

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

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

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

Index Terms—Hypothesis-testing, minimax, detection, stochastic resonance, noise-enhanced detection.

I. INTRODUCTION

Although noise commonly degrades performance of a system, outputs of some nonlinear systems can be improved by injecting additional noise to their inputs [1]-[13]. Such improvements can be considered in the framework of stochastic resonance (SR), which can be regarded as the observation of “noise benefits” related to signal transmission in nonlinear systems [13]-[17].

Improvements that can be obtained via additional independent noise can be in various forms, such as an increase in output signal-to-noise ratio (SNR) [1], [4], [5] or mutual information [6]-[11], a decrease in probability of decision error [18], or an increase in probability of detection under a constraint on probability of false-alarm [12], [13], [15], [19], [20]. In [19], it is shown by an example that detection performance of a suboptimal detector can be improved by adding white Gaussian noise for the problem of detecting a constant signal in Gaussian mixture noise. Also, it is shown in [18] that the optimal noise that minimizes the probability of decision error has a constant value, and a Gaussian mixture example is used to illustrate the improvability of a detector. In [12], a theoretical framework for investigating the effects of additional independent noise on suboptimal detectors is established according to the Neyman-Pearson criterion. Sufficient conditions on improvability and non-improvability of a suboptimal detector via additional independent noise are derived, and it is proven that optimal additional 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 [13] for the same problem, which also proves the two mass point structure of the optimal additional noise PDF, and, in addition, states that an optimal additional noise may not exist in certain scenarios.

The study in [12] is extended to variable detectors in [20], and similar observations as in the fixed detector case are made. In addition, the theoretical framework in [12] is applied to sequential detection and parameter estimation problems in [21] and [22], respectively. In [21], a binary sequential detection problem is studied, and additional noise that reduces at least one of the expected sample sizes for the sequential detection system is obtained. In [22], improvability of estimation performance via additional noise is illustrated under certain conditions for various estimation criteria, and the form of the optimal noise PDF is obtained in each case. The effects of additional noise are investigated also for detection of weak sinusoidal signals and for locally optimally detectors. In [23] and [24], detection of a weak sinusoidal signal is studied, and improvements on detection performance are investigated. In addition, [25] studies the optimization of noise and detector parameters of locally optimal detectors for the problem of detecting a small amplitude sinusoid in non-Gaussian noise.

The study in [20] utilizes the results in [12] and [18] in order to investigate optimal additional noise for suboptimal variable detectors in the Bayesian and minimax frameworks. Although the formulation of optimal additional noise is studied for a binary hypothesis-testing problem in [20], no studies have investigated M -ary hypothesis problems according to the minimax criterion. The main contributions of our study can be summarized as follows:

- Formulation of a generic optimization problem for obtaining optimal additional independent noise in an M -ary hypothesis-testing problem according to the minimax-criterion.
- Characterization of optimal additional independent noise as a discrete random variable with at most M mass points under certain conditions.
- Derivation of sufficient conditions to determine when additional independent noise can or cannot improve detection performance in the the minimax sense.
- Convex relaxation [26] of the optimal additional indepen-

dent noise problem in order to obtain close-to-optimal solutions in polynomial time.

The remainder of the paper is organized as follows. In Section II, the formulation of optimal additional noise is provided for an M -ary hypothesis-testing problem according to the minimax criterion. Then, it is shown in Section III that the optimal additional noise can be represented by a randomization of no more than M signal levels under certain conditions. In addition, a convex relaxation approach is proposed to obtain an accurate approximation to the noise PDF in polynomial time. Also, sufficient conditions are provided regarding the improbability and non-improbability of a given detector via additional independent noise. Finally, numerical examples are presented in Section IV and concluding remarks are made in Section V.

II. PROBLEM FORMULATION AND MOTIVATION

Consider the following M -ary hypothesis-testing problem:

$$\mathcal{H}_i : p_i^{\mathbf{X}}(\mathbf{x}), \quad i = 0, 1, \dots, M-1, \quad (1)$$

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

A generic decision rule can be defined as

$$\phi(\mathbf{x}) = i, \quad \text{if } \mathbf{x} \in \Gamma_i, \quad (2)$$

for $i = 0, 1, \dots, M-1$, where $\Gamma_0, \Gamma_1, \dots, \Gamma_{M-1}$ form a partition of the observation space Γ [27], [28].

In the minimax approach, the prior probabilities of the hypotheses are unknown. However, each decision is associated with a known cost value, and the aim is to minimize the maximum of the average costs of the decision rule conditioned on different hypotheses [27]. More formally, let $C_{ji} \geq 0$ represent the cost of choosing \mathcal{H}_j when \mathcal{H}_i is true. Then, the average cost of decision rule ϕ conditioned on \mathcal{H}_i being the true hypothesis is calculated as

$$R_i(\phi) = \sum_{j=0}^{M-1} C_{ji} P_i(\Gamma_j), \quad (3)$$

where $P_i(\Gamma_j)$ represents the probability of choosing \mathcal{H}_j when \mathcal{H}_i is the true hypothesis. This quantity, $R_i(\phi)$, is called the *conditional risk* of ϕ given \mathcal{H}_i [27]. In the minimax framework, the aim is to reduce the maximum of the conditional risks for different hypotheses as much as possible. In other words, in the minimax framework, the performance metric for a decision rule is specified as $\max\{R_0(\phi), R_1(\phi), \dots, R_{M-1}(\phi)\}$.

In certain scenarios, addition of independent noise to observations, as shown in Fig. 1, can improve the performance of a suboptimal decision rule (detector) [12], [13], [19]. In such cases, instead of the original observation \mathbf{x} , a noise-added version of that, $\mathbf{y} = \mathbf{x} + \mathbf{n}$, is used by the detector, where \mathbf{n} represents the additional noise term. Although a scenario as in Fig. 1 is considered in this study, the results can be extended to the cases in which a nonlinear transformation of

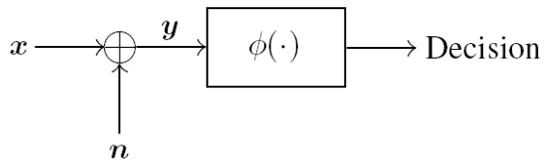


Fig. 1. Independent noise \mathbf{n} is added to observation \mathbf{x} in order to improve the performance of the detector, $\phi(\cdot)$.

the noise-added observation is performed before the detector [12].

The main motivation for observation modification as in Fig. 1 can be explained as follows. In many cases, the optimal detector based on the calculation of likelihood functions is difficult to obtain or requires intense computations [12], [27]. Therefore, a suboptimal detector can be preferred in some practical scenarios. However, the performance of a suboptimal detector may need to be improved in order to meet certain systems requirements. One way to improve the performance of a suboptimal detector without altering the detector structure is to modify its measurements as in Fig. 1 [12]. Although calculation of optimal additional noise results in 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 calculations related to the likelihood functions for each decision, whereas the suboptimal detector with modified observations needs to update the optimal additional noise whenever the statistics of the hypotheses change. For example, 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 during a large number of symbols for slowly varying channels [29].

In this study, the aim is to obtain optimal additional noise PDF $p_{\mathbf{N}}(\cdot)$ that minimizes the maximum of the conditional risks for a given decision rule. In other words, the optimal additional noise is searched for according to the minimax criterion. This problem can be formulated as

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \min_{p_{\mathbf{N}}(\mathbf{n})} \max_{i \in \{0, 1, \dots, M-1\}} R_i^{\mathbf{y}}(\phi), \quad (4)$$

where $R_i^{\mathbf{y}}(\phi)$ represents the conditional risk of ϕ given \mathcal{H}_i when the noise-modified observation \mathbf{y} is used; that is,

$$R_i^{\mathbf{y}}(\phi) = \sum_{j=0}^{M-1} C_{ji} P_i^{\mathbf{y}}(\Gamma_j), \quad (5)$$

with $P_i^{\mathbf{y}}(\Gamma_j)$ representing the probability that $\mathbf{y} \in \Gamma_j$ when \mathcal{H}_i is true.

III. NOISE-ENHANCED HYPOTHESIS-TESTING

In this section, calculation of the optimal additional noise in (4) is studied, and its statistical characterization is provided. In

addition, sufficient conditions on the improvability and non-improvability of detection via additional independent noise are presented.

In order to investigate the solution of the optimization problem in (4), we first express the conditional risk $R_i^y(\phi)$ in (5) as follows:

$$R_i^y(\phi) = \sum_{j=0}^{M-1} C_{ji} \int_{\Gamma_j} p_i^Y(\mathbf{z}) d\mathbf{z} . \quad (6)$$

Since \mathbf{X} and \mathbf{N} are independent, the PDF of $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ can be obtained as the convolution of the PDFs of \mathbf{X} and \mathbf{N} . Then, (6) can be manipulated to derive the following expressions:

$$R_i^y(\phi) = \sum_{j=0}^{M-1} C_{ji} \int_{\Gamma_j} \int_{\mathbb{R}^K} p_N(\mathbf{n}) p_i^X(\mathbf{z} - \mathbf{n}) d\mathbf{n} d\mathbf{z} \quad (7)$$

$$= \sum_{j=0}^{M-1} C_{ji} \int_{\mathbb{R}^K} p_N(\mathbf{n}) \int_{\Gamma_j} p_i^X(\mathbf{z} - \mathbf{n}) d\mathbf{z} d\mathbf{n} \quad (8)$$

$$= \sum_{j=0}^{M-1} C_{ji} E\{F_{ij}(\mathbf{N})\} \quad (9)$$

$$= E\{F_i(\mathbf{N})\} , \quad (10)$$

where

$$F_{ij}(\mathbf{n}) \doteq \int_{\Gamma_j} p_i^X(\mathbf{z} - \mathbf{n}) d\mathbf{z} \quad (11)$$

and

$$F_i(\mathbf{n}) \doteq \sum_{j=0}^{M-1} C_{ji} F_{ij}(\mathbf{n}) . \quad (12)$$

Then, the optimization problem in (4) becomes

$$\min_{p_N(\cdot)} \max_{i \in \{0, 1, \dots, M-1\}} E\{F_i(\mathbf{N})\} . \quad (13)$$

Note that under uniform cost assignment (UCA); that is, when $C_{ji} = 1$ for $j \neq i$, and $C_{ji} = 0$ for $j = i$ [27], the conditional risk can be evaluated from (9) as

$$R_i^y(\phi) = 1 - E\{F_{ii}(\mathbf{N})\} . \quad (14)$$

Then, (13) can be expressed as

$$\max_{p_N(\cdot)} \min_{i \in \{0, 1, \dots, M-1\}} E\{F_{ii}(\mathbf{N})\} . \quad (15)$$

Although it is quite difficult to perform a search over all possible noise PDFs in (13), the following proposition states that the search can be performed over the set of discrete probability distributions with at most M mass points in many practical scenarios.

Proposition 1: Define set U as

$$U = \{(u_0, u_1, \dots, u_{M-1}) : u_0 = F_0(\mathbf{n}), u_1 = F_1(\mathbf{n}), \dots, u_{M-1} = F_{M-1}(\mathbf{n}), \text{ for } \mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}\} , \quad (16)$$

where $\mathbf{n} \in \mathbb{R}^K$, and $\mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}$ means that $a_j \leq n_j \leq b_j$ for $j = 1, \dots, K$.

Assume that the additional noise \mathbf{n} satisfies $\mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}$ and

U is a closed subset of \mathbb{R}^M . Then, the optimal additional noise PDF in (4) can be expressed as

$$p_N^{\text{opt}}(\mathbf{n}) = \sum_{i=0}^{M-1} \lambda_i \delta(\mathbf{n} - \mathbf{n}_i) , \quad (17)$$

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

Proof: Please see Appendix A.

The first assumption in the proposition, which states that the additional noise values satisfy $\mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b}$, is realistic for practical systems since arbitrarily large or arbitrarily small signal levels cannot be generated at the detector. In other words, the maximum and minimum possible noise values determine \mathbf{b} and \mathbf{a} , respectively, in practice. Regarding the assumption that U is a closed set, one sufficient condition is to have $F_0(\mathbf{n}), F_1(\mathbf{n}), \dots, F_{M-1}(\mathbf{n})$ as continuous functions. In that case, the mapping from $[\mathbf{a}, \mathbf{b}]$ to \mathbb{R}^M defined by $G(\mathbf{n}) = (F_0(\mathbf{n}), F_1(\mathbf{n}), \dots, F_{M-1}(\mathbf{n}))$, becomes continuous. Hence, U becomes a closed set. For example, when the PDFs are continuous for all hypotheses, (11) and (12) imply that $G(\mathbf{n})$ is continuous.

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

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

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

Based on Proposition 1, the optimization problem in (13) can be expressed as

$$\begin{aligned} & \min_{\{\mathbf{n}_j, \lambda_j\}_{j=0}^{M-1}} \max_{i \in \{0, 1, \dots, M-1\}} \sum_{j=0}^{M-1} \lambda_j F_i(\mathbf{n}_j) \\ & \text{subject to } \sum_{j=0}^{M-1} \lambda_j = 1 \\ & \lambda_j \geq 0, \quad j = 0, 1, \dots, M-1 . \end{aligned} \quad (18)$$

Although (18) is significantly simpler than (13), it can still be a non-convex optimization problem in general. Therefore, global optimization techniques, such as particle-swarm optimization (PSO) [30], [31], genetic algorithms and differential evolution [32] can be applied to obtain the optimal additional noise PDF. As an alternative approach, we provide an approximate formulation that results in a convex optimization problem. Assume that additional noise \mathbf{n} can take only finitely many known values specified by $\tilde{\mathbf{n}}_1, \dots, \tilde{\mathbf{n}}_L$, and the aim is

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

to determine the weights $\tilde{\lambda}_1, \dots, \tilde{\lambda}_L$ of those possible noise values. Then, (13) can be expressed, after some manipulation, as the following optimization problem:

$$\begin{aligned} & \min_{t, \{\tilde{\lambda}_j\}_{j=1}^L} t \\ & \text{subject to } \sum_{j=1}^L \tilde{\lambda}_j F_i(\tilde{\mathbf{n}}_j) \leq t, \quad i = 0, 1, \dots, M-1 \\ & \sum_{j=0}^L \tilde{\lambda}_j = 1, \\ & \tilde{\lambda}_j \geq 0, \quad j = 1, \dots, L. \end{aligned} \quad (19)$$

The optimization problem in (19) is a linearly constrained linear programming (LCLP) problem, which can be solved in polynomial time [26]. Also, as L is increased (as the optimization is performed over more noise values), the solution of the optimization problem in (19) gets closer to the optimal solution of (13).

Finally, the issue of determining whether additional independent noise can improve the performance of a given detector without actually solving the optimization problem in (13) is addressed. In the following, sufficient conditions are presented for the improvability and the non-improvability of a given detector via the use of additional independent noise.

Proposition 2: Define $J(\mathbf{n}) = \max_{i \in \{0, 1, \dots, M-1\}} F_i(\mathbf{n})$. If $\mathbf{n}_0 = \arg \min_{\mathbf{n}} J(\mathbf{n})$ is non-zero, then the detector is improvable.

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

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

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

The results in Proposition 2 and Proposition 3 can be used to determine when it is necessary to tackle the optimization problem in (13) to obtain the optimal additional noise PDF. For example, when the non-improvability condition in Proposition 3 is satisfied, it is directly concluded that $p_N^{\text{opt}}(\mathbf{n}) = \delta(\mathbf{n})$.

IV. NUMERICAL RESULTS

In this section, numerical examples are provided in order to investigate the theoretical results obtained in the previous

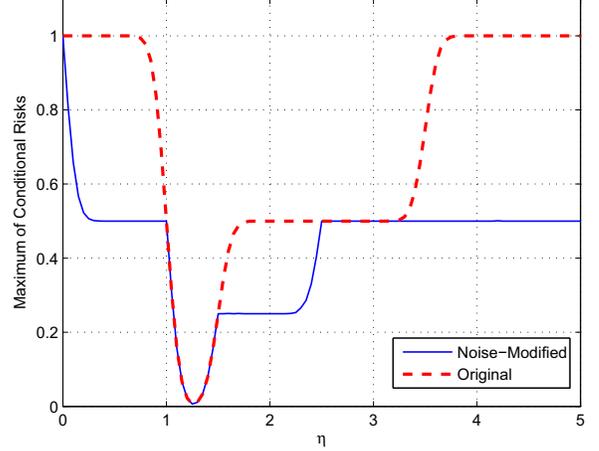


Fig. 2. Maximum of the conditional risks versus η for the original and the noise-modified detectors for $A = 1$, $B = 2.5$, $\sigma = 0.1$, $w_1 = 0.5$ and $w_2 = 0.5$.

section. A ternary hypothesis-testing problem is considered with the following PDFs:

$$\begin{aligned} p_0^X(x) &= w_1 \gamma(x; -A, \sigma^2) + w_2 \gamma(x; A, \sigma^2) \\ p_1^X(x) &= w_1 \gamma(x; -A + B, \sigma^2) + w_2 \gamma(x; A + B, \sigma^2) \\ p_2^X(x) &= w_1 \gamma(x; -A - B, \sigma^2) + w_2 \gamma(x; A - B, \sigma^2) \end{aligned} \quad (20)$$

where

$$\gamma(x; \mu, \sigma^2) \doteq \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \quad (21)$$

The decision rule is described as follows:

$$\phi(x) = \begin{cases} 0, & -\eta < x < \eta \\ 1, & x \geq \eta \\ 2, & x \leq -\eta \end{cases}, \quad (22)$$

where η is a constant. Under UCA, the conditional risks can be obtained from (3), after some manipulation, as

$$\begin{aligned} R_0(\phi) &= 1 - w_1 \left[Q\left(\frac{-\eta + A}{\sigma}\right) - Q\left(\frac{\eta + A}{\sigma}\right) \right] \\ &\quad - w_2 \left[Q\left(\frac{-\eta - A}{\sigma}\right) - Q\left(\frac{\eta - A}{\sigma}\right) \right] \\ R_1(\phi) &= 1 - w_1 Q\left(\frac{\eta + A - B}{\sigma}\right) - w_2 Q\left(\frac{\eta - A - B}{\sigma}\right) \\ R_2(\phi) &= 1 - w_1 Q\left(\frac{\eta - A - B}{\sigma}\right) - w_2 Q\left(\frac{\eta + A - B}{\sigma}\right). \end{aligned}$$

Similarly, $F_{ii}(\mathbf{n})$ can be calculated from (11) for $i = 0, 1, 2$ and the optimization problem in (15) can be solved to obtain optimal additional noise.

Fig. 2 plots the maximum of conditional risks for the original and the noise-modified detectors with respect to η in (22) when the parameters are taken as $A = 1$, $B = 2.5$, $w_1 = 0.5$, $w_2 = 0.5$ and $\sigma = 0.1$. From the figure, it is

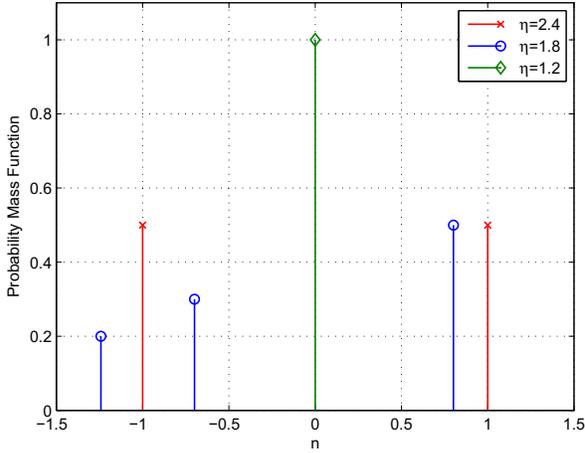


Fig. 3. Probability mass function of optimal additional noise for various threshold values when the parameters are taken as $A = 1$, $B = 2.5$, $\sigma = 0.1$, $w_1 = 0.5$ and $w_2 = 0.5$.

observed that for certain values of η , the performance can be improved via the addition of independent noise. For example, for $\eta = 1.8$, the improvement ratio, defined as the ratio between $\max_{i \in \{0,1,2\}} R_i(\phi)$ and $\max_{i \in \{0,1,2\}} R_i^y(\phi)$, is equal to 2. As another example, for $\eta = 2.4$, the improvement ratio is calculated as 1.52.

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

Finally, Fig. 4 and Fig. 5 illustrate the performance of the original and the noise-modified detectors for $\eta = 1.8$ and $\eta = 2.4$, respectively, versus the standard deviation parameter in (20). The other parameters are set to $A = 1$, $B = 2.5$, $w_1 = 0.5$ and $w_2 = 0.5$. It is observed that as the standard deviation increases, the improvement ratios become smaller, and after a certain value, the detectors become non-improvable.

V. CONCLUSIONS

In this study, the effects of adding independent noise to observations have been investigated for M -ary hypothesis-testing problems in the minimax framework. First, the calculation of optimal additional noise has been formulated as an optimization problem, and it has been proven that the optimal additional noise can be represented as a discrete random variable with at most M mass points under certain conditions. In addition, an approximate technique to calculate the optimal additional noise has been presented as a convex optimization problem. Finally, sufficient conditions have been presented to specify when additional independent noise can

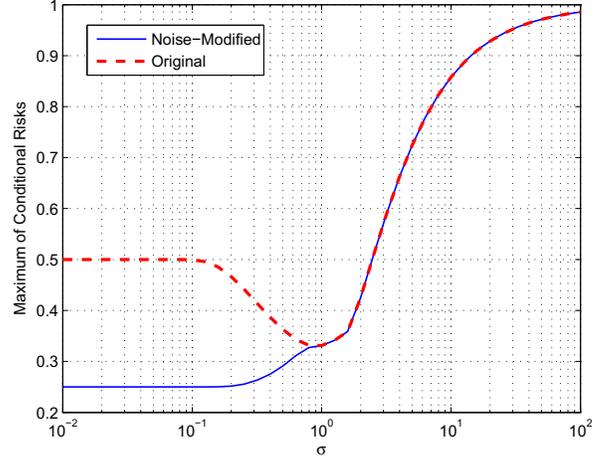


Fig. 4. Maximum of the conditional risks versus σ for the original and the noise-modified detectors when the parameters are taken as $\eta = 1.8$, $A = 1$, $B = 2.5$, $w_1 = 0.5$ and $w_2 = 0.5$.

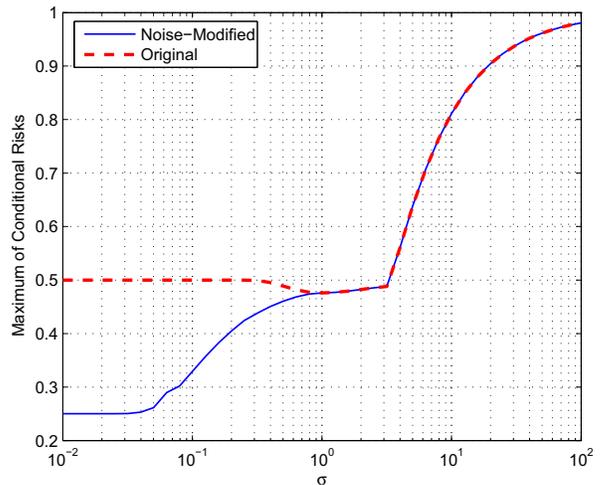


Fig. 5. Maximum of the conditional risks versus σ for the original and the noise-modified detectors when the parameters are taken as $\eta = 2.4$, $A = 1$, $B = 2.5$, $w_1 = 0.5$ and $w_2 = 0.5$.

or cannot improve the performance of a given detector, and a numerical example has been presented.

APPENDIX

A. Proof of Proposition 1

An approach similar those in [12] and [33] is employed in the proof of the proposition. Let V represent the convex hull of U in (16) [34]. From (11) and (12), it is observed that U is a bounded set. Since it is also closed by the assumption in the proposition, U is a compact set. Therefore, its convex hull, V , is a closed subset of \mathbb{R}^M [34].

Next, define W as

$$W = \left\{ (w_0, w_1, \dots, w_{M-1}) : w_i = E\{F_i(\mathbf{n})\}, \right. \\ \left. i = 0, 1, \dots, M-1, \forall p_N(\mathbf{n}), \mathbf{a} \preceq \mathbf{n} \preceq \mathbf{b} \right\}, \quad (23)$$

where $p_N(\mathbf{n})$ is the PDF of the additional independent noise. Since for any vector random variable Θ taking values in set Ω , its expected value $E\{\Theta\}$ is in the convex hull of Ω [33], it is concluded from (16) and (23) that W is in the convex hull V of U ; that is, $V \supseteq W$. In addition, since V is defined as the convex hull of U , each element of V can be expressed as $\mathbf{v} = \sum_{l=1}^{N_L} \lambda_l (F_0(\mathbf{n}_l), F_1(\mathbf{n}_l), \dots, F_{M-1}(\mathbf{n}_l))$, where $\sum_{l=1}^{N_L} \lambda_l = 1$, and $\lambda_l \geq 0 \forall l$. However, each \mathbf{v} is also an element of W since it can be obtained for $p_N(\mathbf{n}) = \sum_{l=1}^{N_L} \lambda_l \delta(\mathbf{n} - \mathbf{n}_l)$. Hence, $V \subseteq W$. Since $V \subseteq W$ and $V \supseteq W$, it is concluded that $W = V$. Therefore, Carathéodory's theorem [35], [36] implies that any point in V (or, W) can be expressed as the convex combination of at most $(M+1)$ points in U as the dimension of U is smaller than or equal to M (c.f. (16)). Since the aim is to minimize the maximum of the conditional risks, the optimal solution must correspond to the boundary of W . Since W (or, V) is a closed set as mentioned at the beginning of the proof, it contains its own boundary. Since any point at the boundary of W can be expressed as the convex combination of at most M elements in U [35], an optimal noise PDF can be represented by a discrete random variable with M mass points as in (17). \square

REFERENCES

- [1] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A: Math. General*, vol. 14, pp. 453–457, 1981.
- [2] P. Makra and Z. Gingl, "Signal-to-noise ratio gain in non-dynamical and dynamical bistable stochastic resonators," *Fluctuat. Noise Lett.*, vol. 2, no. 3, pp. L145–L153, 2002.
- [3] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223–287, Jan. 1998.
- [4] G. P. Harmer, B. R. Davis, and D. Abbott, "A review of stochastic resonance: Circuits and measurement," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 2, pp. 299–309, Apr. 2002.
- [5] K. Loerincz, Z. Gingl, and L. Kiss, "A stochastic resonator is able to greatly improve signal-to-noise ratio," *Phys. Lett. A*, vol. 224, pp. 63–67, 1996.
- [6] I. Goychuk and P. Hanggi, "Stochastic resonance in ion channels characterized by information theory," *Phys. Rev. E*, vol. 61, no. 4, pp. 4272–4280, 2000.
- [7] S. Mitaïm and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1526–1540, Nov. 2004.
- [8] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Mar. 2000.
- [9] X. Godivier and F. Chapeau-Blondeau, "Stochastic resonance in the information capacity of a nonlinear dynamic system," *Int. J. Bifurc. Chaos*, vol. 8, no. 3, pp. 581–589, 1998.
- [10] B. Kosko and S. Mitaïm, "Stochastic resonance in noisy threshold neurons," *Neural Netw.*, vol. 16, pp. 755–761, 2003.
- [11] —, "Robust stochastic resonance for simple threshold neurons," *Phys. Rev. E*, vol. 70, no. 031911, 2004.
- [12] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I—Fixed detectors," *IEEE Trans. Sig. Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.
- [13] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [14] F. Chapeau-Blondeau and D. Rousseau, "Raising the noise to improve performance in optimal processing," *Journal of Statistical Mechanics: Theory and Experiment*, no. P01003, pp. 1–15, Jan. 2009.
- [15] S. Bayram and S. Gezici, "On the improbability and non-improvability of detection via additional independent noise," *IEEE Sig. Processing Lett.*, 2009.
- [16] P. Hanggi, M. E. Inchiosa, D. Fogliatti, and A. R. Bulsara, "Nonlinear stochastic resonance: The saga of anomalous output-input gain," *Physical Review E*, vol. 62, no. 5, pp. 6155–6163, Nov. 2000.
- [17] V. Galdi, V. Pierro, and I. M. Pinto, "Evaluation of stochastic-resonance-based detectors of weak harmonic signals in additive white gaussian noise," *Physical Review E*, vol. 57, no. 6, pp. 6470–6479, June 1998.
- [18] S. M. Kay, J. H. Michels, H. Chen, and P. K. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Sig. Processing Lett.*, vol. 13, no. 11, pp. 695–698, Nov. 2006.
- [19] S. M. Kay, "Can detectability be improved by adding noise?" *IEEE Sig. Processing Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.
- [20] H. Chen and P. K. Varshney, "Theory of the stochastic resonance effect in signal detection: Part II—Variable detectors," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.
- [21] H. Chen, P. K. Varshney, and J. H. Michels, "Improving sequential detection performance via stochastic resonance," *IEEE Sig. Processing Lett.*, vol. 15, pp. 685–688, Dec. 2008.
- [22] —, "Noise enhanced parameter estimation," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.
- [23] A. Asdi and A. Tewfik, "Detection of weak signals using adaptive stochastic resonance," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 2, Detroit, Michigan, May 1995, pp. 1332–1335.
- [24] S. Zozor and P.-O. Amblard, "On the use of stochastic resonance in sine detection," *Signal Process.*, vol. 7, pp. 353–367, Mar. 2002.
- [25] —, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177–3181, Dec. 2003.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [27] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [28] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, Inc., 1998.
- [29] A. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005.
- [30] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems*. IOS Press, 2002, pp. 214–220, in *Intelligent Technologies—Theory and Applications: New Trends in Intelligent Technologies*.
- [31] A. I. F. Vaz and E. M. G. P. Fernandes, "Optimization of nonlinear constrained particle swarm," *Baltic Journal on Sustainability*, vol. 12, no. 1, pp. 30–36, 2006.
- [32] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. New York: Springer, 2005.
- [33] L. Huang and M. J. Neely, "The optimality of two prices: Maximizing revenue in a stochastic network," in *Proc. 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2007.
- [34] C. C. Pugh, *Real Mathematical Analysis*. New York: Springer-Verlag, 2002.
- [35] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.
- [36] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Boston, MA: Athena Specific, 2003.