

# Effects of Additional Independent Noise in Binary Composite Hypothesis-Testing Problems

Suat Bayram and Sinan Gezici  
Department of Electrical and Electronics Engineering  
Bilkent University  
Bilkent, Ankara 06800, Turkey  
{sbayram, gezici}@ee.bilkent.edu.tr

**Abstract**—Performance of some suboptimal detectors can be improved by adding independent noise to their observations. In this paper, the effects of adding independent noise to observations of a detector are investigated for binary composite hypothesis-testing problems in a generalized Neyman-Pearson framework. Sufficient conditions are derived to determine when performance of a detector can or cannot be improved via additional independent noise. Also, upper and lower limits are derived on the performance of a detector in the presence of additional noise, and statistical characterization of optimal additional noise is provided. In addition, two optimization techniques are proposed to calculate the optimal additional noise. Finally, simulation results are presented to investigate the theoretical results.

**Index Terms**—Binary hypothesis-testing, Neyman-Pearson, composite hypothesis-testing, stochastic resonance.

## I. INTRODUCTION

In binary hypothesis testing problems, the aim is to determine the true hypothesis based on a number of observations and, if exists, on prior information about the hypotheses [1], [2]. In the presence of prior information and a specific cost assignment on each decision, the Bayesian approach aims to design a decision rule that minimizes the *Bayes risk*, which is defined as the average of the expected costs for the two hypotheses. The Bayesian approach is employed in various fields, such as digital communications, image processing, robotics, control, and biomedicine [2], [3]. In the absence of prior information about the hypotheses, the minimax approach can be taken, which minimizes the maximum of the expected costs for the two hypotheses [1]. The minimax approach can be considered as an algorithm that tries to optimize the worst-case performance. On the other hand, the Neyman-Pearson approach assumes neither prior information nor specific cost assignments, and aims to maximize the detection probability (probability of correctly selecting the first hypothesis) under a constraint on the false alarm probability (probability of deciding the first hypothesis when the null hypothesis is true) [2]. Neyman-Pearson detectors take into account the tradeoff between detection and false alarm probabilities, and are commonly employed for detecting the presence of signals based on noisy observations [4], [5].

Binary hypothesis-testing problems can be classified into *simple* and *composite* problems [1]. In a simple hypothesis-testing problem, each hypothesis corresponds to a single probability distribution for the observation under that hypothesis. However, in composite hypothesis-testing problems, a hypothesis corresponds to multiple possible distributions. For

example, in radar problems, when the target is present, the observation has multiple unknown parameters, such as range and velocity; hence, the observation can have multiple possible distributions. Another example of a composite hypothesis-testing problem is non-coherent detection of communications signals, where the unknown phase value at the receiver results in composite hypotheses [1].

In this paper, composite hypothesis-testing problems are studied in a Neyman-Pearson framework, and the effects of adding independent noise to observations of a detector are investigated. Recently, it has been shown that adding specific noise to observations of a detector can improve detector performance under certain conditions [6]–[18]. This effect, called *stochastic resonance* (SR), may improve performance of suboptimal detectors according to the Bayesian [19], minimax [20], [21] or Neyman-Pearson criteria [17], [18], [22]–[24]. The studies in [17] and [18] establish a theoretical framework to provide sufficient conditions for improvability or non-improvability of a suboptimal detector via additional independent noise, and propose techniques to obtain the optimal noise distribution in the Neyman-Pearson framework. In [25] and [26], a weak sinusoidal signal is considered and improvements on detection performance are studied. In addition, [27] investigates the optimization of noise and detector parameters for locally optimal detectors.

Although the effects of additional independent noise are studied for simple hypothesis-testing problems [17], [18], no studies have considered composite hypothesis-testing problems to provide a theoretical framework for the effects of additional noise on various detectors. In this paper, the effects of additional independent noise are studied for binary composite hypothesis-testing problems in the generalized Neyman-Pearson framework. First, sufficient conditions are obtained to specify whether additional noise can or cannot improve detection performance for a given detector. Then, statistical characterization of optimal additional noise is provided and upper and lower performance limits are derived. In addition, optimization theoretic approaches are proposed to obtain exact and approximate solutions for the optimal additional noise. Finally, a numerical example is presented to investigate the theoretical results.

The remainder of the paper is organized as follows. In Section II, the problem formulation is presented and the main motivations for this study are explained. In Section III, improvability and non-improvability of detection via additional

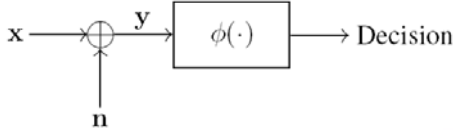


Fig. 1. Independent noise  $\mathbf{n}$  is added to observation  $\mathbf{x}$  to improve the performance of the detector,  $\phi(\cdot)$ .

independent noise are investigated. Then, properties of optimal additional noise are studied in Section IV and various algorithms to obtain exact and approximate optimal solutions are proposed in Section V. Finally, a detection example is presented in Section VI, followed by the concluding remarks in Section VII.

## II. PROBLEM FORMULATION AND MOTIVATION

A binary composite hypothesis-testing problem is studied in this paper, which can be stated as [1]

$$\begin{aligned} \mathcal{H}_0 &: p_{\theta_0}(\mathbf{x}) \quad , \quad \theta_0 \in \Lambda_0 \\ \mathcal{H}_1 &: p_{\theta_1}(\mathbf{x}) \quad , \quad \theta_1 \in \Lambda_1 \end{aligned} \quad (1)$$

where  $\mathcal{H}_i$  denotes the  $i$ th hypothesis for  $i = 0, 1$ . Under hypothesis  $\mathcal{H}_i$ , observation  $\mathbf{x}$ , which is a  $K$ -dimensional vector, has a probability density function (PDF) indexed by  $\theta_i \in \Lambda_i$ , where  $\Lambda_i$  is the set of possible parameter values under hypothesis  $\mathcal{H}_i$ . It is assumed that parameter sets  $\Lambda_0$  and  $\Lambda_1$  are disjoint, and their union forms the parameter space [1]. In addition, the prior probability distributions of the parameters are unknown.

Composite hypothesis-testing problems are encountered in various problems, such as in non-coherent communications receivers and radar systems [1], [4]. When both  $\Lambda_0$  and  $\Lambda_1$  consist of single elements, the problem in (1) reduces to a *simple* hypothesis-testing problem.

A generic detector (decision rule), denoted by  $\phi(\mathbf{x})$ , is considered, which maps the observation into a real number in  $[0, 1]$  that represents the probability of selecting  $\mathcal{H}_1$  [1]. The aim is to investigate the effects of adding independent noise to the observation of a given detector, as shown in Fig. 1, where  $\mathbf{y}$  represents the modified observation expressed as

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad , \quad (2)$$

with  $\mathbf{n}$  denoting the additional noise term that is independent of  $\mathbf{x}$ .

A generalized Neyman-Pearson framework [28], [29] is considered in this study, and performance of a detector is quantified in terms of its worst-case detection probability under a constraint on the maximum probability of false alarm. Before explaining the details of this performance metric, the probabilities of detection and false alarm for specific parameter values are obtained first. Since the additional noise is independent of the observation, the probabilities of detection and false alarm can be expressed, conditioned on  $\theta_1$  and  $\theta_0$ ,

respectively, as

$$P_D^{\mathbf{y}}(\theta_1) = \int_{\mathbb{R}^K} \phi(\mathbf{y}) \left[ \int_{\mathbb{R}^K} p_{\theta_1}(\mathbf{y} - \mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} \right] d\mathbf{y} \quad , \quad (3)$$

$$P_F^{\mathbf{y}}(\theta_0) = \int_{\mathbb{R}^K} \phi(\mathbf{y}) \left[ \int_{\mathbb{R}^K} p_{\theta_0}(\mathbf{y} - \mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} \right] d\mathbf{y} \quad , \quad (4)$$

where  $p_{\mathbf{n}}(\cdot)$  represents the PDF of the additional noise. After some manipulation, (3) and (4) becomes [17]

$$P_D^{\mathbf{y}}(\theta_1) = E_{\mathbf{n}}\{F_{\theta_1}(\mathbf{n})\} \quad , \quad (5)$$

$$P_F^{\mathbf{y}}(\theta_0) = E_{\mathbf{n}}\{G_{\theta_0}(\mathbf{n})\} \quad , \quad (6)$$

for  $\theta_1 \in \Lambda_1$  and  $\theta_0 \in \Lambda_0$ , where

$$F_{\theta_1}(\mathbf{n}) \doteq \int_{\mathbb{R}^K} \phi(\mathbf{y}) p_{\theta_1}(\mathbf{y} - \mathbf{n}) d\mathbf{y} \quad , \quad (7)$$

$$G_{\theta_0}(\mathbf{n}) \doteq \int_{\mathbb{R}^K} \phi(\mathbf{y}) p_{\theta_0}(\mathbf{y} - \mathbf{n}) d\mathbf{y} \quad . \quad (8)$$

It is noted that  $F_{\theta_1}(\mathbf{n})$  and  $G_{\theta_0}(\mathbf{n})$  define, respectively, the probability of detection conditioned on  $\theta_1$  and the probability of false alarm conditioned on  $\theta_0$ , when a constant noise  $\mathbf{n}$  is added to the observation. In the absence of additional noise, i.e.,  $\mathbf{n} = \mathbf{0}$ , the probabilities of detection and false alarm are given, respectively, by  $P_D^{\mathbf{x}}(\theta_1) = F_{\theta_1}(\mathbf{0})$  and  $P_F^{\mathbf{x}}(\theta_0) = G_{\theta_0}(\mathbf{0})$  for given values of the parameters.

In the Neyman-Pearson framework, the main constraint is to have the probability of false alarm under a certain threshold for all possible parameter values  $\theta_0$ ; that is,

$$\max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leq \tilde{\alpha} \quad . \quad (9)$$

In most practical cases, the detectors are designed to operate at the maximum allowed false alarm probability  $\tilde{\alpha}$  in order to obtain maximum detection probabilities [1]. Hence, the constraint on the probability of false alarm can be defined as  $\tilde{\alpha} = \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{x}}(\theta_0) = \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{0})$  for practical scenarios; that is, the detectors commonly operate at the limit for the probability of false alarm.

According to the generalized Neyman-Pearson framework [28], [29], the aim is to maximize the worst-case detection probability,  $\min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1)$ , under the false alarm constraint in (9). The worst-case detection probability corresponds to considering the *least-favorable* distribution for parameter  $\theta_1$  [28]. Therefore, this performance criterion guarantees a detection performance under a given false alarm constraint for all possible parameter distributions. The generalized Neyman-Pearson criterion is commonly employed in composite hypothesis-testing problems in which the prior distributions of the parameters are unknown [29], [30].

Based on the performance criterion described above, the PDF of the optimal additional noise  $\mathbf{n}$  in (2) can be obtained from the solution of the following optimization problem:

$$\max_{p_{\mathbf{n}}(\cdot)} \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) \quad (10)$$

$$\text{subject to } \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leq \tilde{\alpha} \quad (11)$$

where  $P_D^{\mathbf{y}}(\theta_1)$  and  $P_F^{\mathbf{y}}(\theta_0)$  are as in (5)-(8).

There are two main motivations for investigating the prob-

lem in (10) and (11). First, it is important to quantify the performance improvements that can be obtained via additional independent noise, and to determine when additional noise can improve detection performance. In other words, theoretical investigation of the effects of additional independent noise is of interest. Second, in some cases, a suboptimal detector with additional noise as in Fig. 1 can provide a low complexity solution compared to the optimal detector, which is commonly quite complex [1], [28]. It should be noted that although the calculation of the optimal additional noise requires certain computations, the overall computational complexity can still be considerably lower than that of the optimal detector, since the optimal detector needs to perform intense computations for each decision whereas the suboptimal detector with additional noise needs to update the optimal additional noise only when the statistics of the hypotheses change.

### III. IMPROVABILITY & NON-IMPROVABILITY CONDITIONS

In this section, sufficient conditions are specified to determine whether additional independent noise can improve detection performance according to the generalized Neyman-Pearson criterion without actually solving the optimization problem in (10) and (11). A detector is called *improvable* if there exists additional noise  $\mathbf{n}$  that satisfies

$$\min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) > \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1) = \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{0}) \doteq P_{D,\min}^{\mathbf{x}} \quad (12)$$

under the false alarm constraint in (9). Otherwise, the detector is called *non-improvable*.

Based on the improvability definition in (12), a simple observation reveals that if there exists a noise component  $\tilde{\mathbf{n}}$  that satisfies  $\min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\tilde{\mathbf{n}}) > \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{0})$  and  $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\tilde{\mathbf{n}}) \leq \tilde{\alpha}$ , (5) and (6) implies that addition of noise  $\tilde{\mathbf{n}}$  to the observation increases the probability of detection under the false alarm constraint for all  $\theta_1$  values; hence,  $\min_{\theta_1 \in \Lambda_1} P_D^{\tilde{\mathbf{y}}}(\theta_1) > \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1)$  is satisfied under the false alarm constraint, where  $\tilde{\mathbf{y}} = \mathbf{x} + \tilde{\mathbf{n}}$ . However, in some scenarios, improvability may not be obtained by using such a fixed noise component, and a more generic improvability condition can be required.

In order to derive a more generic improvability condition, the approach in [17] for simple hypothesis-testing problems is extended to composite hypothesis-testing problems in the following manner. First, we introduce the following function:

$$H_{\min}(t) \doteq \sup \left\{ \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}) \mid t = \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}), \mathbf{n} \in \mathbb{R}^K \right\} \quad (13)$$

which defines the maximum value of the minimum detection probability for a given value of the maximum false alarm probability. From (13), it is observed that if there exists  $t_0 \leq \tilde{\alpha}$  such that  $H_{\min}(t_0) > P_{D,\min}^{\mathbf{x}}$ , the system is improvable, because under such a condition there exists a noise component  $\mathbf{n}_0$  such that  $\min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}_0) > P_{D,\min}^{\mathbf{x}}$  and  $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_0) \leq \tilde{\alpha}$ . Therefore, the detector performance can be improved by using an additional noise component with  $p_{\mathbf{n}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{n}_0)$ . However, as stated previously, improvability may not be obtained

with fixed noise components in some scenarios. Hence, a more generic improvability condition is derived in the following proposition.

**Proposition 1:** Let  $\alpha = \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{x}}(\theta_0)$  denote the maximum probability of false alarm in the absence of additional noise. If  $H_{\min}(t)$  in (13) is second-order continuously differentiable around  $t = \alpha$  and satisfies  $H_{\min}''(\alpha) > 0$ , then the detector is improvable.

**Proof:** Since  $H_{\min}''(\alpha) > 0$  and  $H_{\min}(t)$  is second-order continuously differentiable around  $t = \alpha$ , there exist  $\epsilon > 0$ ,  $\mathbf{n}_1$  and  $\mathbf{n}_2$  such that  $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_1) = \alpha + \epsilon$  and  $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_2) = \alpha - \epsilon$ . Then, it is proven in the following that additional noise with  $p_{\mathbf{n}}(\mathbf{x}) = 0.5 \delta(\mathbf{x} - \mathbf{n}_1) + 0.5 \delta(\mathbf{x} - \mathbf{n}_2)$  improves the detection performance under the false alarm constraint. First, the maximum false alarm probability in the presence of additional noise is shown not to exceed  $\alpha$ .

$$\begin{aligned} \max_{\theta_0 \in \Lambda_0} E_{\mathbf{n}} \{G_{\theta_0}(\mathbf{n})\} &\leq E_{\mathbf{n}} \left\{ \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}) \right\} \\ &= 0.5(\alpha + \epsilon) + 0.5(\alpha - \epsilon) \\ &= \alpha \end{aligned} \quad (14)$$

Then, the increase in the probability of detection is proven as follows. Since

$$\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}} \{F_{\theta_1}(\mathbf{n})\} \geq E_{\mathbf{n}} \left\{ \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}) \right\} \quad (15)$$

is valid for all noise PDFs,

$$\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}} \{F_{\theta_1}(\mathbf{n})\} \geq 0.5 H_{\min}(\alpha + \epsilon) + 0.5 H_{\min}(\alpha - \epsilon) \quad (16)$$

is satisfied. From the assumptions in the proposition,  $H_{\min}(t)$  is convex in an interval around  $t = \alpha$ . Hence, (16) becomes

$$\begin{aligned} \min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}} \{F_{\theta_1}(\mathbf{n})\} &\geq 0.5 H_{\min}(\alpha + \epsilon) + 0.5 H_{\min}(\alpha - \epsilon) \\ &> H_{\min}(\alpha) . \end{aligned} \quad (17)$$

Because  $H_{\min}(\alpha) \geq P_{D,\min}^{\mathbf{x}}$  by definition, (17) implies that  $\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}} \{F_{\theta_1}(\mathbf{n})\} > P_{D,\min}^{\mathbf{x}}$ . Hence, the detector is improvable.  $\square$

Proposition 1 provides a convenient sufficient condition that deals with a scalar function  $H_{\min}(t)$  irrespective of the dimension of the observation vector, which facilitates simple evaluations of the conditions in the proposition. However, the main complexity can be to obtain an expression for  $H_{\min}(t)$  in certain scenarios. Numerical results are provided in Section VI to illustrate an example.

Next, sufficient conditions for non-improvability are obtained in the following. To that aim, the following function is defined first.

$$J_{\theta_0, \theta_1}(t) \doteq \sup \{F_{\theta_1}(\mathbf{n}) \mid G_{\theta_0}(\mathbf{n}) = t, \mathbf{n} \in \mathbb{R}^K\} . \quad (18)$$

Then, the following proposition can be obtained as an extension of the non-improvability condition in [17].

**Proposition 2:** Let  $\theta_1^{\min}$  represent the value of  $\theta_1 \in \Lambda_1$  that has the minimum detection probability in the absence of

additional noise; that is,

$$\theta_1^{\min} \doteq \arg \min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1). \quad (19)$$

If there exists  $\theta_0 \in \Lambda_0$  and a nondecreasing concave function  $\Psi(t)$  such that  $\Psi(t) \geq J_{\theta_0, \theta_1^{\min}}(t) \forall t$  and  $\Psi(\tilde{\alpha}) = P_D^x(\theta_1^{\min})$ , then the detector is non-improvable.

**Proof:** First, the non-improvability of the detector is proven for  $\theta_1 = \theta_1^{\min}$  in the following. For  $\theta_1 = \theta_1^{\min}$ , the objective function in (10) can be expressed from (5) as follows:

$$\begin{aligned} E_{\mathbf{n}}\{F_{\theta_1^{\min}}(\mathbf{n})\} &= \int p_{\mathbf{n}}(\mathbf{x}) F_{\theta_1^{\min}}(\mathbf{x}) d\mathbf{x} \\ &\leq \int p_{\mathbf{n}}(\mathbf{x}) J_{\theta_0, \theta_1^{\min}}(G_{\theta_0}(\mathbf{x})) d\mathbf{x} \end{aligned} \quad (20)$$

where the inequality is obtained from the definition in (18).

Since  $\Psi(t)$  satisfies  $\Psi(t) \geq J_{\theta_0, \theta_1^{\min}}(t) \forall t$ , and is concave, (20) becomes

$$\begin{aligned} E_{\mathbf{n}}\{F_{\theta_1^{\min}}(\mathbf{n})\} &\leq \int p_{\mathbf{n}}(\mathbf{x}) \Psi(G_{\theta_0}(\mathbf{x})) d\mathbf{x} \\ &\leq \Psi\left(\int p_{\mathbf{n}}(\mathbf{x}) G_{\theta_0}(\mathbf{x}) d\mathbf{x}\right). \end{aligned} \quad (21)$$

Then, the nondecreasing property of  $\Psi(t)$  together with  $\int p_{\mathbf{n}}(\mathbf{x}) G_{\theta_0}(\mathbf{x}) d\mathbf{x} \leq \tilde{\alpha}$  implies that

$$E_{\mathbf{n}}\{F_{\theta_1^{\min}}(\mathbf{n})\} \leq \Psi(\tilde{\alpha}). \quad (22)$$

Since  $\Psi(\tilde{\alpha}) = P_D^x(\theta_1^{\min})$ ,  $E_{\mathbf{n}}\{F_{\theta_1^{\min}}(\mathbf{n})\} \leq P_D^x(\theta_1^{\min})$  is obtained for any additional noise  $\mathbf{n}$ . Hence, the detector is non-improvable at  $\theta_1 = \theta_1^{\min}$ . Since the detector is non-improvable for  $\theta_1 = \theta_1^{\min}$ , it is non-improvable according to the generalized Neyman-Pearson criterion in (10), since its minimum can never increase by using any additional noise.  $\square$

The conditions in Proposition 2 can be used to determine the cases in which the detector performance cannot be improved via additional noise. In that way, unnecessary efforts for solving the optimization problem in (10) and (11) can be prevented.

#### IV. PROPERTIES OF OPTIMAL ADDITIONAL NOISE

In this section, performance limits are obtained for detectors that employ additional independent noise, and statistical characteristics of optimal additional noise are specified.

In order to obtain upper and lower limits on the performance of the detector that employs the additional noise specified by the optimization problem in (10) and (11), consider a separate optimization problem for each  $\theta_1 \in \Lambda_1$  as follows:

$$\begin{aligned} &\max_{p_{\mathbf{n}}(\cdot)} P_D^y(\theta_1) \\ &\text{subject to } \max_{\theta_0 \in \Lambda_0} P_F^y(\theta_0) \leq \tilde{\alpha} \end{aligned} \quad (23)$$

Let  $P_{D, \text{opt}}^y(\theta_1)$  represent the solution of (23), and  $p_{\mathbf{n}_{\theta_1}}(\cdot)$  be the corresponding optimal PDF. In addition, let  $\bar{\theta}_1$  denote the parameter value with the minimum  $P_{D, \text{opt}}^y(\theta_1)$  among all  $\theta_1 \in \Lambda_1$ ; that is,

$$\bar{\theta}_1 = \arg \min_{\theta_1 \in \Lambda_1} P_{D, \text{opt}}^y(\theta_1). \quad (24)$$

Then, the following proposition provides performance limits for the detector in the presence of additional independent noise.

**Proposition 3:** Let  $P_{D, \text{mm}}^y$  denote solution of the optimization problem specified by (10) and (11). It has the following lower and upper limits:

$$\max \left\{ \min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1), \min_{\theta_1 \in \Lambda_1} P_D^{y_{\bar{\theta}_1}}(\theta_1) \right\} \leq P_{D, \text{mm}}^y \leq \min_{\theta_1 \in \Lambda_1} P_{D, \text{opt}}^y(\theta_1) \quad (25)$$

where  $P_{D, \text{opt}}^y(\theta_1)$  is the solution of the optimization problem in (23),  $P_D^x(\theta_1)$  is the probability of detection in the absence of additional noise, and  $P_D^{y_{\bar{\theta}_1}}(\theta_1)$  is the probability of detection in the presence of additional noise  $\mathbf{n}_{\bar{\theta}_1}$ , which is specified by the PDF  $p_{\mathbf{n}_{\bar{\theta}_1}}(\cdot)$  that is the optimizer of (23) for  $\bar{\theta}_1$  that is given by (24).

**Proof:** The upper limit in (25) directly follows from (10), (11) and (23), since  $\max_{\theta_1 \in \Lambda_1} P_D^y(\theta_1) \geq \max_{p_{\mathbf{n}}(\cdot)} \min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$  for all  $\theta_1 \in \Lambda_1$ . To obtain the lower limit, it is first noted that the detector in the presence of additional independent noise can never have lower minimum detection probability than that in the absence of noise, i.e.,  $\min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1)$ . In addition, using an additional noise with PDF  $p_{\mathbf{n}_{\bar{\theta}_1}}(\cdot)$ , which is the optimal noise for the problem in (23) for a specific  $\theta_1$  value, can never result in a larger minimum probability  $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$  than that obtained from the solution of (10) and (11), since the latter directly maximizes the  $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$  metric. Therefore,

$\min_{\theta_1 \in \Lambda_1} P_D^{y_{\bar{\theta}_1}}(\theta_1)$  provides another lower limit.  $\square$

The result in Proposition 3 can be explained as follows. It is noted that  $P_{D, \text{opt}}^y(\theta_1)$  represents the maximum detection probability when an additional noise component that is optimized for a specific value of  $\theta_1$  is used. Therefore, for each  $\theta_1 \in \Lambda_1$ ,  $P_{D, \text{opt}}^y(\theta_1)$  is larger than or equal to  $\max_{p_{\mathbf{n}}(\cdot)} \min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$ , since the latter involves a single additional noise component that is optimized for the minimum detection probability metric and is used for all  $\theta_1$  values. In other words, the upper limit is obtained by assuming a more flexible optimization problem in which a different optimal noise component can be used for each  $\theta_1$  value. Regarding the lower limit, the first lower limit expression is obtained from the fact that the optimal value can never be smaller than  $\min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1)$ , which is the minimum detection probability in the absence of additional noise. The second lower limit is obtained from the observation that the optimal additional noise PDF that maximizes the minimum detection probability,  $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$ , is calculated from the optimization problem in (10) and (11); hence, the resulting optimal value,  $P_{D, \text{mm}}^y$ , is larger than or equal to all other  $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)$  values that are obtained by using a different noise PDF.

For statistical characterization of optimal additional noise, it can be shown that when parameter sets  $\Lambda_0$  and  $\Lambda_1$  in (1) consist of a finite number of parameters, the optimal additional noise can be represented by a discrete random variable with a finite number of mass points as specified below.

**Proposition 4:** Let  $\theta_0 \in \Lambda_0 = \{\theta_{01}, \theta_{02}, \dots, \theta_{0M}\}$  and

$\theta_1 \in \Lambda_1 = \{\theta_{11}, \theta_{12}, \dots, \theta_{1N}\}$ . Assume that the additional noise components can take finite values specified by  $n_i \in [a_i, b_i]$ ,  $i = 1, \dots, K$ , for any finite  $a_i$  and  $b_i$ . Define set  $U$  as

$$U = \{(u_1, \dots, u_{N+M}) : u_1 = F_{\theta_{11}}(\mathbf{n}), \dots, u_N = F_{\theta_{1N}}(\mathbf{n}), \\ u_{N+1} = G_{\theta_{01}}(\mathbf{n}), \dots, u_{N+M} = G_{\theta_{0M}}(\mathbf{n}), \text{ for } \mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}\}, \quad (26)$$

where  $\mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}$  means that  $n_i \in [a_i, b_i]$  for  $i = 1, \dots, K$ . If  $U$  is a closed subset of  $\mathbb{R}^{N+M}$ , an optimal solution to (10) and (11) has the following form

$$p_{\mathbf{n}}(\mathbf{x}) = \sum_{i=1}^{N+M} \lambda_i \delta(\mathbf{x} - \mathbf{n}_i), \quad (27)$$

where  $\sum_{i=1}^{N+M} \lambda_i = 1$  and  $\lambda_i \geq 0$  for  $i = 1, 2, \dots, N+M$ .

**Proof:** Please see Appendix A.

Regarding the first assumption in the proposition, constraining the additional noise values as  $\mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}$  is quite realistic as arbitrarily large/small values cannot be realized in practical systems. The assumption that  $U$  is a closed set makes sure the existence of the optimal solution [18], and it holds, for example, when  $F_{\theta_{1i}}$  and  $G_{\theta_{0j}}$  are continuous functions.

The main implication of Proposition 4 is that when the parameter sets consist of finite numbers of elements, the optimal additional noise can be represented, under certain conditions, by a discrete random variable with a total number of mass points at most equal to the number of possible parameter values. In such a case, the optimization problem in (10) and (11) simplifies significantly (c.f. Section V) since the search space reduces from the set of all probability distributions to the discrete probability distributions with no more than a specified number of mass points.

## V. CALCULATION OF OPTIMAL ADDITIONAL NOISE

In this section, various optimization algorithms are studied in order to obtain the optimal noise PDF from (10) and (11). Let  $p_{\mathbf{n}, f_{\theta_1}}(\cdot)$  denote the PDF of  $f_{\theta_1} = F_{\theta_1}(\mathbf{n})$ , where  $F_{\theta_1}(\mathbf{n})$  is given by (7). Note that  $p_{\mathbf{n}, f_{\theta_1}}(\cdot)$  can be obtained from the noise PDF  $p_{\mathbf{n}}(\cdot)$ , and it is more convenient to work with since it is the PDF of a scalar random variable [17].

Assume that there exists at least one value of  $\theta_1 \in \Lambda_1$ , for which  $F_{\theta_1}(\mathbf{n})$  is one-to-one. Let one of these values be represented by  $\tilde{\theta}_1$ . Then, for a given value  $\mathbf{n}$  of noise,  $f = F_{\tilde{\theta}_1}(\mathbf{n})$  can be used to express  $g_{\theta_0} = G_{\theta_0}(\mathbf{n})$  and  $f_{\theta_1} = F_{\theta_1}(\mathbf{n})$  as  $g_{\theta_0} = G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(f))$  and  $f_{\theta_1} = F_{\tilde{\theta}_1}(F_{\tilde{\theta}_1}^{-1}(f))$ , respectively. Therefore, the optimization problem in (10) and (11) can be reformulated as

$$\max_{p_{\mathbf{n}, f_{\tilde{\theta}_1}}(\cdot)} \min_{\theta_1 \in \Lambda_1} \int_0^1 f_{\theta_1} p_{\mathbf{n}, f_{\tilde{\theta}_1}}(f) df, \\ \text{subject to } \max_{\theta_0 \in \Lambda_0} \int_0^1 g_{\theta_0} p_{\mathbf{n}, f_{\tilde{\theta}_1}}(f) df \leq \tilde{\alpha}. \quad (28)$$

Depending on the nature of the parameter sets, (28) can be solved in different manners.

### A. Case-1: $\Lambda_0$ and $\Lambda_1$ with finite number of elements

Assume that the parameters can take finitely many values specified by  $\theta_0 \in \Lambda_0 = \{\theta_{01}, \theta_{02}, \dots, \theta_{0M}\}$  and  $\theta_1 \in \Lambda_1 = \{\theta_{11}, \theta_{12}, \dots, \theta_{1N}\}$ . In this case, the optimal noise PDF can be represented by  $(N+M)$  mass points, under the conditions in Proposition 4. Then, (28) can be expressed as

$$\max_{\{\lambda_i, f_i\}_{i=1}^{N+M}} \min_{\theta_1 \in \Lambda_1} \sum_{i=1}^{N+M} \lambda_i f_{\theta_1, i} \\ \text{subject to } \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^{N+M} \lambda_i g_{\theta_0, i} \leq \tilde{\alpha} \\ \sum_{i=1}^{N+M} \lambda_i = 1 \\ \lambda_i \geq 0, \quad i = 1, \dots, N+M \quad (29)$$

where  $f_i = F_{\tilde{\theta}_1}(\mathbf{n}_i)$ ,  $f_{\theta_1, i} = F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(f_i))$ ,  $g_{\theta_0, i} = G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(f_i))$ , and  $\mathbf{n}_i$  and  $\lambda_i$  are, respectively, the optimal mass points and their weights as specified in Proposition 4. Since the optimization problem in (29) is not a convex optimization problem in general, global optimization techniques, such as particle-swarm optimization (PSO) [31]-[34], genetic algorithms and differential evolution [35], can be used to obtain the optimal solution. In Section VI, the PSO algorithm is used to obtain the optimal noise PDF from (29).

Since the optimization problem in (29) can have high computational complexity, an approximate and efficient solution can be obtained via convex formulation of the problem. To that aim, suppose that  $f = F_{\tilde{\theta}_1}(\mathbf{n})$  can take only finitely many known values, specified by  $\tilde{f}_1, \dots, \tilde{f}_{\tilde{M}}$ . In that case, the optimization can be performed only over the weights  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{\tilde{M}}$  corresponding to those values. Then, (29) becomes

$$\max_{\tilde{\lambda}} \min_{\theta_1 \in \Lambda_1} \tilde{\mathbf{f}}_{\theta_1}^T \tilde{\lambda} \\ \text{subject to } \tilde{\mathbf{g}}_{\theta_0}^T \tilde{\lambda} \leq \tilde{\alpha}, \quad \forall \theta_0 \in \Lambda_0 \\ \mathbf{1}^T \tilde{\lambda} = 1 \\ \tilde{\lambda} \succeq \mathbf{0} \quad (30)$$

where  $\tilde{\mathbf{f}}_{\theta_1} = [F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_1)) \cdots F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_{\tilde{M}}))]^T$ ,  $\tilde{\mathbf{g}}_{\theta_0} = [G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_1)) \cdots G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_{\tilde{M}}))]^T$ , and  $\tilde{\lambda} = [\tilde{\lambda}_1 \cdots \tilde{\lambda}_{\tilde{M}}]^T$ . The optimization problem (30) can be expressed as a convex problem as follows when an auxiliary optimization variable  $t$  is defined.

$$\max_{\tilde{\lambda}, t} t \\ \text{subject to } \tilde{\mathbf{f}}_{\theta_1}^T \tilde{\lambda} \geq t, \quad \forall \theta_1 \in \Lambda_1 \\ \tilde{\mathbf{g}}_{\theta_0}^T \tilde{\lambda} \leq \tilde{\alpha}, \quad \forall \theta_0 \in \Lambda_0 \\ \mathbf{1}^T \tilde{\lambda} = 1 \\ \tilde{\lambda} \succeq \mathbf{0} \quad (31)$$

The problem in (31) is a linearly constrained linear optimization problem.

tion problem; hence, it can be solved very efficiently [36]. Although (31) provides an approximate solution to (29), it gets very close to the exact solution as more values of  $f = F_{\tilde{\theta}_1}(\mathbf{n})$  are included in the optimization.

#### B. Case-2: $\Lambda$ with infinitely many elements

Now consider the case in which at least one of  $\theta_0$  or  $\theta_1$  can take infinitely many values. Then, the parameter set  $\Lambda = \Lambda_0 \cup \Lambda_1$  includes infinitely many elements. In that case, the optimal noise may not be represented by the randomization of a finite number of mass points as in Proposition 4. Since the optimization over the space of all PDFs is quite complex, one approach to solving the optimization problem in (28) involves the use of PDF approximations. Let the optimal PDF be approximated by

$$p_{\mathbf{n}, f_{\tilde{\theta}_1}}(f) = \sum_{i=1}^L \mu_i \psi_i(f - f_i), \quad (32)$$

where  $\mu_i \geq 0$ ,  $\sum_{i=1}^L \mu_i = 1$ , and  $\psi_i(\cdot)$  is a window function that satisfies  $\psi_i(\mathbf{x}) \geq 0 \forall \mathbf{x}$  and  $\int \psi_i(\mathbf{x}) d\mathbf{x} = 1$ , for  $i = 1, \dots, L$ . The PDF approximation technique in (32) is called *Parzen window density estimation*, which has the property of mean-square convergence to the true PDF under certain conditions [37]. From (32), the optimization problem in (28) can be stated as

$$\begin{aligned} \max_{\{\mu_i, f_i, \sigma_i\}_{i=1}^L} \quad & \min_{\theta_1 \in \Lambda_1} \sum_{i=1}^L \mu_i \tilde{f}_{\theta_1, i} \\ \text{subject to} \quad & \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^L \mu_i \tilde{g}_{\theta_0, i} \leq \tilde{\alpha} \\ & \sum_{i=1}^L \mu_i = 1 \\ & \mu_i \geq 0, \quad i = 1, \dots, L \end{aligned} \quad (33)$$

where  $\sigma_i$  represents the parameter of the  $i$ th window function  $\psi_i(\cdot)$ ,  $\tilde{f}_{\theta_1, i} = \int f_{\theta_1} \psi_i(f - f_i) df$ , and  $\tilde{g}_{\theta_0, i} = \int g_{\theta_0} \psi_i(f - f_i) df$ . Similar to the solution of (29), the PSO approach, for example, can be used to obtain the optimal solution of (33). Also, the approximate convex solution technique can be employed as in (30) and (31) when  $\sigma_i = \sigma \forall i$  is considered as a pre-determined value. Numerical examples are provided in the next section.

### VI. NUMERICAL RESULTS

In this section, a composite version of the detection example in [17] and [22] is studied in order to illustrate the theoretical results obtained in the previous sections. Namely, the following composite hypothesis-testing problem is considered:

$$\begin{aligned} \mathcal{H}_0 : x &= w \\ \mathcal{H}_1 : x &= A + w \end{aligned} \quad (34)$$

where  $A$  is a known constant, and  $w$  is the noise term that has a Gaussian mixture distribution specified as

$$p_w(w) = \frac{1}{2} \gamma(w; -\theta, \sigma^2) + \frac{1}{2} \gamma(w; \theta, \sigma^2), \quad (35)$$

with  $\gamma(w; \theta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(w-\theta)^2}{2\sigma^2}\right\}$ . The PDF of noise  $w$  has an unknown parameter  $\theta$ , which belongs to  $\Lambda_0$  under hypothesis  $\mathcal{H}_0$  and to  $\Lambda_1$  under  $\mathcal{H}_1$ .

From (34) and (35), the probability distributions of observation  $x$  under hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are given, respectively, by

$$p_{\theta_0}(x) = \frac{1}{2} \gamma(x; -\theta_0, \sigma^2) + \frac{1}{2} \gamma(x; \theta_0, \sigma^2), \quad (36)$$

$$p_{\theta_1}(x) = \frac{1}{2} \gamma(x; -\theta_1 + A, \sigma^2) + \frac{1}{2} \gamma(x; \theta_1 + A, \sigma^2). \quad (37)$$

Since additional independent noise can improve the performance of suboptimal detectors only [22], a suboptimal sign detector, as in [17], is considered as the decision rule for the problem in (34), which is given by

$$\phi(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}. \quad (38)$$

Then, from (36)-(38), the probabilities of detection and false alarm when constant noise is added can be calculated, respectively, as (c.f. (7) and (8))

$$\begin{aligned} F_{\theta_1}(x) &= \frac{1}{2} Q\left(\frac{-x + \theta_1 - A}{\sigma}\right) + \frac{1}{2} Q\left(\frac{-x - \theta_1 - A}{\sigma}\right) \\ G_{\theta_0}(x) &= \frac{1}{2} Q\left(\frac{-x + \theta_0}{\sigma}\right) + \frac{1}{2} Q\left(\frac{-x - \theta_0}{\sigma}\right), \end{aligned} \quad (39)$$

where  $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$  is the  $Q$ -function. It is noted that both  $F_{\theta_1}(x)$  and  $G_{\theta_0}(x)$  are monotone increasing functions of  $x$  for all parameter values.

The aim is to add noise  $n$  to observation  $x$  in (34), and to improve the detection performance of the sign detector in (38) under a false alarm constraint. The noise-modified observation is denoted as  $y = x + n$ , and the probabilities of detection and false alarm are given by

$$\begin{aligned} P_D^y(\theta_1) &= \int_{-\infty}^\infty F_{\theta_1}(x) p_n(x) dx, \\ P_F^y(\theta_0) &= \int_{-\infty}^\infty G_{\theta_0}(x) p_n(x) dx, \end{aligned} \quad (40)$$

where  $p_n(\cdot)$  represents the PDF of the additional noise.

#### A. Scenario-1: $\Lambda_0$ and $\Lambda_1$ have finite number of elements

In the first scenario, the parameter sets under  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are specified as  $\theta_0 \in \Lambda_0 = \{0.1, 0.4\}$  and  $\theta_1 \in \Lambda_1 = \{2, 2.5, 4\}$ . According to Proposition 4, the optimal additional noise has a PDF of the form  $p_n(x) = \sum_{i=1}^5 \lambda_i \delta(x - n_i)$ . Then, the probabilities of detection and false alarm in (40) become

$$\begin{aligned} P_D^y(\theta_1) &= \sum_{i=1}^5 \frac{\lambda_i}{2} \left[ Q\left(\frac{-n_i + \theta_1 - A}{\sigma}\right) + Q\left(\frac{-n_i - \theta_1 - A}{\sigma}\right) \right] \\ P_F^y(\theta_0) &= \sum_{i=1}^5 \frac{\lambda_i}{2} \left[ Q\left(\frac{-n_i + \theta_0}{\sigma}\right) + Q\left(\frac{-n_i - \theta_0}{\sigma}\right) \right]. \end{aligned} \quad (41)$$

For the first simulations,  $A = 1$  and  $\sigma = 1$  are used. The original detection probability (i.e., in the absence of additional

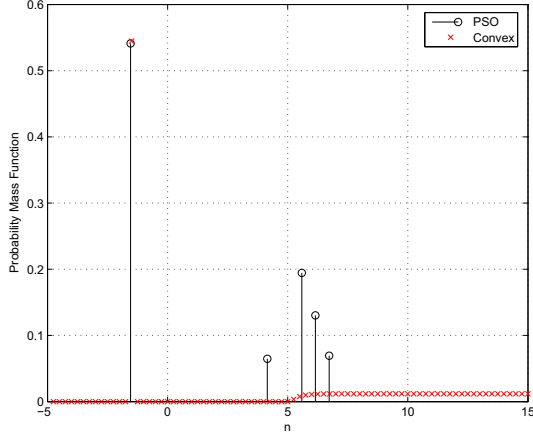


Fig. 2. Probability mass functions of the optimal additional noise based on the PSO and the convex optimization techniques for  $A = 1$  and  $\sigma = 1$ .

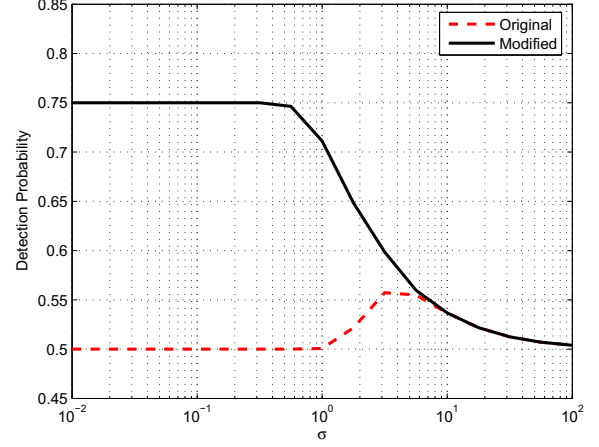


Fig. 3. Comparison of the original and the modified detection probabilities for various values of  $\sigma$ .

noise) can be calculated from (39) as  $P_{D,\min}^x = 0.5007$ , with  $\max_{\theta_0} P_F^x(\theta_0) = \alpha = \tilde{\alpha} = 0.5$ . Then, the PSO and the convex optimization techniques are applied as described in Section V, and the optimal additional noise PDFs are calculated as illustrated in Fig. 2. For the convex solution, the optimization is performed over the noise values that are specified as  $-15 + 0.25l$  for  $l = 0, 1, \dots, 120$ . The resulting detection probability when the PSO algorithm is used is calculated as  $P_{D,\text{mm}}^y = 0.711$  under the constraint that  $\max_{\theta_0} P_F^y(\theta_0) = 0.5$ . In other words, an improvement ratio of  $0.711/0.5007 = 1.42$  is obtained. When the convex relaxation approach is employed, the detection probability becomes  $P_{D,\text{mm}}^y = 0.711$ , which is the same as that obtained by the PSO technique. It is noted from Fig. 2 that the convex solution approximates the optimal PSO solution with 5 mass points with a larger number of non-zero mass points.

Next,  $A = 1$  is used, and the detection probabilities are plotted in Fig. 3 for various values of  $\sigma$  in (35) in the absence and in the presence of additional noise.<sup>1</sup> It is observed that the improvement via additional independent noise increases as  $\sigma$  decreases, and the detector becomes non-improvable for large  $\sigma$  values.

Fig. 4 illustrates the sufficient condition in Proposition 1 with respect to  $\sigma$ . It is obtained that the improvement is guaranteed in the interval  $\sigma \in [0.3981, 3.978]$ , where  $H_{\min}''(\alpha)$  is positive. Comparison of Fig. 4 with Fig. 3 reveals that whenever the second derivative is positive, the detector is improvable as stated in Proposition 1; however, it also indicates that the condition in Proposition 1 is not necessary, as the detector can be improved also for smaller  $\sigma$  values.

#### B. Scenario-2: $\Lambda_0$ and $\Lambda_1$ are continuous intervals

In the second scenario,  $\Lambda_0 = [0.1, 0.4]$  and  $\Lambda_1 = [2, 5]$  are used. As discussed in Section V, an approximation to

<sup>1</sup>The PSO technique is employed in this case.

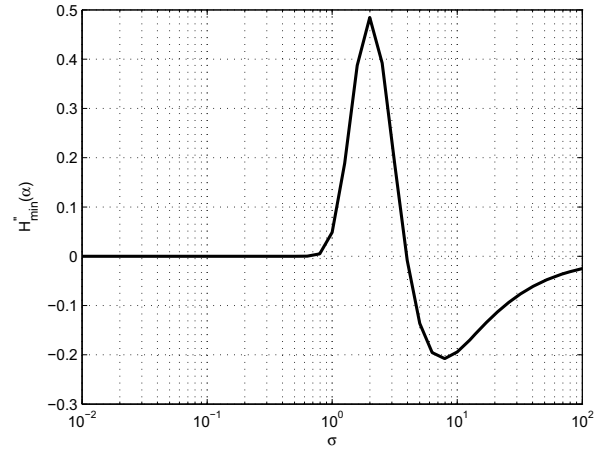


Fig. 4. The second-order derivative  $H_{\min}''(t)$  in (13) at  $t = \alpha$  for various values of  $\sigma$ . Proposition 1 implies that the detector is improvable whenever the second-order derivative at  $t = \alpha$  is positive.

the optimal noise PDF as in (32) can be used to obtain an approximate solution in such a scenario. Considering Gaussian window functions for PDF approximation, the noise PDF can be expressed as<sup>2</sup>

$$p_n(x) = \sum_{i=1}^L \mu_i \gamma(x; \eta_i, \sigma_i^2). \quad (42)$$

Then, the probabilities of detection and false alarm can be calculated from (40), after some manipulation, as

<sup>2</sup>Since scalar observations are considered in this example, the optimization problem can also be solved in the original noise domain, instead of the detection probability domain as in (28).



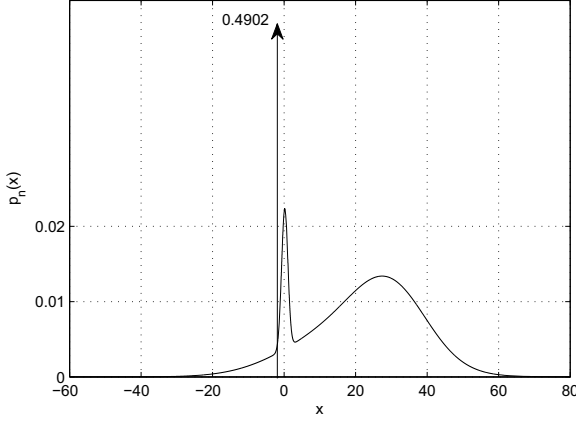


Fig. 5. The optimal noise PDF in (42) for  $A = 1$  and  $\sigma = 1$ . The optimal parameters in (42) obtained via the PSO algorithm are  $\mu = [0.0067 \ 0.1797 \ 0.0411 \ 0.2262 \ 0.0064 \ 0.0498 \ 0 \ 0.4902]$ ,  $\eta = [20.10 \ 15.03 \ 0.1815 \ 29.97 \ 17.27 \ 22.81 \ -0.7561 \ -1.448]$ , and  $\sigma = [16.52 \ 15.14 \ 0.8805 \ 10.16 \ 12.91 \ 17.42 \ 19.19 \ 0.0102]$ . The mass center  $\eta_i = -1.448$  is marked by an arrow for convenience as it has a very small variance.

$$\begin{aligned} P_D^y(\theta_1) &= \sum_{i=1}^L \frac{\mu_i}{2} \left[ Q \left( \frac{-\theta_1 - \eta_i - A}{\sqrt{\sigma^2 + \sigma_i^2}} \right) + Q \left( \frac{\theta_1 - \eta_i - A}{\sqrt{\sigma^2 + \sigma_i^2}} \right) \right] \\ P_F^y(\theta_0) &= \sum_{i=1}^L \frac{\mu_i}{2} \left[ Q \left( \frac{-\theta_0 - \eta_i}{\sqrt{\sigma^2 + \sigma_i^2}} \right) + Q \left( \frac{\theta_0 - \eta_i}{\sqrt{\sigma^2 + \sigma_i^2}} \right) \right]. \end{aligned} \quad (43)$$

For the following simulations,  $L = 8$  is considered, and the parameters  $\{\mu_i, \eta_i, \sigma_i\}_{i=1}^8$  are obtained via the PSO algorithm. First,  $A = 1$  and  $\sigma = 1$  are used. In the absence of additional noise, the probability of detection is given by  $\min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1) = \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(0) = 0.5$  with  $\max_{\theta_0 \in \Lambda_0} P_F^x(\theta_0) = \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(0) = \alpha = \tilde{\alpha} = 0.5$ . When the optimal additional noise PDFs are calculated via the PSO algorithm, the probability of detection becomes  $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1) = 0.6943$ . In other words, an improvement ratio of 1.389 is obtained. The optimal noise PDF is illustrated in Fig. 5.

In Fig. 6, the probabilities of detection are plotted for both the original detector (i.e., without additional noise) and the noise-modified one for  $A = 1$ . Similar to the first scenario, more improvement can be achieved as  $\sigma$  decreases, and no improvement is observed for large values of  $\sigma$ .

Finally, the improvability condition in Proposition 1 is investigated in Fig. 7. It is observed that the detector is improvable in the interval  $\sigma \in [0.5012, 4.996]$ , which together with Fig. 6 imply that the conditions in the propositions are sufficient but not necessary.

## VII. CONCLUDING REMARKS

In this paper, the effects of additional independent noise have been investigated for composite hypothesis-testing prob-

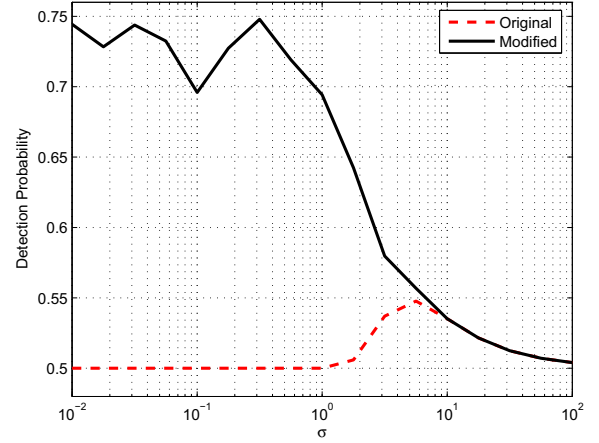


Fig. 6. Comparison of the detection probabilities for the original and the modified detectors for various values of  $\sigma$ .

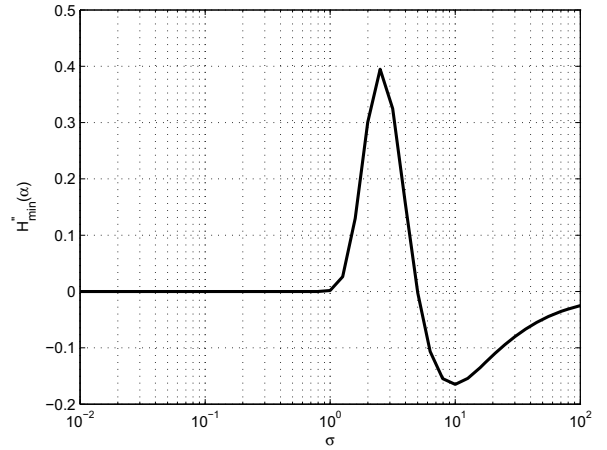


Fig. 7. The second-order derivative  $H_{\min}(t)$  in (13) at  $t = \alpha$  for various values of  $\sigma$ . Proposition 1 implies that the detector is improvable whenever the second-order derivative at  $t = \alpha$  is positive.

lems in the generalized Neyman-Pearson framework. Improvability and non-improvability conditions have been derived, and the statistical characterization of optimal additional noise PDFs has been provided. A detection example has been presented to explain the theoretical results.

## APPENDIX

### A. Proof of Proposition 4

The proof extends the results in [17] and [38] for the two-dimensional case to the  $(M + N)$ -dimensional case. Since the possible additional noise values are specified by  $n_i \in [a_i, b_i]$  for  $i = 1, \dots, K$ ,  $U$  in (26) represents the set of all possible combinations of  $F_{\theta_{1i}}(\mathbf{n})$  and  $G_{\theta_{0j}}(\mathbf{n})$  for  $i = 1, \dots, N$  and  $j = 1, \dots, M$ . Let the convex hull of  $U$  be denoted by set  $V$ . Since  $F_{\theta_{1i}}(\mathbf{n})$  and  $G_{\theta_{0j}}(\mathbf{n})$  are bounded by definition,  $U$  is a



bounded and closed subset of  $\mathbb{R}^{N+M}$  by the assumption in the proposition. Therefore,  $U$  is compact, and the convex hull  $V$  of  $U$  is closed [39]. Also, since  $V \subseteq \mathbb{R}^{N+M}$ , the dimension of  $V$  is smaller than or equal to  $(N + M)$ . Define

$$\begin{aligned} W = \{ & (w_1, \dots, w_{N+M}) : w_1 = E_n\{F_{\theta_{11}}(\mathbf{n})\}, \dots, \\ & w_N = E_n\{F_{\theta_{1N}}(\mathbf{n})\}, w_{N+1} = E_n\{G_{\theta_{01}}(\mathbf{n})\}, \dots, \\ & w_{N+M} = E_n\{G_{\theta_{0M}}(\mathbf{n})\}, \forall p_n(\cdot), \mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b} \} . \end{aligned} \quad (44)$$

Based on [17] and [40], it can be shown that  $W = V$ . Therefore, Carathéodory's theorem [41], [42] implies that any point in  $V$  (hence, in  $W$ ) can be expressed as the convex combination of  $(N + M + 1)$  points in  $U$ . Since an optimal noise PDF must maximize the minimum probability of detection, it corresponds to the boundary of  $V$  [17]. Since  $V$  is closed as discussed above, it always contains its boundary. Therefore, the optimal noise PDF can be expressed as the convex combination of  $(N + M)$  elements in  $U$  [41], [42].  $\square$

## REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, Inc., 1998.
- [3] A. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005.
- [4] M. A. Richards, *Fundamentals of Radar Signal Processing*. USA: McGraw-Hill, Electronic Engineering Series, 2005.
- [5] N. Levanon and E. Mozeson, *Radar Signals*. Wiley-IEEE Pres, 2004.
- [6] R. Benzi, A. Suter, and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A: Math. General*, vol. 14, pp. 453–457, 1981.
- [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, no. 1, 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, no. 2, 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. Mitaïm and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1526–1540, Nov. 2004.
- [13] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, 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. Mitaïm, "Stochastic resonance in noisy threshold neurons," *Neural Netw.*, vol. 16, pp. 755–761, 2003.
- [16] —, "Robust stochastic resonance for simple threshold neurons," *Phys. Rev. E*, vol. 70, no. 031911, 2004.
- [17] 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, no. 7, pp. 3172–3184, July 2007.
- [18] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [19] 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, no. 11, pp. 695–698, Nov. 2006.
- [20] H. Chen and P. K. Varshney, "Theory of the stochastic resonance effect in signal detection: Part II—Variable detectors," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.
- [21] S. Bayram and S. Gezici, "Noise-enhanced  $M$ -ary hypothesis-testing in the minimax framework," in *Proc. 3rd International Conference on Signal Processing and Communication Systems*, Omaha, Nebraska, Sep. 2009.
- [22] S. M. Kay, "Can detectability be improved by adding noise?" *IEEE Sig. Processing Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.
- [23] H. Chen, P. K. Varshney, J. H. Michels, and S. M. Kay, "Approaching near optimal detection performance via stochastic resonance," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, May 2006.
- [24] S. Bayram and S. Gezici, "On the improvability and non-improvability of detection via additional independent noise," *IEEE Sig. Processing Lett.*, 2009.
- [25] S. Zozor and P.-O. Amblard, "On the use of stochastic resonance in sine detection," *Signal Process.*, vol. 7, pp. 353–367, Mar. 2002.
- [26] A. Asdi and A. Tewfik, "Detection of weak signals using adaptive stochastic resonance," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 2, Detroit, Michigan, May 1995, pp. 1332–1335.
- [27] S. Zozor and P.-O. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177–3181, Dec. 2003.
- [28] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Chapman & Hall, 1986.
- [29] J. Cvitanic and I. Karatzas, "Generalized Neyman-Pearson lemma via convex duality," *Bernoulli*, vol. 7, no. 1, pp. 79–97, 2001.
- [30] B. Rudloff and I. Karatzas, "Testing composite hypotheses via convex duality," <http://arxiv.org/abs/0809.4297>, Sep. 2008.
- [31] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems*. IOS Press, 2002, pp. 214–220, in *Intelligent Technologies—Theory and Applications: New Trends in Intelligent Technologies*.
- [32] 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.
- [33] S. Koziel and Z. Michalewicz, "Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization," *Evolutionary Computation*, vol. 7, no. 1, pp. 19–44, 1999.
- [34] X. Hu and R. Eberhart, "Solving constrained nonlinear optimization problems with particle swarm optimization," in *Proc. Sixth World Multiconference on Systemics, Cybernetics and Informatics 2002 (SCI 2002)*, Orlando, FL, 2002.
- [35] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. New York: Springer, 2005.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [37] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley-Interscience, 2000.
- [38] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson signal detection," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, Nevada, Apr. 2008, pp. 3889–3892.
- [39] C. C. Pugh, *Real Mathematical Analysis*. New York: Springer-Verlag, 2002.
- [40] L. Huang and M. J. Neely, "The optimality of two prices: Maximizing revenue in a stochastic network," in *Proc. 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2007.
- [41] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.
- [42] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Boston, MA: Athena Specific, 2003.