Stochastic resonance in binary composite hypothesis-testing problems in the Neyman–Pearson framework ☆

Suat Bayram, Sinan Gezici *

Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey

A B S T R A C T

Performance of some suboptimal detectors can be enhanced by adding independent noise to their inputs via the stochastic resonance (SR) effect. In this paper, the effects of SR are studied for binary composite hypothesis-testing problems. A Neyman–Pearson framework is considered, and the maximization of detection performance under a constraint on the maximum probability of false-alarm is studied. The detection performance is quantified in terms of the sum, the minimum, and the maximum of the detection probabilities corresponding to possible parameter values under the alternative hypothesis. Sufficient conditions under which detection performance can or cannot be improved are derived for each case. Also, statistical characterization of optimal additive noise is provided, and the resulting false-alarm probabilities and bounds on detection performance are investigated. In addition, optimization theoretic approaches to obtaining the probability distribution of optimal additive noise are discussed. Finally, a detection example is presented to investigate the theoretical results.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Stochastic resonance (SR) refers to a physical phenomenon that is observed as an improvement in the output of a nonlinear system when noise level is increased or specific noise is added to the system input [1–15]. Although noise commonly degrades performance of a system, it can also improve performance of some nonlinear systems under certain circumstances. Improvements that can be obtained via noise can be in various forms, such as an increase in output signal-to-noise ratio (SNR) [1–3] or mutual information [8–13], a decrease in the Bayes risk [16–18], or an increase in probability of detection under a constraint on probability of false-alarm [14,15,19–21]. The first study on the SR phenomenon was performed in [1] to explain the periodic recurrence of ice gases. In that work, presence of noise was taken into account in order to explain a natural phenomenon. Since then, the SR concept has been considered in numerous nonlinear systems, such as optical, electronic, magnetic, and neuronal systems [7].

The SR phenomenon has been investigated for hypothesis-testing (detection) problems in recent studies such as [14–30]. By injecting additive noise to the system or by adjusting the noise parameters, performance of some suboptimal detectors can be improved under certain conditions [19,24]. The phenomenon

of improving performance of a detector via noise is also called noise-enhanced detection (NED) [31,32]. Depending on detection performance metrics, additive noise can improve performance of suboptimal detectors according to the Bayesian [16], minimax [20], and Neyman–Pearson [14,15,19,25] criteria. The effects of additive noise on performance of suboptimal detectors are investigated in [16] according to the Bayesian criterion under uniform cost assignment. It is proven that the optimal noise that minimizes the probability of decision error has a constant value, and a Gaussian mixture example is presented to illustrate the improvability of a suboptimal detector via adding constant "noise", which is equivalent to shifting the decision region of the detector. The study in [20] investigates optimal additive noise for suboptimal variable detectors according to the Bayesian and minimax criteria based on the results in [14] and [16].

In the Neyman–Pearson framework, additive noise can be utilized to increase probability of detection under a constraint on probability of false-alarm. In [24], noise effects are investigated for sine detection and it is shown that the conventional incoherent detector can be improved under non-Gaussian noise. In [19], an example is presented to illustrate the effects of additive noise for the problem of detecting a constant signal in Gaussian mixture noise. In [14], a theoretical framework for investigating the effects of additive noise on suboptimal detectors is established according to the Neyman–Pearson criterion. Sufficient conditions are derived for improvability and nonimprovability of a suboptimal detector via additive noise, and it is proven that optimal additive noise can be generated by a randomization of at most two discrete signals, which is an important result since it greatly simplifies the

calculation of the optimal noise probability density function (PDF). An optimization theoretic framework is provided in [15] for the same problem, which also proves the two mass point structure of the optimal additive noise PDF, and, in addition, states that an optimal additive noise may not exist in certain cases.

The results in [14] are extended to variable detectors in [20], and similar conclusions as in the fixed detector case are made. In addition, the theoretical framework in [14] is employed for sequential detection and parameter estimation problems in [33] and [34], respectively. In [33], a binary sequential detection problem is considered, and additive noise that reduces at least one of the expected sample sizes for the sequential detection system is obtained. In [34], improvability of estimation performance via additive noise is illustrated under certain conditions for various estimation criteria, and the form of the optimal noise PDF is derived in each case. The effects of additive noise are studied also for detection of weak sinusoidal signals and for locally optimally detectors. In [26] and [27], detection of a weak sinusoidal signal is considered, and improvements on detection performance are investigated. In addition, [28] focuses on the optimization of noise and detector parameters of locally optimal detectors for the detection a small-amplitude sinusoid in non-Gaussian noise.

The theoretical studies in [14] and [15] on the effects of additive noise on signal detection in the Neyman–Pearson framework consider *simple* binary hypothesis-testing problems in the sense that there exists a single probability distribution (equivalently, one possible value of the unknown parameter) under each hypothesis. The main purpose of this paper is to study *composite* binary hypothesis-testing problems, in which there can be multiple possible distributions, hence, multiple parameter values, under each hypothesis [35]. The Neyman–Pearson framework is considered by imposing a constraint on the *maximum* probability of false-alarm, and three detection criteria are studied [36]. In the first one, the aim is to maximize the sum of the detection probabilities for all possible parameter values under the first (alternative) hypothesis $\mathcal{H}_1$ (*max-sum* criterion), whereas the second one focuses on the maximization of the minimum detection probability among all parameter values under $\mathcal{H}_1$ (*max-min* criterion). Although it is not commonly used in practice, the maximization of the maximum detection probability among all parameter values under $\mathcal{H}_1$ is also studied briefly for theoretical completeness (*max-max* criterion). For all detection criteria, sufficient conditions under which performance of a suboptimal detector can or cannot be improved via additive noise are derived. Also, statistical characterization of optimal additive noise is provided in terms of its PDF structure in each case. In addition, the probability of false-alarm in the presence of optimal additive noise is investigated for the max-sum criterion, and upper and lower bounds on detection performance are obtained for the max-min criterion. Furthermore, optimization theoretic approaches to obtaining the optimal additive noise PDF are discussed for each detection criterion. Both particle swarm optimization (PSO) [37–40] and approximate solutions based on convex relaxation [41] are considered. Finally, a detection example is provided to investigate the theoretical results.

The main contributions of the paper can be summarized as follows:

- Theoretical investigation of the effects of additive noise in binary *composite* hypothesis-testing problems in the Neyman–Pearson framework.
- Extension of the improvability and nonimprovability conditions in [14] for simple hypothesis-testing problems to the composite hypothesis-testing problems.
- Statistical characterization of optimal additive noise according to various detection criteria.
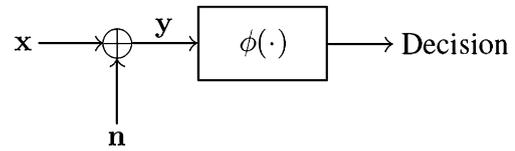

Fig. 1. Independent noise **n** is added to data vector **x** in order to improve the performance of the detector, $\phi(\cdot)$.

- Derivation of upper and lower bounds on the detection performance of suboptimal detectors according to the max-min criterion.
- Optimization theoretic approaches to the calculation of optimal additive noise.

The remainder of the paper is organized as follows. Section 2 describes the composite hypothesis-testing problem, and introduces the detection criteria. Then, Sections 3 and 4 study the effects of additive noise according to the max-sum and the max-min criteria, respectively. In Section 5, the results in the previous sections are extended to the max-max case, and the main implications are briefly summarized. A detection example in provided in Section 6, which is followed by the concluding remarks.

2. Problem formulation and motivation

Consider a binary composite hypothesis-testing problem described as

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

where $\mathcal{H}_i$ denotes the ith hypothesis for $i = 0, 1$. Under hypothesis $\mathcal{H}_i$, data (observation) $\mathbf{x} \in \mathbb{R}^K$ has a PDF indexed by $\theta_i \in \Lambda_i$, namely, $p_{\theta_i}(\mathbf{x})$, where Λ_i is the set of possible parameter values under hypothesis $\mathcal{H}_i$. Parameter sets Λ_0 and Λ_1 are disjoint, and their union forms the parameter space, $\Lambda = \Lambda_0 \cup \Lambda_1$ [35]. In addition, it is assumed that the probability distributions of the parameters are not known *a priori*.

The expressions in (1) present a generic formulation of a binary composite hypothesis-testing problem. Such problems are encountered in various scenarios, such as in radar systems and noncoherent communications receivers [35,42]. In the case that both Λ_0 and Λ_1 consist of single elements, the problem in (1) reduces to a *simple* hypothesis-testing problem [35].

A generic detector (decision rule), denoted by $\phi(\mathbf{x})$, is considered, which maps the data vector into a real number in [0, 1] that represents the probability of selecting $\mathcal{H}_1$ [35]. The aim is to investigate the effects of additive independent noise to the original data, $\mathbf{x}$, of a given detector, as shown in Fig. 1, where $\mathbf{y}$ represents the modified data vector expressed as

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{2}$$

with **n** denoting the additive noise term that is independent of **x**.

The Neyman–Pearson framework is considered in this study, and performance of a detector is specified by its probabilities of detection and false-alarm [35,36,43]. Since the additive noise is independent of the data, the probabilities of detection and false-alarm can be expressed, conditioned on θ_1 and θ_0, respectively, as

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

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

where $p_{\mathbf{n}}(\cdot)$ denotes the PDF of the additive noise. After some manipulation, (3) and (4) can be expressed as [14]

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

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

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

$$F_{\theta_1}(\mathbf{n}) \triangleq \int\limits_{\mathbb{R}^K} \phi(\mathbf{y}) p_{\theta_1}(\mathbf{y} - \mathbf{n}) \, d\mathbf{y}, \tag{7}$$

$$G_{\theta_0}(\mathbf{n}) \triangleq \int\limits_{\mathbb{R}^K} \phi(\mathbf{y}) p_{\theta_0}(\mathbf{y} - \mathbf{n}) \, d\mathbf{y}. \tag{8}$$

Note that $F_{\theta_1}(\mathbf{n})$ and $G_{\theta_0}(\mathbf{n})$ define, respectively, the probability of detection conditioned on θ_1 and the probability of false-alarm conditioned on θ_0 when a constant noise $\mathbf{n}$ is added to the data. Also, in the absence of additive noise, i.e., for $\mathbf{n} = \mathbf{0}$, the probabilities of detection and false-alarm are expressed as $P_D^{\mathbf{x}}(\theta_1) = F_{\theta_1}(\mathbf{0})$ and $P_F^{\mathbf{x}}(\theta_0) = G_{\theta_0}(\mathbf{0})$, respectively, for given values of the parameters.

Various performance metrics can be defined for composite hypothesis-testing problems [35,36]. In the Neyman–Pearson framework, the main constraint is to keep the probability of false-alarm below a certain threshold for all possible parameter values θ_0; i.e.,

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

In most practical cases, the detectors are designed in such a way that they operate at the maximum allowed false-alarm probability $\tilde{\alpha}$ in order to obtain maximum detection probabilities. Therefore, the constraint on the false-alarm probability 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. In other words, in the absence of additive noise $\mathbf{n}$, the detectors commonly operate at the false-alarm probability limit.

Under the constraint in (9), the aim is to maximize a function of the detection probabilities for possible parameter values $\theta_1 \in \Lambda_1$. In this study, the following performance criteria are considered [36]:

- **Max-sum criterion:** In this case, the aim is to maximize $\int_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) \, d\theta_1$, which can be regarded as the "sum" of the detection probabilities for different θ_1 values. This is equivalent to assuming uniform distribution for θ_1 and maximizing the average detection probability [36].
- **Max-min criterion:** According to this criterion, the aim is to maximize the worst-case detection probability, defined as $\min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1)$ [36,43,44]. The worst-case detection probability corresponds to considering the *least-favorable* distribution for θ_1 [36].
- **Max-max criterion:** This criterion maximizes the best-case detection probability, $\max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1)$. This criterion is not very common in practice, since maximizing the detection probability for a single parameter can result in very low detection probabilities for the other parameters. Therefore, this criterion will only be briefly analyzed in Section 5 for completeness of the theoretical results.

There are two main motivations for investigating the effects of additive independent noise in (2) for binary composite hypothesis-testing problems. First, it is important to quantify performance improvements that can be achieved via additive noise, and to determine when additive noise can improve detection performance. In other words, theoretical investigation of SR in binary composite hypothesis-testing problems is of interest. Second, in many cases,

the optimal detector based on the calculation of likelihood functions is challenging to obtain or requires intense computations [14, 35,43,45]. Therefore, a suboptimal detector can be preferable in some practical scenarios. However, the performance of a suboptimal detector may need to be enhanced in order to meet certain system requirements. One way to enhance the performance of a suboptimal detector without changing the detector structure is to modify its original data as in Fig. 1 [14]. Even though calculation of optimal additive noise causes a complexity increase for the suboptimal detector, the overall computational complexity is still considerably lower than that of an optimal detector based on likelihood function calculations. This is because the optimal detector needs to perform intense calculations for each decision whereas the suboptimal detector with modified data needs to update the optimal additive noise whenever the statistics of the hypotheses change. For instance, in a binary communications system, the optimal detector needs to calculate the likelihood ratio for each symbol, whereas a suboptimal detector as in Fig. 1 needs to update $\mathbf{n}$ only when the channel statistics change, which can be constant over a large number of symbols for slowly varying channels [46].

3. Max-sum criterion

In this section, the aim is to determine the optimal additive noise $\mathbf{n}$ in (2) that solves the following optimization problem.

$$\max_{p_{\mathbf{n}}(\cdot)} \int\limits_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) \, d\theta_1, \tag{10}$$

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

where $P_D^{\mathbf{y}}(\theta_1)$ and $P_F^{\mathbf{y}}(\theta_0)$ are as in (5)–(8). Note that the problem in (10) and (11) can also be regarded as a *max-mean* problem since the objective function in (10) can be normalized appropriately so that it defines the average detection probability assuming that all θ_1 parameters are equally likely [36].[1]

From (5) and (6), the optimization problem in (10) and (11) can also be expressed as

$$\max_{p_{\mathbf{n}}(\cdot)} E_{\mathbf{n}}\{F(\mathbf{n})\}, \tag{12}$$

$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} E_{\mathbf{n}}\{G_{\theta_0}(\mathbf{n})\} \leqslant \tilde{\alpha} \tag{13}$$

where $F(\mathbf{n})$ is defined by

$$F(\mathbf{n}) \triangleq \int\limits_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}) \, d\theta_1. \tag{14}$$

Note that $F(\mathbf{n})$ defines the total detection probability for a specific value of additive noise $\mathbf{n}$.

In the following sections, the effects of additive noise are investigated for this max-sum problem, and various results related to optimal solutions are presented.

3.1. Improvability and nonimprovability conditions

According to the max-sum criterion, the detector is called *improvable* if there exists additive independent noise $\mathbf{n}$ that satisfies

$$P_{D,sum}^{\mathbf{y}} \triangleq \int\limits_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) \, d\theta_1 > \int\limits_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1) \, d\theta_1 \triangleq P_{D,sum}^{\mathbf{x}} \tag{15}$$

---

[1] When Λ_1 does not have a finite volume, the max-mean formulation should be used since $\int_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) \, d\theta_1$ may not be finite.

under the false-alarm constraint. From (5) and (14), the condition in (15) can also be expressed as

$$P^{\mathbf{y}}_{D,sum} = E_{\mathbf{n}}\{F(\mathbf{n})\} > F(\mathbf{0}) = P^{\mathbf{x}}_{D,sum}. \tag{16}$$

If the detector cannot be improved, it is called *nonimprovable*.

In order to determine the improvability of a detector according to the max-sum criterion without actually solving the optimization problem in (12) and (13), the approach in [14] for simple hypothesis-testing problems can be extended to composite hypothesis-testing problems in the following manner. First, we introduce the following function

$$H(t) \triangleq \sup\Big\{F(\mathbf{n}) \,\Big|\, \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}) = t, \ \mathbf{n} \in \mathbb{R}^K\Big\}, \tag{17}$$

which defines the maximum value of the total detection probability for a given value of the maximum false-alarm probability. In other words, among all constant noise components $\mathbf{n}$ that achieve a maximum false-alarm probability of t, $H(t)$ defines the maximum probability of detection.

From (17), it is observed that if there exists $t_0 \leqslant \tilde{\alpha}$ such that $H(t_0) > P^{\mathbf{x}}_{D,sum}$, then the system is improvable, since under such a condition there exists a noise component $\mathbf{n}_0$ such that $F(\mathbf{n}_0) > P^{\mathbf{x}}_{D,sum}$ and $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_0) \leqslant \tilde{\alpha}$. Hence, the detector performance can be improved by using an additive noise with $p_{\mathbf{n}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{n}_0)$. However, that condition may not hold in many practical scenarios since, for constant additive noise values, larger total detection probabilities than $P^{\mathbf{x}}_{D,sum}$ are commonly accompanied by false-alarm probabilities that exceed the false-alarm limit. Therefore, a more generic improvability condition is derived in the following theorem.

Theorem 1. *Define the maximum false-alarm probability in the absence of additive noise as $\alpha \triangleq \max_{\theta_0 \in \Lambda_0} P^{\mathbf{x}}_F(\theta_0)$. If $H(t)$ in (17) is second-order continuously differentiable around $t = \alpha$ and satisfies $H''(\alpha) > 0$, then the detector is improvable.*

Proof. Since $H''(\alpha) > 0$ and $H(t)$ in (17) 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 an additive 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 additive noise is shown not to exceed α.

$$\max_{\theta_0 \in \Lambda_0} E_{\mathbf{n}}\{G_{\theta_0}(\mathbf{n})\} \leqslant E_{\mathbf{n}}\Big\{\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n})\Big\}$$
$$= 0.5(\alpha + \epsilon) + 0.5(\alpha - \epsilon) = \alpha. \tag{18}$$

Then, the increase in the detection probability is proven as follows. Due to the assumptions in the theorem, $H(t)$ is convex in an interval around $t = \alpha$. Since $E_{\mathbf{n}}\{F(\mathbf{n})\}$ can attain the value of $0.5H(\alpha + \epsilon) + 0.5H(\alpha - \epsilon)$, which is always larger than $H(\alpha)$ due to convexity, it is concluded that $E_{\mathbf{n}}\{F(\mathbf{n})\} > H(\alpha)$. As $H(\alpha) \geqslant P^{\mathbf{x}}_{D,sum}$ by definition of $H(t)$ in (17), $E_{\mathbf{n}}\{F(\mathbf{n})\} > P^{\mathbf{x}}_{D,sum}$ is satisfied; hence, the detector is improvable. □

Theorem 1 provides a simple condition that guarantees the improvability of a detector according to the max-sum criterion. Note that $H(t)$ is always a single-variable function irrespective of the dimension of the data vector, which facilitates simple evaluations of the conditions in the theorem. However, the main complexity may come into play in obtaining an expression for $H(t)$ in (17) in certain scenarios. An example is presented in Section 6 to illustrate the use of Theorem 1.

In addition to the improvability conditions in Theorem 1, sufficient conditions for nonimprovability can be obtained by defining the following function:

$$J_{\theta_0}(t) \triangleq \sup\big\{F(\mathbf{n}) \,\big|\, G_{\theta_0}(\mathbf{n}) = t, \ \mathbf{n} \in \mathbb{R}^K\big\}. \tag{19}$$

This function is similar to that in [14], but it is defined for each $\theta_0 \in \Lambda_0$ here, since a composite hypothesis-testing problem is considered. Therefore, Theorem 2 in [14] can be extended in the following manner.

Theorem 2. *If there exits $\theta_0 \in \Lambda_0$ and a nondecreasing concave function $\Psi(t)$ such that $\Psi(t) \geqslant J_{\theta_0}(t)$ $\forall t$ and $\Psi(\tilde{\alpha}) = P^{\mathbf{x}}_{D,sum}$, then the detector is nonimprovable.*

Proof. For the θ_0 value in the theorem, the objective function in (12) can be expressed as

$$E_{\mathbf{n}}\{F(\mathbf{n})\} = \int p_{\mathbf{n}}(\mathbf{x})F(\mathbf{x})\,d\mathbf{x} \leqslant \int p_{\mathbf{n}}(\mathbf{x})J_{\theta_0}\big(G_{\theta_0}(\mathbf{x})\big)\,d\mathbf{x}, \tag{20}$$

where the inequality is obtained by the definition in (19).

Since $\Psi(t)$ satisfies $\Psi(t) \geqslant J_{\theta_0}(t)$ $\forall t$, and is concave, (20) becomes

$$E_{\mathbf{n}}\{F(\mathbf{n})\} \leqslant \int p_{\mathbf{n}}(\mathbf{x})\Psi\big(G_{\theta_0}(\mathbf{x})\big)\,d\mathbf{x}$$
$$\leqslant \Psi\bigg(\int p_{\mathbf{n}}(\mathbf{x})G_{\theta_0}(\mathbf{x})\,d\mathbf{x}\bigg). \tag{21}$$

Finally, the nondecreasing property of $\Psi(t)$ together with $\int p_{\mathbf{n}}(\mathbf{x})G_{\theta_0}(\mathbf{x})\,d\mathbf{x} \leqslant \tilde{\alpha}$ implies that $E_{\mathbf{n}}\{F(\mathbf{n})\} \leqslant \Psi(\tilde{\alpha})$. Since $\Psi(\tilde{\alpha}) = P^{\mathbf{x}}_{D,sum}$, $E_{\mathbf{n}}\{F(\mathbf{n})\} \leqslant P^{\mathbf{x}}_{D,sum}$ is obtained for any additive noise $\mathbf{n}$. Hence, the detector is nonimprovable. □

The conditions in Theorem 2 can be used to determine that the detector performance cannot be improved via additive noise, which prevents efforts for solving the optimization problem in (10) and (11).[2] However, it should also be noted that the detector can still be nonimprovable although the conditions in the theorem are not satisfied; that is, Theorem 2 does not provide necessary conditions for nonimprovability.

3.2. Characterization of optimal solution

In this section, the statistical characterization of optimal additive noise components is provided. First, the maximum false-alarm probabilities of optimal solutions are specified. Then, the structures of the optimal noise PDFs are investigated.

In order to investigate the false-alarm probabilities of the optimal solution obtained from (10) and (11) without actually solving the optimization problem, $H(t)$ in (17) can be utilized. Let F_{max} represent the maximum value of $H(t)$, i.e., $F_{max} = \max_t H(t)$. Assume that this maximum is attained at $t = t_m$.[3] Then, one immediate observation is that if t_m is smaller than or equal to the false-alarm limit, i.e., $t_m \leqslant \tilde{\alpha}$, then the noise component $\mathbf{n}_m$ that results in $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_m) = t_m$ is the optimal noise component; i.e., $p_{\mathbf{n}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{n}_m)$. However, in many practical scenarios, the maximum of $H(t)$ is attained for $t_m > \tilde{\alpha}$, since larger detection probabilities can be achieved for larger false-alarm probabilities. In such cases, the following theorem specifies the false-alarm probability achieved by the optimal solution.

---

[2] The optimization problem yields $p_{\mathbf{n}}(\mathbf{x}) = \delta(\mathbf{x})$ when the detector is nonimprovable.

[3] If there are multiple t values that result in the maximum value F_{max}, then the minimum of those values is selected.

Theorem 3. *If $t_m > \tilde{\alpha}$, then the optimal solution of* (10) *and* (11) *satisfies* $\max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) = \tilde{\alpha}$.

Proof. Assume that the optimal solution to (10) and (11) is given by $p_{\tilde{\mathbf{n}}}(\mathbf{x})$ with $\beta \triangleq \max_{\theta_0 \in \Lambda_0} P_F^{\tilde{\mathbf{y}}}(\theta_0) < \tilde{\alpha}$. Define another noise $\mathbf{n}$ with the following PDF:

$$p_{\mathbf{n}}(\mathbf{x}) = \frac{\tilde{\alpha} - \beta}{t_m - \beta} \delta(\mathbf{x} - \mathbf{n}_m) + \frac{t_m - \tilde{\alpha}}{t_m - \beta} p_{\tilde{\mathbf{n}}}(\mathbf{x}), \qquad (22)$$

where $\mathbf{n}_m$ is the noise component that results in the maximum total detection probability; that is, $F(\mathbf{n}_m) = F_{max}$, and t_m is the maximum false-alarm probability when noise $\mathbf{n}_m$ is employed; i.e., $t_m = \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_m)$.

For the noise PDF in (22), the false-alarm and detection probabilities can be obtained as

$$P_{D,sum}^{\mathbf{y}} = E_{\mathbf{n}}\{F(\mathbf{n})\} = \frac{\tilde{\alpha} - \beta}{t_m - \beta} F(\mathbf{n}_m) + \frac{t_m - \tilde{\alpha}}{t_m - \beta} P_{D,sum}^{\tilde{\mathbf{y}}}, \qquad (23)$$

$$P_F^{\tilde{\mathbf{y}}}(\theta_0) = E_{\mathbf{n}}\{G_{\theta_0}(\mathbf{n})\} = \frac{\tilde{\alpha} - \beta}{t_m - \beta} G_{\theta_0}(\mathbf{n}_m) + \frac{t_m - \tilde{\alpha}}{t_m - \beta} P_F^{\tilde{\mathbf{y}}}(\theta_0), \qquad (24)$$

for all $\theta_0 \in \Lambda_0$. Since $F(\mathbf{n}_m) > P_{D,sum}^{\tilde{\mathbf{y}}}$, (23) implies $P_{D,sum}^{\mathbf{y}} > P_{D,sum}^{\tilde{\mathbf{y}}}$. On the other hand, as $G_{\theta_0}(\mathbf{n}_m) \leqslant t_m$ and $P_F^{\tilde{\mathbf{y}}}(\theta_0) \leqslant \beta$, $P_F^{\tilde{\mathbf{y}}}(\theta_0) \leqslant \tilde{\alpha}$ is obtained. Therefore, $\tilde{\mathbf{n}}$ cannot be an optimal solution, which indicates a contradiction. In other words, any noise PDF that satisfies $\max_{\theta_0 \in \Lambda_0} P_F^{\tilde{\mathbf{y}}}(\theta_0) < \tilde{\alpha}$ cannot be optimal. □

The main implication of Theorem 3 is that, in most practical scenarios, the false-alarm probabilities are set to the maximum false-alarm probability limit; i.e., $\max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) = \tilde{\alpha}$, in order to optimize the detection performance according to the max-sum criterion.

Another important characterization of the optimal noise involves the specification of the optimal noise PDF. In [14] and [15], it is shown for simple hypothesis-testing problems that an optimal noise PDF, if exists, can be represented by a randomization of at most two discrete signals. In general, the optimal noise specified by (10) and (11) for the composite hypothesis-testing problem can have more than two mass points. The following theorem specifies the structure of the optimal noise PDF under certain conditions.

Theorem 4. *Let $\theta_0 \in \Lambda_0 = \{\theta_{01}, \theta_{02}, \ldots, \theta_{0M}\}$. Assume that the additive noise components can take finite values specified by $n_i \in [a_i, b_i]$, $i = 1, \ldots, K$, for any finite a_i and b_i. Define set U as*

$$U = \{(u_0, u_1, \ldots, u_M): u_0 = F(\mathbf{n}), u_1 = G_{\theta_{01}}(\mathbf{n}), \ldots, $$
$$u_M = G_{\theta_{0M}}(\mathbf{n}), \text{ for } \boldsymbol{a} \preccurlyeq \mathbf{n} \preccurlyeq \boldsymbol{b}\}, \qquad (25)$$

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

$$p_{\mathbf{n}}(\mathbf{x}) = \sum_{i=1}^{M+1} \lambda_i \delta(\mathbf{x} - \mathbf{n}_i), \qquad (26)$$

where $\sum_{i=1}^{M+1} \lambda_i = 1$ and $\lambda_i \geqslant 0$ for $i = 1, 2, \ldots, M + 1$.

Proof. The proof extends the results in [14] and [15] for the two mass point probability distributions to the $(M + 1)$ mass point ones. Since the possible additive noise components are specified by $n_i \in [a_i, b_i]$ for $i = 1, \ldots, K$, U in (25) represents the set of all possible combinations of $F(\mathbf{n})$ and $G_{\theta_{0i}}(\mathbf{n})$ for $i = 1, \ldots, M$. Let the convex hull of U be denoted by set V. Since $F(\mathbf{n})$ and $G_{\theta_{0i}}(\mathbf{n})$ are bounded by definition, U is a bounded and closed subset of $\mathbb{R}^{M+1}$

by the assumption in the theorem. Therefore, U is compact, and the convex hull V of U is closed [47]. In addition, since $V \subseteq \mathbb{R}^{M+1}$, the dimension of V is smaller than or equal to $(M + 1)$. In addition, define W as the set of all possible total detection and false-alarm probabilities; i.e.,

$$W = \{(w_0, w_1, \ldots, w_M): w_0 = E_{\mathbf{n}}\{F(\mathbf{n})\}, w_1 = E_{\mathbf{n}}\{G_{\theta_{01}}(\mathbf{n})\}, $$
$$\ldots, w_M = E_{\mathbf{n}}\{G_{\theta_{0M}}(\mathbf{n})\}, \forall p_{\mathbf{n}}(\mathbf{n}), \boldsymbol{a} \preccurlyeq \mathbf{n} \preccurlyeq \boldsymbol{b}\}. \qquad (27)$$

Similar to [14] and [48], it can be shown that $W = V$. Therefore, Carathéodory's theorem [49,50] implies that any point in V (hence, in W) can be expressed as the convex combination of $(M + 2)$ points in U. Since an optimal PDF must maximize the total detection probability, it corresponds to the boundary of V [14]. Since V is closed, it always contains its boundary. Therefore, the optimal PDF can be expressed as the convex combination of $(M + 1)$ elements in U. □

In other words, for composite hypothesis-testing problems with a finite number of possible parameter values under hypothesis $\mathcal{H}_0$, the optimal PDF can be expressed as a discrete PDF with a finite number of mass points. Therefore, Theorem 4 generalizes the two mass points result for simple hypothesis-testing problems [14,15]. It should be noted that the result in Theorem 4 is valid irrespective of the number of parameters under hypothesis $\mathcal{H}_1$; that is, Λ_1 in (1) can be discrete or continuous. However, the theorem does not guarantee a discrete PDF if the parameter space for $\mathcal{H}_0$ includes continuous intervals.

Regarding the first assumption in the proposition, constraining the additive noise values as $\boldsymbol{a} \preccurlyeq \mathbf{n} \preccurlyeq \boldsymbol{b}$ is quite realistic since arbitrarily large/small values cannot be realized in practical systems. In other words, in practice, the minimum and maximum possible values of n_i define a_i and b_i, respectively. In addition, the assumption that U is a closed set guarantees the *existence* of the optimal solution [15], and it holds, for example, when F and $G_{\theta_{0j}}$ are continuous functions.

3.3. Calculation of optimal solution and convex relaxation

After the derivation of the improvability and nonimprovability conditions, and the characterization of optimal additive noise in the previous sections, the calculation of optimal noise PDFs is studied in this section.

Let $p_{\mathbf{n}, f}(\cdot)$ represent the PDF of $f = F(\mathbf{n})$, where $F(\mathbf{n})$ is given by (14). Note that $p_{\mathbf{n}, f}(\cdot)$ can be obtained from the noise PDF, $p_{\mathbf{n}}(\cdot)$. As studied in [14], working with $p_{\mathbf{n}, f}(\cdot)$ is more convenient since it results in an optimization problem in a single-dimensional space. Assume that $F(\mathbf{n})$ is a one-to-one function.[4] Then, for a given value of noise $\mathbf{n}$, the false-alarm probabilities in (8) can be expressed as $g_{\theta_0} = G_{\theta_0}(F^{-1}(f))$, where $f = F(\mathbf{n})$. Therefore, the optimization problem in (10) and (11) can be stated as

$$\max_{p_{\mathbf{n}, f}(\cdot)} \int_0^\infty f p_{\mathbf{n}, f}(f) \, df,$$

$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} \int_0^\infty g_{\theta_0} p_{\mathbf{n}, f}(f) \, df \leqslant \tilde{\alpha}. \qquad (28)$$

Note that since $p_{\mathbf{n}, f}(\cdot)$ specifies a PDF, the optimization problem in (28) has also implicit constraints that $p_{\mathbf{n}, f}(f) \geqslant 0 \ \forall f$ and $\int p_{\mathbf{n}, f}(f) \, df = 1$.

---

[4] Similar to the approach in [14], the one-to-one assumption can be removed. However, it is employed in this study to obtain convenient expressions.

In order to solve the optimization problem in (28), first consider the case in which the unknown parameter θ_0 under hypothesis $\mathcal{H}_0$ can take finitely many values specified by $\theta_0 \in \Lambda_0 = \{\theta_{01}, \theta_{02}, \ldots, \theta_{0M}\}$. Then, the optimal noise PDF has $(M+1)$ mass points, under the conditions in Theorem 4. Hence, (28) can be expressed as

$$\max_{\{\lambda_i, f_i\}_{i=1}^{M+1}} \sum_{i=1}^{M+1} \lambda_i f_i,$$

$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^{M+1} \lambda_i g_{\theta_0, i} \leqslant \tilde{\alpha},$$

$$\sum_{i=1}^{M+1} \lambda_i = 1,$$

$$\lambda_i \geqslant 0, \quad i = 1, \ldots, M+1 \tag{29}$$

where $f_i = F(\mathbf{n}_i)$, $g_{\theta_0, i} = G_{\theta_0}(F^{-1}(f_i))$, and $\mathbf{n}_i$ and λ_i are the optimal mass points and their weights as specified in Theorem 4. Note that the optimization problem in (29) may not be formulated as a convex optimization problem in general since $g_{\theta_0, i} = G_{\theta_0}(F^{-1}(f_i))$ may be non-convex. Therefore, global optimization algorithms, such as PSO [37–40], genetic algorithms and differential evolution [51], can be employed to obtain the optimal solution. In this study, the PSO approach is used since it is based on simple iterations with low computational complexity and has been successfully applied to numerous problems in various fields [52–56]. In Section 6, the PSO technique is applied to this optimization problem, which results in accurate calculation of the optimal additive noise in the specified scenario (please refer to [37–40] for detailed descriptions of the PSO algorithm).

Another approach to solve the optimization problem in (29) is to perform *convex relaxation* [41] of the problem. To that end, assume that $f = F(\mathbf{n})$ can take only finitely many known (predetermined) values $\tilde{f}_1, \ldots, \tilde{f}_{\tilde{M}}$. In that case, the optimization can be performed only over the weights $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{\tilde{M}}$ corresponding to those values. Then, (29) can be expressed as

$$\max_{\tilde{\lambda}} \tilde{\mathbf{f}}^T \tilde{\lambda},$$

$$\text{subject to} \quad \tilde{\mathbf{g}}_{\theta_0}^T \tilde{\lambda} \leqslant \tilde{\alpha}, \quad \forall \theta_0 \in \Lambda_0,$$

$$\mathbf{1}^T \tilde{\lambda} = 1,$$

$$\tilde{\lambda} \succcurlyeq \mathbf{0} \tag{30}$$

where $\tilde{\mathbf{f}} = [\tilde{f}_1 \cdots \tilde{f}_{\tilde{M}}]^T$, $\tilde{\lambda} = [\tilde{\lambda}_1 \cdots \tilde{\lambda}_{\tilde{M}}]^T$, and $\tilde{\mathbf{g}}_{\theta_0} = [G_{\theta_0}(F^{-1}(\tilde{f}_1)) \cdots G_{\theta_0}(F^{-1}(\tilde{f}_{\tilde{M}}))]^T$. The optimization problem in (30) is a linearly constrained linear programming (LCLP) problem. Therefore, it can be solved efficiently in polynomial time [41]. Although (30) is an approximation to (29) (since it assumes that $f = F(\mathbf{n})$ can take only specific values), the solutions can get very close to each other as $\tilde{M}$ is increased; i.e., as more values of $f = F(\mathbf{n})$ are included in the optimization problem in (30). Also, it should be noted that the assumption for $F(\mathbf{n})$ to take only finitely many known values can be practical in some cases, since a digital system cannot generate additive noise components with infinite precision due to quantization effects; hence, there can be only finitely many possible values of $\mathbf{n}$. When the computational complexity of the convex problem in (30) is compared with that of (29), which is solved via PSO, it is concluded that the convex relaxation approach can provide significant reductions in the computational complexity. This is mainly because of the fact that functions F and G_{θ_0} need to be evaluated for each particle in each iteration in the PSO algorithm [37–40], which can easily lead to tens of thousands of evaluations in total.

On the other hand, in the convex relaxation approach, these functions are evaluated only once for the possible values of the additive noise, and then the optimal weights are calculated via fast interior point algorithms [41].

For the case in which the unknown parameter θ_0 under hypothesis $\mathcal{H}_0$ can take infinitely many values, the optimal noise may not be represented by $(M+1)$ mass points as in Theorem 4. In that case, an approximate solution is proposed based on PDF approximation techniques. Let the optimal PDF for the optimization problem in (28) be expressed approximately by

$$p_{\mathbf{n}, f}(f) \approx \sum_{i=1}^{L} \mu_i \psi_i(f - f_i), \tag{31}$$

where $\mu_i \geqslant 0$, $\sum_{i=1}^{L} \mu_i = 1$, and $\psi_i(\cdot)$ is a window function that satisfies $\psi_i(\mathbf{x}) \geqslant 0$ $\forall \mathbf{x}$ and $\int \psi_i(\mathbf{x}) \, d\mathbf{x} = 1$, for $i = 1, \ldots, L$. The PDF approximation technique in (31) is called *Parzen window density estimation*, which has the property of mean-square convergence to the true PDF under certain conditions [57]. In general, a larger L facilitates better approximation to the true PDF. A common example of a window function is the Gaussian window, which is expressed as $\psi_i(f) = \exp\{-f^2/(2\sigma_i^2)\}/(\sqrt{2\pi}\,\sigma_i)$. Compared to other approaches such as vector quantization and data clustering, the Parzen window density estimation technique has the advantage that it both provides an explicit expression for the density function and can approximate any density function as accurately as desired as the number of windows are increased.

Based on the approximate PDF in (31), the optimization problem in (28) can be stated as

$$\max_{\{\mu_i, f_i, \sigma_i\}_{i=1}^{L}} \sum_{i=1}^{L} \mu_i \tilde{f}_i,$$

$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^{L} \mu_i \tilde{g}_{\theta_0, i} \leqslant \tilde{\alpha},$$

$$\sum_{i=1}^{L} \mu_i = 1,$$

$$\mu_i \geqslant 0, \quad i = 1, \ldots, L \tag{32}$$

where σ_i represents the parameter[5] of the ith window function $\psi_i(\cdot)$, $\tilde{f}_i = \int_0^\infty f \psi_i(f - f_i) \, df$ and $\tilde{g}_{\theta_0, i} = \int_0^\infty g_{\theta_0} \psi_i(f - f_i) \, df$. Similar to the solution of (29), the PSO approach can be applied to obtain the optimal solution. Also, convex relaxation can be employed as in (30) when $\sigma_i = \sigma$ $\forall i$ is considered as a pre-determined value, and the optimization problem is considered as determining the weights for a number of pre-determined f_i values.

4. Max-min criterion

In this section, the aim is to determine the optimal additive noise $\mathbf{n}$ in (2) that solves the following optimization problem.

$$\max_{p_{\mathbf{n}}(\cdot)} \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1), \tag{33}$$

$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leqslant \tilde{\alpha} \tag{34}$$

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

---

[5] If there are constraints on this parameter, they should be added to the set of constraints in (32).

4.1. Improvability and nonimprovability conditions

According to this criterion, the detector is called *improvable* if there exists additive noise **n** that satisfies

$$\min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) > \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1) = \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{0}) \triangleq P_{D,min}^{\mathbf{x}} \qquad (35)$$

under the false-alarm constraint. Otherwise, the detector is *nonimprovable*.

A simple sufficient condition for improvability can be obtained from the improvability definition in (35). 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 $G_{\theta_0}(\tilde{\mathbf{n}}) \leqslant \tilde{\alpha} \; \forall \theta_0 \in \Lambda_0$, (5) and (6) implies that addition of noise $\tilde{\mathbf{n}}$ to the data vector increases the probability of detection under the false-alarm constraint for all θ_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, where $\tilde{\mathbf{y}} = \mathbf{x} + \tilde{\mathbf{n}}$. However, such a noise component may not be available in many practical scenarios. Therefore, a more generic improvability condition is obtained in the following.

Similar to the max-sum case, the following function is defined for deriving generic improvability conditions:

$$H_{min}(t) \triangleq \sup \left\{ \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}) \; \middle| \; t = \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}), \; \mathbf{n} \in \mathbb{R}^K \right\}, \qquad (36)$$

which defines the maximum value of the minimum detection probability for a given value of the maximum false-alarm probability. From (36), it is observed that if there exists $t_0 \leqslant \tilde{\alpha}$ such that $H_{min}(t_0) > P_{D,min}^{\mathbf{x}}$, the system is improvable, since 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 $\max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}_0) \leqslant \tilde{\alpha}$. Hence, the detector performance can be improved by using an additive noise with $p_{\mathbf{n}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{n}_0)$. However, as stated previously, such a condition may not hold in many practical scenarios. Therefore, a more generic improvability condition is derived in the following theorem.

Theorem 5. *Let* $\alpha = \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{x}}(\theta_0)$ *denote the maximum false-alarm probability in the absence of additive noise. If* $H_{min}(t)$ *in (36) 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 additive 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 additive noise is shown not to exceed α.

$$\max_{\theta_0 \in \Lambda_0} E_{\mathbf{n}}\{G_{\theta_0}(\mathbf{n})\} \leqslant E_{\mathbf{n}}\left\{ \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}) \right\}$$
$$= 0.5(\alpha + \epsilon) + 0.5(\alpha - \epsilon) = \alpha. \qquad (37)$$

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

$$\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}}\{F_{\theta_1}(\mathbf{n})\} \geqslant E_{\mathbf{n}}\left\{ \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{n}) \right\} \qquad (38)$$

is valid for all noise PDFs,

$$\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}}\{F_{\theta_1}(\mathbf{n})\} \geqslant 0.5H_{min}(\alpha + \epsilon) + 0.5H_{min}(\alpha - \epsilon) \qquad (39)$$

can be obtained. Due to the assumptions in the theorem, $H_{min}(t)$ is convex in an interval around $t = \alpha$. Therefore, (39) becomes

$$\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}}\{F_{\theta_1}(\mathbf{n})\} \geqslant 0.5H_{min}(\alpha + \epsilon) + 0.5H_{min}(\alpha - \epsilon)$$
$$> H_{min}(\alpha). \qquad (40)$$

Since $H_{min}(\alpha) \geqslant P_{D,min}^{\mathbf{x}}$ by definition, (40) implies $\min_{\theta_1 \in \Lambda_1} E_{\mathbf{n}}\{F_{\theta_1}(\mathbf{n})\} > P_{D,min}^{\mathbf{x}}$. Therefore, the detector is improvable. □

Similar to Theorem 1 in Section 3.1, Theorem 5 provides a convenient sufficient condition that deals with a scalar function $H_{min}(t)$ irrespective of the dimension of the observation vector.

In order to obtain sufficient conditions for nonimprovability, the following function is defined as an extension of that in (19).

$$J_{\theta_0,\theta_1}(t) \triangleq \sup \left\{ F_{\theta_1}(\mathbf{n}) \; \middle| \; G_{\theta_0}(\mathbf{n}) = t, \; \mathbf{n} \in \mathbb{R}^K \right\}. \qquad (41)$$

Then, the following theorem can be obtained as an extension of Theorem 2 in Section 3.1.

Theorem 6. *Let* θ_1^{min} *represent the value of* $\theta_1 \in \Lambda_1$ *that has the minimum detection probability in the absence of additive noise; that is,*

$$\theta_1^{min} \triangleq \arg \min_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1). \qquad (42)$$

If there exits $\theta_0 \in \Lambda_0$ *and a nondecreasing concave function* $\Psi(t)$ *such that* $\Psi(t) \geqslant J_{\theta_0,\theta_1^{min}}(t) \; \forall t$ *and* $\Psi(\tilde{\alpha}) = P_D^{\mathbf{x}}(\theta_1^{min})$, *then the detector is nonimprovable.*

Proof. If the detector is nonimprovable for $\theta_1 = \theta_1^{min}$, it is nonimprovable according to the max-min criterion, since its minimum can never increase by using additive noise components. Therefore, the result in Theorem 6 directly follows from that in Theorem 2 by considering the nonimprovability conditions at $\theta_1 = \theta_1^{min}$. □

The conditions in Theorem 6 can be used to determine the scenarios in which the detector performance cannot be improved via additive noise. Hence, unnecessary efforts for solving the optimization problem in (33) and (34) can be prevented.

4.2. Characterization of optimal solution

In this section, performance bounds for the detector based on $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where the PDF of $\mathbf{n}$ is obtained from (33) and (34) are derived. In addition, statistical characterization of optimal noise PDFs is provided.

In order to obtain upper and lower bounds on the performance of the detector that employs the noise specified by the optimization problem in (33) and (34), consider a separate optimization problem for each $\theta_1 \in \Lambda_1$ as follows:

$$\max_{p_{\mathbf{n}}(\cdot)} P_D^{\mathbf{y}}(\theta_1),$$
$$\text{subject to} \quad \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leqslant \tilde{\alpha}. \qquad (43)$$

Let $P_{D,opt}^{\mathbf{y}}(\theta_1)$ represent the solution of (43), and $p_{\mathbf{n}_{\theta_1}}(\cdot)$ denote the corresponding optimal PDF. In addition, let $\tilde{\theta}_1$ represent the parameter value with the minimum $P_{D,opt}^{\mathbf{y}}(\theta_1)$ among all $\theta_1 \in \Lambda_1$. That is,

$$\tilde{\theta}_1 = \arg \min_{\theta_1 \in \Lambda_1} P_{D,opt}^{\mathbf{y}}(\theta_1). \qquad (44)$$

Then, the following theorem provides performance bounds for the noise-modified detector according to the max-min criterion.

Theorem 7. *Let* $P_{D,mm}^{\mathbf{y}}$ *represent solution of the optimization problem specified by* (33) *and* (34). *It has the following lower and upper bounds*:

$$\max\left\{\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1),\ \min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)\right\}$$
$$\leqslant P_{D,mm}^{\mathbf{y}}\leqslant\min_{\theta_1\in\Lambda_1}P_{D,opt}^{\mathbf{y}}(\theta_1), \qquad (45)$$

where $P_{D,opt}^{\mathbf{y}}(\theta_1)$ *is the solution of the optimization problem in* (43), $P_D^{\mathbf{x}}(\theta_1)$ *is the probability of detection in the absence of additive noise, and* $P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ *is the probability of detection in the presence of additive noise* $\mathbf{n}_{\tilde{\theta}_1}$, *which is specified by the PDF* $p_{\mathbf{n}_{\tilde{\theta}_1}}(\cdot)$ *that is the optimizer of* (43) *for* $\tilde{\theta}_1$ *defined by* (44).

Proof. The upper bound in (45) directly follows from (33), (34) and (43), since $\max_{p_{\mathbf{n}}(\cdot)}P_D^{\mathbf{y}}(\theta_1)\geqslant\max_{p_{\mathbf{n}}(\cdot)}\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$ for all $\theta_1\in\Lambda_1$. For the lower bound, it is first noted that the noise-modified detector can never have lower minimum detection probability than that in the absence of noise, i.e., $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1)$. In addition, using a noise with PDF $p_{\mathbf{n}_{\tilde{\theta}_1}}(\cdot)$, which is the optimal noise for the problem in (43) for a specific θ_1 value, can never result in a larger minimum probability $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$ than that obtained from the solution of (33) and (34), since the latter directly maximizes the $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$ metric. Therefore, $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ provides another lower bound. $\square$

The main intuition behind the upper and lower bounds in Theorem 7 can be explained as follows. Note that $P_{D,opt}^{\mathbf{y}}(\theta_1)$ represents the maximum detection probability when an additive noise component that is optimized for a specific value of θ_1 is used. Therefore, for each $\theta_1\in\Lambda_1$, $P_{D,opt}^{\mathbf{y}}(\theta_1)$ is larger than $\max_{p_{\mathbf{n}}(\cdot)}\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$, as the latter involves a single additive noise component that is optimized for the minimum detection probability metric and is used for all θ_1 values. In other words, the upper bound is obtained by assuming a more flexible optimization problem in which a different optimal noise component can be used for each θ_1 value. Considering the lower bound, the first lower bound expression is obtained from the fact that the optimal value can never be smaller than $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1)$, which is the minimum detection probability in the absence of additive noise. The second lower bound is obtained from the observation that the optimal noise PDF that maximizes the minimum detection probability, $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$, is obtained from the optimization problem in (33) and (34); hence, the resulting optimal value, $P_{D,mm}^{\mathbf{y}}$, is larger than or equal to all other $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}}(\theta_1)$ values that are obtained by using a different noise PDF.

Both the lower and the upper bounds in Theorem 7 are achievable. For example, when the detector is nonimprovable, the lower bound is achieved since $P_{D,mm}^{\mathbf{y}}=\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1)$ and $P_{D,mm}^{\mathbf{y}}\geqslant\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$. Note that $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ can be smaller than $P_{D,mm}^{\mathbf{y}}$ in certain scenarios since the additive noise $p_{\mathbf{n}_{\tilde{\theta}_1}}(\cdot)$ that is optimized for $\theta_1=\tilde{\theta}_1$ can degrade the detection performance for other θ_1 values. In fact, this is the main reason why a maximum operator in used for the lower bound in Theorem 7. On the other hand, for scenarios in which the detector performance can be improved, $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ can be larger than $\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1)$. Also, in some cases, $\min_{\theta_1\in\Lambda_1}P_{D,opt}^{\mathbf{y}}(\theta_1)=P_{D,mm}^{\mathbf{y}}=\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)\geqslant\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{x}}(\theta_1)$ can be satisfied; that is, the upper and lower bounds in Theorem 7 can be equal. If $P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\tilde{\theta}_1)\leqslant P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ for all $\theta_1\in\Lambda_1$, then $p_{\mathbf{n}_{\tilde{\theta}_1}}(\cdot)$ becomes the optimal PDF for the max-min problem as well, since any other noise PDF will have smaller detection probability than $P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\tilde{\theta}_1)$ at $\theta_1=\tilde{\theta}_1$, and hence will de-

crease the minimum detection probability. In addition, using a different optimal noise for each θ_1 will not improve the max-min performance since $P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\tilde{\theta}_1)$ will be the limiting factor. Therefore, $\min_{\theta_1\in\Lambda_1}P_{D,opt}^{\mathbf{y}}(\theta_1)=\min_{\theta_1\in\Lambda_1}P_D^{\mathbf{y}_{\tilde{\theta}_1}}(\theta_1)$ is satisfied, and the lower and upper bounds become equal in such a case.

Regarding the statistical characterization of the optimal additive noise according to the max-min criterion, it can be shown that when parameter sets Λ_0 and Λ_1 in (1) consist of a finite number of parameters, the optimal additive noise can be represented by a discrete random variable with a finite number of mass points as specified below.

Theorem 8. *Let* $\theta_0\in\Lambda_0=\{\theta_{01},\theta_{02},\ldots,\theta_{0M}\}$ *and* $\theta_1\in\Lambda_1=\{\theta_{11},\theta_{12},\ldots,\theta_{1N}\}$. *Assume that the additive noise components can take finite values specified by* $n_i\in[a_i,b_i]$, $i=1,\ldots,K$, *for any finite* a_i *and* b_i. *Define set* U *as*

$$U=\big\{(u_1,\ldots,u_{N+M}):\ u_1=F_{\theta_{11}}(\mathbf{n}),\ldots,u_N=F_{\theta_{1N}}(\mathbf{n}),$$
$$u_{N+1}=G_{\theta_{01}}(\mathbf{n}),\ldots,u_{N+M}=G_{\theta_{0M}}(\mathbf{n}),\ for\ \boldsymbol{a}\preccurlyeq\mathbf{n}\preccurlyeq\boldsymbol{b}\big\}, \quad (46)$$

where $\boldsymbol{a}\preccurlyeq\mathbf{n}\preccurlyeq\boldsymbol{b}$ *means that* $n_i\in[a_i,b_i]$ *for* $i=1,\ldots,K$. *If* U *is a closed subset of* $\mathbb{R}^{N+M}$, *an optimal solution to* (33) *and* (34) *has the following form*

$$p_{\mathbf{n}}(\mathbf{x})=\sum_{i=1}^{N+M}\lambda_i\delta(\mathbf{x}-\mathbf{n}_i), \qquad (47)$$

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

Proof. The proof is omitted since it is a straightforward extension of that of Theorem 4. $\square$

The main difference of Theorem 8 from Theorem 4 in Section 3.2 is that both Λ_0 and Λ_1 should be discrete for the optimal PDF to have a discrete structure in the max-min framework. However, for the max-sum criterion, it is enough to have a discrete Λ_0 in order to have a discrete PDF as stated in Theorem 4. The reason for this is that according to the max-sum criterion, the objective function to maximize becomes $E_{\mathbf{n}}\{F(\mathbf{n})\}$, where $F(\mathbf{n})=\int_{\theta_1\in\Lambda_1}F_{\theta_1}(\mathbf{n})d\theta_1$ is as defined in (14). In other words, maximization of a single function is considered in the max-sum problem under the false-alarm constraint.

4.3. Calculation of optimal solution and convex relaxation

In this section, possible approaches to solving the optimization problem in (33) and (34) are considered. In order to express the optimization problem as optimization over a single-dimensional PDF, consider a specific value of $\theta_1\in\Lambda_1$, for which $F_{\theta_1}(\mathbf{n})$ is one-to-one. Let this value be represented as $\tilde{\theta}_1$. 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 $f_{\theta_1}=F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(f))$, respectively. Therefore, the optimization problem in (33) and (34) can be reformulated as

$$\max_{p_{\mathbf{n},f_{\tilde{\theta}_1}}(\cdot)}\ \min_{\theta_1\in\Lambda_1}\int_0^1 f_{\theta_1}p_{\mathbf{n},f_{\tilde{\theta}_1}}(f)\,df,$$

$$\text{subject to}\quad \max_{\theta_0\in\Lambda_0}\int_0^1 g_{\theta_0}p_{\mathbf{n},f_{\tilde{\theta}_1}}(f)\,df\leqslant\tilde{\alpha}. \qquad (48)$$

First, consider the case in which the parameters can take finitely many values specified by $\theta_0\in\Lambda_0=\{\theta_{01},\theta_{02},\ldots,\theta_{0M}\}$ and

$\theta_1 \in \Lambda_1 = \{\theta_{11}, \theta_{12}, \ldots, \theta_{1N}\}$. In this case, the optimal noise PDF can be represented by $(N + M)$ mass points under the conditions in Theorem 8. Hence, (48) can be expressed as

$$\max_{\{\lambda_i, f_i\}_{i=1}^{N+M}} \min_{\theta_1 \in \Lambda_1} \sum_{i=1}^{N+M} \lambda_i f_{\theta_1, i},$$

subject to $\quad \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^{N+M} \lambda_i g_{\theta_0, i} \leqslant \tilde{\alpha}$,

$$\sum_{i=1}^{N+M} \lambda_i = 1,$$

$$\lambda_i \geqslant 0, \quad i = 1, \ldots, N + M \qquad (49)$$

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 λ_i are, respectively, the optimal mass points and their weights as specified in Theorem 8. Since the optimization problem in (49) may not be formulated as a convex optimization problem in general, global optimization techniques, such as PSO [37–40] can be employed to obtain the optimal solution, as studied in Section 6.

Due to the complexity of the optimization problem in (49), an approximate and efficient formulation can obtained by the convex relaxation approach as in Section 3.3. Assume that $f = F_{\tilde{\theta}_1}(\mathbf{n})$ can take known values of $\tilde{f}_1, \ldots, \tilde{f}_{\tilde{M}}$ only. In that case, the optimization can be performed only over the weights $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{\tilde{M}}$ corresponding to those values. Hence, (49) becomes

$$\max_{\tilde{\lambda}} \min_{\theta_1 \in \Lambda_1} \tilde{\boldsymbol{f}}_{\theta_1}^T \tilde{\lambda},$$

subject to $\quad \tilde{\boldsymbol{g}}_{\theta_0}^T \tilde{\lambda} \leqslant \tilde{\alpha}, \quad \forall \theta_0 \in \Lambda_0,$

$$\mathbf{1}^T \tilde{\lambda} = 1,$$

$$\tilde{\lambda} \succcurlyeq \mathbf{0} \qquad (50)$$

where $\tilde{\boldsymbol{f}}_{\theta_1} = [F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_1)) \cdots F_{\theta_1}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_{\tilde{M}}))]^T$, $\tilde{\boldsymbol{g}}_{\theta_0} = [G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_1))$ $\cdots G_{\theta_0}(F_{\tilde{\theta}_1}^{-1}(\tilde{f}_{\tilde{M}}))]^T$, and $\tilde{\lambda} = [\tilde{\lambda}_1 \cdots \tilde{\lambda}_{\tilde{M}}]^T$. The optimization problem (50) can be expressed as a convex problem when we define an auxiliary optimization variable t as follows:

$$\max_{\tilde{\lambda}, t} t,$$

subject to $\quad \tilde{\boldsymbol{f}}_{\theta_1}^T \tilde{\lambda} \geqslant t, \quad \forall \theta_1 \in \Lambda_1,$

$$\tilde{\boldsymbol{g}}_{\theta_0}^T \tilde{\lambda} \leqslant \tilde{\alpha}, \quad \forall \theta_0 \in \Lambda_0,$$

$$\mathbf{1}^T \tilde{\lambda} = 1,$$

$$\tilde{\lambda} \succcurlyeq \mathbf{0}. \qquad (51)$$

In fact, (51) can be recognized as an LCLP problem if the new optimization variable is defined as $\mathbf{x} = [\tilde{\lambda}^T \ t]^T$. Therefore, it can be solved efficiently in polynomial time [41]. Although (51) is an approximation to (49), the solutions get very close as more values of $f = F_{\tilde{\theta}_1}(\mathbf{n})$ are included in the optimization.

When at least one of θ_0 or θ_1 can take infinitely many values, the optimal noise may not be represented by a finite number of mass points as in Theorem 8. In such cases, the optimization problem in (48) can be solved over the set of PDF approximations as in Section 3.3. Let the optimal PDF be approximated similarly to (31). Then, the optimization problem in (48) can be stated as

$$\max_{\{\mu_i, f_i, \sigma_i\}_{i=1}^L} \min_{\theta_1 \in \Lambda_1} \sum_{i=1}^L \mu_i \tilde{f}_{\theta_1, i},$$

subject to $\quad \max_{\theta_0 \in \Lambda_0} \sum_{i=1}^L \mu_i \tilde{g}_{\theta_0, i} \leqslant \tilde{\alpha}$,

$$\sum_{i=1}^L \mu_i = 1,$$

$$\mu_i \geqslant 0, \quad i = 1, \ldots, L \qquad (52)$$

where σ_i represents the parameter of the ith 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 (49), the PSO approach can be employed, for example, to obtain the optimal solution of (52). Also, the convex relaxation technique can be employed as in (50) and (51) when $\sigma_i = \sigma \ \forall i$ is considered as a pre-determined value.

5. Max-max criterion

In this section, the aim is to determine the optimal additive noise $\mathbf{n}$ in (2) that solves the following optimization problem.

$$\max_{p_{\mathbf{n}}(\cdot)} \max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1), \qquad (53)$$

subject to $\quad \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leqslant \tilde{\alpha} \qquad (54)$

where $P_D^{\mathbf{y}}(\theta_1)$ and $P_F^{\mathbf{y}}(\theta_0)$ are as in (5)–(8). According to the max-max criterion, the detector is called *improvable* if there exists additive noise $\mathbf{n}$ that satisfies

$$\max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1) > \max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1) = \max_{\theta_1 \in \Lambda_1} F_{\theta_1}(\mathbf{0}) \triangleq P_{D,max}^{\mathbf{x}} \qquad (55)$$

under the false-alarm constraint. Otherwise, the detector is *nonimprovable*.

The results in the previous sections can be extended to cover the max-max case as well. Since the derivations are quite similar, the results for this case are stated without any proofs.

Let θ_1^{max} represent the value of $\theta_1 \in \Lambda_1$ that has the maximum detection probability in the absence of additive noise; that is, $\theta_1^{max} \triangleq \arg\max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1)$. In addition, define

$$H_{\theta_1}(t) \triangleq \sup\left\{ F_{\theta_1}(\mathbf{n}) \ \Big| \ \max_{\theta_0 \in \Lambda_0} G_{\theta_0}(\mathbf{n}) = t, \ \mathbf{n} \in \mathbb{R}^K \right\}. \qquad (56)$$

Then, the detector is improvable if $H_{\theta_1^{max}}(t)$ is second-order continuously differentiable around $t = \alpha$ and satisfies $H_{\theta_1^{max}}''(\alpha) > 0$, where $\alpha \triangleq \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{x}}(\theta_0)$. This result can be proven as in Theorem 1. In fact, it directly follows from the observation that if the detector can be improved for $\theta_1 = \theta_1^{max}$, then the maximum of $\max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{y}}(\theta_1)$ is always larger than $\max_{\theta_1 \in \Lambda_1} P_D^{\mathbf{x}}(\theta_1)$.

A nonimprovability condition can be obtained in a similar way to that in Theorem 6. The detector is nonimprovable if there exits $\theta_0 \in \Lambda_0$ and a nondecreasing concave function $\Psi_{\theta_1}(t)$ such that $\Psi_{\theta_1}(t) \geqslant J_{\theta_0, \theta_1}(t) \ \forall t$ and $\Psi_{\theta_1}(\tilde{\alpha}) = P_D^{\mathbf{x}}(\theta_1)$ for all $\theta_1 \in \Lambda_1$, where $J_{\theta_0, \theta_1}(t)$ is given by (41).

Regarding the structure of the optimal noise PDF for the problem in (53) and (54), consider a composite hypothesis-testing problem with $\theta_0 \in \Lambda_0 = \{\theta_{01}, \theta_{02}, \ldots, \theta_{0M}\}$. Then, it can be concluded that the optimal PDF can be represented by $(M + 1)$ mass points under the conditions in Theorem 4. This follows from the fact that the max-max problem in (53) and (54) can be solved by choosing the PDF that results in the maximum detection probability among the PDFs that solve the following optimization problems:

$$\max_{p_{\mathbf{n}}(\cdot)} P_D^{\mathbf{y}}(\theta_1), \qquad (57)$$

subject to $\quad \max_{\theta_0 \in \Lambda_0} P_F^{\mathbf{y}}(\theta_0) \leqslant \tilde{\alpha} \qquad (58)$

for $\theta_1 \in \Lambda_1$. In other words, the optimal noise PDF can be calculated for each $\theta_1 \in \Lambda_1$ separately, and the noise PDF that yields the maximum detection probability becomes the solution of the max-max problem. Since the structure of each optimization problem is as in the max-sum formulation, Theorem 4 applies to the max-max case as well.

Finally, for the solution of the max-max problem, the approaches in Section 3.3 for the max-sum problem can directly be applied, since the optimization problems in (10)–(11) and (57)–(58) have the same structure.

6. Numerical results

In this section, a composite version of the detection example in [14] and [19] is studied in order to illustrate the theoretical results obtained in the previous sections. Namely, the following composite hypothesis-testing problem is considered:

$$\mathcal{H}_0: \quad x = w,$$
$$\mathcal{H}_1: \quad x = A + w \tag{59}$$

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

$$p_w(w) = \frac{1}{2}\gamma\left(w; -\theta, \sigma^2\right) + \frac{1}{2}\gamma\left(w; \theta, \sigma^2\right), \tag{60}$$

with $\gamma(w; \theta, \sigma^2) = \exp\{-(w - \theta)^2/(2\sigma^2)\}/\sqrt{2\pi\sigma^2}$. The PDF of noise w has an unknown parameter θ, which belongs to Λ_0 under hypothesis $\mathcal{H}_0$ and to Λ_1 under $\mathcal{H}_1$ with $\Lambda_0 \cap \Lambda_1 = \emptyset$.

From (59) and (60), the probability distributions of observation x under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are given, respectively, by

$$p_{\theta_0}(x) = \frac{1}{2}\gamma\left(x; -\theta_0, \sigma^2\right) + \frac{1}{2}\gamma\left(x; \theta_0, \sigma^2\right), \tag{61}$$

$$p_{\theta_1}(x) = \frac{1}{2}\gamma\left(x; -\theta_1 + A, \sigma^2\right) + \frac{1}{2}\gamma\left(x; \theta_1 + A, \sigma^2\right). \tag{62}$$

Since additive noise can improve the performance of suboptimal detectors only [19], a suboptimal sign detector, as in [14], is considered as the decision rule for the problem in (59), which is given by

$$\phi(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leqslant 0. \end{cases} \tag{63}$$

Then, from (61)–(63), detection and false-alarm probabilities when constant noise is added can be calculated as (see (7) and (8))

$$F_{\theta_1}(x) = \int_{-\infty}^{\infty} \phi(y)p_{\theta_1}(y - x)\,dy$$
$$= \frac{1}{2}Q\left(\frac{-x + \theta_1 - A}{\sigma}\right) + \frac{1}{2}Q\left(\frac{-x - \theta_1 - A}{\sigma}\right) \tag{64}$$

and

$$G_{\theta_0}(x) = \int_{-\infty}^{\infty} \phi(y)p_{\theta_0}(y - x)\,dy$$
$$= \frac{1}{2}Q\left(\frac{-x + \theta_0}{\sigma}\right) + \frac{1}{2}Q\left(\frac{-x - \theta_0}{\sigma}\right), \tag{65}$$

respectively, where $Q(x) = (1/\sqrt{2\pi})\int_x^{\infty} e^{-t^2/2}\,dt$ 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 (59), and to improve the detection performance of the sign detector in (63) 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

$$P_D^y(\theta_1) = \int_{-\infty}^{\infty} F_{\theta_1}(x)p_n(x)\,dx,$$

$$P_F^y(\theta_0) = \int_{-\infty}^{\infty} G_{\theta_0}(x)p_n(x)\,dx, \tag{66}$$

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

Remark. In terms of the computational complexity, using a low-complexity suboptimal detector (such as the sign detector) and enhancing its performance via optimal additive noise can be more advantageous than employing the optimal detector in some scenarios. Let the computational complexity of the optimal (suboptimal) detector be denoted by C_o (C_s), and the computational complexity of obtaining the optimal additive noise PDF be represented by C_n. Considering N_d consecutive decisions and assuming that the statistics of the hypotheses do not change over N_s consecutive decisions, the computational complexities of using the optimal decision rule and employing the noise injection approach are given by $N_d C_o$ and $N_d C_s + C_n N_d/N_s$, respectively. Note that the calculation of optimal additive noise PDF needs to be performed only when the statistics of the hypotheses change. Based on these expressions, it is concluded that the noise injection approach is beneficial when $C_o \gg C_s + C_n/N_s$. In the examples below, $C_o \gg C_s$ is satisfied. Therefore, the noise injection approach can be useful if the statistics do not change rapidly (which depends on the specific application scenario).

6.1. Scenario-1: Λ_0 and Λ_1 with 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 Theorem 4 and Theorem 8, the optimal additive noise has a PDF of the form $p_n(x) = \sum_{i=1}^{N_m} \lambda_i \delta(x - n_i)$, where $N_m = 3$ for the max-sum case, and $N_m = 5$ for the max-min case. For the noise PDF specified as $p_n(x) = \sum_{i=1}^{N_m} \lambda_i \delta(x - n_i)$, the detection and false-alarm probabilities in (66) become

$$P_D^y(\theta_1) = \sum_{i=1}^{N_m} \frac{\lambda_i}{2}\left[Q\left(\frac{-n_i + \theta_1 - A}{\sigma}\right) + Q\left(\frac{-n_i - \theta_1 - A}{\sigma}\right)\right],$$

$$P_F^y(\theta_0) = \sum_{i=1}^{N_m} \frac{\lambda_i}{2}\left[Q\left(\frac{-n_i + \theta_0}{\sigma}\right) + Q\left(\frac{-n_i - \theta_0}{\sigma}\right)\right]. \tag{67}$$

For the first simulations, $A = 1$ and $\sigma = 1$ are used. For the max-sum and max-min cases, the original detection probabilities (i.e., in the absence of additive noise) can be calculated from (64) and (65) as $P_{D,sum}^x = 1.613$ and $P_{D,min}^x = 0.5007$, respectively, with $\max_{\theta_0} P_F^x(\theta_0) = \alpha = \tilde{\alpha} = 0.5$. Then, the PSO[6] and the convex relaxation techniques are applied as described in Sections 3.3 and 4.3, and the optimal additive noise PDFs are calculated for both the max-sum and max-min cases, which are illustrated in Fig. 2 and Fig. 3, respectively. For the convex solutions,

---

[6] In the PSO algorithm, 50 particles and 1000 iterations are employed. In addition, the other parameters are set to $c_1 = c_2 = 2.05$ and $\chi = 0.72984$, and the inertia weight ω is changed from 1.2 to 0.1 linearly with the iteration number. Please refer to [37] for the details of the PSO algorithm and the definitions of the parameters. In the considered examples, the use of 50 particles is observed to provide a good tradeoff between accuracy and computational complexity.
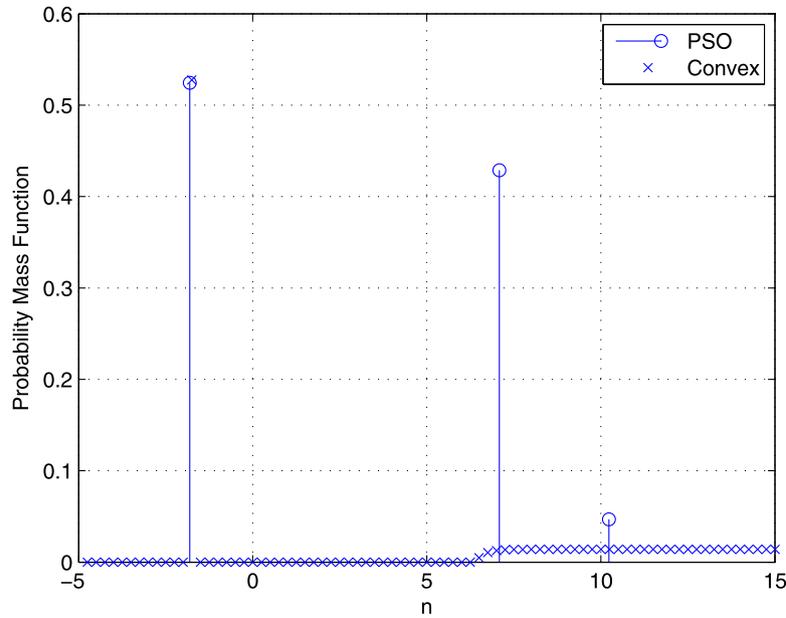
Fig. 2. Probability mass functions of the optimal additive noise based on the PSO and the convex relaxation techniques for the max-sum case when $A = 1$ and $\sigma = 1$.
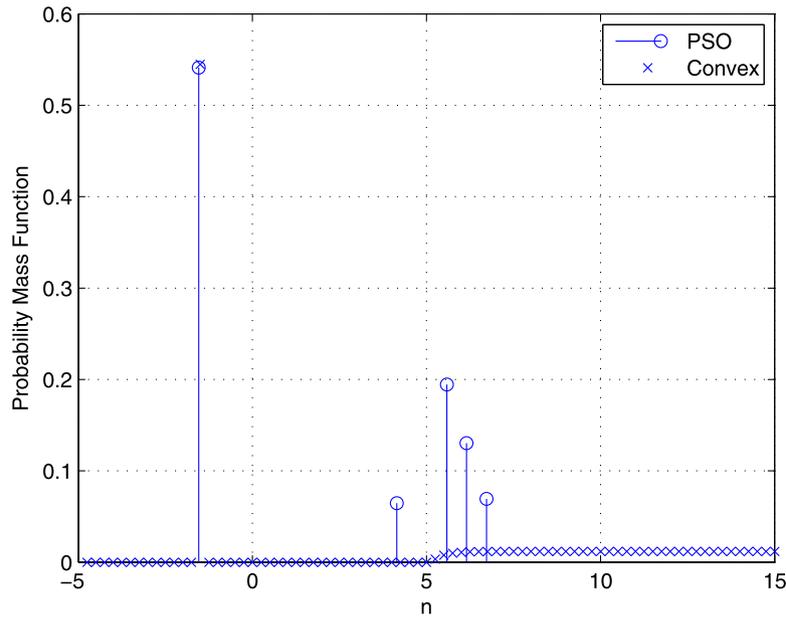

Fig. 3. Probability mass functions of the optimal additive noise based on the PSO and the convex relaxation techniques for the max-min case when $A = 1$ and $\sigma = 1$.

the optimizations are performed over the noise values that are specified as $-15 + 0.25i$ for $i = 0, 1, \ldots, 120$. The resulting detection probabilities when the PSO algorithm is used are calculated as $P_{D,sum}^y = 2.172$ and $P_{D,mm}^y = 0.711$ under the constraint that $\max_{\theta_0} P_F^y(\theta_0) = 0.5$. In other words, improvement ratios of $2.172/1.613 = 1.347$ and $0.711/0.5007 = 1.420$ are obtained according to the max-sum and max-min criteria, respectively. When the convex relaxation approach is employed, the detection probabilities become $P_{D,sum}^y = 2.171$ and $P_{D,mm}^y = 0.711$, which are almost the same as those obtained by the PSO technique. It is noted from Fig. 2 and Fig. 3 that the convex solutions approximate the optimal PSO solutions with 3 and 5 mass points (for the max-sum and max-min cases, respectively) with a larger number of non-zero mass points. This is mainly due to the fact that the possible additive noise values are fixed and the optimization is performed only over the probabilities of those fixed values in the convex approach.

(Hence, Theorems 4 and 8 do not apply in general for the convex solution.)

Next, $A = 1$ is used, and the detection probabilities are plotted versus σ in Fig. 4 in the absence and in the presence of additive noise (labeled as "original" and "SR", respectively) for both the max-sum and the max-min criteria.[7] In the figure, the normalized ('average') detection probabilities are plotted for the max-sum criterion, which is defined as $P_{D,sum}^y/3$ since there are three possible values of θ_1. On the other hand, the minimum (worst-case) detection probabilities are illustrated for the max-min criterion. It is observed from the figure that the improvement via additive noise decreases as σ increases. This is mainly due to the fact that the improvability is commonly caused by the multi-modal nature of the measurement noise PDF in (60), which reduces as σ

---

[7] The PSO technique is employed for calculating the optimal additive noise PDFs.
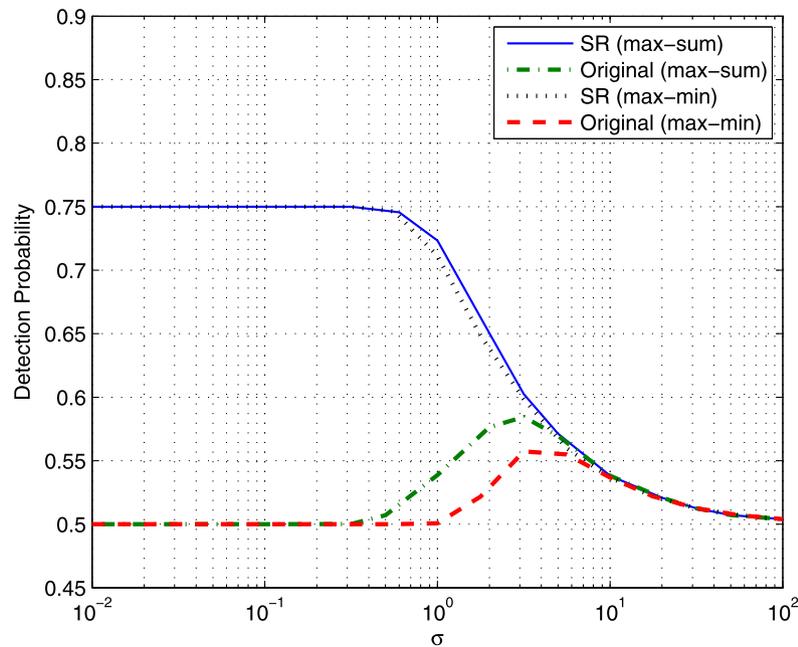
Fig. 4. Detection probabilities in the absence ("original") and presence ("SR") of additive noise according to the max-sum and max-min criteria for various values of σ. Normalized and minimum detection probabilities are illustrated for the max-sum and max-min criteria, respectively.
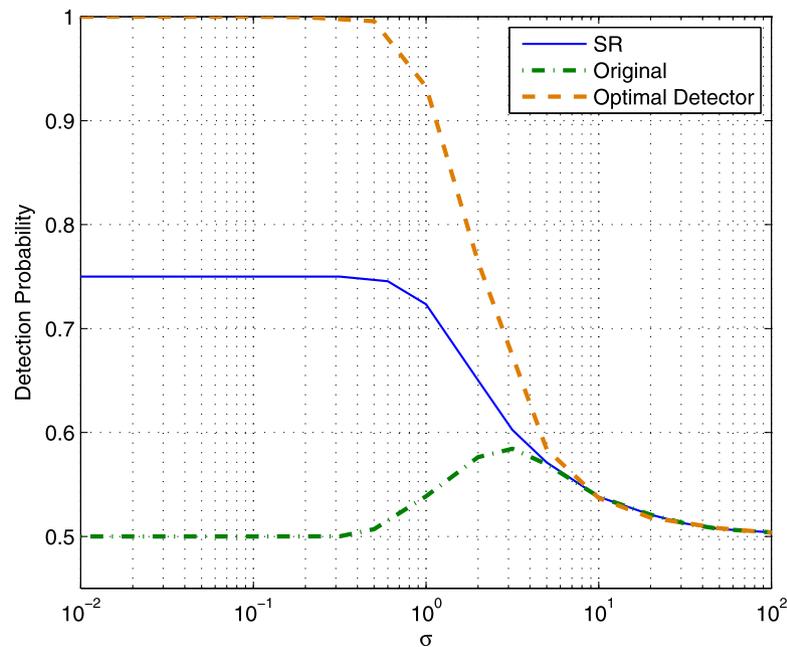

Fig. 5. Normalized detection probabilities versus σ for the max-sum criterion considering the sign detector in the absence ("original") and presence ("SR") of additive noise, and the optimal detector.

increases. In addition, as expected, the worst-case (minimum) detection probabilities in the max-min criterion are lower than or equal to the average detection probabilities in the max-sum criterion.

In order to compare the performance of the sign detector in the presence and absence of optimal additive noise to the performance of the optimal detector, Fig. 5 is presented for the max-sum criterion based on the same parameters as those for Fig. 4. It is observed that the optimal detector achieves higher detection probabilities for low values of σ, and the detection probabilities get close for high values of σ. However, as discussed previously, the noise injection approach can still be preferred in practice in order to reduce the computational complexity since the optimal detector

has significantly higher complexity than the sign detector in this example.

Fig. 6 illustrates the sufficient conditions in Theorem 1 and Theorem 5 for the max-sum and max-min cases with respect to σ. It is obtained that the improvement is guaranteed in the interval $\sigma \in [0.1259, 2.639]$ for the max-sum case and in the interval $\sigma \in [0.3981, 3.978]$ for the max-min case. Comparison of Fig. 6 with Fig. 4 reveals that whenever the second derivative is positive, the detector is improvable as stated in the related theorems; however, it also indicates that the conditions in Theorem 1 and Theorem 5 are not necessary conditions, as the detector can be improved also for smaller σ values.
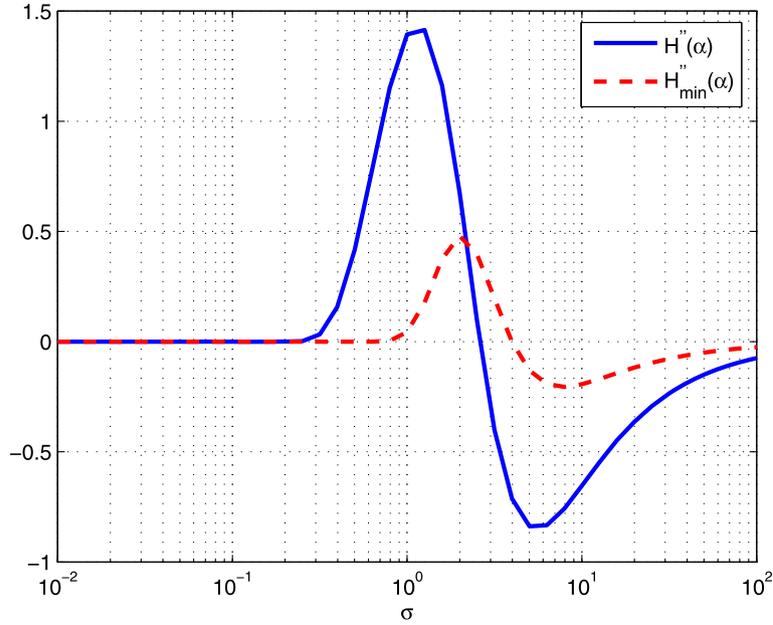
Fig. 6. The second-order derivatives of $H(t)$ in (17) and $H_{\min}(t)$ (36) at $t = \alpha$ for various values of σ. Theorem 1 and Theorem 5 imply that the detector is improvable whenever the second-order derivative at $t = \alpha$ is positive.
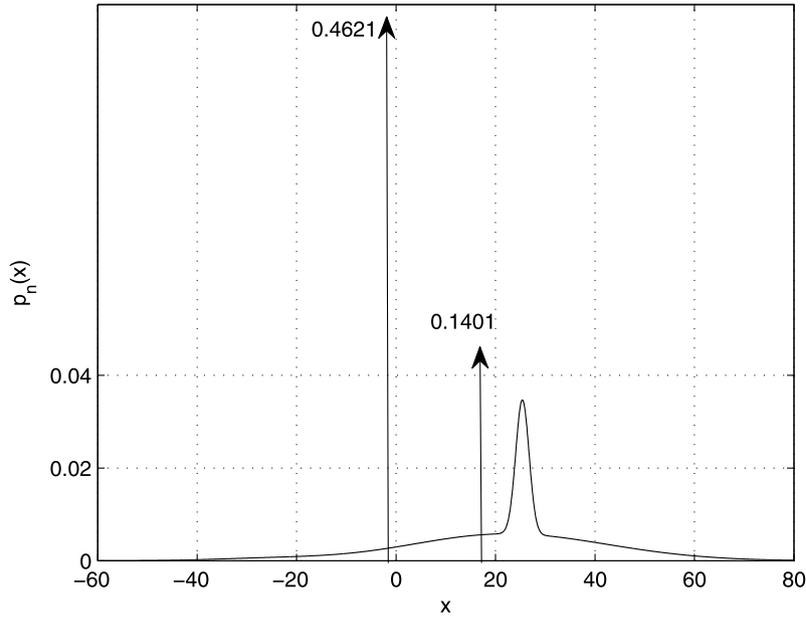

Fig. 7. The optimal additive noise PDF in (68) for $A = 1$ and $\sigma = 1$ according to the max-sum criterion. The optimal parameters in (68) obtained via the PSO algorithm are $\boldsymbol{\mu} = [0.0969\ 0\ 0.0019\ 0.1401\ 0.1377\ 0.0143\ 0.1470\ 0.4621]$, $\boldsymbol{\eta} = [25.4039\ -20.1423\ 13.7543\ 17.0891\ 29.7452\ -25.0785\ 17.6887\ -2.2085]$, and $\boldsymbol{\sigma} = [1.3358\ 26.2930\ 11.3368\ 0\ 19.5556\ 11.5953\ 17.9838\ 0.0001]$. The mass centers with very small variances ($\eta_i = 17.0891$ and $\eta_i = -2.2085$) are marked by arrows for convenience.

6.2. Scenario-2: Λ_0 and Λ_1 are continuous intervals

In the second scenario, $\Lambda_0 = [0.1, 0.4]$ and $\Lambda_1 = [2, 5]$ are used. As discussed in Sections 3.3 and 4.3, an approximation to the optimal additive noise PDF as in (31) can be used to obtain an approximate solution in such a scenario. Considering Gaussian window functions for PDF approximation, the additive noise PDF can be expressed as[8]

$$p_n(x) \approx \sum_{i=1}^{L} \mu_i \gamma\left(x; \eta_i, \sigma_i^2\right). \tag{68}$$

Then, the probabilities of detection and false-alarm can be calculated from (66), after some manipulation, as

$$P_D^y(\theta_1) = \sum_{i=1}^{L} \frac{\mu_i}{2} \left[Q\left(\frac{-\theta_1 - \eta_i - A}{\sqrt{\sigma^2 + \sigma_i^2}}\right) + Q\left(\frac{\theta_1 - \eta_i - A}{\sqrt{\sigma^2 + \sigma_i^2}}\right) \right], \tag{69}$$

$$P_F^y(\theta_0) = \sum_{i=1}^{L} \frac{\mu_i}{2} \left[Q\left(\frac{-\theta_0 - \eta_i}{\sqrt{\sigma^2 + \sigma_i^2}}\right) + Q\left(\frac{\theta_0 - \eta_i}{\sqrt{\sigma^2 + \sigma_i^2}}\right) \right]. \tag{70}$$

---

[8] 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) or (48).
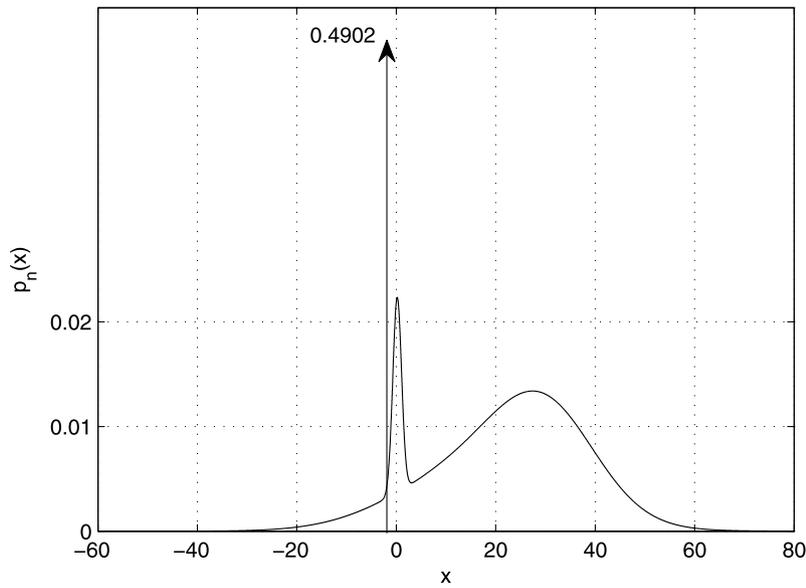
Fig. 8. The optimal additive noise PDF in (68) for the max-min criterion when $A = 1$ and $\sigma = 1$. The optimal parameters in (68) obtained via the PSO algorithm are $\boldsymbol{\mu} = [0.0067\ 0.1797\ 0.0411\ 0.2262\ 0.0064\ 0.0498\ 0\ 0.4902]$, $\boldsymbol{\eta} = [20.1017\ 15.0319\ 0.1815\ 29.9668\ 17.2657\ 22.8092\ -0.7561\ -1.4484]$, and $\boldsymbol{\sigma} = [16.5204\ 15.1445\ 0.8805\ 10.1573\ 12.9094\ 17.4184\ 19.0959\ 0.0102]$. The mass center $\eta_i = -1.4484$ is marked by an arrow for convenience as it has a very small variance.
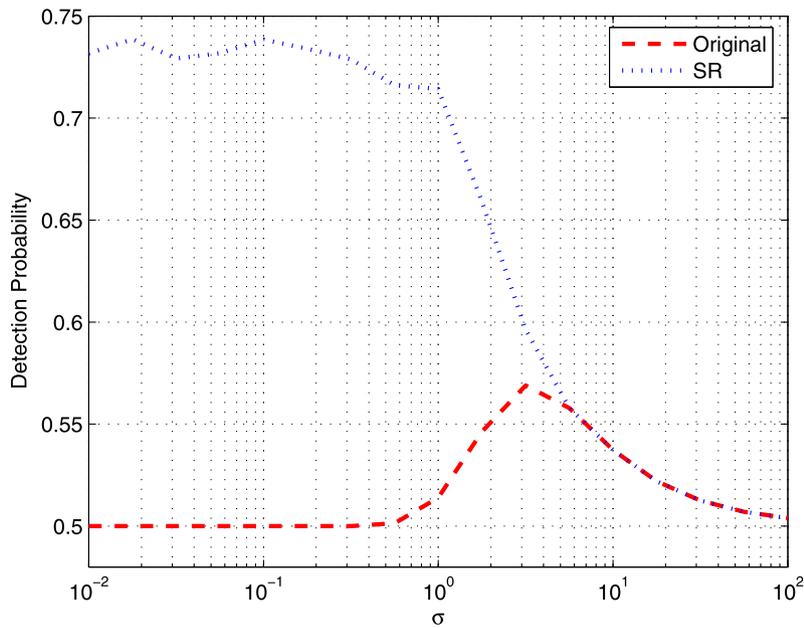

Fig. 9. Comparison of normalized detection probabilities in the absence ("original") and presence ("SR") of additive noise according to the max-sum criterion for various values of σ.

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 for both the max-sum and max-min cases. First, $A = 1$ and $\sigma = 1$ are used. In the absence of additive noise, the detection probabilities in the max-sum and max-min cases are given, respectively, by $\int_{\theta_1 \in \Lambda_1} P_D^x(\theta_1)\, d\theta_1 = \int_{\theta_1 \in \Lambda_1} F_{\theta_1}(0)\, d\theta_1 = 1.5417$ and $\min_{\theta_1 \in \Lambda_1} P_D^x(\theta_1) = \min_{\theta_1 \in \Lambda_1} F_{\theta_1}(0) = 0.5$ 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 additive noise PDFs are calculated via the PSO algorithm, the detection probabilities become $\int_{\theta_1 \in \Lambda_1} P_D^y(\theta_1)\, d\theta_1 = 2.1426$ for the max-sum case, and $\min_{\theta_1 \in \Lambda_1} P_D^y(\theta_1) = 0.6943$ for the max-min case. In other words, improvement ratios of 1.390 and 1.389 are obtained for the max-sum and max-min cases, respectively. The optimal additive noise

PDFs for the max-sum and max-min cases are shown in Figs. 7 and 8, respectively.

In Figs. 9 and 10, the detection probabilities according to the max-sum and max-min criteria are plotted, respectively, for both the original detector (i.e., without additive noise) and the noise-modified one when $A = 1$. For the max-sum case, the detection probability is normalized as $\frac{1}{3} \int_2^5 P_D^y(\theta_1)\, d\theta_1$. Similar to the first scenario, more improvement can be achieved as σ decreases, and no improvement is observed for large values of σ. It is also noted that some fluctuations are present in the detection curves in the noise-modified scenarios, which is mainly due to the fact that the additive noise is optimized based on the approximate model in (68).
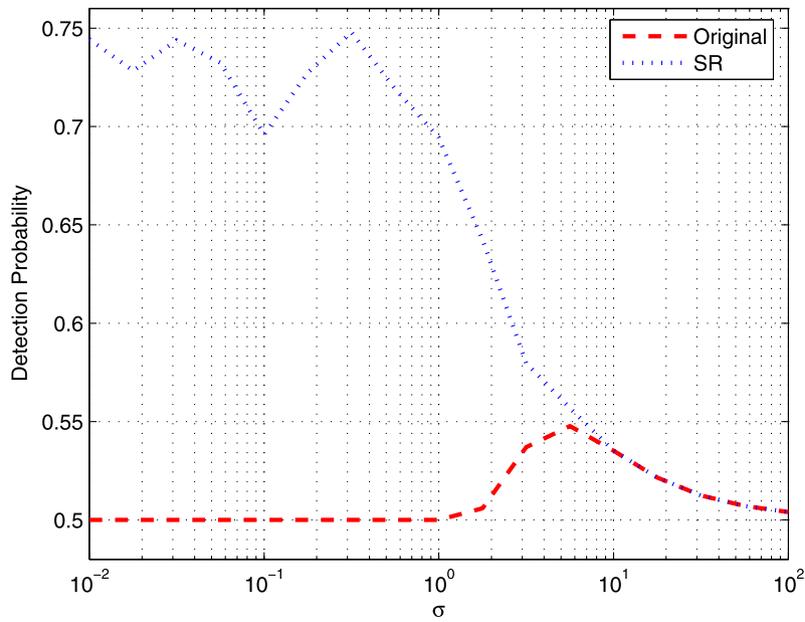
Fig. 10. Comparison of detection probabilities in the absence ("original") and presence ("SR") of additive noise according to the max-min criterion for various values of σ.
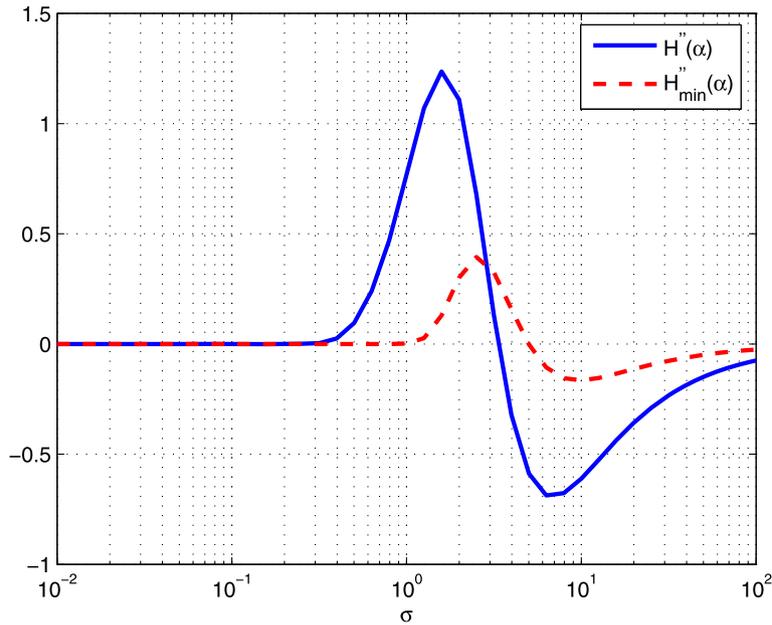

Fig. 11. The second-order derivatives of $H(t)$ in (17) and $H_{\min}(t)$ (36) at $t = \alpha$ for various values of σ. Theorem 1 and Theorem 5 imply that the detector is improvable whenever the second-order derivative at $t = \alpha$ is positive.

Finally, the improvability conditions in Theorem 1 and Theorem 5 are investigated in Fig. 11. It is observed from the figures that the detector is improvable in the interval $\sigma \in [0.1585, 3.398]$ for the max-sum case and in the interval $\sigma \in [0.5012, 4.996]$ for the max-min case, which together with Figs. 9 and 10 imply that the conditions in the theorems are sufficient but not necessary.

7. Concluding remarks and extensions

In this paper, the effects of additive independent noise have been investigated for composite hypothesis-testing problems. The Neyman–Pearson framework has been considered, and performance of noise-modified detectors has been analyzed according to the max-sum, max-min, and max-max criteria. Improvability and nonimprovability conditions have been derived for each case, and the statistical characterization of optimal additive noise PDFs has

been provided. A detection example has been presented in order to explain the theoretical results.

Although the additive independent noise as in Fig. 1 is considered in this study, the results can be extended to other noise injection approaches than the addition operation by considering a nonlinear transformation of the observation, as discussed in [14]. In that case, the nonlinear operator and the original detector can be regarded together as a new detector and the results in this study can directly be applied.

References

[1] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of stochastic resonance, J. Phys. A: Math. Gen. 14 (1981) 453–457.
[2] G.P. Harmer, B.R. Davis, D. Abbott, A review of stochastic resonance: Circuits and measurement, IEEE Trans. Instrum. Meas. 51 (2) (2002) 299–309.

[3] K. Loerincz, Z. Gingl, L. Kiss, A stochastic resonator is able to greatly improve signal-to-noise ratio, Phys. Lett. A 224 (1996) 63–67.

[4] L.G.F. Marchesoni, E. Menichella-Saetta, S. Santucci, Stochastic resonance in bistable systems, Phys. Rev. Lett. 62 (4) (1989) 349–352.

[5] M.E. Inchiosa, A.R. Bulsara, Signal detection statistics of stochastic resonators, Phys. Rev. E 53 (3) (1996) R2021–R2024.

[6] P. Makra, Z. Gingl, Signal-to-noise ratio gain in non-dynamical and dynamical bistable stochastic resonators, Fluct. Noise Lett. 2 (3) (2002) L145–L153.

[7] L. Gammaitoni, P. Hanggi, P. Jung, F. Marchesoni, Stochastic resonance, Rev. Modern Phys. 70 (1) (1998) 223–287.

[8] I. Goychuk, P. Hanggi, Stochastic resonance in ion channels characterized by information theory, Phys. Rev. E 61 (4) (2000) 4272–4280.

[9] S. Mitaim, B. Kosko, Adaptive stochastic resonance in noisy neurons based on mutual information, IEEE Trans. Neural Netw. 15 (6) (2004) 1526–1540.

[10] N.G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, Phys. Rev. Lett. 84 (11) (2000) 2310–2313.

[11] X. Godivier, F. Chapeau-Blondeau, Stochastic resonance in the information capacity of a nonlinear dynamic system, Int. J. Bifurc. Chaos 8 (3) (1998) 581–589.

[12] B. Kosko, S. Mitaim, Stochastic resonance in noisy threshold neurons, Neural Netw. 16 (2003) 755–761.

[13] B. Kosko, S. Mitaim, Robust stochastic resonance for simple threshold neurons, Phys. Rev. E 70 (2004) 031911.

[14] H. Chen, P.K. Varshney, S.M. Kay, J.H. Michels, Theory of the stochastic resonance effect in signal detection: Part I – fixed detectors, IEEE Trans. Signal Process. 55 (7) (2007) 3172–3184.

[15] A. Patel, B. Kosko, Optimal noise benefits in Neyman–Pearson and inequality-constrained signal detection, IEEE Trans. Signal Process. 57 (5) (2009) 1655–1669.

[16] S.M. Kay, J.H. Michels, H. Chen, P.K. Varshney, Reducing probability of decision error using stochastic resonance, IEEE Signal Process. Lett. 13 (11) (2006) 695–698.

[17] S. Bayram, S. Gezici, H.V. Poor, Noise enhanced hypothesis-testing in the restricted Bayesian framework, IEEE Trans. Signal Process. 58 (8) (2010) 3972–3989.

[18] S. Bayram, S. Gezici, Noise enhanced M-ary composite hypothesis-testing in the presence of partial prior information, IEEE Trans. Signal Process. 9 (3) (2011) 1292–1297.

[19] S.M. Kay, Can detectability be improved by adding noise? IEEE Signal Process. Lett. 7 (1) (2000) 8–10.

[20] H. Chen, P.K. Varshney, S.M. Kay, J.H. Michels, Theory of the stochastic resonance effect in signal detection: Part II – variable detectors, IEEE Trans. Signal Process. 56 (10) (2007) 5031–5041.

[21] S. Bayram, S. Gezici, On the improvability and nonimprovability of detection via additional independent noise, IEEE Signal Process. Lett. 16 (11) (2009) 1001–1004.

[22] F. Chapeau-Blondeau, Stochastic resonance and optimal detection of pulse trains by threshold devices, Digital Signal Process. 9 (1999) 162–177.

[23] S. Zozor, P. Amblard, Stochastic resonance in discrete-time nonlinear AR(1) models, IEEE Trans. Signal Process. 47 (1) (1999) 108–122.

[24] S. Zozor, P. Amblard, Can stochastic resonance be used in detection? in: Proc. 10th European Signal Processing Conference (EUSIPCO), Tampere, Finland, 2000.

[25] H. Chen, P.K. Varshney, J.H. Michels, S.M. Kay, Approaching near optimal detection performance via stochastic resonance, in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3, 2006.

[26] A. Asdi, A. Tewfik, Detection of weak signals using adaptive stochastic resonance, in: Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Vol. 2, Detroit, Michigan, 1995, pp. 1332–1335.

[27] S. Zozor, P.-O. Amblard, On the use of stochastic resonance in sine detection, Signal Process. 7 (2002) 353–367.

[28] S. Zozor, P.-O. Amblard, Stochastic resonance in locally optimal detectors, IEEE Trans. Signal Process. 51 (12) (2003) 3177–3181.

[29] A.A. Saha, G. Anand, Design of detectors based on stochastic resonance, Signal Process. 83 (2003) 1193–1212.

[30] D. Rousseau, F. Chapeau-Blondeau, Stochastic resonance and improvement by noise in optimal detection strategies, Digital Signal Process. 15 (2005) 19–32.

[31] I. Lee, X. Liu, C. Zhou, B. Kosko, Noise-enhanced detection of subthreshold signals with carbon nanotubes, IEEE Trans. Nanotechnol. 5 (6) (2006) 613–627.

[32] H. Chen, P.K. Varshney, S. Kay, J.H. Michels, Noise enhanced nonparametric detection, IEEE Trans. Inform. Theory 55 (2) (2009) 499–506.

[33] H. Chen, P.K. Varshney, J.H. Michels, Improving sequential detection performance via stochastic resonance, IEEE Signal Process. Lett. 15 (2008) 685–688.

[34] H. Chen, P.K. Varshney, J.H. Michels, Noise enhanced parameter estimation, IEEE Trans. Signal Process. 56 (10) (2008) 5074–5081.

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

[36] E.L. Lehmann, Testing Statistical Hypotheses, 2nd edition, Chapman & Hall, New York, 1986.

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

[38] A.I.F. Vaz, E.M.G.P. Fernandes, Optimization of nonlinear constrained particle swarm, Baltic J. Sustain. 12 (1) (2006) 30–36.

[39] S. Koziel, Z. Michalewicz, Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization, Evol. Comput. 7 (1) (1999) 19–44.

[40] X. Hu, 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.

[41] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, Cambridge, UK, 2004.

[42] M.A. Richards, Fundamentals of Radar Signal Processing, Electron. Eng. Ser., McGraw–Hill, USA, 2005.

[43] J. Cvitanic, I. Karatzas, Generalized Neyman–Pearson lemma via convex duality, Bernoulli 7 (1) (2001) 79–97.

[44] B. Rudloff, I. Karatzas, Testing composite hypotheses via convex duality, Bernoulli 16 (4) (2010) 1224–1239.

[45] S. Bayram, S. Gezici, On the restricted Neyman–Pearson approach for composite hypothesis-testing in the presence of prior distribution uncertainty, IEEE Trans. Signal Process. 59 (10) (2011) 5056–5065.

[46] A. Goldsmith, Wireless Communications, Cambridge University Press, Cambridge, UK, 2005.

[47] C.C. Pugh, Real Mathematical Analysis, Springer-Verlag, New York, 2002.

[48] L. Huang, 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, 2007.

[49] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1968.

[50] D.P. Bertsekas, A. Nedic, A.E. Ozdaglar, Convex Analysis and Optimization, Athena Specific, Boston, MA, 2003.

[51] K.V. Price, R.M. Storn, J.A. Lampinen, Differential Evolution: A Practical Approach to Global Optimization, Springer, New York, 2005.

[52] Y. Chen, V.K. Dubey, Ultrawideband source localization using a particle-swarm-optimized Capon estimator, in: Proc. IEEE Int. Conf. Commun. (ICC), Vol. 4, Seoul, Korea, 2005, pp. 2825–2829.

[53] Y. Rahmat-Samii, D. Gies, J. Robinson, Particle swarm optimization (PSO): A novel paradigm for antenna designs, Radio Sci. Bull. 305 (2003) 14–22.

[54] Z. Yangyang, J. Chunlin, Y. Ping, L. Manlin, W. Chaojin, W. Guangxing, Particle swarm optimization for base station placement in mobile communication, in: Proc. IEEE International Conference on Networking, Sensing and Control, vol. 1, 2004, pp. 428–432.

[55] W. Jatmiko, K. Sekiyama, T. Fukuda, A PSO-based mobile sensor network for odor source localization in dynamic environment: Theory, simulation and measurement, in: Proc. IEEE Congress on Evolutionary Computation, Vancouver, BC, 2006, pp. 1036–1043.

[56] J. Pugh, A. Martinoli, Y. Zhang, Particle swarm optimization for unsupervised robotic learning, in: Proc. Swarm Intelligence Symposium (SIS), Pasadena, California, 2005, pp. 92–99.

[57] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd edition, Wiley–Interscience, New York, 2000.

Suat Bayram received the B.S. degree from Middle East Technical University, Ankara, Turkey in 2007, and the M.S. and the Ph.D. degrees from Bilkent University, Ankara, Turkey, in 2009 and 2011, respectively. His research interests are in the statistical signal processing and communications fields. Currently, he has particular interest in compressive sensing, adaptive sensing, and classification problems.

Sinan Gezici received the B.S. degree from Bilkent University, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. Since 2007, he has been an Assistant Professor in the Department of Electrical and Electronics Engineering at Bilkent University. Dr. Gezici's research interests are in the areas of detection and estimation theory, wireless communications, and localization systems.