x_{M/2} = \max\{x_i\}$, as stated in Proposition 3. From (2.62), the probability of error for the conventional case can be calculated for $x_i < A < x_{i+1}$ as (c.f. Fig. 2.6-(a))

$$\mathrm{P_{conv}} = \frac{1}{2}\left(1 - 2\sum_{l=1}^{i} w_l\right) = \frac{1}{2} - \sum_{l=1}^{i} w_l \,, \tag{2.66}$$

where the symmetry property of the Gaussian mixture, i.e., $x_i = -x_{M-i+1}$ and $w_i = w_{M-i+1}$ for $i = 1, \ldots, M/2$, is employed.

In order to obtain the maximum improvement that can be obtained via additional noise, the parameter values that result in the minimum $\mathrm{P_{SR}}$ in (2.64)

[5] *Assuming an even M does not reduce the generality of the result due to the symmetry of the Gaussian mixture noise.*

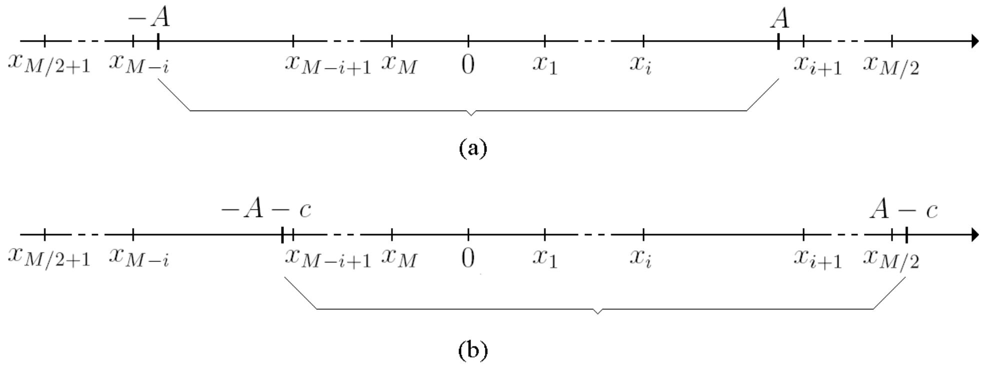

Figure 2.6: (a) In the conventional case, the mean values (x_j's) of the Gaussian mixture noise that are in the interval $[-A, A]$ determine the probability of error. (b) When a constant noise term c is added, the mean values (x_j's) of the Gaussian mixture noise that are in the interval $[-A-c, A-c]$ determine the probability of error.

should be determined. The interpretation of the probability of error calculation related to the weights of x_j's that reside in the interval $[-A-c, A-c]$ (as in the example in Fig. 2.5) implies that the maximum improvement can be obtained for a value of c that results in a shift of the interval $[-A, A]$ such that all the x_j values that are on the shift direction are included in the new interval $[-A-c, A-c]$ in addition to the x_j's that are already included in $[-A, A]$. This scenario is depicted in Fig. 2.6. In the conventional case, $\pm x_1, \ldots, \pm x_i$ are included in the interval $[-A, A]$. The minimum probability of error when noise c is added corresponds to the case in which the interval $[-A-c, A-c]$ includes as many x_j's as possible. Since shifting the interval $[-A, A]$ to one direction (to the right for $c < 0$ and to the left for $c > 0$) guarantees that the at least $M/2 - i$ points will be outside $[-A-c, A-c]$, the best case is obtained when the interval $[-A-c, A-c]$ includes all the remaining $M/2 + i$ points, as in Fig. 2.6-(b). In that case, the probability of error is given by

$$\mathrm{P}_{\mathrm{SR}} = \frac{1}{2}\left(1 - 2\sum_{l=1}^{i} w_l - \sum_{l=i+1}^{M/2} w_l\right) . \tag{2.67}$$

Due to symmetry, $\sum_{l=1}^{M/2} w_l = 1/2$. Therefore, $\sum_{l=i+1}^{M/2} w_l$ in (2.67) can be expressed as $1/2 - \sum_{l=1}^{i} w_l$. Hence, (2.67) becomes

$$\mathrm{P_{SR}} = \frac{1}{2}\left(\frac{1}{2} - \sum_{l=1}^{i} w_l\right) = \frac{\mathrm{P_{conv}}}{2} \ , \tag{2.68}$$

as claimed in the proposition.

Note that the scenario in Fig. 2.6-(b) can be obtained if $-A-c < x_{M-i+1}$ and $A - c > x_{M/2}$. Since $x_{M-i+1} = -x_i$, these inequalities imply $A > (x_i + x_{M/2})/2$. As A is assumed to satisfy $x_i < A < x_{i+1}$, the minimum probability of error can be obtained when $x_{i+1} > A > \frac{x_i + x_{M/2}}{2}$, as stated in the proposition.[6]

To complete the proof, the equality case is considered as well. Let $A = x_i$ for any $i \in \{1, \ldots, M/2\}$. Then, the probability of error in (2.62) can be obtained as

$$\mathrm{P_{conv}} = \frac{1}{2}\left(1 - 2\sum_{l=1}^{i-1} w_l - w_i\right) \ . \tag{2.69}$$

Similar to preceding arguments for calculating the minimum probability of error for the noise enhanced case, (2.64) can be expressed as

$$\mathrm{P_{SR}} = \frac{1}{2}\left(1 - 2\sum_{l=1}^{i-1} w_l - \sum_{l=i}^{M/2} w_i\right) = \frac{1}{2}\left(\frac{1}{2} - \sum_{l=1}^{i-1} w_l\right) \ . \tag{2.70}$$

From (2.69) and (2.70), $\mathrm{P_{SR}} > \mathrm{P_{conv}}/2$ is obtained. Hence, the maximum improvement cannot be obtained for $A = x_i$. $\Box$

The practical importance of Proposition 5 is that it defines an upper bound on the performance improvement that can be obtained by using additional noise, when the variances of the Gaussian components in the mixture noise (c.f. (2.40)) are significantly smaller than the distances between consecutive mean values, x_j's in (2.39). In such a case, Proposition 5 states that the noise enhanced detector cannot have a probability of decision error that is smaller than half of that for the conventional case.

[6]For a leftwards shift, i.e., for $c > 0$, $-A - c < x_{M/2+1} = -x_{M/2}$ and $A - c > x_i$ need to be satisfied for the maximum improvement, which results in the same expression.

The proof of Proposition 5 also leads to derivation of some necessary and sufficient conditions for improvability or non-improvability of detection via SR as $\sigma_i \to 0$ for $i = 1, \ldots, M$. A simple sufficient condition for improvability can be obtained by investigation of Fig. 2.6-(a). If A satisfies $A > (x_i + x_{i+1})/2$, shifting the interval $[-A, A]$ to the right (left) by an amount that is slightly larger than $x_{i+1} - A$; that is, setting $|c| = x_{i+1} - A + \epsilon$ for sufficiently small $\epsilon > 0$, results in including x_{i+1} (x_{M-i}) in the interval $[-A - c, A - c]$, in addition to all the points that are already included in $[-A, A]$. Therefore, smaller probability of error can be obtained in that case. Hence, the detector in (2.42) is improvable if $A > (x_i + x_{i+1})/2$ for $i \in \{1, \ldots, M/2 - 1\}$.

A sufficient condition for non-improvability can be obtained in a similar manner as $\sigma_i \to 0$ for $i = 1, \ldots, M$. First, the previous arguments imply that $A \leq (x_i + x_{i+1})/2$ is a necessary condition for non-improvability. In order to find a condition that guarantees that the detector cannot be improved by any value of c, it is first observed that for any possible improvement, the interval $[-A, A]$ in Fig. 2.6 must be shifted to the right (or, left) direction so that it includes some of $x_{i+1}, \ldots, x_{M/2}$ (or, $x_{M-i}, \ldots, x_{M/2+1}$) in the shifted interval $[-A - c, A - c]$. However, $A \leq (x_i + x_{i+1})/2$ implies that at least x_{M-i+1} (or, x_i) must be excluded from the interval $[-A - c, A - c]$ in order to include at least one of $x_{i+1}, \ldots, x_{M/2}$ (or, $x_{M-i}, \ldots, x_{M/2+1}$). If $w_i \geq \sum_{l=i+1}^{M/2} w_l$, the probability of error can never be lower for a non-zero value of c, since exclusion of x_{M-i+1} (or, x_i) causes an increase in the probability of error which cannot be compensated even if all of $x_{i+1}, \ldots, x_{M/2}$ (or, $x_{M-i}, \ldots, x_{M/2+1}$) are included in $[-A - c, A - c]$. In other words, in the presence of non-zero additional noise ($c \neq 0$), the probability of error can never be smaller than the conventional probability of error ($c = 0$). Therefore, for $x_i < A < x_{i+1}$, $A \leq (x_i + x_{i+1})/2$ and $w_i \geq \sum_{l=i+1}^{M/2} w_l$ are sufficient conditions for non-improvability.

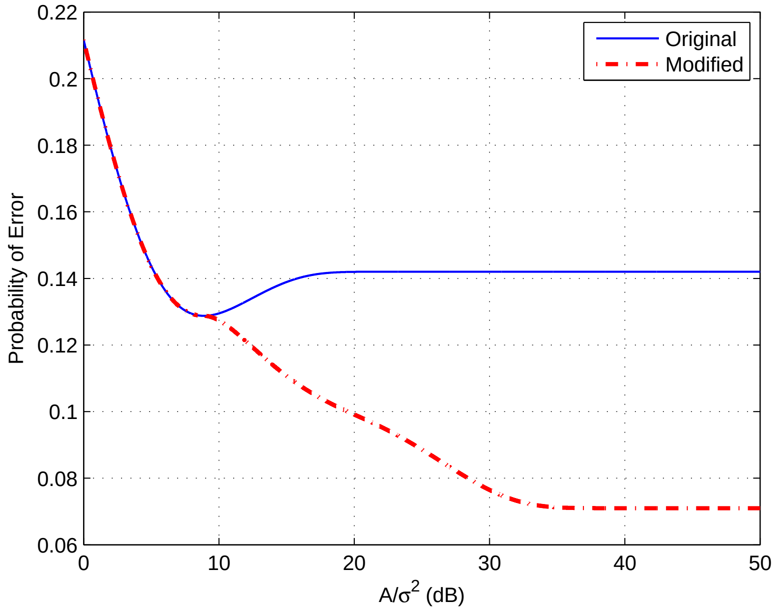

Figure 2.7: Probability of error versus A/σ^2 for symmetric Gaussian mixture noise with $M = 10$, where the center values are $\pm[0.02\ 0.18\ 0.30\ 0.55\ 1.35]$ with corresponding weights of $[0.167\ 0.075\ 0.048\ 0.068\ 0.142]$.

2.2.5 Numerical Results

In this section, numerical examples are provided in order to investigate the theoretical results obtained in the previous sections. For all cases, the variances of the Gaussian components in the mixture noise are assumed to be the same; i.e., $\sigma_i = \sigma$ for $i = 1, \ldots, M$ in (2.40).

First, symmetric Gaussian mixture noise with $M = 10$ is considered, where the mean values of the Gaussian components in the mixture noise in (2.39) are specified as $\pm[0.02\ 0.18\ 0.30\ 0.55\ 1.35]$ with corresponding weights of $[0.167\ 0.075\ 0.048\ 0.068\ 0.142]$. Fig. 2.7 illustrates the probabilities of error for the sign detector with and without additional noise (denoted as "modified" and "original", respectively) for various values of A/σ^2. The signal value A in (2.38) is set to $A = 1$, and σ is varied in order to obtain various A/σ^2 values. It is observed from Fig. 2.7 that the use of additional noise can improve detector

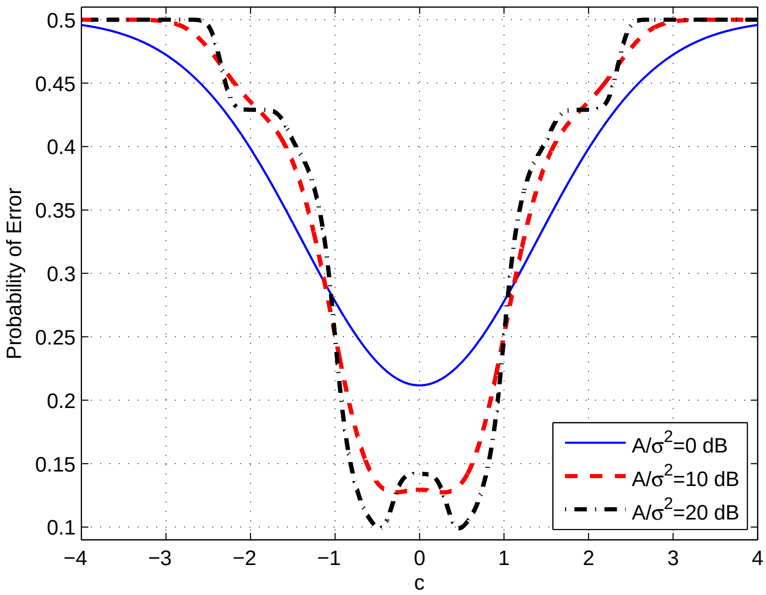

Figure 2.8: Probability of error in (2.47) versus c for various A/σ^2 values for the scenario in Fig. 2.7.

performance significantly for large A/σ^2 values, that is, as σ is decreased. In addition, the probability of error of the noise enhanced detector reduces monotonically with A/σ^2, as predicted by Proposition 4. On the other hand, the conventional sign detector without additional noise exhibits a non-monotonic behavior and experiences an error floor for high A/σ^2 values. Also, it is observed that as the variance increases, the detector becomes non-improvable as can be expected from Proposition 2.

In order to investigate the scenario in Fig. 2.7 in more detail, Fig. 2.8 plots the probability of error in (2.47) versus c for various A/σ^2 values, and Fig. 2.9 plots the improvability function in (2.48) and the optimal additional noise, c_{opt}, obtained from (2.46) versus A/σ^2. As stated by Proposition 1, whenever the function in (2.48) is negative, the detector is improvable, meaning that $c_{\text{opt}} \neq 0$. It is again observed that as σ^2 increases, the system becomes non-improvable.

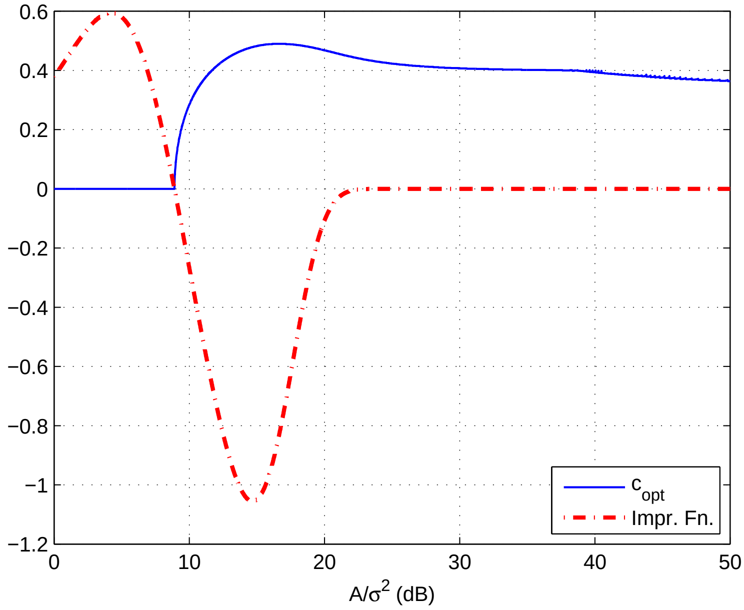

Figure 2.9: The improvability function in (2.48) and the optimal additional signal value c_{opt} in (2.46) versus A/σ^2 for the scenario in Fig. 2.7.

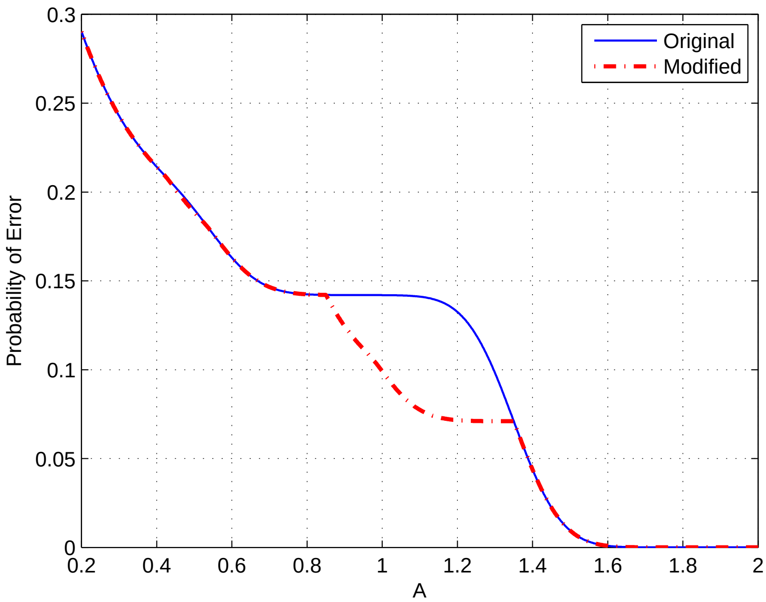

Figure 2.10: Probability of error versus A for symmetric Gaussian mixture noise with $\sigma_i = 0.1$ for $i = 1, \ldots, M$ and $M = 10$, where the center values are $\pm[0.02\ 0.18\ 0.30\ 0.55\ 1.35]$ with corresponding weights of $[0.167\ 0.075\ 0.048\ 0.068\ 0.142]$.

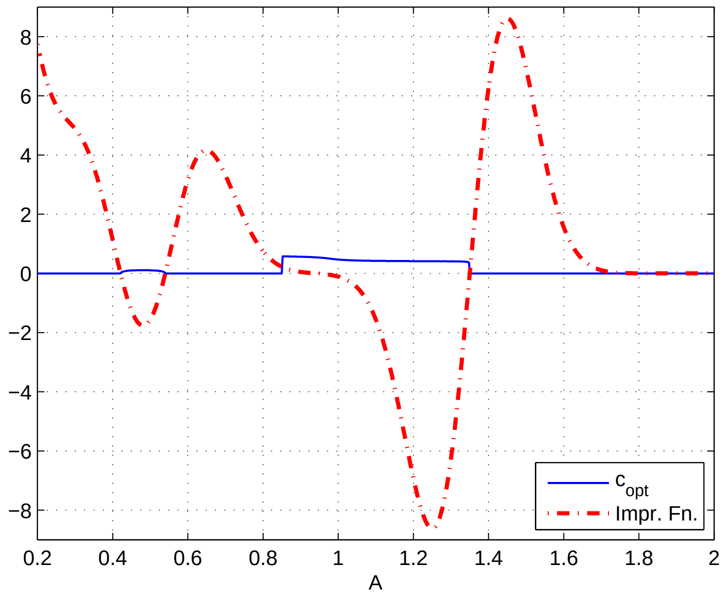

Figure 2.11: The improvability function in (2.48) and the optimal additional signal value $c_{\rm opt}$ in (2.46) versus A for the scenario in Fig. 2.10.

Next, the same Gaussian mixture noise as in the previous scenario is assumed, and the probabilities of error of the conventional and the noise enhanced detectors are plotted versus A for $\sigma = 0.1$ in Fig. 2.10. As stated in Proposition 3, the detector is non-improvable when the signal amplitude A is larger than or equal to the maximum mean value of the Gaussian components in the mixture noise; that is, $A \geq 1.35$ in this case. Fig. 2.11 illustrates the improvability function in (2.48) and the optimal additional noise $c_{\rm opt}$ in (2.46) versus A for this scenario. Again, it is observed that whenever the function in (2.48) is negative there is improvement (i.e., $c_{\rm opt} \neq 0$) in accordance with Proposition 1. It is also noted that the condition in Proposition 1 is a sufficient but not a necessary condition for improvability, which can be observed, for example, at $A = 0.86$, where the function value is positive and the detector is improvable. Investigation

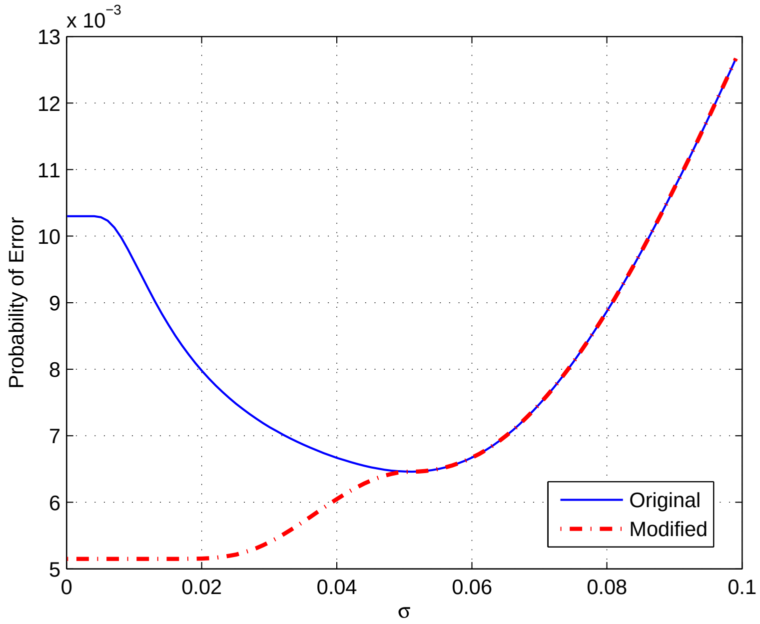

Figure 2.12: Probability of error versus σ for $A = 1$ and for symmetric Gaussian mixture noise with $M = 12$, where the center values are $\pm[0.0965\ 0.2252\ 0.4919\ 0.6372\ 0.8401\ 1.0151]$ with corresponding weights of $[0.1020\ 0.0022\ 0.2486\ 0.0076\ 0.1293\ 0.0103]$.

of Fig. 2.11 also reveals that the detector is improvable for $A \in [0.42, 0.54]$ and $A \in [0.85, 1.35]$.[7]

For the final scenario, a symmetric Gaussian mixture noise with $M = 12$ is considered, where the mean values of the Gaussian components in the mixture noise in (2.39) are specified as $\pm[0.0965\ 0.2252\ 0.4919\ 0.6372\ 0.8401\ 1.0151]$ with corresponding weights of $[0.1020\ 0.0022\ 0.2486\ 0.0076\ 0.1293\ 0.0103]$. Fig. 2.12 plots the probabilities of error for the conventional and noise enhanced detectors versus σ when the signal amplitude A is set to $A = 1$. In accordance with Proposition 2, as σ increases, the detector becomes non-improvable, which is also observed from Fig. 2.13, the plot of c_{opt} versus σ. In addition, as stated in Proposition 4, the probability of error is a monotone increasing function of

[7]Although the improvement for $A \in [0.42, 0.54]$ is difficult to observe from Fig. 2.10, the numerical results show slight reductions in the probabilities of error of the noise enhanced detector in that range.

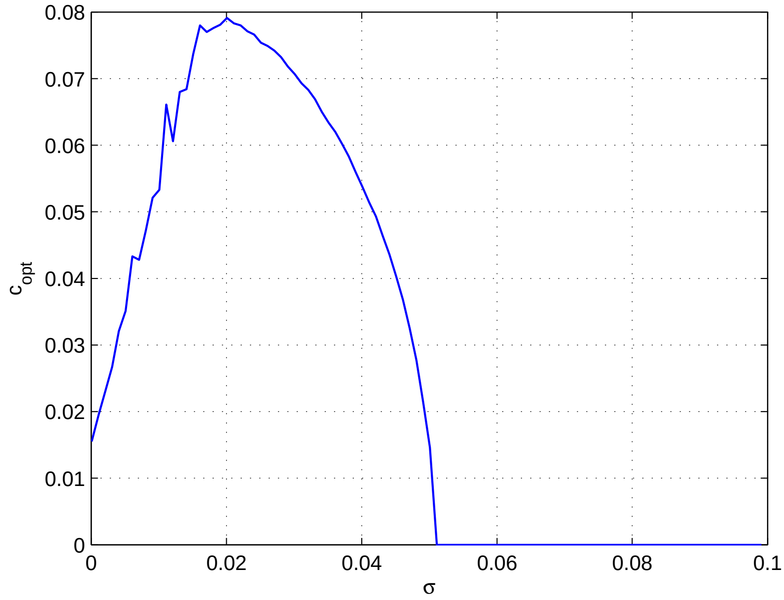

Figure 2.13: The optimal additional noise c_{opt} in (2.46) versus σ for the scenario in Fig. 2.12.

σ for the noise enhanced detector, whereas the conventional one exhibits a non-monotonic behavior. Also, as $\sigma \to 0$, the ratio between the probability of error in the conventional case and the noise enhanced case becomes 2 ($\text{P}_{\text{conv}} = 0.0103$ and $\text{P}_{\text{SR}} = 0.00515$). This is expected from Proposition 5, since A satisfies the condition in the proposition, $x_{i+1} > A > (x_i + x_{M/2})/2$ for $i = 5$ (namely, $1.1051 > 1 > (0.8401 + 1.1051)/2 = 0.9726$). Finally, Fig. 2.14 illustrates the improvability function in (2.48), which indicates that the detector is improvable (i.e., $c_{\text{opt}} \neq 0$) whenever the function takes a negative value.

2.3 Concluding Remarks and Extensions

In this chapter, two main contributions have been provided. First, the effects of independent additional noise have been investigated for M-ary composite

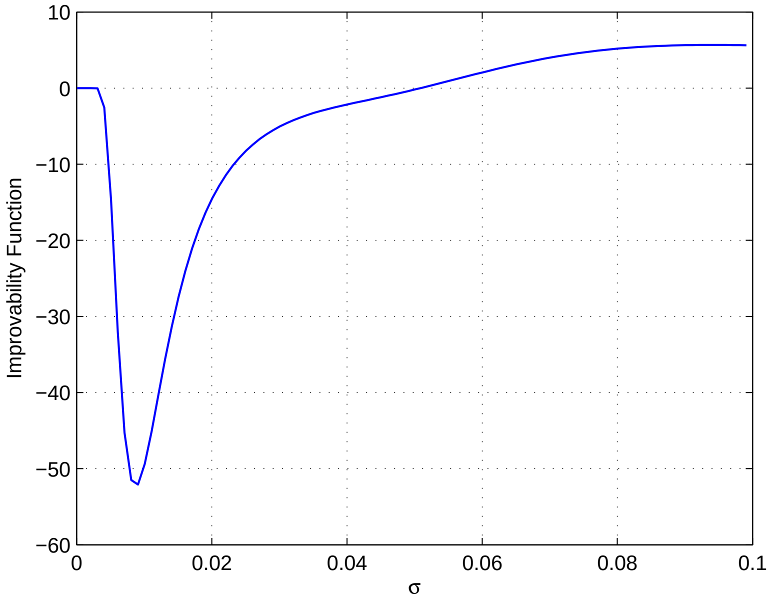

Figure 2.14: The improvability function in (2.48) versus σ for the scenario in Fig. 2.12.

hypothesis-testing problems in the Bayesian framework. It has been shown that the optimal additional noise can be expressed as the shifts of the measurements.

Second, the effects of additional noise on a sign detector have been studied for binary hypothesis testing under symmetric Gaussian mixture noise. Sufficient conditions have been obtained for improvability and non-improvability of the detector in terms of the desired signal amplitude and the parameters of the Gaussian mixture noise. Also, the monotonicity property of the noise enhanced detector has been proven for equal variances of the Gaussian components in the mixture noise. In addition, for infinitesimally small variances of the Gaussian mixture components, the maximum improvement that can be achieved via additional noise has been specified as half of the probability of error for the detector without additional noise. The numerical examples have been provided to support and explain the theoretical results.

It should be noted that the results in Section 2.2 can be extended to Gaussian mixture noise that is symmetric around a non-zero value as well. In that case, the conventional sign detector compares each measurement to the mean of the mixture noise. The framework in Section 2.2 can still be employed, based on modified measurements that are obtained by subtracting the mean value from the original measurements. In addition, the theoretical results on probability of error performance in Section 2.2.4 can be considered for asymmetric Gaussian mixture noise in the following manner. As stated after Proposition 4, the monotonicity result is valid also for the asymmetric case. Considering the theoretical limit in Proposition 5 on performance improvement that can be achieved via additional noise, the maximum ratio between the probabilities of error for the conventional and noise enhanced detectors becomes infinity for asymmetric Gaussian mixture noise as the variances of the Gaussian components converge zero. This is because there can be cases in which the interval $[-A - c, A - c]$ in Fig. 2.6-(b) includes all the mean values (x_j's) while the interval $[-A, A]$ in Fig. 2.6-(a) does not, which is possible due to the asymmetry of the mean values. In that case, the probability of error becomes zero for the noise enhanced case whereas it is non-zero for the conventional one. In other words, the performance improvement that can be obtained via additional noise is unbounded for asymmetric Gaussian mixture measurement noise.

Future work includes extensions of the theoretical framework in Section 2.2 to other noise distributions than the Gaussian mixture noise.

Chapter 3

Noise Enhanced M-ary Hypothesis-Testing in the Minimax Framework

In this chapter, noise benefits are investigated for M-ary hypothesis-testing problems in the minimax framework. In Section 3.1, the formulation of optimal additional noise is provided for an M-ary hypothesis-testing problem according to the minimax criterion. Then, it is shown in Section 3.2 that the optimal additional noise can be represented by a randomization of no more than M signal levels. In addition, a convex relaxation approach is proposed to obtain an accurate approximation to the noise probability density function (PDF) in polynomial time. Also, sufficient conditions are provided regarding the improvability and non-improvability of a given detector via additional noise. Finally, numerical examples and concluding remarks are presented in Section 3.3.

3.1 Problem Formulation

Consider the following M-ary hypothesis-testing problem:

$$\mathcal{H}_i \ : \ p_i^{\mathbf{X}}(\boldsymbol{x}) \ , \qquad i = 0, 1, \ldots, M-1 \ , \tag{3.1}$$

where $p_i^{\mathbf{X}}(\boldsymbol{x})$ represents the PDF of the observation under hypothesis $\mathcal{H}_i$ and the observation (measurement) $\boldsymbol{x}$ is a vector with K components; i.e., $\boldsymbol{x} \in \mathbb{R}^K$.

A generic decision rule can be defined as

$$\phi(\boldsymbol{x}) = i \ , \quad \text{if} \ \ \boldsymbol{x} \in \Gamma_i \ , \tag{3.2}$$

for $i = 0, 1, \ldots, M-1$, where $\Gamma_0, \Gamma_1, \ldots, \Gamma_{M-1}$ form a partition of the observation space Γ.

In the minimax approach, the prior probabilities of the hypotheses are unknown. However, each decision is associated with a known cost value, and the aim is to minimize the maximum of the average costs of the decision rule conditioned on different hypotheses [21]. More formally, let $C_{ji} \geq 0$ represent the cost of choosing $\mathcal{H}_j$ when $\mathcal{H}_i$ is true. Then, the average cost of a decision rule ϕ conditioned on $\mathcal{H}_i$ being the true hypothesis is calculated as

$$\mathrm{R}_i(\phi) = \sum_{j=0}^{M-1} C_{ji} \mathrm{P}_i(\Gamma_j) \ , \tag{3.3}$$

where $\mathrm{P}_i(\Gamma_j)$ represents the probability of choosing $\mathcal{H}_j$ when $\mathcal{H}_i$ is the true hypothesis. This quantity, $\mathrm{R}_i(\phi)$, is called the *conditional risk* of ϕ given $\mathcal{H}_i$. Under the minimax framework, the aim is to reduce the maximum of the conditional risks for different hypotheses as much as possible.

In some cases, addition of independent noise to measurements can improve the performance of a suboptimal decision rule (detector) [1], [17], [31]. In such scenarios, instead of the original measurement (observation) $\boldsymbol{x}$, a noise-added version of that, $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}$, is used by the detector, where $\boldsymbol{n}$ represents the additional noise term. The main motivation for such noise enhanced detection approaches is to use a low-complexity suboptimal detector, and improve its performance via adjusting the measurements.

In this study, we consider a fixed decision rule ϕ, and aim to obtain the optimal additional noise PDF $p_{\boldsymbol{N}}(\cdot)$ that minimizes the maximum of the conditional risks.

$$p_{\boldsymbol{N}}^{\mathrm{opt}}(\boldsymbol{n}) = \arg \min_{p_{\boldsymbol{N}}(\boldsymbol{n})} \max_{i\in\{0,1,\dots,M-1\}} \mathrm{R}_i^{\boldsymbol{y}}(\phi) \;, \tag{3.4}$$

where $\mathrm{R}_i^{\boldsymbol{y}}(\phi)$ represents the conditional risk of ϕ given $\mathcal{H}_i$ when the noise-added measurement $\boldsymbol{y}$ is used; that is, $\mathrm{R}_i^{\boldsymbol{y}}(\phi) = \sum_{j=0}^{M-1} C_{ji}\mathrm{P}_i^{\boldsymbol{y}}(\Gamma_j)$, with $\mathrm{P}_i^{\boldsymbol{y}}(\Gamma_j)$ representing the probability that $\boldsymbol{y} \in \Gamma_j$ when $\mathcal{H}_i$ is true.

3.2 Noise Enhanced Hypothesis-Testing

In order to investigate the solution of the optimization problem in (3.4), we first manipulate the conditional risk $\mathrm{R}_i^{\boldsymbol{y}}(\phi)$ as follows:

$$\begin{aligned}
\mathrm{R}_i^{\boldsymbol{y}}(\phi) &= \sum_{j=0}^{M-1} C_{ji}\mathrm{P}_i^{\boldsymbol{y}}(\Gamma_j) = \sum_{j=0}^{M-1} C_{ji} \int_{\Gamma_j} p_i^{\boldsymbol{Y}}(\mathbf{z})d\mathbf{z} &(3.5)\\
&= \sum_{j=0}^{M-1} C_{ji} \int_{\Gamma_j}\int_{\mathbb{R}^K} p_{\boldsymbol{N}}(\boldsymbol{n}) p_i^{\boldsymbol{X}}(\mathbf{z}-\boldsymbol{n})\, d\boldsymbol{n}\, d\mathbf{z} &(3.6)\\
&= \sum_{j=0}^{M-1} C_{ji} \int_{\mathbb{R}^K} p_{\boldsymbol{N}}(\boldsymbol{n}) \int_{\Gamma_j} p_i^{\boldsymbol{X}}(\mathbf{z}-\boldsymbol{n})\, d\mathbf{z}\, d\boldsymbol{n} &(3.7)\\
&= \sum_{j=0}^{M-1} C_{ji}\, \mathrm{E}\{F_{ij}(\boldsymbol{N})\} = \mathrm{E}\{F_i(\boldsymbol{N})\} \;, &(3.8)
\end{aligned}$$

with $F_{ij}(\boldsymbol{n}) \doteq \int_{\Gamma_j} p_i^{\boldsymbol{X}}(\mathbf{z}-\boldsymbol{n})d\mathbf{z}$ and $F_i(\boldsymbol{n}) \doteq \sum_{j=0}^{M-1} C_{ji}F_{ij}(\boldsymbol{n})$. From (3.5) to (3.6), the independence of $\boldsymbol{X}$ and $\boldsymbol{N}$ is employed to obtain the PDF of $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N}$. Then, the optimization problem in (3.4) becomes

$$\min_{p_{\boldsymbol{N}}(\cdot)} \max_{i\in\{0,1,\dots,M-1\}} \mathrm{E}\{F_i(\boldsymbol{N})\} \;. \tag{3.9}$$

Note that under uniform cost assignment (UCA); that is, when $C_{ji} = 1$ for $j \neq i$, and $C_{ji} = 0$ for $j = i$, the conditional risk becomes $\mathrm{R}_i^{\boldsymbol{y}}(\phi) = 1 - \mathrm{E}\{F_{ii}(\boldsymbol{N})\}$. Then, (3.9) can be expressed as

$$\max_{p_{\boldsymbol{N}}(\cdot)} \min_{i\in\{0,1,\ldots,M-1\}} \mathrm{E}\{F_{ii}(\boldsymbol{N})\}. \tag{3.10}$$

Although it is quite difficult to perform a search over all possible noise PDFs in (3.9), the following proposition states that the search can be performed over the set of discrete probability distributions with at most M mass points in most practical scenarios.

Proposition 1: *Define set U as*

$$\begin{aligned} U = \{(u_0, u_1, \ldots, u_{M-1}) \; : \; u_0 = F_0(\boldsymbol{n}),\; u_1 = F_1(\boldsymbol{n}), \\ \ldots,\; u_{M-1} = F_{M-1}(\boldsymbol{n}) \;,\; \text{for } \boldsymbol{a} \preceq \boldsymbol{n} \preceq \boldsymbol{b}\} \;, \end{aligned} \tag{3.11}$$

where $\boldsymbol{n} \in \mathbb{R}^K$, and $\boldsymbol{a} \preceq \boldsymbol{n} \preceq \boldsymbol{b}$ means that $n_j \in [a_j, b_j]$ for $j = 1, \ldots, K$. If U is a closed subset of $\mathbb{R}^K$, then the optimal noise PDF in (3.4) *can be expressed as*

$$p_{\boldsymbol{N}}^{\text{opt}}(\boldsymbol{n}) = \sum_{i=0}^{M-1} \lambda_i \, \delta(\boldsymbol{n} - \boldsymbol{n}_i) \;, \tag{3.12}$$

where $\sum_{i=0}^{M-1} \lambda_i = 1$ and $\lambda_i \geq 0$ for $i = 0, 1, \ldots, M-1$.

Proof: The proof can be viewed as an extension of the results in [17] and [1] for the two-dimensional case to the M-dimensional case. The details of a similar proof can be found in [32]. $\square$

The main implication of Proposition 1 is that an optimal additional noise can be represented by a randomization of no more than M different signal levels. Under certain conditions, such as the following one, the optimal noise PDF can be guaranteed to include less than M mass points.

Corollary 1: *Let S_1 and S_2 represent two sets such that $S_1 \cap S_2 = \emptyset$ and $S_1 \cup S_2 = \{0, 1, \ldots, M-1\}$. If $\max_{i \in S_2} F_i(\boldsymbol{n}) \leq \min_{i \in S_1} F_i(\boldsymbol{n})\ \forall \boldsymbol{n}$, then the optimal noise PDF contains at most $|S_1|$ mass points.*[1]

Proof: Under the conditions in the corollary, the conditional risks indexed by S_2 do not have any effect on the minimax risk, since the other conditional risks determine the maximum risk for all possible noise values. Therefore, the result in the corollary directly follows from Proposition 1. □

Based on Proposition 1, the optimization problem in (3.9) can be expressed as

$$\min_{\{\boldsymbol{n}_j, \lambda_j\}_{j=0}^{M-1}} \ \max_{i \in \{0,1,\ldots,M-1\}} \sum_{j=0}^{M-1} \lambda_j \, F_i(\boldsymbol{n}_j) \ , \tag{3.13}$$

subject to $\sum_{j=0}^{M-1} \lambda_j = 1$ and $\lambda_j \geq 0$ for $j = 0, 1, \ldots, M-1$. Although (3.13) is significantly simpler than (3.9), it can still be a non-convex optimization problem in general. Therefore, global optimization techniques, such as particle-swarm optimization (PSO), can be applied to obtain the optimal noise PDF [33]. As an alternative approach, we provide an approximate formulation that results in a convex optimization problem. Assume that the additional noise $\boldsymbol{n}$ can take only finitely many known values specified by $\tilde{\boldsymbol{n}}_1, \ldots, \tilde{\boldsymbol{n}}_L$, and the aim is to determine the weights $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_L$ of those possible noise values. Then, (3.9) can be expressed, after some manipulation, as the following optimization problem:

$$\begin{aligned} &\min_{t, \{\tilde{\lambda}_j\}_{j=1}^{L}} \ t \\ &\text{subject to} \ \sum_{j=1}^{L} \tilde{\lambda}_j \, F_i(\tilde{\boldsymbol{n}}_j) \leq t \ , \ \text{for} \ \ i = 0, 1, \ldots, M-1 \\ &\qquad\qquad \sum_{j=0}^{L} \tilde{\lambda}_j = 1 \ , \quad \tilde{\lambda}_j \geq 0 \ , \ j = 1, \ldots, L \end{aligned} \tag{3.14}$$

The optimization problem in (3.14) is a linearly constrained linear programming (LCLP) problem, which can be solved in polynomial time [34]. Also, as L is

[1] *Here, $|S_1|$ denotes the number of elements in set S_1.*

increased (as the optimization is performed over more noise values), the solution of the optimization problem in (3.14) gets close to the optimal solution of (3.9).

Finally, the issue of determining whether additional noise can improve the performance of a given detector without actually solving the optimization problem in (3.9) is addressed. In the following, sufficient conditions are presented for the improvability and the non-improvability of a given detector via the use of additional noise.

Proposition 2: *Define* $J(\boldsymbol{n}) = \max_{i\in\{0,1,\ldots,M-1\}} F_i(\boldsymbol{n})$. *If* $\boldsymbol{n}_0 = \arg\min_{\boldsymbol{n}} J(\boldsymbol{n})$ *is not equal to zero, then the detector is improvable.*

Proof: Consider that the noise with PDF $p_{\boldsymbol{N}}(\boldsymbol{n}) = \delta(\boldsymbol{n} - \boldsymbol{n}_0)$ is added to the observation $\boldsymbol{x}$. Then, the maximum of the conditional risks become $\max_i \mathrm{R}_i^{\boldsymbol{y}}(\phi) = \max_i F_i(\boldsymbol{n}_0) = J(\boldsymbol{n}_0)$. Since $\boldsymbol{n}_0 = \arg\min_{\boldsymbol{n}} J(\boldsymbol{n}) \neq \mathbf{0}$, $J(\boldsymbol{n}_0) < J(\mathbf{0}) = \max_i F_i(\mathbf{0}) = \max_i \mathrm{R}_i(\phi)$. In other words, $\max_i \mathrm{R}_i^{\boldsymbol{y}}(\phi) < \max_i \mathrm{R}_i(\phi)$; hence, the detector is improvable. □

Proposition 3: *Let* $k = \arg\max_i F_i(\mathbf{0})$. *If* $\arg\min_{\boldsymbol{n}} F_k(\boldsymbol{n})$ *is equal to zero, then the detector is non-improvable.*

Proof: The statement $k = \arg\max_i F_i(\mathbf{0})$ means that in the absence of additional noise, the kth conditional risk is the maximum one; hence, it determines the overall risk in the minimax framework. If $\arg\min_{\boldsymbol{n}} F_k(\boldsymbol{n})$ is equal to zero, it means that addition of noise cannot reduce the kth conditional risk. Since the kth conditional risk cannot be reduced by any additional noise and it is the maximum one among all the conditional risks, the performance of the detector cannot be improved. □

The results in Proposition 2 and Proposition 3 can be used to determine when it is necessary to tackle the optimization problem in (3.9) to obtain the optimal

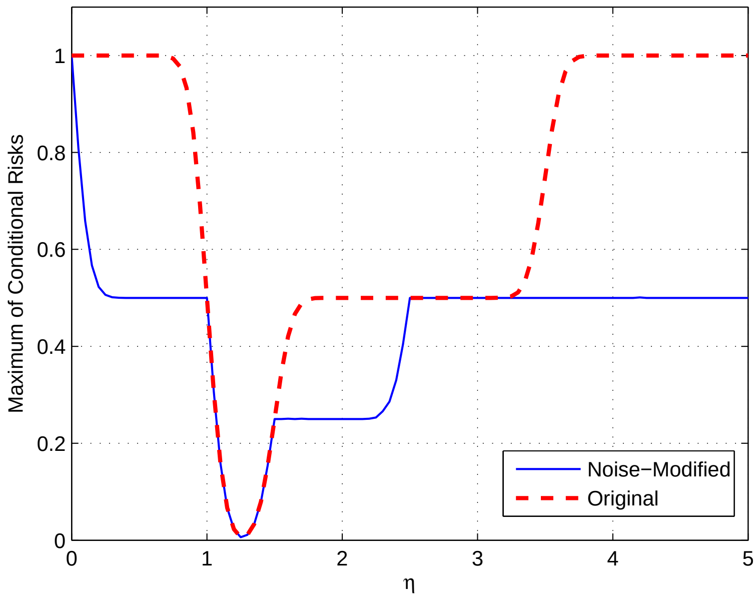

Figure 3.1: Maximum of the conditional risks versus η for the original and the noise-modified detectors for $A = 1$, $B = 2.5$, $\sigma = 0.1$, $w_1 = 0.5$ and $w_2 = 0.5$.

additional noise PDF. For example, when the non-improvability condition in Proposition 3 is satisfied, it is directly concluded that $p_{\boldsymbol{N}}^{\text{opt}}(\boldsymbol{n}) = \delta(\boldsymbol{n})$.

3.3 Numerical Results

In this section, numerical examples are provided in order to investigate the theoretical results obtained in the previous section. A ternary hypothesis-testing problem is considered with the following PDFs:

$$
\begin{aligned}
p_0^X(x) &= w_1\gamma(x; -A, \sigma^2) + w_2\gamma(x; A, \sigma^2) \\
p_1^X(x) &= w_1\gamma(x; -A + B, \sigma^2) + w_2\gamma(x; A + B, \sigma^2) \\
p_2^X(x) &= w_1\gamma(x; -A - B, \sigma^2) + w_2\gamma(x; A - B, \sigma^2)
\end{aligned} \tag{3.15}
$$

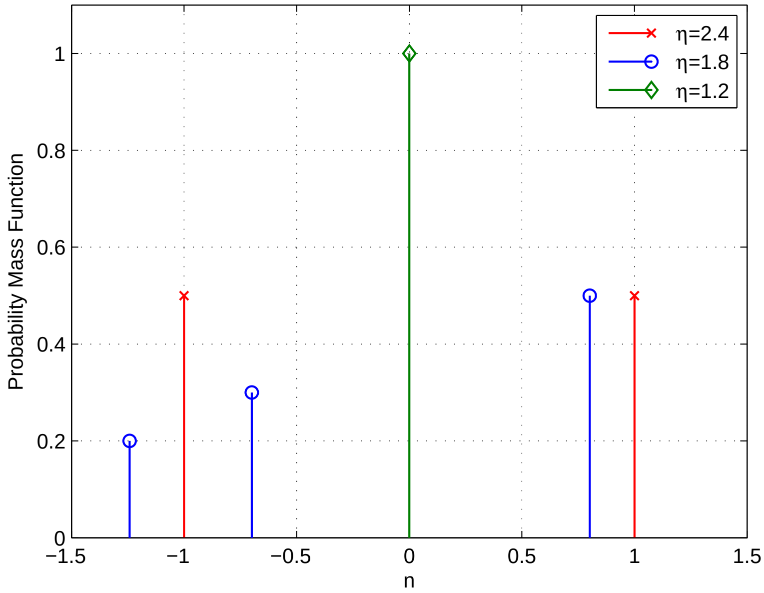

Figure 3.2: Probability mass function of optimal additional noise for various threshold values when the parameters are taken as $A = 1$, $B = 2.5$, $\sigma = 0.1$, $w_1 = 0.5$ and $w_2 = 0.5$.

where $\gamma(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. The decision rule is described as follows:

$$\phi(x) = \begin{cases} 0\,, & -\eta < x < \eta \\ 1\,, & x \geq \eta \\ 2\,, & x \leq -\eta \end{cases}\,, \tag{3.16}$$

where η is a constant. Under UCA, the conditional risks can be obtained, after some manipulation, as

$$\begin{aligned} \mathrm{R}_0(\phi) &= 1 - w_1 \left[Q\left(\frac{-\eta + A}{\sigma}\right) - Q\left(\frac{\eta + A}{\sigma}\right) \right] \\ &\quad - w_2 \left[Q\left(\frac{-\eta - A}{\sigma}\right) - Q\left(\frac{\eta - A}{\sigma}\right) \right] \\ \mathrm{R}_i(\phi) &= 1 - w_1\, Q\left(\frac{\eta + s_i A - B}{\sigma}\right) - w_2\, Q\left(\frac{\eta - s_i A - B}{\sigma}\right) \end{aligned}$$

for $i = 1, 2$, where $s_1 = 1$ and $s_2 = -1$.

Fig. 3.1 plots the maximum of conditional risks for the original and the noise-modified detectors with respect to η in (3.16) when the parameters are taken as $A = 1$, $B = 2.5$, $w_1 = 0.5$, $w_2 = 0.5$ and $\sigma = 0.1$. From the figure, it is observed that for certain values of η, the performance can be improved via the addition of noise. For example, for $\eta = 1.8$, the improvement ratio, defined as the ratio between $\max_{i\in\{1,2,3\}} \mathrm{R}_i(\phi)$ and $\max_{i\in\{1,2,3\}} \mathrm{R}_i^y(\phi)$, is equal to 2. As another example, for $\eta = 2.4$, the improvement ratio is calculated as 1.52.

In Fig. 3.2, the probability distributions of the optimal additional noise components are illustrated for $\eta = 1.2$, $\eta = 1.8$ and $\eta = 2.4$ based on the parameter settings for Fig. 3.1. It is observed that the optimal noise PDFs for $\eta = 2.4$, $\eta = 1.8$ and $\eta = 1.2$ contain 2, 3 and 1 mass points, respectively, in accordance with Proposition 1. Also, it is noted that since the detector is non-improvable for $\eta = 1.2$, the optimal noise turns out to be zero.

Finally, Fig. 3.3 illustrates the performance of the original and the noise-modified detectors versus the standard deviation parameter in (3.15) for $\eta = 1.8$, $A = 1$, $B = 2.5$, $w_1 = 0.5$ and $w_2 = 0.5$. As the standard deviation increases, the improvement ratio becomes smaller, and after a certain value, the detector becomes non-improvable.

3.4 Concluding Remarks

In this chapter, the effects of adding independent noise to measurements have been studied for M-ary hypothesis-testing problems according to the minimax criterion. It has been shown that the optimal additional noise can be represented by a randomization of at most M signal values. In addition, a convex relaxation approach has been proposed to obtain an accurate approximation to the noise probability distribution in polynomial time. Furthermore, sufficient conditions have been presented to determine when additional noise can or cannot improve

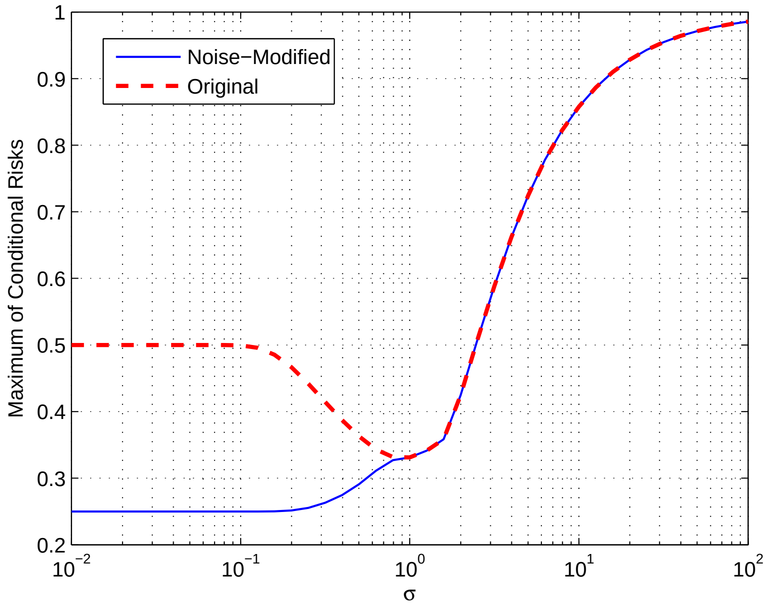

Figure 3.3: Original and noise-modified maximum of the conditional risks vs σ graph for the parameters taken as $\eta = 1.8$, $A = 1$, $B = 2.5$, $w_1 = 0.5$ and $w_2 = 0.5$.

the performance of a given detector. Finally, a numerical example has been presented.

Chapter 4

On the Improvability and Non-improvability of Detection in the Neyman-Pearson Framework

In this chapter, noise benefits are investigated in the Neyman-Pearson framework [1], [17]; that is, improvements in detection probability under a constraint on the probability of false-alarm are considered. Section 4.1 introduces the detection problem and the formal definitions of improvability and non-improvability. Then, a non-improvability condition is presented in Section 4.2. In Section 4.3, a generic improvability condition, as well as more specific and explicit ones are derived. Finally, a numerical example is provided in Section 4.4, and concluding remarks are made in Section 4.5.

4.1 Signal Model

Consider a binary hypothesis-testing problem described as

$$\begin{aligned} \mathcal{H}_0 &: p_0(\boldsymbol{x}) \ , \\ \mathcal{H}_1 &: p_1(\boldsymbol{x}) \ , \end{aligned} \tag{4.1}$$

where $\boldsymbol{x}$ is the K-dimensional data (measurement) vector, and $p_0(\boldsymbol{x})$ and $p_1(\boldsymbol{x})$ represent the PDFs of $\boldsymbol{x}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.

The decision rule (detector) is denoted by $\phi(\boldsymbol{x})$, which maps the data vector into a real number in $[0,1]$, which represents the probability of selecting $\mathcal{H}_1$ [21]. Under certain circumstances, detector performance can be improved by adding independent noise to the data vector $\boldsymbol{x}$ [1], [17]. Let $\boldsymbol{y}$ represent the modified data vector expressed as

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n} \ , \tag{4.2}$$

where $\boldsymbol{n}$ represents the additional independent noise term.

The Neyman-Pearson framework is considered in this study, and performance of a detector is specified by its probability of detection and probability of false alarm [21]. Since the additional noise is independent of the data, the probabilities of detection and false alarm are given, respectively, by

$$\mathrm{P}_{\mathrm{D}}^{\mathbf{y}} = \int_{\mathbb{R}^K} \phi(\boldsymbol{y}) \left[\int_{\mathbb{R}^K} p_1(\boldsymbol{y}-\boldsymbol{x}) p_{\boldsymbol{N}}(\boldsymbol{x}) d\boldsymbol{x} \right] d\boldsymbol{y} \ , \tag{4.3}$$

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} = \int_{\mathbb{R}^K} \phi(\boldsymbol{y}) \left[\int_{\mathbb{R}^K} p_0(\boldsymbol{y}-\boldsymbol{x}) p_{\boldsymbol{N}}(\boldsymbol{x}) d\boldsymbol{x} \right] d\boldsymbol{y} \ , \tag{4.4}$$

where K is the dimension of the data vector. After some manipulation, (4.3) and (4.4) can be expressed as [1]

$$\mathrm{P}_{\mathrm{D}}^{\mathbf{y}} = \mathrm{E}\{F_1(\boldsymbol{N})\} \ , \tag{4.5}$$

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} = \mathrm{E}\{F_0(\boldsymbol{N})\} \ , \tag{4.6}$$

where $\boldsymbol{N}$ is the random variable representing the additional noise term and

$$F_i(\boldsymbol{n}) \doteq \int_{\mathbb{R}^K} \phi(\boldsymbol{y}) p_i(\boldsymbol{y}-\boldsymbol{n}) d\boldsymbol{y} \ , \quad i = 0, 1 \ . \tag{4.7}$$

Note that in the absence of additional noise, i.e., $\boldsymbol{n} = \mathbf{0}$, the probabilities of detection and false alarm are given by $\mathrm{P}_\mathrm{D}^{\mathbf{x}} = F_1(\mathbf{0})$ and $\mathrm{P}_\mathrm{F}^{\mathbf{x}} = F_0(\mathbf{0})$, respectively. The detector $\phi(\cdot)$ is called *improvable* if there exists additional noise[1] $\boldsymbol{n}$ that satisfies $\mathrm{P}_\mathrm{D}^{\mathbf{y}} > \mathrm{P}_\mathrm{D}^{\mathbf{x}} = F_1(\mathbf{0})$ and $\mathrm{P}_\mathrm{F}^{\mathbf{y}} \le \mathrm{P}_\mathrm{F}^{\mathbf{x}} = F_0(\mathbf{0})$. Otherwise, the detector is called *non-improvable*.

4.2 Non-improvability Conditions

In [1], sufficient conditions for improvability and non-improvability are derived based on the following function:

$$J(t) = \sup \left\{ F_1(\boldsymbol{n}) \mid F_0(\boldsymbol{n}) = t \ , \ \boldsymbol{n} \in \mathbb{R}^K \right\} \ , \tag{4.8}$$

which defines the maximum probability of detection, obtained by adding constant noise $\boldsymbol{n}$, for a given probability of false alarm. It is stated that if there exists a non-decreasing concave function $\Psi(t)$ that satisfies $\Psi(t) \ge J(t)\ \forall t$ and $\Psi(\mathrm{P}_\mathrm{F}^{\mathbf{x}}) = J(\mathrm{P}_\mathrm{F}^{\mathbf{x}}) = F_1(\mathbf{0})$, then the detector is non-improvable [1]. The main advantage of this result is that it is based on single-variable functions $J(t)$ and $\Psi(t)$ irrespective of the dimension of the data vector. However, in certain cases, it may be difficult to calculate $J(t)$ in (4.8) or to obtain $\Psi(t)$. Therefore, we aim to derive a non-improvability condition that depends directly on F_0 and F_1 in (4.7).

The following proposition provides a sufficient condition for non-improvability based on convexity and concavity arguments for F_0 and F_1.

Proposition 1: *Assume that $F_0(\boldsymbol{n}) \le F_0(\mathbf{0})$ implies $F_1(\boldsymbol{n}) \le F_1(\mathbf{0})$ for all $\boldsymbol{n} \in \mathcal{S}_\mathrm{n}$, where $\mathcal{S}_\mathrm{n}$ is a convex set*[2] *consisting of all possible values of additional*

[1] In this thesis, additional noise that is independent of the original data is considered.

[2] *Since convex combination of individual noise components can be obtained via randomization [35], $\mathcal{S}_\mathrm{n}$ can be modeled as convex.*

noise $\boldsymbol{n}$. If $F_0(\boldsymbol{n})$ is a convex function and $F_1(\boldsymbol{n})$ is a concave function over $\mathcal{S}_{\mathrm{n}}$, then the detector is non-improvable.

Proof: Due to the convexity of F_0, the probability of false alarm in (4.6) can be bounded, via the Jensen's inequality, as

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} = \mathrm{E}\{F_0(\boldsymbol{N})\} \geq F_0\left(\mathrm{E}\{\boldsymbol{N}\}\right) \ . \tag{4.9}$$

Since $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} \leq \mathrm{P}_{\mathrm{F}}^{\mathbf{x}} = F_0(\mathbf{0})$ is a necessary condition for improvability, (4.9) implies that $F_0\left(\mathrm{E}\{\boldsymbol{N}\}\right) \leq F_0(\mathbf{0})$ is required. Since $\mathrm{E}\{\boldsymbol{N}\} \in \mathcal{S}_{\mathrm{n}}$, $F_0\left(\mathrm{E}\{\boldsymbol{N}\}\right) \leq F_0(\mathbf{0})$ implies that $F_1\left(\mathrm{E}\{\boldsymbol{N}\}\right) \leq F_1(\mathbf{0})$ due to the assumption in the proposition. Therefore,

$$\mathrm{P}_{\mathrm{D}}^{\mathbf{y}} = \mathrm{E}\{F_1(\boldsymbol{N})\} \leq F_1\left(\mathrm{E}\{\boldsymbol{N}\}\right) \leq F_1(\mathbf{0}) \ , \tag{4.10}$$

where the first inequality results from the concavity of F_1. Then, from (4.9) and (4.10), it is concluded that $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} \leq F_0(\mathbf{0}) = \mathrm{P}_{\mathrm{F}}^{\mathbf{x}}$ implies $\mathrm{P}_{\mathrm{D}}^{\mathbf{y}} \leq F_1(\mathbf{0}) = \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}$. Therefore, the detector is non-improvable.[3] $\square$

Consider the assumption in the proposition, which states that $F_0(\boldsymbol{n}) \leq F_0(\mathbf{0})$ implies $F_1(\boldsymbol{n}) \leq F_1(\mathbf{0})$ for all possible values of $\boldsymbol{n}$. This assumption is realistic in most practical scenarios, since decreasing the probability of false alarm by using a constant additional noise $\boldsymbol{n}$ does not usually result in an increase in the probability of detection. In fact, if there exists a noise component $\tilde{\boldsymbol{n}}$ such that $F_0(\tilde{\boldsymbol{n}}) \leq F_0(\mathbf{0})$ and $F_1(\tilde{\boldsymbol{n}}) > F_1(\mathbf{0})$, the detector can be improved simply by adding $\tilde{\boldsymbol{n}}$ to the original data, i.e., for $p_{\boldsymbol{N}}(\boldsymbol{x}) = \delta(\boldsymbol{x} - \tilde{\boldsymbol{n}})$. Therefore, the assumption in the proposition is in fact a necessary condition for non-improvability.

As an example application of Proposition 1, consider a hypothesis-testing problem in which $\mathcal{H}_0$ is represented by a zero-mean Gaussian distribution with variance σ^2 and $\mathcal{H}_1$ by a Gaussian distribution with mean $\mu > 0$ and variance

[3] It is shown in [17] and [1] that the optimal noise PDF is in the form of $p_{\boldsymbol{N}}(\boldsymbol{x}) = \lambda\,\delta(\boldsymbol{x} - \boldsymbol{n}_1) + (1-\lambda)\delta(\boldsymbol{x} - \boldsymbol{n}_2)$. Hence, it would be sufficient to perform the proof for $\mathrm{E}\{F(\boldsymbol{N})\} = \lambda\,F(\boldsymbol{n}_1) + (1-\lambda)F(\boldsymbol{n}_2)$, although we provide a more generic one.

σ^2. The decision rule selects $\mathcal{H}_1$ if $y \geq 0.5\mu$ and $\mathcal{H}_0$ otherwise. Let $\mathcal{S}_\mathrm{n} = (-0.5\mu, 0.5\mu)$ represent the set of additional noise values for possible performance improvement. From (4.7), F_0 and F_1 can be obtained as $F_0(x) = Q\left(\frac{0.5\mu - x}{\sigma}\right)$ and $F_1(x) = Q\left(\frac{-0.5\mu - x}{\sigma}\right)$. It is observed that F_0 is convex and F_1 is concave over $\mathcal{S}_\mathrm{n}$. Therefore, Proposition 1 implies that the detector is non-improvable.

Comparison of the non-improvability condition in Proposition 1 with that in [1], stated at the beginning of this section, reveals that the former provides a more direct way of evaluating the non-improvability since there is no need to obtain auxiliary functions, such as $\Psi(t)$ and $J(t)$ in (4.8). However, if $J(t)$ can be obtained easily, then the result in [1] can be more advantageous since it always deals with a function of a single variable irrespective of the dimension of the data vector. Therefore, for multi-dimensional measurements, the result in [1] can be preferred if the calculation of $J(t)$ in (4.8) is tractable.

In fact, even for multi-dimensional measurements, the problem can be considered as a one-dimensional problem in some cases if the measurement noise components are independent and identically distributed. Hence, the result of Proposition 1 can still be more advantageous in such scenarios.

4.3 Improvability Conditions

Based on the definition in (4.8), it is stated in [1] that the detector is improvable if $J(\mathrm{P}_\mathrm{F}^\mathbf{x}) > \mathrm{P}_\mathrm{D}^\mathbf{x}$ or $J''(\mathrm{P}_\mathrm{F}^\mathbf{x}) > 0$ when $J(t)$ is second-order continuously differentiable around $\mathrm{P}_\mathrm{F}^\mathbf{x}$.[4] Similar to the previous section, the aim is to obtain improvability conditions that directly depend on F_0 and F_1 in (4.7) instead of J in (4.8).

First, it can be observed from (4.5) and (4.6) that if there exists a noise component $\tilde{\boldsymbol{n}}$ such that $F_1(\tilde{\boldsymbol{n}}) > F_1(\mathbf{0})$ and $F_0(\tilde{\boldsymbol{n}}) \leq F_0(\mathbf{0})$, then the detector

[4] In this thesis, $J'(a)$ and $J''(a)$ are used to represent, respectively, the first and second derivatives of $J(t)$ at $t = a$.

can be improved by using $p_{\boldsymbol{N}}(\boldsymbol{x}) = \delta(\boldsymbol{x} - \tilde{\boldsymbol{n}})$. From (4.8), it is concluded that this result provides a generalization of the $J(\mathrm{P}_{\mathrm{F}}^{\mathbf{x}}) > \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}$ condition [1].

In practical scenarios, $F_0(\boldsymbol{n}) \leq F_0(\boldsymbol{0})$ commonly implies $F_1(\boldsymbol{n}) \leq F_1(\boldsymbol{0})$. Therefore, the previous result cannot be applied in many cases. Therefore, a more generic improvability condition is presented in the following proposition.

Proposition 2: *The detector is improvable if there exist $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ that satisfy*

$$\frac{[F_0(\boldsymbol{0}) - F_0(\boldsymbol{n}_2)][F_1(\boldsymbol{n}_1) - F_1(\boldsymbol{n}_2)]}{F_0(\boldsymbol{n}_1) - F_0(\boldsymbol{n}_2)} > F_1(\boldsymbol{0}) - F_1(\boldsymbol{n}_2)\ . \tag{4.11}$$

Proof: Consider additional noise $\boldsymbol{n}$ with $p_{\boldsymbol{N}}(\boldsymbol{x}) = \lambda\,\delta(\boldsymbol{x}-\boldsymbol{n}_1)+(1-\lambda)\,\delta(\boldsymbol{x}-\boldsymbol{n}_2)$. The detector is improvable if $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, and $\lambda \in [0,1]$ satisfy

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} = \mathrm{E}_{\boldsymbol{n}}\{F_0(\boldsymbol{n})\} = \lambda\,F_0(\boldsymbol{n}_1) + (1-\lambda)\,F_0(\boldsymbol{n}_2) \leq F_0(\boldsymbol{0})\ , \tag{4.12}$$

$$\mathrm{P}_{\mathrm{D}}^{\mathbf{y}} = \mathrm{E}_{\boldsymbol{n}}\{F_1(\boldsymbol{n})\} = \lambda\,F_1(\boldsymbol{n}_1) + (1-\lambda)\,F_1(\boldsymbol{n}_2) > F_1(\boldsymbol{0})\ . \tag{4.13}$$

Although $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} \leq F_0(\boldsymbol{0})$ is sufficient for improvability, the equality condition in (4.12), i.e., $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} = F_0(\boldsymbol{0})$, is satisfied in most practical cases. As studied in Theorem 4 in [1], $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}} < F_0(\boldsymbol{0})$ implies a trivial case in which the detector can be improved by using a constant noise value. Therefore, the equality condition in (4.12) can be considered, although it is not a necessary condition. Then, λ can be expressed as $\lambda = [F_0(\boldsymbol{0}) - F_0(\boldsymbol{n}_2)]/[F_0(\boldsymbol{n}_1) - F_0(\boldsymbol{n}_2)]$, which can be inserted in (4.13) to obtain (4.11). $\square$

Although the condition in Proposition 2 can directly be evaluated based on F_0 and F_1 functions in (4.7), finding suitable $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ values can be time consuming in some cases. In fact, it may not always be simpler to check the condition in Proposition 2 than to calculate the optimal noise PDF as in [1]. Therefore, more explicit and simpler improvability conditions are derived in the following.

Proposition 3: *Assume that $F_0(\boldsymbol{x})$ and $F_1(\boldsymbol{x})$ are second-order continuously differentiable around $\boldsymbol{x} = \mathbf{0}$. The detector is improvable if there exists a K-dimensional vector $\mathbf{z}$ such that $\sum_{i=1}^{K} z_i \frac{\partial F_j(\boldsymbol{x})}{\partial x_i} > 0$ for $j = 0, 1$ and*

$$\left(\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_1(\boldsymbol{x})}{\partial x_l \partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i}\right) > \left(\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_0(\boldsymbol{x})}{\partial x_l \partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i}\right) \tag{4.14}$$

are satisfied at $\boldsymbol{x} = \mathbf{0}$, where x_i and z_i represent the ith components of $\boldsymbol{x}$ and $\mathbf{z}$, respectively.

Proof: Consider the improvability conditions in (4.12) and (4.13) with infinitesimally small noise components, $\boldsymbol{n}_j = \boldsymbol{\epsilon}_j$ for $j = 1, 2$. Then, $F_i(\boldsymbol{\epsilon}_j)$ can be approximated by using the Taylor series expansion as $F_i(\mathbf{0}) + \boldsymbol{\epsilon}_j^T \mathbf{f}_i + 0.5\, \boldsymbol{\epsilon}_j^T \mathbf{H}_i \boldsymbol{\epsilon}_j$, where $\mathbf{H}_i$ and $\mathbf{f}_i$ are the Hessian and the gradient of $F_i(\boldsymbol{x})$ at $\boldsymbol{x} = \mathbf{0}$, respectively. Therefore, (4.12) and (4.13) require

$$\begin{aligned}
&\lambda\, \boldsymbol{\epsilon}_1^T \mathbf{H}_0 \boldsymbol{\epsilon}_1 + (1-\lambda)\boldsymbol{\epsilon}_2^T \mathbf{H}_0 \boldsymbol{\epsilon}_2 + 2[\lambda\, \boldsymbol{\epsilon}_1 + (1-\lambda)\boldsymbol{\epsilon}_2]^T \mathbf{f}_0 < 0 \ , \\
&\lambda\, \boldsymbol{\epsilon}_1^T \mathbf{H}_1 \boldsymbol{\epsilon}_1 + (1-\lambda)\boldsymbol{\epsilon}_2^T \mathbf{H}_1 \boldsymbol{\epsilon}_2 + 2[\lambda\, \boldsymbol{\epsilon}_1 + (1-\lambda)\boldsymbol{\epsilon}_2]^T \mathbf{f}_1 > 0 \ .
\end{aligned} \tag{4.15}$$

Let $\boldsymbol{\epsilon}_1 = \kappa\, \mathbf{z}$ and $\boldsymbol{\epsilon}_2 = \nu\, \mathbf{z}$, where κ and ν are infinitesimally small real numbers, and $\mathbf{z}$ is a K-dimensional real vector. Then, the conditions in (4.15) can be simplified, after some manipulation, as

$$\left[\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_0(\boldsymbol{x})}{\partial x_l \partial x_i} + c \sum_{i=1}^{K} z_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i}\right]\Bigg|_{\boldsymbol{x}=\mathbf{0}} < 0 \ , \tag{4.16}$$

$$\left[\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_1(\boldsymbol{x})}{\partial x_l \partial x_i} + c \sum_{i=1}^{K} z_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i}\right]\Bigg|_{\boldsymbol{x}=\mathbf{0}} > 0 \ . \tag{4.17}$$

where

$$c \doteq \frac{2\,[\lambda\,\kappa + (1-\lambda)\,\nu]}{\lambda\,\kappa^2 + (1-\lambda)\,\nu^2} \ . \tag{4.18}$$

Since $\sum_{i=1}^{K} z_i \frac{\partial F_j(\boldsymbol{x})}{\partial x_i} > 0$ at $\boldsymbol{x} = \boldsymbol{0}$ for $j = 0, 1$, (4.16) and (4.17) can also be expressed as

$$\Bigg[\left(\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_0(\boldsymbol{x})}{\partial x_l \partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i}\right) + c\left(\sum_{i=1}^{K} z_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i}\right)\Bigg]\Bigg|_{\boldsymbol{x}=\boldsymbol{0}} < 0\ , \tag{4.19}$$

$$\Bigg[\left(\sum_{l=1}^{K}\sum_{i=1}^{K} z_l z_i \frac{\partial^2 F_1(\boldsymbol{x})}{\partial x_l \partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i}\right) + c\left(\sum_{i=1}^{K} z_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i}\right)\left(\sum_{i=1}^{K} z_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i}\right)\Bigg]\Bigg|_{\boldsymbol{x}=\boldsymbol{0}} > 0\ . \tag{4.20}$$

It is noted from (4.18) that c can take any value in $(-\infty, \infty)$ by selecting appropriate $\lambda \in [0, 1]$ and infinitesimally small κ and ν values. Therefore, under the condition in (4.14), which states that the first term in (4.19) is smaller than the first term in (4.20), there always exists c that satisfies the improvability conditions in (4.19) and (4.20). $\square$

Note that Proposition 3 employs only the first and second derivatives of F_0 and F_1 without requiring the calculation of $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ as in Proposition 2. In [17], an improvability condition is obtained for scalar observations (i.e., for $K = 1$) based only on $\frac{\partial F_j(x)}{\partial x}$ and $\frac{\partial^2 F_j(x)}{\partial x^2}$ terms for $j = 0, 1$. Hence, Proposition 3 extends the improvability result in [17] not only to the case of vector observations but also to a more generic condition that involves partial derivatives ("interactions" among additional noise components), $\frac{\partial^2 F_j(x)}{\partial x_l x_i}$, as well.

Another improvability condition that depends directly on F_0 and F_1 is provided in the following proposition.

Proposition 4: *The detector is improvable if $F_1(\boldsymbol{x})$ and $-F_0(\boldsymbol{x})$ are strictly convex at $\boldsymbol{x} = \boldsymbol{0}$.*

Proof: Consider the improvability conditions in (4.15). Let $\boldsymbol{\epsilon}_1 = -\boldsymbol{\epsilon}_2 = \boldsymbol{\epsilon}$ and $\lambda = 0.5$. Then, (4.15) becomes

$$\boldsymbol{\epsilon}^T \mathbf{H}_0 \boldsymbol{\epsilon} < 0 \ , \quad \boldsymbol{\epsilon}^T \mathbf{H}_1 \boldsymbol{\epsilon} > 0 \ . \tag{4.21}$$

Since $F_1(\boldsymbol{x})$ is strictly convex and $F_0(\boldsymbol{x})$ is strictly concave at $\boldsymbol{x} = \mathbf{0}$, $\mathbf{H}_1$ is positive definite and $\mathbf{H}_0$ is negative definite. Hence, there exists $\boldsymbol{\epsilon}$ that guarantees improvability. □

Finally, an improvability condition that depends on the first-order partial derivatives of $F_0(\boldsymbol{x})$ and $F_1(\boldsymbol{x})$ is derived in the following proposition, which can be considered as an extension of the improvability condition in [17].

Proposition 5: *Assume that $F_0(\boldsymbol{x})$ and $F_1(\boldsymbol{x})$ are continuously differentiable around $\boldsymbol{x} = \mathbf{0}$. The detector is improvable if there exists a K-dimensional vector* $\mathbf{s}$ *such that*

$$\left(\sum_{i=1}^{K} s_i \frac{\partial F_1(\boldsymbol{x})}{\partial x_i} \right) \left(\sum_{i=1}^{K} s_i \frac{\partial F_0(\boldsymbol{x})}{\partial x_i} \right) < 0 \tag{4.22}$$

is satisfied at $\boldsymbol{x} = \mathbf{0}$, where s_i represents the ith component of $\mathbf{s}$.

Proof: Consider the improvability conditions in (4.15). Let $\boldsymbol{\epsilon}_1 = \varsigma\, \mathbf{s}_1$ and $\boldsymbol{\epsilon}_2 = \varsigma\, \mathbf{s}_2$ where $\mathbf{s}_1$ and $\mathbf{s}_2$ are any K-dimensional real vectors and ς is an infinitesimally small positive real number. Then, it can be shown that when

$$[\lambda\, \mathbf{s}_1 + (1-\lambda)\, \mathbf{s}_2]^T \mathbf{f}_0 < 0 \ \text{ and } \ [\lambda\, \mathbf{s}_1 + (1-\lambda)\, \mathbf{s}_2]^T \mathbf{f}_1 > 0 \tag{4.23}$$

are satisfied, one can find an infinitesimally small positive ς such that the conditions in (4.15) are satisfied. Let $\mathbf{s} \doteq \lambda\, \mathbf{s}_1 + (1-\lambda)\, \mathbf{s}_2$. Note that $\mathbf{s}$ can be any K-dimensional real vector for suitable values of $\mathbf{s}_1$, $\mathbf{s}_2$ and $\lambda \in [0,1]$. Based on the definition of $\mathbf{s}$, (4.23) can be expressed as $\mathbf{s}^T \mathbf{f}_0 < 0$ and $\mathbf{s}^T \mathbf{f}_1 > 0$.

For $\varsigma < 0$, similar argument can be used to show that $\mathbf{s}^T \mathbf{f}_0 > 0$ and $\mathbf{s}^T \mathbf{f}_1 < 0$ are sufficient conditions for improvability. Hence, $(\mathbf{s}^T \mathbf{f}_1)(\mathbf{s}^T \mathbf{f}_0) < 0$ can be obtained as the overall improvability condition. □

Comparison of the improvability conditions in this section with those in [1] reveals that the results in this section all depend on functions F_0 and F_1 in (4.7) directly, whereas those in [1] are obtained based on $J(t)$ defined in (4.8). Therefore, this study provides a direct way of evaluating the improvability of a detector. However, the approach in [1] can be more advantageous in certain cases, since it always deals with a single-variable function irrespective of the dimension of the data vector. Also, it is shown in the next section that under certain circumstances, the improvability condition in [1] is equivalent to that in Proposition 3.

4.4 Numerical Results

In this section, a binary hypothesis-testing problem is studied in order to provide an example of the results presented in the previous sections. The hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are defined as

$$\begin{aligned} \mathcal{H}_0 &: \boldsymbol{x} = \mathbf{w} \,, \\ \mathcal{H}_1 &: \boldsymbol{x} = A\mathbf{1} + \mathbf{w} \,, \end{aligned} \tag{4.24}$$

where $\boldsymbol{x} \in \mathbb{R}^2$, $\mathbf{1}$ denotes a vector of ones, $A > 0$ is a known scalar value, and $\mathbf{w}$ is Gaussian mixture noise with the following PDF

$$\begin{aligned} p_{\mathbf{w}}(\boldsymbol{x}) = \frac{1}{4\pi} \Bigg[& \frac{1}{|\boldsymbol{\Sigma}_1|^{0.5}} \exp\left(-\frac{1}{2}\,(\boldsymbol{x}+\boldsymbol{\mu})^T \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{x}+\boldsymbol{\mu})\right) \\ & + \frac{1}{|\boldsymbol{\Sigma}_2|^{0.5}} \exp\left(-\frac{1}{2}\,(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right) \Bigg] \,, \end{aligned} \tag{4.25}$$

where $\boldsymbol{\Sigma}_1 = \begin{bmatrix} \sigma^2 & \rho_1\sigma^2 \\ \rho_1\sigma^2 & \sigma^2 \end{bmatrix}$, $\boldsymbol{\Sigma}_2 = \begin{bmatrix} \sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \sigma^2 \end{bmatrix}$, $\boldsymbol{x} = [x_1 \;\; x_2]^T$, and $\boldsymbol{\mu} = [\mu_1 \;\; \mu_2]^T$. In addition, the detector is described by

$$\phi(\boldsymbol{y}) = \begin{cases} 1\,, & y_1 + y_2 \geq A/2 \\ 0\,, & y_1 + y_2 < A/2 \end{cases} \,, \tag{4.26}$$

where $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}$, with $\boldsymbol{n}$ representing the additional independent noise term.

Based on (4.25), $F_0(\boldsymbol{x})$ and $F_1(\boldsymbol{x})$ can be calculated as follows:

$$F_i(\boldsymbol{x}) = \frac{1}{2}\, Q\left(\frac{A/2 - x_1 - x_2 + \mu_1 + \mu_2 - s_i}{\sigma\sqrt{2(1+\rho_1)}}\right) + \frac{1}{2}\, Q\left(\frac{A/2 - x_1 - x_2 - \mu_1 - \mu_2 - s_i}{\sigma\sqrt{2(1+\rho_2)}}\right), \tag{4.27}$$

for $i = 0, 1$, where $s_0 = 0$, $s_1 = 2A$, and $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \mathrm{e}^{-t^2/2}dt$ denotes the Q-function. From (4.27), the first and second derivatives can be obtained as

$$\frac{\partial F_i(\boldsymbol{x})}{\partial x_1} = \frac{\partial F_i(\boldsymbol{x})}{\partial x_2} = \frac{1}{4\sqrt{\pi}\,\sigma}\left(\frac{1}{\sqrt{1+\rho_1}}\,\mathrm{e}^{-\frac{(A/2-\gamma_2-s_i)^2}{4\sigma^2(1+\rho_1)}} + \frac{1}{\sqrt{1+\rho_2}}\,\mathrm{e}^{-\frac{(A/2-\gamma_1-s_i)^2}{4\sigma^2(1+\rho_2)}}\right),$$

$$\frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_1^2} = \frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_2^2} = \frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_1 \partial x_2}$$
$$= \frac{\sigma^{-3}}{8\sqrt{\pi}}\left(\frac{(A/2-\gamma_2-s_i)}{\sqrt{(1+\rho_1)^3}}\,\mathrm{e}^{-\frac{(A/2-\gamma_2-s_i)^2}{4\sigma^2(1+\rho_1)}} + \frac{(A/2-\gamma_1-s_i)}{\sqrt{(1+\rho_2)^3}}\,\mathrm{e}^{-\frac{(A/2-\gamma_1-s_i)^2}{4\sigma^2(1+\rho_2)}}\right), \tag{4.28}$$

for $i = 0, 1$, where $\gamma_1 \doteq x_1 + x_2 + \mu_1 + \mu_2$ and $\gamma_2 \doteq x_1 + x_2 - \mu_1 - \mu_2$. It is noted from (4.28) that the first-order derivatives are always positive and all the first-order derivatives and the second-order derivatives are the same. Therefore, the improvability condition in (4.14) becomes independent of $\mathbf{z}$ for this example. Therefore, the improvability condition in Proposition 3 can be stated as *when* $g(\sigma) \doteq \left[\frac{\partial^2 F_1(\boldsymbol{x})}{\partial x_1^2}\frac{\partial F_0(\boldsymbol{x})}{\partial x_1} - \frac{\partial^2 F_0(\boldsymbol{x})}{\partial x_1^2}\frac{\partial F_1(\boldsymbol{x})}{\partial x_1}\right]\Big|_{\boldsymbol{x}=\mathbf{0}}$ *is positive, the detector is improvable.* Fig. 4.1 plots the *improvability function* $g(\sigma)$ for various values of A. It is observed that the detector performance can be improved for $A = 1$ if $\sigma \in [0.55, 3.24]$, for $A = 2$ if $\sigma \in [0.42, 3.09]$, for $A = 4$ if $\sigma \in [0.29, 2.38]$. On the other hand, when the more generic result in Proposition 2 is applied to the same example, it is obtained that the detector is improvable for $A = 1$ if $\sigma \le 3.24$, for $A = 2$ if $\sigma \le 3.14$, and for $A = 4$ if $\sigma \le 2.59$. Hence, Proposition 2 provides more generic improvability conditions as expected.

Fig. 4.2 plots the detection probabilities of the original (no additional noise) and the noise modified detectors with respect to σ for $A = 2$. From the figure, it

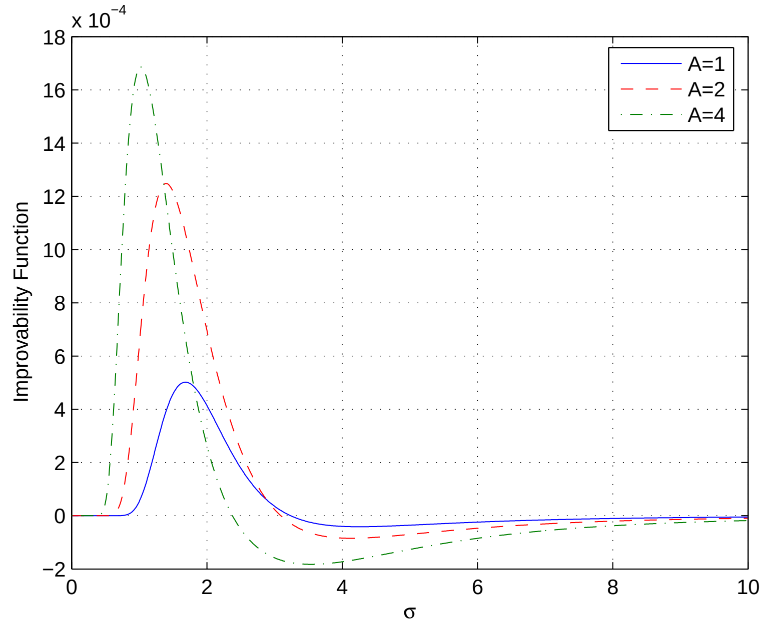

Figure 4.1: The improvability function obtained from Proposition 3 for various values of A, where $\rho_1 = 0.1$, $\rho_2 = 0.2$, $\mu_1 = 2$, and $\mu_2 = 3$.

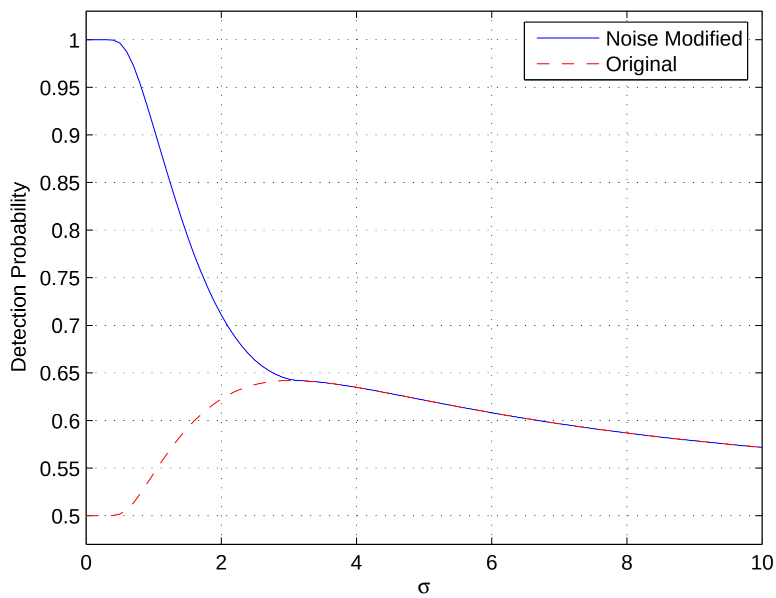

Figure 4.2: Detection probabilities of the original and noise modified detectors versus σ for $A = 2$, $\rho_1 = 0.1$, $\rho_2 = 0.2$, $\mu_1 = 2$, and $\mu_2 = 3$.

is observed that for smaller values of σ, more improvement is obtained, and after $\sigma = 3.14$ there is no improvement as expected from the improvability conditions.

In this specific example, it can be shown that the improvability conditions in Proposition 3 and in [1] are equivalent. Since the functions F_0 and F_1 defined in (4.27) are both monotone increasing functions of $x_1 + x_2$, $J(t) = \sup\{F_1(\boldsymbol{x}) \mid F_0(\boldsymbol{x}) = t\}$ can be obtained as $J(t) = \tilde{F}_1\left(\tilde{F}_0^{-1}(t)\right)$, where $\tilde{F}_i(m) \doteq F_i(\boldsymbol{x})\big|_{x_1+x_2=m}$. Then, $J''(t)$ can be obtained as

$$\begin{aligned} J''(t) &= \frac{d}{dt}\left\{\frac{\tilde{F}_1'\left(\tilde{F}_0^{-1}(t)\right)}{\tilde{F}_0'\left(\tilde{F}_0^{-1}(t)\right)}\right\} \\ &= \frac{\tilde{F}_1''\left(\tilde{F}_0^{-1}(t)\right) - \tilde{F}_1'\left(\tilde{F}_0^{-1}(t)\right)\tilde{F}_0''\left(\tilde{F}_0^{-1}(t)\right)/\tilde{F}_0'\left(\tilde{F}_0^{-1}(t)\right)}{\left[\tilde{F}_0'\left(\tilde{F}_0^{-1}(t)\right)\right]^2}\,. \end{aligned} \tag{4.29}$$

At $t = \mathrm{P}_{\mathrm{F}}^{\mathbf{x}} = F_0(\mathbf{0}) = \tilde{F}_0(0)$, $\tilde{F}_0^{-1}(t)$ becomes equal to 0; hence, $J''(\mathrm{P}_{\mathrm{F}}^{\mathbf{x}}) > 0$ implies $\tilde{F}_1''(0) - \tilde{F}_0''(0)\tilde{F}_1'(0)/\tilde{F}_0'(0) > 0$. For this specific problem, it can be shown that $\left.\frac{d\tilde{F}_i(m)}{dm}\right|_{m=0} = \left.\frac{\partial F_i(\boldsymbol{x})}{\partial x_1}\right|_{\boldsymbol{x}=\mathbf{0}} = \left.\frac{\partial F_i(\boldsymbol{x})}{\partial x_2}\right|_{\boldsymbol{x}=\mathbf{0}}$ and $\left.\frac{d^2\tilde{F}_i(m)}{dm^2}\right|_{m=0} = \left.\frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_1^2}\right|_{\boldsymbol{x}=\mathbf{0}} = \left.\frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_2^2}\right|_{\boldsymbol{x}=\mathbf{0}} = \left.\frac{\partial^2 F_i(\boldsymbol{x})}{\partial x_1 \partial x_2}\right|_{\boldsymbol{x}=\mathbf{0}}$ for $i = 0, 1$, and $\left.\frac{d\tilde{F}_0(m)}{dm}\right|_{m=0}$ is a positive constant. Therefore, the improvability conditions in Proposition 3 and that in [1] are equivalent in this specific example. However, it should be noted that the two conditions are not equivalent in general, and the calculation of $J(t)$ can be difficult in the absence of monotonicity properties related to F_0.

For the same measurement noise distribution, if we use a sign detector instead of the detector in (4.26), then, from the improvability function, it is obtained that the detector performance can be improved for $A = 1$ if $\sigma \in [0.57, 3.1628]$, for $A = 2$ if $\sigma \in [0.46, 2.7150]$, for $A = 3$ if $\sigma \in [0.43, 0.9203]$, as shown in Fig. 4.3. On the other hand, when the result in Proposition 2 is applied to this example, it is obtained that the detector is improvable for $A = 1$ if $\sigma \leq 3.20$, for $A = 2$ if $\sigma \leq 3.01$, and for $A = 3$ if $\sigma \leq 2.60$.

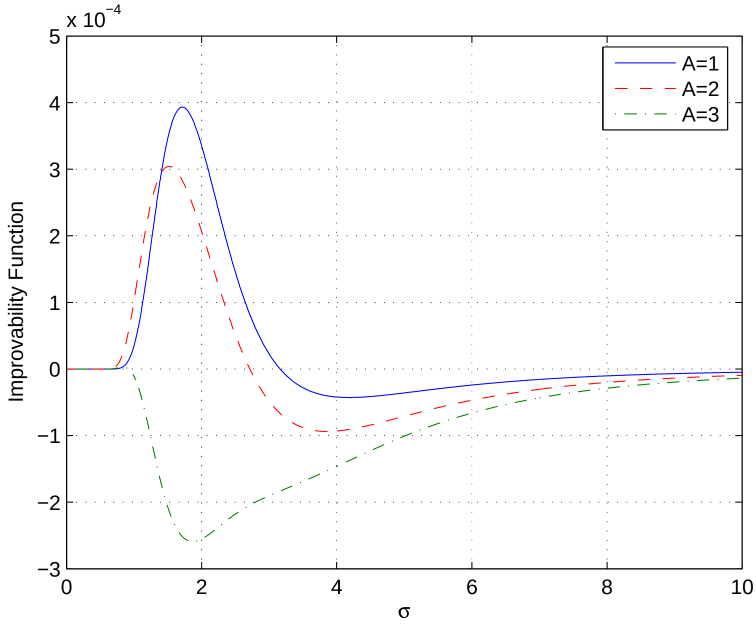

Figure 4.3: The improvability function obtained from Proposition 3 for various values of A in the case of sign detector where $\rho_1 = 0.1$, $\rho_2 = 0.2$, $\mu_1 = 2$, $\mu_2 = 3$.

4.5 Concluding Remarks

In this chapter, improvability and non-improvability conditions have been proposed to specify when detection performance of a suboptimal detector can be improved via additional noise under a constraint on probability of false alarm. The proposed results are defined in terms of the probabilities of detection and false alarm for specific additional noise values (cf. (4.7)) without the need for any other auxiliary functions as in [1]. However, for multi-dimensional measurements with dependent noise, the conditions in [1] can still be advantageous in some cases if the calculation of the auxiliary function in (4.8) is not challenging. In addition, the improvability results in [17] have been extended to both more generic conditions and to multi-dimensional measurements. All in all, this study has provided new improvability and non-improvability conditions that can be useful in various scenarios.

Bibliography

[1] 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. Sig. Processing*, vol. 55, pp. 3172–3184, July 2007.

[2] H. Chen, P. K. Varshney, S. Kay, and J. H. Michels, "Noise enhanced nonparametric detection," *IEEE Trans. Inform. Theory*, vol. 55, pp. 499–506, Feb. 2009.

[3] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A: Math. General*, vol. 14, pp. 453–457, 1981.

[4] F. Chapeau-Blondeau and D. Rousseau, "Raising the noise to improve performance in optimal processing," *Journal of Statistical Mechanics: Theory and Experiment*, pp. 1–15, Jan. 2009.

[5] P. Hanggi, M. E. Inchiosa, D. Fogliatti, and A. R. Bulsara, "Nonlinear stochastic resonance: The saga of anomalous output-input gain," *Physical Review E*, vol. 62, pp. 6155–6163, Nov. 2000.

[6] V. Galdi, V. Pierro, and I. M. Pinto, "Evaluation of stochastic-resonance-based detectors of weak harmonic signals in additive white gaussian noise," *Physical Review E*, vol. 57, pp. 6470–6479, June 1998.

[7] P. Makra and Z. Gingl, "Signal-to-noise ratio gain in non-dynamical and dynamical bistable stochastic resonators," *Fluctuat. Noise Lett.*, vol. 2, no. 3, pp. L145–L153, 2002.

[8] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, pp. 223–287, Jan. 1998.

[9] G. P. Harmer, B. R. Davis, and D. Abbott, "A review of stochastic resonance: Circuits and measurement," *IEEE Trans. Instrum. Meas*, vol. 51, pp. 299–309, Apr. 2002.

[10] K. Loerincz, Z. Gingl, and L. Kiss, "A stochastic resonator is able to greatly improve signal-to-noise ratio," *Phys. Lett. A*, vol. 224, pp. 63–67, 1996.

[11] I. Goychuk and P. Hanggi, "Stochastic resonance in ion channels characterized by information theory," *Phys. Rev. E*, vol. 61, no. 4, pp. 4272–4280, 2000.

[12] S. Mitaim and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Trans. Neural Netw.*, vol. 15, pp. 1526–1540, Nov. 2004.

[13] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, pp. 2310–2313, Mar. 2000.

[14] X. Godivier and F. Chapeau-Blondeau, "Stochastic resonance in the information capacity of a nonlinear dynamic system," *Int. J. Bifurc. Chaos*, vol. 8, no. 3, pp. 581–589, 1998.

[15] B. Kosko and S. Mitaim, "Stochastic resonance in noisy threshold neurons," *Neural Netw.*, vol. 16, pp. 755–761, 2003.

[16] B. Kosko and S. Mitaim, "Robust stochastic resonance for simple threshold neurons," *Phys. Rev. E*, vol. 70, no. 031911, 2004.

[17] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, pp. 1655–1669, May 2009.

[18] S. M. Kay, J. H. Michels, H. Chen, and P. K. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Sig. Processing Lett.*, vol. 13, pp. 695–698, Nov. 2006.

[19] S. Fauve and F. Heslot, "Stochastic resonance in a bistable system," *Phys. Lett.*, vol. 97A, pp. 5–7, Aug. 1983.

[20] M. A. Richards, *Fundamentals of Radar Signal Processing.* USA: McGraw-Hill, Electronic Engineering Series, 2005.

[21] H. V. Poor, *An Introduction to Signal Detection and Estimation.* New York: Springer-Verlag, 1994.

[22] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 4th ed., 2001.

[23] S. Verdu, *Multiuser Detection.* 1st ed. Cambridge, UK: Cambridge University Press, 1998.

[24] T. Erseghe and S. Tomasin, "Optimized demodulation for MAI resilient UWB W-PAN receivers," in *Proc. IEEE Int. Conf. Commun. (ICC)*, (Beijing, China), pp. 4867–4871, May 2008.

[25] T. Erseghe, V. Cellini, and G. Dona, "On UWB impulse radio receivers derived by modeling MAI as a Gaussian mixture process," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2388–2396, June 2008.

[26] M. McGuire, "Location of mobile terminals with quantized measurements," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Commun. (PIMRC)*, vol. 3, (Berlin, Germany), pp. 2045–2049, Sep. 2005.

[27] E. Ekrem, M. Koca, and H. Delic, "Robust ultra-wideband signal acquisition," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4656–4669, Nov. 2008.

[28] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Upper Saddle River, NJ: Prentice Hall, Inc., 1998.

[29] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part II–Variable detectors," *IEEE Trans. Sig. Processing*, vol. 56, pp. 5031–5041, Oct. 2007.

[30] R. Pradeepa and G. V. Anand, "Estimation of signals in colored non Gaussian noise based on Gaussian mixture models," in *Proc. IEEE Nonlinear Statistical Signal Processing Workshop*, vol. 1, (Cambridge, UK), pp. 17–20, Sep. 2006.

[31] S. M. Kay, "Can detectability be improved by adding noise?," *IEEE Sig. Processing Lett.*, vol. 7, pp. 8–10, Jan. 2000.

[32] S. Bayram and S. Gezici, "Effects of stochastic resonance on composite hypothesis-testing," *submitted to IEEE Trans. Sig. Processing*, March 2009. Available: http://www.ee.bilkent.edu.tr/∼gezici/TSP09.pdf.

[33] A. I. F. Vaz and E. M. G. P. Fernandes, "Optimization of nonlinear constrained particle swarm," *Baltic Journal on Sustainability*, vol. 12, no. 1, pp. 30–36, 2006.

[34] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, UK: Cambridge University Press, 2004.

[35] S. M. Kay, "Noise enhanced detection as a special case of randomization," *IEEE Sig. Processing Lett.*, vol. 15, pp. 709–712, 2008.