NOISE ENHANCED DETECTION IN RESTRICTED NEYMAN-PEARSON FRAMEWORK

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

By
Şan Gültekin

July, 2013

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.

Assoc. Prof. Dr. Sinan Gezici (Advisor)

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.

Assoc. Prof. Dr. Savaş Dayanık

Approved for the Graduate School of Engineering and Science:

Prof. Dr. Levent Onural
Director of the Graduate School

ABSTRACT

NOISE ENHANCED DETECTION IN RESTRICTED NEYMAN-PEARSON FRAMEWORK

Şan Gültekin
M.S. in Electrical and Electronics Engineering
Supervisor: Assoc. Prof. Dr. Sinan Gezici
July, 2013

Hypothesis tests frequently arise in many different engineering problems. Among the most frequently used tests are Bayesian, minimax, and Neyman-Pearson. Even though these tests are capable of addressing many real-life problems, they can be insufficient in certain scenarios. For this reason, developing new hypothesis tests is an important objective. One such developed test is the restricted Neyman-Pearson test, where one tries to maximize the average detection probability while keeping the worst-case detection and false-alarm probabilities bounded.

Finding the best hypothesis testing approach for a problem-at-hand is an important point. Another important one is to employ a detector with an acceptable performance. In particular, if the employed detector is suboptimal, it is crucial that it meets the performance requirements. Previous research has proven that performance of some suboptimal detectors can be improved by adding noise to their inputs, which is known as *noise enhancement*.

In this thesis we investigate noise enhancement according to the restricted Neyman-Pearson framework. To that aim, we formulate an optimization problem for optimal additive noise. Then, generic improvability and nonimprovability conditions are derived, which specify if additive noise can result in performance improvements. We then analyze the special case in which the parameter space is discrete and finite, and show that the optimal noise probability density function is discrete with a certain number of point masses. The improvability results are also extended and more precise conditions are derived. Finally, a numerical example is provided which illustrates the theoretical results and shows the benefits of applying noise enhancement to a suboptimal detector.

Keywords: Detection, hypothesis-testing, Neyman-Pearson, noise enhanced detection.

ÖZET

KISITLANDIRILMIŞ NEYMAN-PEARSON ÇERÇEVESİNDE GÜRÜLTÜ İYİLEŞTİRMELİ SEZİM

Şan Gültekin
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Yöneticisi: Doç. Dr. Sinan Gezici
Temmuz, 2013

Hipotez testleri pek çok mühendislik probleminin önemli bir bileşenidir. Sık kullanılan testler arasında *Bayesian*, *minimax* ve *Neyman-Pearson* testleri yer alır. Her ne kadar bu testler pek çok problemin modellenmesi için yeterli olsalar da, bazı senaryolarda yetersiz kalabilmektedirler. Bu sebeple, yeni hipotez testleri geliştirmek önemli bir araştırma konusudur. Bu şekilde geliştirilmiş testlerden biri de kısıtlandırılmış Neyman-Pearson testidir; bu testteki amaç ortalama sezim olasılığını en yüksek düzeye çıkarırken en kötü sezim ve yanlış alarm olasılıklarını sınırlamaktır.

Verilen bir problem için en iyi testi seçebilmek önemlidir. Bir o kadar önemli olan diğer bir nokta ise yüksek performanslı bir detektör seçebilmektir. Bilhassa, eğer seçilen detektör optimal değilse, onun performans kriterlerini sağlayıp sağlamadığını tespit etmek son derece önemlidir. Yapılan araştırmalar optimal olmayan bazı detektörlerin girişine gürültü eklenerek performanslarının artırılabileceğini ortaya koymuştur. Buna *gürültü iyileştirmesi* denmektedir.

Bu tezde kısıtlandırılmış Neyman-Pearson çerçevesinde gürültü iyileştirmesi incelenmektedir. Bunun için öncelikle optimal ek gürültü için bir optimizasyon problemi formüle edilmektedir. Bu formülasyonu takiben iyileştirilebilirlik ve iyileştirilemezlik için yeter koşullar sunulmaktadır ki bu koşullar ek gürültünün performans gelişimi sağlayıp sağlamayacağını belirlemektedir. Bundan sonra parametre uzayının ayrık ve sonlu olduğu bir özel durum incelenmektedir. Bu incelemede optimal gürültü olasılık dağılımı fonksiyonunun ayrık ve belirli bir sayıda kütle noktasından oluştuğu ortaya konmaktadır. Önceden elde edilmiş olan iyileştirilebilirlik sonuçları da buraya uyarlanmaktadır. En son olarak da, elde edilmiş olan kuramsal sonuçları örnekleyen ve optimal olmayan detektörlere

gürültü iyileştirmesi uygulamanın faydalarını gösteren bir sayısal örneğe yer verilmektedir.

Anahtar sözcükler: Sezim, hipotez testi, Neyman-Pearson, gürültü iyileştirilmeli sezim.

Acknowledgement

I would like to thank Assoc. Prof. Dr. Sinan Gezici for his supervision throughout this study. It was a great experience for me to work with him. I am also indebted to Prof. Dr. Orhan Arıkan and Assoc. Prof. Dr. Savaş Dayanık for their support as my thesis committee members.

Also I would like to thank all of my friends for enhancing my Bilkent experience and making the thesis process fun.

Special thanks to TÜBİTAK for their prestigious fellowship support which made this study possible.

Finally I would like to thank my parents and grandparents for their support throughout my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Objectives and Contributions of the Thesis

Within the detection theory framework, one can specify three main hypothesis testing approaches: Bayesian, minimax, and Neyman-Pearson [1]. The Bayesian approach can be employed when the prior probabilities of hypotheses as well as the costs of decisions under all scenarios are known. On the other hand, minimax approach can be used if the prior probabilities of hypotheses are unknown but the costs are known. Finally, the Neyman-Pearson approach is a common choice if neither the prior probabilities nor the costs are known. All of these approaches are applicable to both simple and composite hypothesis testing problems.

Bayesian, minimax, and Neyman-Pearson tests are optimum for certain scenarios. Yet, there are many cases in which the actual scenario does not correspond to one of these optimality scenarios. For example, the costs may be known but there can be some uncertainty on the knowledge of the prior probabilities. In that case, the Bayesian approach will be too optimistic whereas the minimax approach will be too conservative. In such cases the need for alternative hypothesis testing criteria arises. For example, for the case above, there are alternative methods such as the restricted Bayesian approach [2], which aims to compromise the Bayesian and minimax approaches.

In [3], the restricted Neyman-Pearson (RNP) approach is investigated. RNP is similar to the restricted Bayesian approach in the sense that, this time, the prior distribution of the parameter in the composite hypothesis testing framework is assumed to be known with some uncertainty. In the composite Neyman-Pearson hypothesis testing framework, both of the null and alternative hypotheses can be composite. For the null hypothesis, the general approach is to apply a false alarm constraint for all values of the parameter. For the alternative hypothesis, however, there are a number of methods to employ. Two of these methods are as follows: In the first one, the average detection probability is maximized, which is called the max-mean approach. In the second one, the minimum of the detection probabilities is maximized, which is called the max-min approach. Note that the former assumes perfect knowledge of the prior distribution of the parameter whereas the latter assumes no such knowledge. In this respect, the former and latter are similar to the Bayesian and minimax approaches respectively. What the RNP approach does is to compromise the max-mean and max-min approaches. In other words, in the RNP framework, the aim is to maximize the average detection probability under constraints on the worst-case detection and false-alarm probabilities [2, 3]. It is worthwhile to note that since this approach uses the available prior distribution information to a degree, it encompasses both probabilistic and nonprobabilistic descriptions of uncertainty.

Recently performance improvements that can be obtained via "noise" have been investigated for various problems in the literature ([4]-[19]). Although increasing noise levels or injecting additive noise to a system usually results in degraded performance, it can also lead to performance enhancements in some cases [20]-[30]. Enhancements obtained via noise can, for instance, be in the form of increased signal-to-noise ratio (SNR), mutual information or detection probability, or in the form of reduced average probability of error [5]-[18].

In hypothesis-testing problems, additive noise can be used to improve performance of a suboptimal detector according to Bayesian, minimax, and Neyman-Pearson criteria. In [20], the Bayesian criterion is considered under uniform cost assignment, and it is shown that the optimal noise that minimizes the probability of decision error has a constant value. The study in [17] obtains optimal additive noise for suboptimal variable detectors according to the Bayesian and minimax criteria based on the results in [20] and [16]. In [32], noise enhanced M-ary composite hypothesis-testing is studied in the presence of partial prior information, and optimal additive noise is investigated according to average and worst-case Bayes risk criteria. In [31], noise enhanced hypothesis-testing is investigated in the restricted Bayesian framework, which generalizes the Bayesian and minimax criteria and cover them as special cases [2, 33].

In the Neyman-Pearson framework, additive noise can be utilized to increase detection probability of a suboptimal detector under a constraint on false-alarm probability [18]. In [21], an example is provided to illustrate improvements in detection probability due to additive independent noise for the problem of detecting a constant signal in Gaussian mixture noise. A theoretical framework is established in [16] for noise enhanced hypothesis-testing according to the Neyman-Pearson criterion, and sufficient conditions are obtained for improvability and nonimprovability of a suboptimal detector via additive noise. In addition, it is shown that optimal additive noise can be realized by a randomization between at most two different signal levels. Noise enhanced detection in the Neyman-Pearson framework is studied also in [18], which provides an optimization theoretic framework, and proves the two point mass structure of the optimal additive noise probability distribution.

Noise benefits are also studied for *composite* hypothesis-testing problems. Such problems are encountered in various scenarios such as radar systems, noncoherent communications receivers, and spectrum sensing in cognitive radio networks [1]-[35]. Noise enhanced hypothesis-testing is investigated for composite hypothesis-testing problems according to the Bayesian, Neyman-Pearson, and Restricted Bayesian criteria in [31, 32, 36]. However, no studies have considered the noise enhanced hypothesis-testing problem according to RNP criterion.

In this thesis, noise enhancement is investigated for composite hypothesis-testing problems according to the RNP criterion [37]. A formulation is provided for obtaining the probability distribution of the optimal additive noise in the RNP framework. Also, sufficient conditions of improvability and nonimprovability are derived in order to determine when the use of additive noise can or cannot improve performance of a given detector. In addition, a special case in which there exist finitely many possible values of the unknown parameter under each hypothesis is considered, and the optimal additive noise is shown to correspond to a discrete random variable with a certain number of point masses in that scenario. Furthermore, particular improvability conditions are derived for that special case. Finally, numerical examples are presented in order to illustrate the improvements obtained via additive noise and to provide applications of the improvability conditions. Since a generic composite hypothesis-testing problem with prior distribution uncertainty is investigated in this thesis, the results can be considered to generalize the previous studies in the literature [16, 18, 36].

1.2 Organization of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, the noise enhanced hypothesis-testing problem is formulated according to the RNP approach. Chapter 3 presents the theoretical results; in particular, it presents improvability and nonimprovability conditions as well as the special case of discrete and finite parameter space. The results here are illustrated and supported by the numerical evaluations in Chapter 4. Finally, concluding remarks are made in Section 5.

Chapter 2

Problem Formulation

Consider a binary composite hypothesis-testing problem formulated as

$$\mathcal{H}_0 \;:\; p_\theta^{\mathbf{X}}(\mathbf{x})\,,\;\; \theta \in \Lambda_0\;, \qquad \mathcal{H}_1 \;:\; p_\theta^{\mathbf{X}}(\mathbf{x})\,,\;\; \theta \in \Lambda_1 \tag{2.1}$$

where $p_\theta^{\mathbf{X}}(\cdot)$ denotes the probability density function (PDF) of observation $\mathbf{X}$ for a given value of the parameter, $\Theta = \theta$, the observation (measurement), $\mathbf{x}$, is a K-dimensional vector (i.e., $\mathbf{x} \in \mathbb{R}^K$), and Λ_i is the set of possible parameter values under $\mathcal{H}_i$ for $i = 0, 1$ [1]. Parameter sets Λ_0 and Λ_1 are disjoint, and their union forms the parameter space Λ; that is, $\Lambda = \Lambda_0 \cup \Lambda_1$.

In this thesis, we consider a practical scenario in which there exists *imperfect* prior information about the parameter. In particular, we assume that the prior probability distribution of the parameter under each hypothesis is known with some *uncertainty* [38]. Let $w_0(\theta)$ and $w_1(\theta)$ represent the *imperfect* prior probability distributions of parameter θ under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. These probability distributions may differ from the true prior probability distributions, which are not known by the designer. For instance, $w_0(\theta)$ and $w_1(\theta)$ can be obtained via estimation based on previous decisions (experience). Then, uncertainty is related to estimation errors, and higher amount of uncertainty is observed as estimation errors increase [3].

For theoretical analysis, we consider a generic decision rule (detector), which is expressed as

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

for $i = 0, 1$, where Γ_0 and Γ_1 form a partition of the observation space Γ. The aim in this thesis is to investigate the effects of adding independent "noise" to inputs of given generic detectors as in (2.2) and to obtain optimal probability distributions of such additive "noise" in the restricted NP framework. As investigated in recent studies such as [5, 31, 17, 21, 16, 18], addition of independent noise to observations can improve detection performance of suboptimal detectors in some cases.

Let $\mathbf{n}$ denote the "noise" component that is added to original observation $\mathbf{x}$. Then, the noise modified observation is formed as $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ has a p.d.f. denoted by $p_{\mathbf{N}}(\cdot)$. The detector in (2.2) uses the noise modified observation $\mathbf{y}$ in order to make a decision. As in [31, 16, 18], we assume that the detector in (2.2) is fixed, and that the only way of enhancing the performance of the detector is to optimize the additive noise component, $\mathbf{n}$.

According to the RNP criterion [2, 3], the optimal additive noise should maximize the average detection probability under constraints on the worst-case detection and false-alarm probabilities. Therefore, the probability distribution of the optimal additive noise can be obtained from the solution of the following optimization problem:

$$\begin{aligned} \max_{p_{\mathbf{N}}(\cdot)} \;\; & \int_{\Lambda_1} \mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta)\, w_1(\theta)\, d\theta \\ \text{subject to} \;\; & \mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta) \geq \beta, \quad \forall \theta \in \Lambda_1 \\ & \mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta) \leq \alpha, \quad \forall \theta \in \Lambda_0 \end{aligned} \tag{2.3}$$

where $\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta)$ and $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta)$ denote respectively the detection and false-alarm probabilities of a given decision rule ϕ, which employs the noise modified observation $\mathbf{y}$, for a given value of $\Theta = \theta$, β is the lower limit on the worst-case detection probability, α is the false-alarm constraint, and $w_1(\theta)$ is the imperfect prior distribution of the parameter under hypothesis $\mathcal{H}_1$. The objective function in (2.3) corresponds to the average detection probability based on the imperfect prior distribution; that is, $\int_{\Lambda_1} \mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta)\, w_1(\theta)\, d\theta = \mathrm{E}\{\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\Theta)\} \triangleq \mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi)$. In addition, $\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta)$ and $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta)$ can be expressed as

$$\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta) = \mathrm{E}\left\{\phi(\boldsymbol{Y}) \,|\, \Theta = \theta\right\} = \int_{\Gamma} \phi(\mathbf{y})\, p_{\theta}^{\boldsymbol{Y}}(\mathbf{y})\, d\mathbf{y} \qquad \theta \in \Lambda_1 \tag{2.4}$$

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta) = \mathrm{E}\left\{\phi(\boldsymbol{Y}) \,|\, \Theta = \theta\right\} = \int_{\Gamma} \phi(\mathbf{y})\, p_{\theta}^{\boldsymbol{Y}}(\mathbf{y})\, d\mathbf{y} \qquad \theta \in \Lambda_0 \tag{2.5}$$

where $p_{\theta}^{\boldsymbol{Y}}(\cdot)$ is the PDF of the noise modified observation for a given value of $\Theta = \theta$.

In order to express the optimization problem in (2.3) more explicitly, we first manipulate $\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta)$ in (2.4) as follows:

$$\begin{aligned}
\mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta) &= \int_{\Gamma}\int_{\mathbb{R}^K} \phi(\mathbf{y})\, p_{\theta}^{\mathbf{X}}(\mathbf{y}-\mathbf{n})\, p_{\mathbf{N}}(\mathbf{n})\, d\mathbf{n}\, d\mathbf{y} &(2.6)\\
&= \int_{\mathbb{R}^K} p_{\mathbf{N}}(\mathbf{n}) \left[\int_{\Gamma} \phi(\mathbf{y}) p_{\theta}^{\mathbf{X}}(\mathbf{y}-\mathbf{n})\, d\mathbf{y}\right] d\mathbf{n} &(2.7)\\
&\triangleq \int_{\mathbb{R}^K} p_{\mathbf{N}}(\mathbf{n})\, F_{\theta}(\mathbf{n})\, d\mathbf{n} &(2.8)\\
&= \mathrm{E}\{F_{\theta}(\mathbf{N})\} &(2.9)
\end{aligned}$$

for $\theta \in \Lambda_1$, where the independence of $\mathbf{X}$ and $\mathbf{N}$ is used to obtain (2.6) from (2.4), and F_θ is defined as

$$F_{\theta}(\mathbf{n}) \triangleq \int_{\Gamma} \phi(\mathbf{y})\, p_{\theta}^{\mathbf{X}}(\mathbf{y}-\mathbf{n})\, d\mathbf{y}. \tag{2.10}$$

Note that $F_{\theta}(\mathbf{n})$ corresponds to the detection probability for a given value of $\theta \in \Lambda_1$ and for a constant value of additive noise, $\mathbf{N} = \mathbf{n}$. Therefore, for $\mathbf{n} = \mathbf{0}$, $F_{\theta}(\mathbf{0}) = \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}(\phi;\theta)$ is obtained; that is, $F_{\theta}(\mathbf{0})$ is equal to the detection probability of the decision rule for a given value of $\theta \in \Lambda_1$ and for the original observation $\mathbf{x}$.

Based on similar manipulations as in (2.6)-(2.9), $\mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta)$ in (2.5) can be expressed as

$$\mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta) = \mathrm{E}\{G_{\theta}(\mathbf{N})\} \tag{2.11}$$

for $\theta \in \Lambda_0$, where

$$G_\theta(\mathbf{n}) \triangleq \int_\Gamma \phi(\mathbf{y})\, p_\theta^{\mathbf{X}}(\mathbf{y}-\mathbf{n})\, d\mathbf{y}. \tag{2.12}$$

Note that $G_\theta(\mathbf{n})$ defines the false alarm probability for a given value of $\theta \in \Lambda_0$ and for a constant value of additive noise, $\mathbf{N} = \mathbf{n}$. Hence, for $\mathbf{n} = \mathbf{0}$, $G_\theta(\mathbf{0}) = \mathrm{P}_{\mathrm{F}}^{\mathbf{x}}(\phi; \theta)$ is obtained; that is, $G_\theta(\mathbf{0})$ is equal to the false alarm probability of the decision rule for a given value of $\theta \in \Lambda_0$ and for the original observation $\mathbf{x}$.

From (2.9) and (2.11), the optimization problem in (2.3) can be reformulated as

$$\begin{aligned} &\max_{p_{\mathbf{N}}(\cdot)} \int_{\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\}\, w_1(\theta)\, d\theta \\ \text{subject to } &\min_{\theta\in\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\} \geq \beta \\ &\max_{\theta\in\Lambda_0} \mathrm{E}\{G_\theta(\mathbf{N})\} \leq \alpha \end{aligned} \tag{2.13}$$

In addition, based on the following definition,

$$F(\mathbf{n}) \triangleq \int_{\Lambda_1} F_\theta(\mathbf{n}) w_1(\theta)\, d\theta\ , \tag{2.14}$$

the optimization problem in (2.13) can be expressed in the following simpler form:

$$\begin{aligned} &\max_{p_{\mathbf{N}}(\cdot)} \mathrm{E}\{F(\mathbf{N})\}, \\ \text{subject to } &\min_{\theta\in\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\} \geq \beta \\ &\max_{\theta\in\Lambda_0} \mathrm{E}\{G_\theta(\mathbf{N})\} \leq \alpha. \end{aligned} \tag{2.15}$$

Based on the definitions in (2.10) and (2.14), it is noted that $F(\mathbf{0}) = \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}(\phi)$; that is, $F(\mathbf{0})$ is equal to the average detection probability for the original observation $\mathbf{x}$ (i.e., the average detection probability in the absence of additive noise).

The exact solution of the optimization problem in (2.15) is very difficult to obtain in general as it requires a search over all possible additive noise PDFs.

Hence, an approximate solution can be obtained based on the Parzen window density estimation technique [31, 36, 39]. In particular, the additive noise PDF can be parameterized as

$$p_{\mathbf{N}}(\mathbf{n}) \approx \sum_{l=1}^{L} \mu_l \, \varphi_l(\mathbf{n}) \tag{2.16}$$

where $\mu_l \geq 0$, $\sum_{l=1}^{L} \mu_l = 1$, and $\varphi_l(\cdot)$ is a window function that satisfies $\varphi_l(\mathbf{x}) \geq 0$ $\forall \mathbf{x}$ and $\int \varphi_l(\mathbf{x}) d\mathbf{x} = 1$, for $l = 1, \ldots, L$. A common window function is the Gaussian window, for which $\varphi_l(\mathbf{n})$ is given by the PDF of a Gaussian random vector with a certain mean vector and a covariance matrix. Based on (2.16), the optimization problem in (2.15) can be solved over a number of parameters instead of PDFs, which significantly reduces the computational complexity. However, even in that case, the problem is nonconvex in general; hence, global optimization algorithms such as particle swarm optimization (PSO) need to be used [31, 40].

Chapter 3

Theoretical Results

3.1 Improvability and Nonimprovability Conditions

Since the optimization problem in (2.15) is complex to solve in general, it can be useful to determine *beforehand* if additive noise can or cannot improve the performance of a given detector. For that purpose, we obtain sufficient conditions for which the use of additive noise can or cannot provide any performance improvements compared to the case of not employing any additive noise. To that aim, we first define *improvability* and *nonimprovability* in the RNP framework as follows:

Definition 1: *According to the RNP criterion, a detector is called improvable if there exists additive noise* $\mathbf{N}$ *such that* $\mathrm{E}\{F(\mathbf{N})\} > \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}(\phi) = F(\mathbf{0})$ *and* $\min_{\theta\in\Lambda_1} \mathrm{P}_{\mathrm{D}}^{\mathbf{y}}(\phi;\theta) = \min_{\theta\in\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\} \geq \beta$*, and* $\max_{\theta\in\Lambda_0} \mathrm{P}_{\mathrm{F}}^{\mathbf{y}}(\phi;\theta) = \max_{\theta\in\Lambda_0} \mathrm{E}\{G_\theta(\mathbf{N})\} \leq \alpha$*. Otherwise, the detector is called nonimprovable.*

In other words, for improvability of a detector, there must exist additive noise that increases the average detection probability under the worst-case detection and false-alarm constraints.

According to Definition 1, we first obtain the following nonimprovability condition based on the properties of $F_\theta(\cdot)$ in (2.10), $G_\theta(\cdot)$ in (2.12), and $F(\cdot)$ in (2.14).

Proposition 1: *Assume that there exits $\theta^* \in \Lambda_0$ ($\theta^* \in \Lambda_1$) such that $G_{\theta^*}(\mathbf{n}) \leq \alpha$ ($F_{\theta^*}(\mathbf{n}) \geq \beta$) implies $F(\mathbf{n}) \leq F(\mathbf{0})$ for all $\mathbf{n} \in \mathcal{S}_\mathrm{n}$, where $\mathcal{S}_\mathrm{n}$ is a convex set*[1] *consisting of all possible values of additive noise $\mathbf{n}$. If $G_{\theta^*}(\mathbf{n})$ is a convex function ($F_{\theta^*}(\mathbf{n})$ is a concave function), and $F(\mathbf{n})$ is a concave function over $\mathcal{S}_\mathrm{n}$, then the detector is nonimprovable.*

Proof: The proof is similar to those in [31] and [22]. The convexity of $G_{\theta^*}(\cdot)$ implies that the false alarm probability in (2.9) is bounded, via Jensen's inequality, as

$$\mathrm{P}_\mathrm{F}^{\mathbf{y}}(\phi;\theta^*) = \mathrm{E}\{G_{\theta^*}(\mathbf{N})\} \geq G_{\theta^*}\left(\mathrm{E}\{\mathbf{N}\}\right). \tag{3.1}$$

As $\mathrm{P}_\mathrm{F}^{\mathbf{y}}(\phi;\theta^*) \leq \alpha$ must hold for improvability, (3.1) requires that $G_{\theta^*}\left(\mathrm{E}\{\mathbf{N}\}\right) \leq \alpha$ must be satisfied. Since $\mathrm{E}\{\mathbf{N}\} \in \mathcal{S}_\mathrm{n}$, $G_{\theta^*}\left(\mathrm{E}\{\mathbf{N}\}\right) \leq \alpha$ implies that $F\left(\mathrm{E}\{\mathbf{N}\}\right) \leq F(\mathbf{0})$ due to the assumption in the proposition. Hence,

$$\mathrm{P}_\mathrm{D}^{\mathbf{y}}(\phi) = \mathrm{E}\{F(\mathbf{N})\} \leq F\left(\mathrm{E}\{\mathbf{N}\}\right) \leq F(\mathbf{0}), \tag{3.2}$$

where the first inequality results from the concavity of F. Then, from (3.1) and (3.2), it is concluded that whenever the false-alarm constraint is satisfied, the average detection probability can never be higher than that in the absence of additive noise; that is, $\mathrm{P}_\mathrm{F}^{\mathbf{y}}(\phi;\theta^*) \leq \alpha$ implies $\mathrm{P}_\mathrm{D}^{\mathbf{y}}(\phi;\theta^*) \leq F(\mathbf{0}) = \mathrm{P}_\mathrm{D}^{\mathbf{x}}(\phi)$. For this reason, the detector is nonimprovable. Based on similar arguments, the alternative nonimprovability condition in terms of F_θ (stated in the parentheses in the proposition) can be proven as well. $\square$

The nonimprovability conditions in Proposition 1 can be useful in determining when it is unnecessary to solve the optimization problem in (2.15). When these conditions are satisfied, additive noise should not be employed in the system at all

[1]It is reasonable to model $\mathcal{S}_\mathrm{n}$ as a convex set since convex combination of individual noise components can be obtained via randomization [31, 41].

since it cannot provide any performance improvements according to the restricted NP criterion.

In addition to the nonimprovability conditions in Proposition 1, we obtain sufficient conditions for improvability in the remainder of this section. Assume that $F(\mathbf{x})$, $F_\theta(\mathbf{x})\ \forall\,\theta \in \Lambda_1$, and $G_\theta(\mathbf{x})\ \forall\,\theta \in \Lambda_0$ are second-order continuously differentiable around $\mathbf{x} = \mathbf{0}$. Then, we define the following functions for notational convenience:

$$g_\theta^{(1)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T \nabla G_\theta(\mathbf{x}) \tag{3.3}$$

$$f_\theta^{(1)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T \nabla F_\theta(\mathbf{x}) \tag{3.4}$$

$$f^{(1)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T \nabla F(\mathbf{x}) \tag{3.5}$$

$$g_\theta^{(2)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T H(G_\theta(\mathbf{x}))\,\mathbf{z} \tag{3.6}$$

$$f_\theta^{(2)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T H(F_\theta(\mathbf{x}))\,\mathbf{z} \tag{3.7}$$

$$f^{(2)}(\mathbf{x},\mathbf{z}) \triangleq \mathbf{z}^T H(F(\mathbf{x}))\,\mathbf{z} \tag{3.8}$$

where ∇ and H represent the Gradient and Hessian operators, respectively. For example, $\nabla G_\theta(\mathbf{x})$ is a K-dimensional column vector with its ith element being equal to $\frac{\partial G_\theta(\mathbf{x})}{\partial x_i}$, where x_i denotes the ith component of $\mathbf{x}$, and $H(G_\theta(\mathbf{x}))$ is a $K \times K$ matrix with its element in row l and column i being given by $\frac{\partial^2 G_\theta(\mathbf{x})}{\partial x_l \partial x_i}$.

Based on the preceding definitions, the following proposition provides sufficient conditions for improvability.

Proposition 2: *Let $\mathcal{L}_0$ and $\mathcal{L}_1$ denote the sets of θ values that maximize $G_\theta(\mathbf{0})$ and minimize $F_\theta(\mathbf{0})$, respectively. Then the detector is improvable if there exists a K-dimensional vector $\mathbf{z}$ such that one of the following conditions is satisfied for all $\theta_0 \in \mathcal{L}_0$ and $\theta_1 \in \mathcal{L}_1$:*

- $f^{(1)}(\mathbf{x},\mathbf{z}) > 0$, $f_{\theta_1}^{(1)}(\mathbf{x},\mathbf{z}) > 0$, and $g_{\theta_0}^{(1)}(\mathbf{x},\mathbf{z}) < 0$ at $\mathbf{x} = \mathbf{0}$.

- $f^{(1)}(\mathbf{x},\mathbf{z}) < 0$, $f_{\theta_1}^{(1)}(\mathbf{x},\mathbf{z}) < 0$, and $g_{\theta_0}^{(1)}(\mathbf{x},\mathbf{z}) > 0$ at $\mathbf{x} = \mathbf{0}$.

- $f^{(2)}(\mathbf{x},\mathbf{z}) > 0$, $f_{\theta_1}^{(2)}(\mathbf{x},\mathbf{z}) > 0$, and $g_{\theta_0}^{(2)}(\mathbf{x},\mathbf{z}) < 0$ at $\mathbf{x} = \mathbf{0}$.

Proof: For the improvability of a detector in the RNP framework, there must exist a noise PDF $p_{\mathbf{N}}(\mathbf{n})$ that satisfies $\mathrm{E}\{F(\mathbf{N})\} > F(\mathbf{0})$, $\min_{\theta\in\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\} \geq \beta$, and $\max_{\theta\in\Lambda_0} \mathrm{E}\{G_\theta(\mathbf{N})\} \leq \alpha$, which can be expressed as $\int_{\mathbb{R}^K} p_{\mathbf{N}}(\mathbf{n})\, F(\mathbf{n})\, d\mathbf{n} > F(\mathbf{0})$, $\int_{\mathbb{R}^K} p_{\mathbf{N}}(\mathbf{n})\, F_\theta(\mathbf{n})\, d\mathbf{n} \geq \beta$, $\forall\theta \in \Lambda_1$, and $\int_{\mathbb{R}^K} p_{\mathbf{N}}(\mathbf{n})\, G_\theta(\mathbf{n})\, d\mathbf{n} \leq \alpha$, $\forall\theta \in \Lambda_0$. Employing a similar approach to that in the proof of Theorem 2 in [31], we consider a noise PDF with L infinitesimal noise components, $p_{\mathbf{N}}(\mathbf{n}) = \sum_{j=1}^{L} \lambda_j\, \delta(\mathbf{n} - \boldsymbol{\epsilon}_j)$. Then, the conditions above become

$$\sum_{j=1}^{L} \lambda_j\, F(\boldsymbol{\epsilon}_j) > F(\mathbf{0}) \ , \quad \sum_{j=1}^{L} \lambda_j\, F_\theta(\boldsymbol{\epsilon}_j) \geq \beta \ , \ \forall\theta \in \Lambda_1 \ , \quad \sum_{j=1}^{L} \lambda_j\, G_\theta(\boldsymbol{\epsilon}_j) \leq \alpha \ , \ \forall\theta \in \Lambda_0 \ . \tag{3.9}$$

As $\boldsymbol{\epsilon}_j$'s are infinitesimally small, $F(\boldsymbol{\epsilon}_j)$, $F_\theta(\boldsymbol{\epsilon}_j)$, and $G_\theta(\boldsymbol{\epsilon}_j)$ can be approximated via the Taylor series expansion as $F(\mathbf{0}) + \boldsymbol{\epsilon}_j^T \mathbf{f} + 0.5\, \boldsymbol{\epsilon}_j^T \mathbf{H} \boldsymbol{\epsilon}_j$, $F_\theta(\mathbf{0}) + \boldsymbol{\epsilon}_j^T \mathbf{f}_\theta + 0.5\, \boldsymbol{\epsilon}_j^T \mathbf{H}_\theta^{\mathbf{f}} \boldsymbol{\epsilon}_j$, and $G_\theta(\mathbf{0}) + \boldsymbol{\epsilon}_j^T \mathbf{g}_\theta + 0.5\, \boldsymbol{\epsilon}_j^T \mathbf{H}_\theta^{\mathbf{g}} \boldsymbol{\epsilon}_j$, respectively, where $\mathbf{f}$ ($\mathbf{f}_\theta$, $\mathbf{g}_\theta$) and $\mathbf{H}$ ($\mathbf{H}_\theta^{\mathbf{f}}$, $\mathbf{H}_\theta^{\mathbf{g}}$) are the Gradient and Hessian of $F(\mathbf{x})$ ($F_\theta(\mathbf{x})$, $G_\theta(\mathbf{x})$) at $\mathbf{x} = \mathbf{0}$, respectively. Hence, (3.9) leads to

$$\begin{aligned}
\sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{H} \boldsymbol{\epsilon}_j + 2 \sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{f} > 0 \ , \\
\sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{H}_\theta^{\mathbf{f}} \boldsymbol{\epsilon}_j + 2 \sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{f}_\theta \geq 2\,(\beta - F_\theta(\mathbf{0})) \ \ , \ \forall\theta \in \Lambda_1 \ , \\
\sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{H}_\theta^{\mathbf{g}} \boldsymbol{\epsilon}_j + 2 \sum_{j=1}^{L} \lambda_j\, \boldsymbol{\epsilon}_j^T \mathbf{g}_\theta \leq 2\,(\alpha - G_\theta(\mathbf{0})) \ \ , \ \forall\theta \in \Lambda_0 \ .
\end{aligned} \tag{3.10}$$

Express $\boldsymbol{\epsilon}_j$ as $\boldsymbol{\epsilon}_j = \rho_j\, \mathbf{z}$ for $j = 1, 2, \ldots, L$, where ρ_j for $j = 1, 2, \ldots, L$ are infinitesimal real numbers, and $\mathbf{z}$ is a K-dimensional real vector. Then, based on the definitions in (3.3)-(3.8), the conditions in (3.10) can be simplified to the following:

$$\left(f^{(2)}(\mathbf{x},\mathbf{z}) + c\, f^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0\ , \tag{3.11}$$

$$\left(f_\theta^{(2)}(\mathbf{x},\mathbf{z}) + c\, f_\theta^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > \frac{2\,(\beta - F_\theta(\mathbf{0}))}{\sum_{j=1}^{L}\lambda_j\,\rho_j^2}\ ,\ \forall\theta\in\Lambda_1\ , \tag{3.12}$$

$$\left(g_\theta^{(2)}(\mathbf{x},\mathbf{z}) + c\, g_\theta^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} < \frac{2\,(\alpha - G_\theta(\mathbf{0}))}{\sum_{j=1}^{L}\lambda_j\,\rho_j^2}\ ,\ \forall\theta\in\Lambda_0\ , \tag{3.13}$$

where $c \triangleq 2\sum_{j=1}^{L}\lambda_j\,\rho_j \big/ \sum_{j=1}^{L}\lambda_j\,\rho_j^2$. Because $\beta = F_\theta(\mathbf{0})$ for $\theta\in\mathcal{L}_1$ ($\alpha = G_\theta(\mathbf{0})$ for $\theta\in\mathcal{L}_0$) and $\beta < \min\limits_{\theta\in\Lambda_1\setminus\mathcal{L}_1} F_\theta(\mathbf{0})$ $\big(\alpha > \max\limits_{\theta\in\Lambda_0\setminus\mathcal{L}_0} G_\theta(\mathbf{0})\big)$, the right-hand-side of (3.12) ((3.13)) goes to minus infinity for $\{\theta\in\Lambda_1\,|\,\theta\notin\mathcal{L}_1\}$ (plus infinity for $\{\theta\in\Lambda_0\,|\,\theta\notin\mathcal{L}_0\}$). Hence, we should consider only the $\theta\in\mathcal{L}_1$ case for $\theta\in\Lambda_1$ and the $\theta\in\mathcal{L}_0$ case for $\theta\in\Lambda_0$. Thus, (3.11), (3.12), and (3.13) can be expressed as

$$\left(f^{(2)}(\mathbf{x},\mathbf{z}) + c\, f^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0 \tag{3.14}$$

$$\left(f_{\theta_1}^{(2)}(\mathbf{x},\mathbf{z}) + c\, f_{\theta_1}^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0 \tag{3.15}$$

$$\left(g_{\theta_0}^{(2)}(\mathbf{x},\mathbf{z}) + c\, g_{\theta_0}^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} < 0\ . \tag{3.16}$$

Note that c can take any real value by definition via the selection of appropriate λ_i and infinitesimal ρ_i values for $i = 1, 2, \ldots, L$. Then, based on (3.14)-(3.16), the following conclusions are made for the three bullets in the proposition:

- If the conditions in the first bullet of Proposition 2 are satisfied, c can be set to a sufficiently large positive number to satisfy the inequalities in (3.14)-(3.16).
- If the conditions in the second bullet of Proposition 2 are satisfied, c can be set to a sufficiently large negative number to satisfy the inequalities in (3.14)-(3.16).
- If the conditions in the first bullet of Proposition 2 are satisfied, c can be set to zero to satisfy the inequalities in (3.14)-(3.16). □

Proposition 2 implies that under the stated conditions, one can always find a noise PDF that increases the average detection probability under the constraints on the worst case detection and false alarm probabilities. In other words, the conditions in the proposition guarantee the existence of additive noise that improves the detection performance according to the RNP criterion.

In addition to the improvability conditions in Proposition 2, we can obtain alternative sufficient conditions for improvability based on the approaches in [31, 16]. For that purpose, we first define two new functions $J(t)$ and $H(t)$ as follows:

$$J(t) \triangleq \sup \left\{ F(\mathbf{n}) \mid \max_{\theta \in \Lambda_0} G_\theta(\mathbf{n}) = t \right\} \tag{3.17}$$

$$H(t) \triangleq \inf \left\{ \min_{\theta \in \Lambda_1} F_\theta(\mathbf{n}) \mid \max_{\theta \in \Lambda_0} G_\theta(\mathbf{n}) = t \right\} \tag{3.18}$$

which represent, respectively, the maximum average detection probability and the minimum worst-case detection probability for a given value of the maximum false-alarm probability considering constant values of additive noise. As an initial observation from (3.17) and (3.18), one can conclude that if there exists $t_0 \leq \alpha$ such that $J(t_0) > F(\mathbf{0})$ and $H(t_0) \geq \beta$, then the detector is improvable, since under such a condition there exists a noise component $\mathbf{n}_0$ that satisfies $F(\mathbf{n}_0) > F(\mathbf{0})$, $\min_{\theta \in \Lambda_1} F_\theta(\mathbf{n}_0) \geq \beta$ and $\max_{\theta \in \Lambda_0} G_\theta(\mathbf{n}_0) \leq \alpha$ (i.e., performance improvement can be achieved by adding a constant noise component $\mathbf{n}_0$ to the observation).

Since improvability of a detector via constant noise component is not very common in practice, the following improvability condition is presented for more practical scenarios.

Proposition 3: *Define the minimum value of the detection probability and the maximum value of the false alarm probability in the absence of additive noise as* $\tilde{\beta} \triangleq \min_{\theta \in \Lambda_1} \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}(\phi; \theta)$ *and* $\tilde{\alpha} \triangleq \max_{\theta \in \Lambda_0} \mathrm{P}_{\mathrm{F}}^{\mathbf{x}}(\phi; \theta)$, *respectively, where* $\tilde{\beta} \geq \beta$ *and* $\tilde{\alpha} \leq \alpha$. *Assume that* $H(\tilde{\alpha}) = \tilde{\beta}$, *where* H *is as defined in* (3.18). *Then the detector is improvable if* $J(t)$ *in* (3.17) *and* $H(t)$ *in* (3.18) *are second-order continuously differentiable around* $t = \tilde{\alpha}$, *and satisfy* $J''(\tilde{\alpha}) > 0$ *and* $H''(\tilde{\alpha}) \geq 0$.

Proof: As $J(t)$ in (3.17) and $H(t)$ in (3.18) are second-order continuously differentiable around $t = \tilde{\alpha}$, one can find $\epsilon > 0$, $\mathbf{n}_1$, and $\mathbf{n}_2$ such that $\max_{\theta\in\Lambda_0} G_\theta(\mathbf{n}_1) = \tilde{\alpha} + \epsilon$ and $\max_{\theta\in\Lambda_0} G_\theta(\mathbf{n}_2) = \tilde{\alpha} - \epsilon$ [31]. Then, in the following, it is proved that an additive noise component with $p_{\mathbf{N}}(\mathbf{n}) = 0.5\,\delta(\mathbf{x} - \mathbf{n}_1) + 0.5\,\delta(\mathbf{x} - \mathbf{n}_2)$ improves the detector performance according to the restricted NP criterion (i.e., under the worst-case detection and false alarm constraints). First, under the condition of $H''(\tilde{\alpha}) \geq 0$, the minimum value of the detection probability and the maximum value of the false alarm probability in the presence of additive noise are shown not to remain below β and exceed α, respectively:

$$\min_{\theta\in\Lambda_1} \mathrm{E}\{F_\theta(\mathbf{N})\} \geq \mathrm{E}\left\{\min_{\theta\in\Lambda_1} F_\theta(\mathbf{N})\right\} \geq 0.5H(\tilde{\alpha}+\epsilon) + 0.5H(\tilde{\alpha}-\epsilon) \geq H(\tilde{\alpha}) = \tilde{\beta} \geq \beta \tag{3.19}$$

$$\max_{\theta\in\Lambda_0} \mathrm{E}\{G_\theta(\mathbf{N})\} \leq \mathrm{E}\left\{\max_{\theta\in\Lambda_0} G_\theta(\mathbf{N})\right\} = 0.5(\tilde{\alpha}+\epsilon) + 0.5(\tilde{\alpha}-\epsilon) = \tilde{\alpha} \leq \alpha\ . \tag{3.20}$$

In addition, due to the assumptions in the proposition, $J(t)$ is convex in an interval around $t = \tilde{\alpha}$. As $\mathrm{E}\{F(\mathbf{N})\}$ can achieve the value of $0.5\,J(\tilde{\alpha}+\epsilon) + 0.5\,J(\tilde{\alpha}-\epsilon)$, which is always larger than $J(\tilde{\alpha})$ due to convexity, it is concluded that $\mathrm{E}\{F(\mathbf{N})\} > J(\tilde{\alpha})$. Since $J(\tilde{\alpha}) \geq F(\mathbf{0})$ by definition of $J(t)$ in (3.17), $\mathrm{E}\{F(\mathbf{N})\} > F(\mathbf{0})$ is satisfied. Therefore, the detector is improvable. $\square$

Proposition 3 can be employed in a similar manner to Proposition 2 in order to determine if a given detector is improvable according to the RNP framework. The main advantage of Proposition 3 is that $J(t)$ and $H(t)$ are always single-variable functions irrespective of the dimension of the observation vector, which facilitates simple evaluation of the conditions in the proposition. However, in some cases, it can be challenging to obtain an expression for $J(t)$ in (3.17) and $H(t)$ in (3.18). On the other hand, Proposition 2 deals directly with $G_\theta(\cdot)$, $F_\theta(\cdot)$, and $F(\cdot)$ without defining auxiliary functions as in Proposition 3; hence, it can be employed more efficiently in some cases. However, it should also be noted that the functions in Proposition 2 are always K-dimensional, which can make the evaluation of the conditions more complex than those in Proposition 3 in some other cases.

3.2 Special Case: Discrete and Finite Parameter Space

The results obtained in the previous section are generic in the sense that there are no specific restrictions on the parameter sets Λ_0 and Λ_1 corresponding to hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. In this section, we provide more detailed theoretical analysis for the special case in which the parameter sets consist of finitely many elements. Let $\Lambda_0 = \{\theta_{01}, \theta_{02}, \ldots, \theta_{0M}\}$ and $\Lambda_1 = \{\theta_{11}, \theta_{12}, \ldots, \theta_{1N}\}$.

The most important simplification in this case is that the optimal p.d.f. of additive noise can be represented by a discrete probability distribution with at most $M + N$ point masses under mild conditions as specified in the following proposition.

Proposition 4: *Suppose that each component of additive noise is upper and lower bounded by two finite values; that is, $n_j \in [a_j, b_j]$ for $j = 1, \ldots, K$ where a_j and b_j are finite.*[2] *If $F_\theta(\cdot)$ and $G_\theta(\cdot)$ are continuous functions for all θ in Λ_1 and Λ_0 respectively, then the PDF of an optimal additive noise can be expressed as*

$$p_{\mathbf{N}}(\mathbf{n}) = \sum_{l=1}^{M+N} \lambda_l \, \delta(\mathbf{n} - \mathbf{n}_l) \ , \tag{3.21}$$

where $\sum_{l=1}^{M+N} \lambda_l = 1$ and $\lambda_l \geq 0$ for $l = 1, 2, \ldots, M + N$.

Proof: The proof is omitted since it can be obtained similarly to the proofs of Theorem 4 in [31] and Theorem 8 in [36], which are based on the approach in [16]. □

Based on Proposition 4, the optimization problem in (2.15) can be expressed

[2]This is a reasonable assumption because additive noise cannot take infinitely large values in practice.

as

$$\begin{aligned}
&\max_{\{\lambda_l,\,\mathbf{n}_l\}_{l=1}^{M+N}} \; \sum_{l=1}^{M+N} \lambda_l\, F(\mathbf{n}_l) \\
&\text{subject to} \;\; \min_{\theta\in\Lambda_1} \sum_{l=1}^{M+N} \lambda_l\, F_\theta(\mathbf{n}_l) \geq \beta \\
&\qquad\qquad\;\; \max_{\theta\in\Lambda_0} \sum_{l=1}^{M+N} \lambda_l\, G_\theta(\mathbf{n}_l) \leq \alpha \\
&\qquad\qquad\;\; \sum_{l=1}^{M+N} \lambda_l = 1 \;, \quad \lambda_l \geq 0 \text{ for } l = 1,2,\ldots,M+N
\end{aligned} \tag{3.22}$$

Compared to (2.15), the optimization problem in (3.22) has much lower computational complexity in general since it requires optimization over a number of variables instead of over all possible PDFs. However, depending on the number of possible parameter values, $M+N$, the computational complexity can still be high in some cases.

Next, we obtain sufficient conditions for improvability according to the RNP criterion. Let $\mathcal{S}_\beta$ ($\mathcal{S}_\alpha$) denote the set of indices for which $F_{\theta_{1i}}(\mathbf{0})$ ($G_{\theta_{0i}}(\mathbf{0})$) achieves the minimum value of β (maximum value of α), and let $\bar{\mathcal{S}}_\beta$ ($\bar{\mathcal{S}}_\alpha$) represent the set of indices with $F_{\theta_{1i}}(\mathbf{0}) > \beta$ ($G_{\theta_{0i}}(\mathbf{0}) < \alpha$); that is,

$$\mathcal{S}_\beta = \{i \in \{1,2,\ldots,N\} \mid F_{\theta_{1i}}(\mathbf{0}) = \beta\} \tag{3.23}$$

$$\bar{\mathcal{S}}_\beta = \{i \in \{1,2,\ldots,N\} \mid F_{\theta_{1i}}(\mathbf{0}) > \beta\} \tag{3.24}$$

$$\mathcal{S}_\alpha = \{i \in \{1,2,\ldots,M\} \mid G_{\theta_{0i}}(\mathbf{0}) = \alpha\} \tag{3.25}$$

$$\bar{\mathcal{S}}_\alpha = \{i \in \{1,2,\ldots,M\} \mid G_{\theta_{0i}}(\mathbf{0}) < \alpha\}\,. \tag{3.26}$$

Note that $\mathcal{S}_\beta \cup \bar{\mathcal{S}}_\beta = \{1,2,\ldots,N\}$ ($\mathcal{S}_\alpha \cup \bar{\mathcal{S}}_\alpha = \{1,2,\ldots,M\}$); hence, $F_{\theta_{1i}}(\mathbf{0}) = \mathrm{P}_{\mathrm{D}}^{\mathbf{x}}(\phi;\theta_{1i}) \geq \beta$ for $i = 1,2,\ldots,N$ ($G_{\theta_{0i}}(\mathbf{0}) = \mathrm{P}_{\mathrm{F}}^{\mathbf{x}}(\phi;\theta_{0i}) \leq \alpha$ for $i = 1,2,\ldots,M$).

Based on the functions in (3.3)-(3.8), we define new functions as $f_i^{(n)}(\mathbf{x},\mathbf{z}) \triangleq f_{\theta_{1i}}^{(n)}(\mathbf{x},\mathbf{z})$ and $g_i^{(n)}(\mathbf{x},\mathbf{z}) \triangleq g_{\theta_{1i}}^{(n)}(\mathbf{x},\mathbf{z})$. Also let $\mathcal{F}_n$ and $\mathcal{G}_n$ ($n = 1,2$) represent the sets that consist of $f^{(n)}(\mathbf{x},\mathbf{z})$, $f_i^{(n)}(\mathbf{x},\mathbf{z})$ for $i \in \mathcal{S}_\beta$, and $g_i^{(n)}(\mathbf{x},\mathbf{z})$ for $i \in \mathcal{S}_\alpha$;

namely,

$$\mathcal{F}_n = \left\{ f^{(n)}(\mathbf{x},\mathbf{z}), f_i^{(n)}(\mathbf{x},\mathbf{z}) \text{ for } i \in \mathcal{S}_\beta \right\} \tag{3.27}$$

$$\mathcal{G}_n = \left\{ g_i^{(n)}(\mathbf{x},\mathbf{z}) \text{ for } i \in \mathcal{S}_\alpha \right\}, \tag{3.28}$$

for $n = 1, 2$. Note that $\mathcal{F}_n$ ($\mathcal{G}_n$) has $|\mathcal{S}_\beta| + 1$ ($|\mathcal{S}_\alpha|$) elements, where $|\mathcal{S}_\beta|$ ($|\mathcal{S}_\alpha|$) denotes the number of elements in $\mathcal{S}_\beta$ ($\mathcal{S}_\alpha$). Representing by $\mathcal{F}_n(j)$ ($\mathcal{G}_n(j)$) the jth element of $\mathcal{F}_n$ ($\mathcal{G}_n$), it is noted that $\mathcal{F}_n(1) = f^{(n)}(\mathbf{x},\mathbf{z})$ and $\mathcal{F}_n(j) = f^{(n)}_{\mathcal{S}_\beta(j-1)}(\mathbf{x},\mathbf{z})$ for $j = 2, \ldots, |\mathcal{S}_\beta| + 1$ ($\mathcal{G}_n(j) = g^{(n)}_{\mathcal{S}_\alpha(j)}(\mathbf{x},\mathbf{z})$ for $j = 2, \ldots, |\mathcal{S}_\alpha|$), where $\mathcal{S}_\beta(j-1)$ is the $(j-1)$th element of $\mathcal{S}_\beta$ ($\mathcal{S}_\alpha(j)$ is the jth element of $\mathcal{S}_\alpha$). Furthermore, the following sets are defined for the indices $j \in \mathcal{S}_\beta$ ($j \in \mathcal{S}_\alpha$) for which $\mathcal{F}_1(j)$ ($\mathcal{G}_1(j)$) is zero, negative or positive:

$$\mathcal{S}_\beta^z = \left\{ j \in \{1_\beta^z, 2_\beta^z, \ldots, (|\mathcal{S}_\beta| + 1)_\beta^z\} \mid \mathcal{F}_1(j) = 0 \right\} \tag{3.29}$$

$$\mathcal{S}_\beta^n = \left\{ j \in \{1_\beta^n, 2_\beta^n, \ldots, (|\mathcal{S}_\beta| + 1)_\beta^n\} \mid \mathcal{F}_1(j) < 0 \right\} \tag{3.30}$$

$$\mathcal{S}_\beta^p = \left\{ j \in \{1_\beta^p, 2_\beta^p, \ldots, (|\mathcal{S}_\beta| + 1)_\beta^p\} \mid \mathcal{F}_1(j) > 0 \right\} \tag{3.31}$$

$$\mathcal{S}_\alpha^z = \{ j \in \{1_\alpha^z, 2_\alpha^z, \ldots, (|\mathcal{S}_\alpha|)_\alpha^z\} \mid \mathcal{G}_1(j) = 0 \} \tag{3.32}$$

$$\mathcal{S}_\alpha^n = \{ j \in \{1_\alpha^n, 2_\alpha^n, \ldots, (|\mathcal{S}_\alpha|)_\alpha^n\} \mid \mathcal{G}_1(j) < 0 \} \tag{3.33}$$

$$\mathcal{S}_\alpha^p = \{ j \in \{1_\alpha^p, 2_\alpha^p, \ldots, (|\mathcal{S}_\alpha|)_\alpha^p\} \mid \mathcal{G}_1(j) > 0 \} \tag{3.34}$$

where we denote j as j_α (j_β) in order to emphasize that j is coming from set $\mathcal{S}_\alpha$ (is not coming from set $\mathcal{S}_\alpha$) and we use z, n, and p to denote the subsets.

In the following proposition, an indicator function $\mathcal{I}_A(x)$ is used, which is defined as $\mathcal{I}_A(x) = 1$ if $x \in A$ and $\mathcal{I}_A(x) = 0$ otherwise. Based on the definitions in (3.23)-(3.34), the following proposition provides sufficient conditions for improvability in the RNP framework.

Proposition 5: *When Λ consists of a finite number of elements, a detector is improvable according to the RNP criterion if there exists a K-dimensional vector* $\mathbf{z}$ *such that the following two conditions are satisfied at* $\mathbf{x} = \mathbf{0}$*:*

1. $\mathcal{F}_2(j) > 0$, $\forall j \in \mathcal{S}_\beta^z$ *and* $\mathcal{G}_2(j) < 0$, $\forall j \in \mathcal{S}_\alpha^z$.

2. *One of the following is satisfied:*

- *Any three of* $|\mathcal{S}_\beta^n|$, $|\mathcal{S}_\beta^p|$, $|\mathcal{S}_\alpha^n|$ *and* $|\mathcal{S}_\alpha^p|$ *is zero, or* $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^p|=0$*, or* $|\mathcal{S}_\alpha^n|+|\mathcal{S}_\beta^p|=0$.
- $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^n|$ *is an odd number,* $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^p|>0$, $|\mathcal{S}_\alpha^n|+|\mathcal{S}_\beta^p|>0$ *and*

$$\min_{j\in\mathcal{S}_\beta^n\cup\mathcal{S}_\alpha^p}\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^n}(j)+\mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^p}(j)\Big)\prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\backslash\{j\}}\Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l)+\mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)$$
$$>\max_{j\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n}\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^p}(j)+\mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^n}(j)\Big)\prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\backslash\{j\}}\Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l)+\mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)\,. \tag{3.35}$$

- $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^n|$ *is an even number,* $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^p|>0$, $|\mathcal{S}_\alpha^n|+|\mathcal{S}_\beta^p|>0$ *and*

$$\min_{j\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n}\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^p}(j)+\mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^n}(j)\Big)\prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\backslash\{j\}}\Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l)+\mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)$$
$$>\max_{j\in\mathcal{S}_\beta^n\cup\mathcal{S}_\alpha^p}\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^n}(j)+\mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^p}(j)\Big)\prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\backslash\{j\}}\Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l)+\mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)\,. \tag{3.36}$$

Proof: According to Proposition 4, the optimal additive noise has a discrete probability distribution with at most $M+N$ point masses. Then, a detector is improvable if there exists a noise PDF $p_{\mathbf{N}}(\mathbf{n})=\sum_{l=1}^{M+N}\lambda_l\,\delta(\mathbf{n}-\mathbf{n}_l)$that satisfies $\mathrm{E}\{F(\mathbf{N})\}>F(\mathbf{0})$, $\min\limits_{i\in\{1,2,\ldots,N\}}\mathrm{E}\{F_{\theta_{1i}}(\mathbf{N})\}\geq\beta$, and $\max\limits_{i\in\{1,2,\ldots,M\}}\mathrm{E}\{G_{\theta_{0i}}(\mathbf{N})\}\leq\alpha$, which can be stated as

$$\begin{aligned}\sum_{l=1}^{M}\lambda_l\,F(\mathbf{n}_l)&>F(\mathbf{0})\\ \min_{i\in\{1,2,\ldots,N\}}\sum_{l=1}^{M+N}\lambda_l\,F_{\theta_{1i}}(\mathbf{n}_l)&\geq\beta\\ \max_{i\in\{1,2,\ldots,M\}}\sum_{l=1}^{M+N}\lambda_l\,G_{\theta_{0i}}(\mathbf{n}_l)&\leq\alpha\ .\end{aligned} \tag{3.37}$$

Similarly to the approach in the proof of Proposition 2, consider the improvability conditions in (3.37) for infinitesimal noise components, $\mathbf{n}_l=\boldsymbol{\epsilon}_l=\rho_l\,\mathbf{z}$

for $l = 1, 2, \ldots, M+N$, where ρ_l's are infinitesimal real numbers, and $\mathbf{z}$ is a K-dimensional real vector. Then, based on similar manipulations, the following conditions are obtained:

$$\left(f^{(2)}(\mathbf{x},\mathbf{z}) + c\, f^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0 \tag{3.38}$$

$$\left(f_i^{(2)}(\mathbf{x},\mathbf{z}) + c\, f_i^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > \frac{2\,(\beta - F_{\theta_{1i}}(\mathbf{0}))}{\sum_{j=1}^{M} \lambda_j\, \rho_j^2}\ , \quad i = 1, 2, \ldots, N \tag{3.39}$$

$$\left(g_i^{(2)}(\mathbf{x},\mathbf{z}) + c\, g_i^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} < \frac{2\,(\alpha - G_{\theta_{0i}}(\mathbf{0}))}{\sum_{j=1}^{M} \lambda_j\, \rho_j^2}\ , \quad i = 1, 2, \ldots, M \tag{3.40}$$

where $c \triangleq 2\sum_{j=1}^{M} \lambda_j\, \rho_j \big/ \sum_{j=1}^{M} \lambda_j\, \rho_j^2$.

Because $F_{\theta_{1i}}(0) > \beta$, $\forall i \in \bar{\mathcal{S}}_\beta$ and $G_{\theta_{0i}}(0) < \alpha$, $\forall i \in \bar{\mathcal{S}}_\alpha$, the right-hand-side of (3.39) and (3.40) becomes minus infinity for $i \in \bar{\mathcal{S}}_\beta$ and plus infinity for $i \in \bar{\mathcal{S}}_\alpha$, respectively. Therefore, it is sufficient to consider $i \in \mathcal{S}_\beta$ and $i \in \mathcal{S}_\alpha$ only. Hence, (3.38)-(3.40) can be expressed as

$$\left(f^{(2)}(\mathbf{x},\mathbf{z}) + c\, f^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0 \tag{3.41}$$

$$\left(f_i^{(2)}(\mathbf{x},\mathbf{z}) + c\, f_i^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0,\ \ \forall i \in \mathcal{S}_\beta \tag{3.42}$$

$$\left(g_i^{(2)}(\mathbf{x},\mathbf{z}) + c\, g_i^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} < 0,\ \ \forall i \in \mathcal{S}_\alpha. \tag{3.43}$$

From the definitions in (3.27) and (3.28), (3.41)-(3.43) can be written as

$$\Big(\mathcal{F}_2(j) + c\,\mathcal{F}_1(j)\Big)\Big|_{\mathbf{x}=\mathbf{0}} > 0 \ \text{ for } \ j = 1, 2, \ldots, |\mathcal{S}_\beta| + 1 \tag{3.44}$$

$$\Big(\mathcal{G}_2(j) + c\,\mathcal{G}_1(j)\Big)\Big|_{\mathbf{x}=\mathbf{0}} < 0 \ \text{ for } \ j = 1, 2, \ldots, |\mathcal{S}_\alpha|\,. \tag{3.45}$$

It is again observed that c can take any real value by selecting appropriate λ_i and infinitesimal ρ_i values for $i = 1, 2, \ldots, M+N$. Therefore, from (3.29) and (3.32), it is concluded that for the conditions in (3.44) and (3.45) to hold,

$$\mathcal{F}_2(j)\,\Big|_{\mathbf{x}=\mathbf{0}} > 0 \ \ \forall j \in \mathcal{S}_\beta^z \ \text{ and } \ \mathcal{G}_2(j)\,\Big|_{\mathbf{x}=\mathbf{0}} < 0 \ \ \forall j \in \mathcal{S}_\alpha^z \tag{3.46}$$

must be satisfied, which is the first condition in the proposition.

In addition to (3.46), one of the following conditions must be satisfied for the improvability conditions in (3.44) and (3.45) to hold:

- When any three of $|\mathcal{S}_\beta^n|$, $|\mathcal{S}_\beta^p|$, $|\mathcal{S}_\alpha^n|$, and $|\mathcal{S}_\alpha^p|$ are zero, as stated in the first part of the second condition in Proposition 5, all the second terms that are nonzero in (3.44) and (3.45) are either all non-negative or all non-positive and the corresponding signs of the inequalities are the same. Therefore, there always exists a c that satisfies the improvability conditions in (3.44) and (3.45) when the first condition in Proposition 5 (cf. (3.46)) is satisfied.

  When $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^p| = 0$, as stated in the first part of the second condition in Proposition 5, assume that $|\mathcal{S}_\alpha^n|$ is an odd number (this does not reduce the generality of the result in the proposition). Then, (3.44) and (3.45) can be stated after some manipulations as

$$\mathcal{F}_2(j)\Big|_{\mathbf{x}=\mathbf{0}} > 0,\ \forall j \in \mathcal{S}_\beta^z \tag{3.47}$$

$$\mathcal{G}_2(j)\Big|_{\mathbf{x}=\mathbf{0}} < 0,\ \forall j \in \mathcal{S}_\alpha^z \tag{3.48}$$

$$\Bigg(\mathcal{F}_2(j) \prod_{l\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n}(l)\Big)$$
$$+c \prod_{l\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n}(l)\Big)\Bigg)\Bigg|_{\mathbf{x}=\mathbf{0}} < 0,\ \forall j \in \mathcal{S}_\beta^p \tag{3.49}$$

$$\Bigg(\mathcal{G}_2(j) \prod_{l\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n}(l)\Big)$$
$$+c \prod_{l\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n}(l)\Big)\Bigg)\Bigg|_{\mathbf{x}=\mathbf{0}} < 0,\ \forall j \in \mathcal{S}_\alpha^n. \tag{3.50}$$

  In obtaining (3.49) and (3.50), (3.44) and (3.45) are multiplied by $\prod_{l\in\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n}(l)\Big)$, which is a positive (negative) quantity when $j \in \mathcal{S}_\alpha^n$ ($j \in \mathcal{S}_\beta^p$) since $|\mathcal{S}_\alpha^n|$ is an odd number. The conditions in (3.47) and (3.48) are satisfied from the first condition in Proposition 5. Therefore, there always exists a c that satisfies the improvability conditions in (3.49) and (3.50) as the second terms and the sign of the inequalities in (3.49) and (3.50) are the same. When $|\mathcal{S}_\alpha^n|$ is an even number, only the sign of the inequalities (3.49) and (3.50) change; hence, the same result is valid as well.

When $|\mathcal{S}_\beta^p| + |\mathcal{S}_\alpha^n| = 0$, as stated in the first part of the second condition in Proposition 5, via similar manipulations as in the previous paragraph, it can be proved that the detector is improvable with the first condition in Proposition 5.

- When $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^n|$ is an odd number, $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^p| > 0$, $|\mathcal{S}_\alpha^n| + |\mathcal{S}_\beta^p| > 0$, (3.44) and (3.45) can be written as

$$\mathcal{F}_2(j)\Big|_{\mathbf{x}=\mathbf{0}} > 0,\ \forall j \in \mathcal{S}_\beta^z \tag{3.51}$$

$$\mathcal{G}_2(j)\Big|_{\mathbf{x}=\mathbf{0}} < 0,\ \forall j \in \mathcal{S}_\alpha^z \tag{3.52}$$

$$\Bigg(\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^n}(j) + \mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^p}(j))\Big) \prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)$$
$$+c \prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)\Bigg)\Bigg|_{\mathbf{x}=\mathbf{0}} > 0,\ \forall j \in \mathcal{S}_\beta^n \cup \mathcal{S}_\alpha^p \tag{3.53}$$

$$\Bigg(\Big(\mathcal{F}_2(j)\mathcal{I}_{\mathcal{S}_\beta^p}(j) + \mathcal{G}_2(j)\mathcal{I}_{\mathcal{S}_\alpha^n}(j))\Big) \prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)$$
$$+c \prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l) + \mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)\Bigg)\Bigg|_{\mathbf{x}=\mathbf{0}} < 0,\ \forall j \in \mathcal{S}_\beta^p \cup \mathcal{S}_\alpha^n\ . \tag{3.54}$$

In obtaining (3.53) and (3.54), (3.44) and (3.45) are multiplied by $\prod_{l\in\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p\cup\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p\setminus\{j\}} \Big(\mathcal{F}_1(l)\mathcal{I}_{\mathcal{S}_\beta^n\cup\mathcal{S}_\beta^p}(l)+\mathcal{G}_1(l)\mathcal{I}_{\mathcal{S}_\alpha^n\cup\mathcal{S}_\alpha^p}(l)\Big)$, which is a positive (negative) quantity when $j \in \mathcal{S}_\beta^n \cup \mathcal{S}_\alpha^n$ ($j \in \mathcal{S}_\beta^p \cup \mathcal{S}_\alpha^p$) since $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^n|$ is an odd number. The conditions in (3.51) and (3.52) are satisfied from the first condition in the proposition. Also, under the condition in (3.35), there always exists a c that satisfies the improvability conditions in (3.53) and (3.54).

- When $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^n|$ is an even number, $|\mathcal{S}_\beta^n| + |\mathcal{S}_\alpha^p| > 0$, and $|\mathcal{S}_\alpha^n| + |\mathcal{S}_\beta^p| > 0$ (3.44) and (3.45) can be expressed by four conditions similar to those in (3.51)-(3.54) with the only difference being that the signs of the inequalities in (3.53) and (3.54) are switched. In that scenario, the first and the second conditions are satisfied from the first condition in the proposition. In addition, under the condition in (3.36), there always exists a c that satisfies the third and the fourth conditions. $\square$

Whenever the two conditions in Proposition 5 are satisfied, it is guaranteed that the detection performance can be improved via additive noise. Although the expression in the proposition may seem complicated at first, it is noted that, after defining the sets in (3.23)-(3.34), it is simple to check the conditions. An example application of Proposition 5 is provided in the next section.

The following improvability condition can be obtained as a corollary of Proposition 5.

Corollary 1: *Assume that* $F(\mathbf{x})$, $F_{\theta_{1i}}(\mathbf{x})$, $i = 1, 2, \ldots, N$, *and* $G_{\theta_{0i}}(\mathbf{x})$, $i = 1, 2, \ldots, M$ *are second-order continuously differentiable around* $\mathbf{x} = \mathbf{0}$ *and that* $\min_{i\in\{1,2,\ldots,N\}} F_{\theta_{1i}}(\mathbf{0}) > \beta$ *and* $\max_{i\in\{1,2,\ldots,M\}} G_{\theta_{0i}}(\mathbf{0}) < \alpha$. *Let* $\mathbf{f}$ *denote the gradient of* $F(\mathbf{x})$ *at* $\mathbf{x} = \mathbf{0}$. *Then, the detector is improvable*

- *if* $\mathbf{f} \neq \mathbf{0}$*; or,*
- *if* $F(\mathbf{x})$ *is not concave around* $\mathbf{x} = \mathbf{0}$.

Proof: Because $\min_{i\in\{1,2,\ldots,N\}} F_{\theta_{1i}}(\mathbf{0}) > \beta$ and $\max_{i\in\{1,2,\ldots,M\}} G_{\theta_{0i}}(\mathbf{0}) < \alpha$, the right-hand-side of (3.39) and (3.40) in the proof of Proposition 5 become minus infinity and plus infinity for any i, respectively. Then, it is sufficient to consider the condition in (3.38) only; namely,

$$\left(f^{(2)}(\mathbf{x},\mathbf{z}) + c\,f^{(1)}(\mathbf{x},\mathbf{z})\right)\Big|_{\mathbf{x}=\mathbf{0}} > 0\ . \tag{3.55}$$

This condition can be expressed as $\mathbf{z}^T\mathbf{H}\mathbf{z} + c\,\mathbf{z}^T\mathbf{f} > 0$ in terms of the Gradient $\mathbf{f}$ and the Hessian $\mathbf{H}$ of $F(\mathbf{x})$ at $\mathbf{x} = \mathbf{0}$. As c can take any real value by definition as discussed before and as $\mathbf{z}$ can be chosen arbitrarily small, the improvability condition is always satisfied if $\mathbf{f} \neq \mathbf{0}$. On the other hand, if $\mathbf{f} = \mathbf{0}$, the improvability condition becomes $\mathbf{z}^T\mathbf{H}\mathbf{z} > 0$. In that case, if $F(\mathbf{x})$ is not concave around $\mathbf{x} = \mathbf{0}$, $\mathbf{H}$ is not negative semidefinite. Then, there exists $\mathbf{z}$ such that $\mathbf{z}^T\mathbf{H}\mathbf{z} > 0$ is satisfied. Therefore, the detector is improvable. $\square$

Chapter 4

Numerical Results

In this chapter, the binary hypothesis-testing problem considered in [3] is studied in order to illustrate theoretical results in the previous chapter. The hypotheses are specified as follows:

$$\mathcal{H}_0 \; : \; X = V \quad , \quad \mathcal{H}_1 \; : \; X = \Theta + V \tag{4.1}$$

where $X \in \mathbb{R}$, Θ is the unknown parameter, and V is symmetric Gaussian mixture noise that has the following PDF

$$p_V(v) = \sum_{i=1}^{N_m} \omega_i \, \psi_i(v - m_i) \ , \tag{4.2}$$

where $\omega_i \geq 0$ for $i = 1, \ldots, N_m$, $\sum_{i=1}^{N_m} \omega_i = 1$, and $\psi_i(x) = 1/(\sqrt{2\pi}\,\sigma_i) \, \exp\left(-x^2/(2\,\sigma_i^2)\right)$ for $i = 1, \ldots, N_m$. Since noise V is symmetric, its parameters satisfy $m_l = -m_{N_m-l+1}$, $\omega_l = \omega_{N_m-l+1}$ and $\sigma_l = \sigma_{N_m-l+1}$ for $l = 1, \ldots, \lfloor N_m/2 \rfloor$, where $\lfloor y \rfloor$ denotes the largest integer smaller than or equal to y. (If N_m is an odd number, $m_{(N_m+1)/2}$ is set to zero for symmetry.)

The unknown parameter Θ in (4.1) is modeled as a random variable with the following PDF.

$$w_1(\theta) = \rho\,\delta(\theta - A) + (1-\rho)\,\delta(\theta + A) \tag{4.3}$$

where A is a positive constant that is known exactly, whereas ρ is known with some uncertainty. (Please see [3] for the motivations of this model.)

Based on the preceding problem formulation, the parameter sets under $\mathcal{H}_0$ and $\mathcal{H}_1$ as specified as $\Lambda_0 = \{0\}$ and $\Lambda_1 = \{-A, A\}$, respectively. Also, the conditional PDF of the original observation X for a given value of $\Theta = \theta$ is obtained as

$$p_\theta^{\mathbf{X}}(x) = \sum_{i=1}^{N_m} \frac{\omega_i}{\sqrt{2\pi}\,\sigma_i} \exp\left(\frac{-(x-\theta-m_i)^2}{2\,\sigma_i^2}\right) . \tag{4.4}$$

Suppose that the following detector is employed.

$$\phi(y) = \begin{cases} 0\,, & A/2 > y > -A/2 \\ 1\,, & \text{otherwise} \end{cases} , \tag{4.5}$$

where $y = x+n$, with n representing the additive noise term. This is a reasonable detector for the model in (4.1) since noise V is zero mean, and Θ is either A of $-A$. Although it is not the optimal detector for the specified problem, it can be employed in practical scenarios due to its simplicity.

From (2.10), (2.12), and (2.14), $F_{\theta_{1i}}$ for $\theta_{11} = A$ and $\theta_{12} = -A$, $G_{\theta_{0i}}$ for

$\theta_{01} = 0$, and F can be calculated as follows:

$$
\begin{aligned}
F_A(n) &= \sum_{i=1}^{N_m} w_i \left(Q\left(\frac{-A/2 - m_i - n}{\sigma_i}\right) + Q\left(\frac{3A/2 + m_i + n}{\sigma_i}\right) \right) , \\
F_{-A}(n) &= \sum_{i=1}^{N_m} w_i \left(Q\left(\frac{3A/2 - m_i - n}{\sigma_i}\right) + Q\left(\frac{-A/2 + m_i + n}{\sigma_i}\right) \right) , \\
G_0(n) &= \sum_{i=1}^{N_m} w_i \left(Q\left(\frac{A/2 - m_i - n}{\sigma_i}\right) + Q\left(\frac{A/2 + m_i + n}{\sigma_i}\right) \right) , \\
F(n) &= \rho\, F_A(n) + (1-\rho)\, F_{-A}(n) \ ,
\end{aligned}
\tag{4.6}
$$

where $Q(x) = (1/\sqrt{2\pi}\,) \int_x^{\infty} \mathrm{e}^{-t^2/2} dt$ is the Q-function.

In the numerical example, $N_m = 4$ is considered for the symmetric Gaussian mixture noise, and the mean values of the Gaussian components in the mixture noise are specified as [0.01 0.6 −0.6 −0.01] with the corresponding weights of [0.25 0.25 0.25 0.25]. Also, the variances of the Gaussian components in the mixture noise are assumed to be the same; i.e., $\sigma_i = \sigma$ for $i = 1, \ldots, N_m$.

In Figures 4.1, 4.2, and 4.3, average detection probabilities are plotted with respect to σ for various values of β in the cases of $\alpha = 0.35$, $\alpha = 0.4$, and $\alpha = 0.45$, respectively, where $A = 1$ and $\rho = 0.8$. It is observed that the use of additive noise enhances the average detection probability, and significant improvements can be achieved via additive noise for low values of the standard deviation, σ. As the standard deviation increases, the amount of improvement in the average detection probability reduces. In fact, after some values of σ, the constraints on the minimum detection probability or the false alarm probability are not satisfied; hence, the RNP solution does not exist after certain values of σ. (Therefore, the curves are plotted up to those specific values in the figures.) Another observation from the figures is that the average detection probabilities decrease as β increases. This is expected since a larger value of β imposes a stricter constraint on the worst-case detection probability (see (2.3)), which in turn reduces the average detection probability. In other words, there is a tradeoff between β and the average detection probability, which is an essential characteristic of the RNP approach [3].

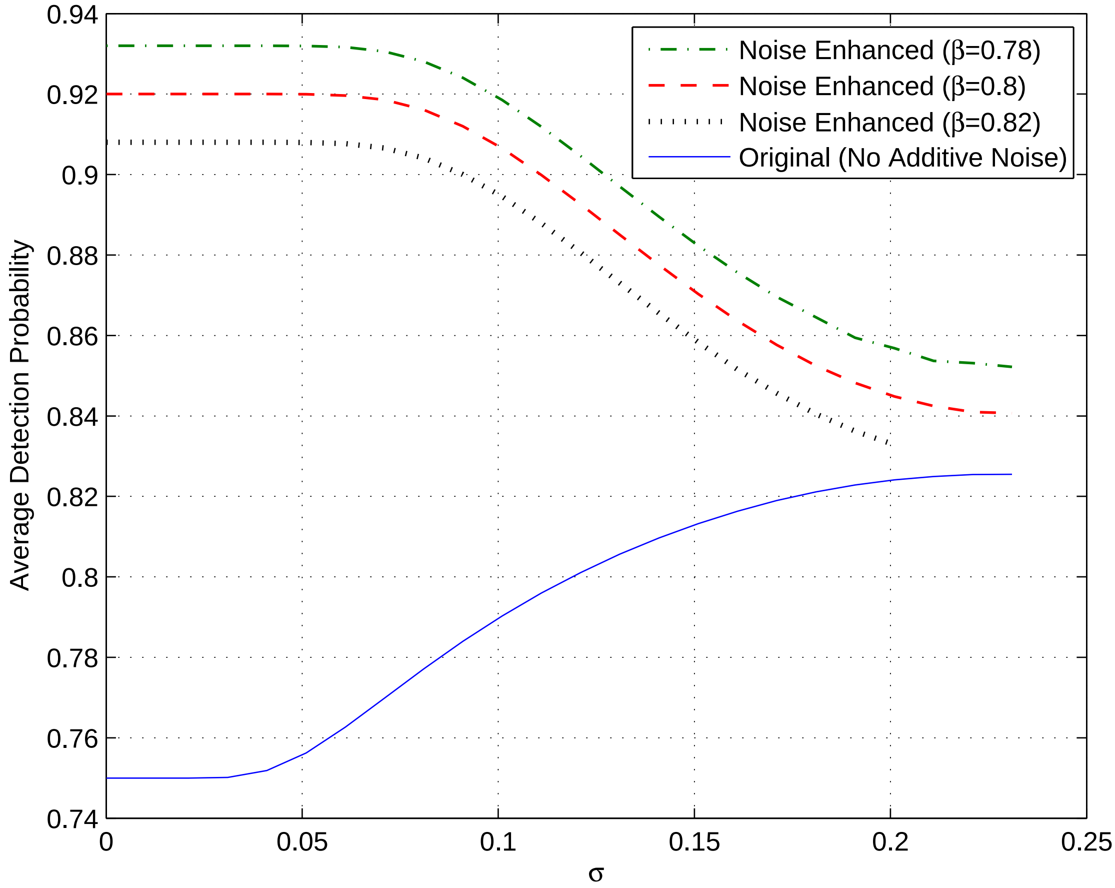

Figure 4.1: Average detection probability versus σ for various values of β, where $\alpha = 0.35$, $A = 1$ and $\rho = 0.8$.

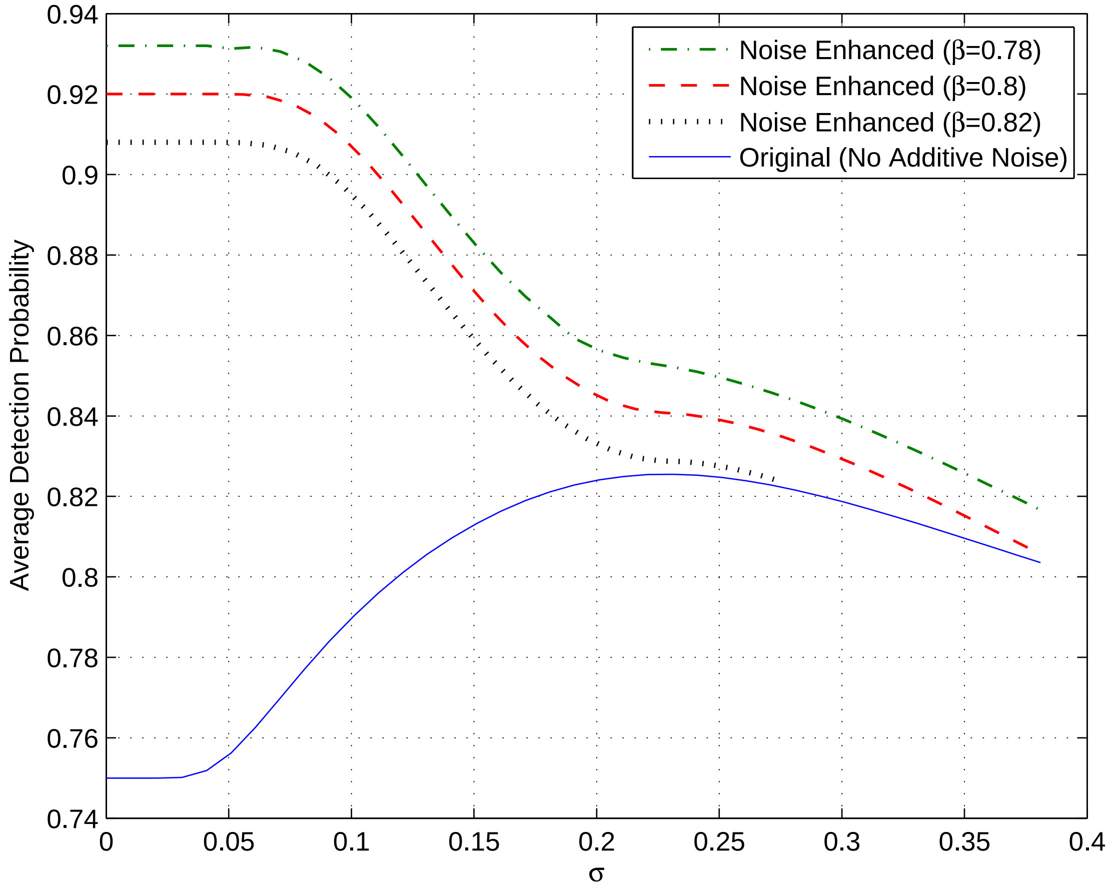

Figure 4.2: Average detection probability versus σ for various values of β, where $\alpha = 0.4$, $A = 1$ and $\rho = 0.8$.

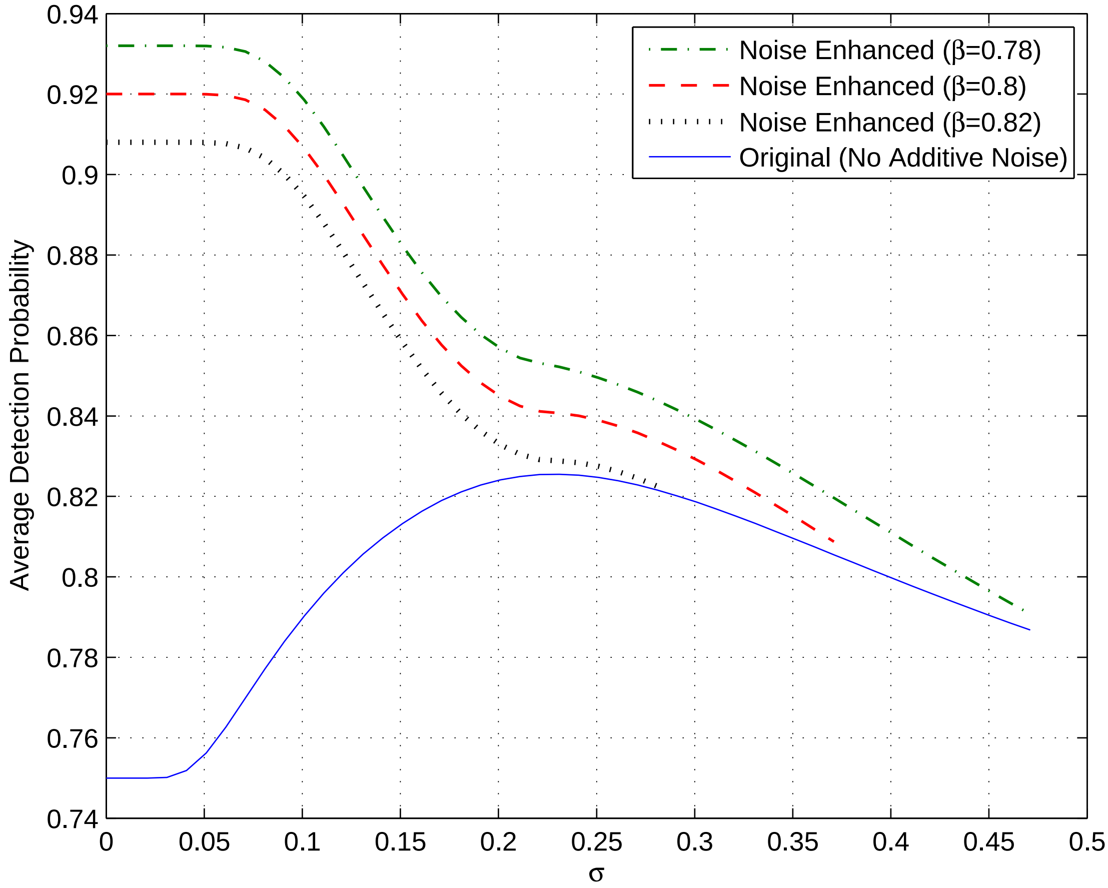

Figure 4.3: Average detection probability versus σ for various values of β, where $\alpha = 0.45$, $A = 1$ and $\rho = 0.8$.

Table 4.1: Optimal additive noise PDFs, in the form of $p_N(n) = \lambda_1\,\delta(n-n_1) + \lambda_2\,\delta(n-n_2) + (1-\lambda_1-\lambda_2)\,\delta(n-n_3)$, for various values of σ, where $\beta = 0.82$, $\alpha = 0.35$, $A = 1$ and $\rho = 0.8$.

σ	λ_1	λ_2	n_1	n_2	n_3
0	0.4181	0.3019	0.1136	0.4887	-0.4807
0.01	0.5043	0.2157	0.4146	0.1718	-0.4115
0.1	0.6886	0.3114	0.2818	-0.2818	–
0.15	0.6032	0.3968	0.2544	-0.2544	–
0.2	0.5481	0.4519	0.1796	-0.1796	–

Table 4.2: Optimal additive noise PDFs, in the form of $p_N(n) = \lambda_1\,\delta(n-n_1) + \lambda_2\,\delta(n-n_2) + (1-\lambda_1-\lambda_2)\,\delta(n-n_3)$, for various values of σ, where $\beta = 0.8$, $\alpha = 0.4$, $A = 1$ and $\rho = 0.8$.

σ	λ_1	λ_2	n_1	n_2	n_3
0	0.6098	0.1902	0.4750	0.2088	-0.2804
0.05	0.5375	0.2624	0.3002	0.2956	-0.2755
0.1	0.7689	0.2311	0.2821	-0.2821	–
0.2	0.6653	0.3347	0.1796	-0.1796	–
0.3	1	–	0.0384	–	–

Tables 4.1, 4.2, and 4.3 illustrate the optimal additive noise PDFs for various values of σ in the cases of $\beta = 0.82$ with $\alpha = 0.35$, $\beta = 0.80$ with $\alpha = 0.40$, and $\beta = 0.78$ with $\alpha = 0.45$ respectively, where $A = 1$ and $\rho = 0.8$. From Proposition 4, it is known that the optimal additive noise in this example can be represented by a discrete probability distribution with at most three point masses (since $\Lambda_0 = \{0\}$ and $\Lambda_1 = \{-A, A\}$; i.e., $M = 1$ and $N = 2$). Therefore, it can be expressed as $p_N(n) = \lambda_1\,\delta(n-n_1) + \lambda_2\,\delta(n-n_2) + (1-\lambda_1-\lambda_2)\,\delta(n-n_3)$. It is observed from the tables that the optimal additive noise PDFs have three point masses for certain values of σ, whereas they have two point masses or a single point mass for other σ's. These results are in accordance with Proposition 4, which states that an optimal PDF can be represented by a probability distribution with at most three point masses for the considered scenario.

In order to determine if any of the conditions in Proposition 2 are satisfied for the example above, the numerical values of $f^{(2)}$, $f^{(2)}_{\theta_1}$, and $g^{(2)}_{\theta_0}$ are calculated and

Table 4.3: Optimal additive noise PDFs, in the form of $p_N(n) = \lambda_1\,\delta(n-n_1) + \lambda_2\,\delta(n-n_2)+(1-\lambda_1-\lambda_2)\,\delta(n-n_3)$, for various values of σ for $\beta = 0.78$, $\alpha = 0.45$, $A = 1$ and $\rho = 0.8$.

σ	λ_1	λ_2	n_1	n_2	n_3
0	0.4510	0.12	0.2209	-0.2763	0.4344
0.05	0.5888	0.2912	0.2955	0.2848	-0.2895
0.15	0.7734	0.2266	0.2547	-0.2547	–
0.35	1	–	0.0608	–	–
0.45	1	–	0.0238	–	–

Table 4.4: Numerical values of the auxiliary functions defined for Proposition 2.

σ	$f^{(1)}$	$f_A^{(1)}$	$f_{-A}^{(1)}$	$g_0^{(1)}$	$f^{(2)}$	$f_A^{(2)}$	$f_{-A}^{(2)}$	$g_0^{(2)}$
0.05	0.1614	0.2694	-0.2705	0.0011	10.8	10.8	10.8	-21.6
0.10	0.3627	0.6046	-0.6052	6.049×10^{-4}	6.0489	6.0489	6.049	-12.1
0.15	0.3225	0.5376	-0.5378	2.25×10^{-4}	2.25	2.25	2.25	-4.5
0.20	0.2905	0.4841	-0.4842	5.502×10^{-5}	0.5507	0.5507	0.5507	-1.1
0.25	0.2856	0.4759	-0.4759	-2.758×10^{-5}	-0.2669	-0.2669	-0.2669	0.5515
0.30	0.2683	0.4772	-0.4771	-5.764×10^{-5}	-0.5395	-0.5395	-0.5395	1.153

tabulated in Table 4.4.[1] It is observed that, in this specific example, $F_{\theta_1}(\mathbf{0})$ has two minimizers; one is at $\theta_1 = -A$ and the other is at $\theta_1 = A$. Therefore, sets $\mathcal{L}_1$ and $\mathcal{L}_0$ in Proposition 2 are defined as $\mathcal{L}_1 = \{-A, A\}$ and $\mathcal{L}_0 = \{0\}$, respectively. Hence, the conditions in Proposition 2 must hold for two groups: $f^{(2)}, f_A^{(2)}, g_0^{(2)}$ and $f^{(2)}, f_{-A}^{(2)}, g_0^{(2)}$. From Table 4.4, it is noted that $f^{(2)}$, $f_A^{(2)}$ and $f_{-A}^{(2)}$ are always positive whereas $g_0^{(2)}$ is always negative for the given values of σ. For this reason, the third condition in Proposition 2 is satisfied for both groups for those values of σ, implying that the detector is improvable as a result of the proposition, which is also verified from Fig.s 4.1–4.3.

Finally, the conditions in Proposition 5 are checked in the following. We consider the Gaussian mixture noise in (4.1) with $\sigma = 0.05$, and calculate the values of $f^{(1)}$, $f_A^{(1)}$, $f_{-A}^{(1)}$, $g_0^{(1)}$, $f^{(2)}$, $f_A^{(2)}$, $f_{-A}^{(2)}$, and $g_0^{(2)}$. These values are tabulated in 4.4. From the signs of the first derivatives it is straightforward to construct

[1]Because scalar observations are considered, the signs of $f^{(2)}$, $f_{\theta_1}^{(2)}$, and $g_{\theta_0}^{(2)}$ in (3.6)-(3.8) do not depend on $\mathbf{z}$; hence, $\mathbf{z} = 1$ is used for Table 4.4.

the following sets:

- $\mathcal{S}_\beta^z = \emptyset$
- $\mathcal{S}_\beta^n = \{1_\beta^n\}$
- $\mathcal{S}_\beta^p = \{1_\beta^p, 2_\beta^p\}$
- $\mathcal{S}_\alpha^z = \emptyset$
- $\mathcal{S}_\alpha^n = \emptyset$
- $\mathcal{S}_\alpha^p = \{1_\alpha^p\}$

where $F_1(1_\beta^p) = f^{(1)}$, $F_1(2_\beta^p) = f_A^{(1)}$, $F_1(1_\beta^n) = f_{-A}^{(1)}$, and $G_1(1_\alpha^p) = g_0^{(1)}$. Now the conditions in Proposition 5 are checked.

1. Since both $\mathcal{S}_\beta^z$ and $\mathcal{S}_\alpha^z$ are empty sets, the first condition is automatically satisfied.
2. The first bullet of the second condition is not satisfied. Since $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^n| = 1$ is an odd number, we have to check the condition in the second bullet, which reduces, for this example, to the following:

   $$\min\{f_{-A}^{(2)} f_A^{(1)} g_0^{(1)} f^{(1)}, g_0^{(2)} f_A^{(1)} f_{-A}^{(1)} f^{(1)}\} > \max\{f^{(2)} f_A^{(1)} f_{-A}^{(1)} g_0^{(1)}, f_A^{(2)} f_{-A}^{(1)} g_0^{(1)} f^{(1)}\} \ .$$

   Due to the signs of the derivatives, it turns out that the two inputs of the min function on the left-hand side are positive whereas the two inputs of the max function on the right-hand side are negative so that the inequality is satisfied.

Hence, the detector is improvable as a result of Proposition 5. Moreover, when $\sigma = 0.10$, $\sigma = 0.15$, or $\sigma = 0.20$, the signs of the derivatives are the same as those in the case of $\sigma = 0.05$. Therefore, for all these cases the detector is improvable.

Now consider the case in which $\sigma = 0.25$. Again, the values of $f^{(1)}$, ${f_A}^{(1)}$, ${f_{-A}}^{(1)}$, ${g_0}^{(1)}$, $f^{(2)}$, ${f_A}^{(2)}$, ${f_{-A}}^{(2)}$, and ${g_0}^{(2)}$ are tabulated in Table 4.4. In this scenario, the sets are obtained as follows:

- $\mathcal{S}_\beta^z = \emptyset$
- $\mathcal{S}_\beta^n = \{1_\beta^n\}$
- $\mathcal{S}_\beta^p = \{1_\beta^p, 2_\beta^p\}$
- $\mathcal{S}_\alpha^z = \emptyset$
- $\mathcal{S}_\alpha^n = \{1_\alpha^n\}$
- $\mathcal{S}_\alpha^p = \emptyset$

where $F_1(1_\beta^p) = f^{(1)}$, $F_1(2_\beta^p) = f_A^{(1)}$, $F_1(1_\beta^n) = f_{-A}^{(1)}$, and $G_1(1_\alpha^n) = g_0^{(1)}$. Now the conditions in Proposition 5 are checked.

1. Since both $\mathcal{S}_\beta^z$ and $\mathcal{S}_\alpha^z$ are empty sets, the first condition is satisfied.
2. The first bullet of the second condition is not satisfied. Since $|\mathcal{S}_\beta^n|+|\mathcal{S}_\alpha^n| = 2$ is an even number, we have to check the condition in the third bullet, which, reduces, for this example, to the following:

   $$\min\{f_A^{(2)} f_{-A}^{(1)} g_0^{(1)} f^{(1)}, g_0^{(2)} f_A^{(1)} f_{-A}^{(1)} f^{(1)}, f^{(2)} f_A^{(1)} f_{-A}^{(1)} g_0^{(1)}\} > \max\{f_{-A}^{(2)} f_A^{(1)} g_0^{(1)} f^{(1)}\}$$

   For this case it turns out that all the three inputs of the min function on the left-hand side are positive and the single input to the max function on the right-hand side is negative so that the inequality is not satisfied.

Hence, the improvability conditions in Proposition 5 are not satisfied for this scenario. Similar calculations show that the same holds for $\sigma = 0.30$ as well.

Chapter 5

Conclusion

In this thesis, we have studied noise enhancement in the RNP framework. To that aim, we have formulated an optimization problem for the optimal additive noise PDF. Then, generic improvability and nonimprovability conditions have been derived, which determine if the employing additive noise becomes beneficial. We have also narrowed down the problem such that the parameter space is discrete and finite. In that scenario, we have shown that the PDF of the optimal additive noise is discrete with a certain number of point masses. Moreover, we have derived a more implicit improvability condition. Finally, the theoretical results have been supported by a numerical example in which the benefits of applying noise enhancement can be observed.

Bibliography

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

[2] J. L. Hodges, Jr. and E. L. Lehmann, "The use of previous experience in reaching statistical decisions," *The Annals of Mathematical Statistics*, vol. 23, pp. 396–407, Sep. 1952.

[3] S. Bayram and S. Gezici, "On the restricted Neyman-Pearson approach for composite hypothesis-testing in the presence of prior distribution uncertainty," *IEEE Trans. Sig. Processing*, vol. 59, Sep. 2011.

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

[5] M. D. McDonnell, "Is electrical noise useful?," *Proceedings of the IEEE*, vol. 99, pp. 242–246, Feb. 2011.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

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

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

[16] 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.

[17] 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. 2008.

[18] 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.

[19] M. D. McDonnell and D. Abbott, "What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology," *PLoS Computational Biology*, vol. 5, May 2009.

[20] 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.

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

[22] S. Bayram and S. Gezici, "On the improvability and nonimprovability of detection via additional independent noise," *IEEE Sig. Processing Lett.*, vol. 16, pp. 1001–1004, Nov. 2009.

[23] X. Godivier and F. Chapeau-Blondeau, "Noise-assisted signal transmission in a nonlinear electronic comparator: Experiment and theory," *Signal Processing*, vol. 56, pp. 293–303, Feb. 1997.

[24] S. Zozor and P. Amblard, "Stochastic resonance in discrete time nonlinear AR(1) models," *IEEE Trans. Signal Process.*, vol. 47, pp. 108–122, Jan. 1999.

[25] S. Bayram and S. Gezici, "Noise-enhanced M-ary hypothesis-testing in the minimax framework," in *Proc. International Conference on Signal Processing and Commun. Systems*, (Omaha, Nebraska), pp. 31–36, Sep. 2009.

[26] F. Chapeau-Blondeau, "Nonlinear test statistic to improve signal detection in non-gaussian noise," *IEEE Sig. Processing Lett.*, vol. 7, pp. 205–207, July 2000.

[27] S. Zozor and P. Amblard, "Can stochastic resonance be used in detection?," in *Proc. of the 10th European Signal Processing Conference (EUSIPCO)*, (Tampere, Finland), Sep. 2000.

[28] S. Zozor and P. Amblard, "On the use of stochastic resonance in sine detection," *Signal Process.*, vol. 7, pp. 353–367, Mar. 2002.

[29] S. Zozor and P. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, pp. 3177–3181, Dec. 2003.

[30] A. A. Saha and G. Anand, "Design of detectors based on stochastic resonance," *Signal Process.*, vol. 83, pp. 1193–1212, 2003.

[31] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Processing*, vol. 58, pp. 3972–3989, Aug. 2010.

[32] S. Bayram and S. Gezici, "Noise enhanced M-ary composite hypothesis-testing in the presence of partial prior information," *IEEE Trans. Sig. Processing*, vol. 59, pp. 1292–1297, Mar. 2011.

[33] E. L. Lehmann, *Testing Statistical Hypotheses.* New York: Chapman & Hall, 2 ed., 1986.

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

[35] S. Zarrin and T. J. Lim, "Composite hypothesis testing for cooperative spectrum sensing in cognitive radio," in *Proc. IEEE Int. Conf. Commun. (ICC)*, (Dresden, Germany), June 2009.

[36] S. Bayram and S. Gezici, "Stochastic resonance in binary composite hypothesis-testing problems in the Neyman-Pearson framework," *Digital Sig. Processing*, vol. 22, pp. 391–406, May 2012.

[37] S. Bayram, S. Gultekin, and S. Gezici, "Noise enhanced detection in restricted Neyman-Pearson framework," in *Proc. IEEE Workshop on Sig. Processing Advances in Wireless Commun. (SPAWC)*, June 2012.

[38] F. H. Knight, *Risk, Uncertainty and Profit.* Boston: Houghton Mifflin, 1921.

[39] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification.* New York: Wiley-Interscience, 2nd ed., 2000.

[40] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems*, pp. 214–220. IOS Press, 2002. in Intelligent Technologies–Theory and Applications.

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