

# Noise Benefits in Joint Detection and Estimation Problems

Abdullah Basar Akbay and Sinan Gezici

## Abstract

Adding noise to inputs of some suboptimal detectors or estimators can improve their performance under certain conditions. In the literature, noise benefits have been studied for detection and estimation systems separately. In this study, noise benefits are investigated for joint detection and estimation systems. The analysis is performed under the Neyman-Pearson (NP) and Bayesian detection frameworks and according to the Bayesian estimation criterion. The maximization of the system performance is formulated as an optimization problem. The optimal additive noise is shown to have a specific form, which is derived under both NP and Bayesian detection frameworks. In addition, the proposed optimization problem is approximated as a linear programming (LP) problem, and conditions under which the performance of the system can or cannot be improved via additive noise are obtained. With an illustrative numerical example, performance comparison between the noise enhanced system and the original system is presented to support the theoretical analysis.

*Index Terms*—Detection, parameter estimation, linear programming, noise enhancement.

## I. INTRODUCTION

Although an increase in the noise power is generally associated with performance degradation, addition of noise to a system may introduce performance improvements under certain arrangements and conditions in a number of electrical engineering applications including neural signal processing, biomedical signal processing, lasers, nano-electronics, digital audio and image processing, analog-to-digital converters, control theory, statistical signal processing, and information theory, as exemplified in [1] and references therein. In the field of statistical signal processing, noise benefits are investigated in various studies such

A. B. Akbay is with the Electrical Engineering Department, University of California, Los Angeles, USA. S. Gezici with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara, 06800, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, Email: gezici@ee.bilkent.edu.tr. This research was supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013). Part of this work was presented at IEEE Signal Processing and Communications Applications Conference (SIU 2014).

as [2]– [17]. In [2], it is shown that the detection probability of the optimal detector for a described network with nonlinear elements driven by a weak sinusoidal signal in white Gaussian noise is non-monotonic with respect to the noise power and fixed false alarm probability; hence, detection probability enhancements can be achieved via increasing the noise level in certain scenarios. For an optimal Bayesian estimator, in a given nonlinear setting, with examples of a quantizer [3] and phase noise on a periodic wave [4], a non-monotonic behavior in the estimation mean-square error is demonstrated as the intrinsic noise level increases. In [5], the proposed simple suboptimal nonlinear detector scheme, in which the detector parameters are chosen according to the system noise level and distribution, outperforms the matched filter under non-Gaussian noise in the Neyman-Pearson (NP) framework. In [6], it is noted that the performance of some optimal detection strategies display a non-monotonic behavior with respect to the noise root-mean square amplitude in a binary hypothesis testing problem with a nonlinear setting, where non-Gaussian noise (two different distributions are examined for numerical purposes: Gaussian mixture and uniform distributions) acts on the phase of a periodic signal. In [16] and [17], theoretical conditions are provided related to improvability and nonimprovability of suboptimal detectors for weak signal detection via noise benefits.

One approach for realizing noise benefits is to tune the parameters of a nonlinear system, as employed, e.g., in [8]–[12]. An alternative approach is the injection of a random process independent of both the meaningful information signal (transmitted or hidden signal) and the background noise (undesired signal). It is firstly shown by Kay in [13] that addition of independent randomness may improve suboptimal detectors under certain conditions. Later, it is proved that a suboptimal detector in the Bayesian framework may be improved (i.e., the Bayes risk can be reduced) by adding a constant signal to the observation signal; that is, the optimal probability density function is a single Dirac delta function [14]. This intuition is extended in various directions and it is demonstrated that injection of additive noise to the observation signal at the input of a suboptimal detector can enhance the system performance [15], [18]–[34]. In this manuscript, performance improvements through noise benefits are addressed in the context of *joint detection and estimation* systems by adding an independent noise component to the observation signal at the input of a suboptimal system. Notice that the most critical keyword in this approach is suboptimality. Under non-Gaussian background noise, optimal detectors/estimators are often nonlinear,

difficult to implement, and complex systems [35], [36]. Hence, the main aim is to improve the performance of a fairly simple and practical system by adding specific randomness (“noise”) at the input.

Chen *et al.* revealed that the detection probability of a suboptimal detector in the NP framework can be increased via additive independent noise [15]. They examined the convex structure of the problem and specified the nature of the optimal probability distribution of additive noise as a probability mass function with at most two point masses. This result is generalized for  $M$ -ary composite hypothesis testing problems under NP, restricted NP and restricted Bayes criteria [25], [29], [34]. In estimation problems, additive noise can also be utilized to improve the performance of a given suboptimal estimator [4], [19], [30]. As an example of noise benefits for an estimation system, it is shown that Bayesian estimator performance can be enhanced by adding non-Gaussian noise to the system, and this result is extended to the general parameter estimation problem in [19]. As an alternative example of noise enhancement application, injection of noise to blind multiple error rate estimators in wireless relay networks is presented in [30].

In this study, noise benefits are investigated for a joint detection and estimation system, which is presented in [37]. Without introducing any modification to the structure of the system, the aim is to improve the performance of the joint detection and estimation system by only adding noise to the observation signal at the input. Therefore, the detector and estimator are assumed to be given and fixed. In [37], optimal detectors and estimators are derived for this joint system. However, the optimal structures may be overcomplicated for an implementation. In this study, it is assumed that the given joint detection and estimation system is suboptimal, and the purpose is defined as the examination of the performance improvements via additive noise under this assumption. The main contributions of this study can be summarized as follows:

- Noise benefits are investigated for joint detection and estimation systems for the first time.
- Both Bayesian and NP detection frameworks are considered, and the probability distribution of the optimal additive noise is shown to correspond to a discrete probability mass function with a certain number of point masses under each framework.
- For practical applications in which additive noise can take finitely many different values, a linear programming (LP) problem is formulated to obtain the optimal additive noise.
- Necessary and sufficient conditions are derived to specify the scenarios in which additive noise can

or cannot improve system performance.

In addition, theoretical results are also illustrated on a numerical example and noise benefits are investigated from various perspectives.

## II. PROBLEM FORMULATION

Consider a joint detection and estimation system as illustrated in Fig. 1, where the aim is to investigate possible improvements on the performance of the system by adding “noise”  $N$  to observation  $\mathbf{X}$ . In other words, instead of employing the original observation  $\mathbf{X}$ , the system operates based on the noise modified observation  $\mathbf{Y}$ , which is generated as follows:

$$\mathbf{Y} = \mathbf{X} + N. \quad (1)$$

The problem is defined as the determination of the optimum probability distribution for the additive noise without modifying the given joint detection and estimation system; that is, detector  $\phi(\cdot)$  and estimator  $\hat{\theta}(\cdot)$  are fixed. Also, the additive noise  $N$  is independent of the observation signal  $\mathbf{X}$ .

For the joint detection and estimation system, the model in [37] is adopted. Namely, the system consists of a detector and an estimator subsequent to it, and the detection is based on the following binary composite hypothesis testing problem [37]:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X} \sim f_0^X(\mathbf{x}) \\ \mathcal{H}_1 &: \mathbf{X} \sim f_1^X(\mathbf{x}|\Theta = \boldsymbol{\theta}), \quad \Theta \sim \pi(\boldsymbol{\theta}) \end{aligned} \quad (2)$$

where  $\mathbf{X} \in \mathbb{R}^K$  is the observation signal. Under hypothesis  $\mathcal{H}_0$ , the probability density function of the observation signal is completely known, which is denoted by  $f_0^X(\mathbf{x})$ . On the other hand, under hypothesis  $\mathcal{H}_1$ , the probability density function  $f_1^X(\mathbf{x}|\boldsymbol{\theta})$  of the observation signal  $\mathbf{X}$  is given as a function of the unknown parameter  $\boldsymbol{\theta}$ . It is also assumed that the prior distribution of the unknown parameter  $\Theta$  is available as  $\pi(\boldsymbol{\theta})$  in the parameter space  $\Lambda$  [37].

In Fig. 1, the detector is modeled as a generic decision rule  $\phi(\mathbf{x})$ , which specifies the decision probability in preference to hypothesis  $\mathcal{H}_1$  with  $0 \leq \phi(\mathbf{x}) \leq 1$ .<sup>1</sup> After the decision, there are two scenarios. If the

<sup>1</sup>Since a given (fixed) detection and estimation system is considered, the detector (decision rule) is modeled to be a generic one; that is, any deterministic or randomized decision rule can be considered.

decision is in favor of hypothesis  $\mathcal{H}_0$ , then there is no need for parameter estimation since the unknown parameter value is a known value, say  $\theta_0$ , under  $\mathcal{H}_0$ , and this knowledge is already included in  $f_0^X(x)$  (see (2)). If the decision is hypothesis  $\mathcal{H}_1$ , then the estimator in Fig. 1 estimates the value of the unknown parameter as  $\hat{\theta}(y)$ .

In general, the optimality of the detector and estimator to minimize the decision cost and estimation risk is an important goal in the detection and estimation theory. Optimal detectors and estimators for this joint detection and estimation scheme in the NP hypothesis-testing framework are already obtained in [37]. Since optimal detectors and estimators can have high computational complexity in some scenarios, the focus in this study is to consider fixed (given) detector and estimator structures with low complexity and to improve the performance of the given system only by adding “noise” to the observation as shown in Fig. 1.

With respect to the problem definition, different decision schemes such as Bayesian or NP approaches and estimation functions can be regarded in this context. If the prior probabilities of the hypotheses,  $P(\mathcal{H}_i)$ , are unknown, an NP type hypothesis-testing problem can be defined. On the other hand, if the prior probabilities are given, the Bayesian approach could be adopted [38]. The noise enhanced joint detection and estimation system is analyzed in both of these frameworks in parallel throughout the manuscript. For both frameworks, the aim is described as the optimization of the estimation performance without causing any degradation in the detection performance. Depending upon the application, the problem can be posed differently. It is not possible to cover all cases here, and the provided discussion can be considered to construct and solve similar problems (dual formulations) as well.

#### A. NP Hypothesis-Testing Framework

When the prior probabilities of the hypotheses are unknown, the NP framework can be employed for detection. In the NP detection framework, the main parameters are the probability of false alarm and the probability of detection [38]. Based on the hypotheses in (2) and the system model in Fig. 1, the

probabilities of false alarm and detection can be obtained, respectively, as

$$P_0(\hat{\mathcal{H}}_1) = \int_{\mathbb{R}^K} f^N(\mathbf{n}) \int_{\mathbb{R}^K} \phi(\mathbf{y}) f_0^X(\mathbf{y}-\mathbf{n}) d\mathbf{y} d\mathbf{n} \quad (3)$$

$$P_1(\hat{\mathcal{H}}_1) = \int_{\mathbb{R}^K} f^N(\mathbf{n}) \int_{\Lambda} \int_{\mathbb{R}^K} \phi(\mathbf{y}) \pi(\boldsymbol{\theta}) f_1^X(\mathbf{y}-\mathbf{n}|\boldsymbol{\theta}) d\mathbf{y} d\boldsymbol{\theta} d\mathbf{n} \quad (4)$$

where  $P_i(\hat{\mathcal{H}}_1)$  denotes the probability that the decision is hypothesis  $\mathcal{H}_1$ , denoted by  $\hat{\mathcal{H}}_1$ , when hypothesis  $\mathcal{H}_i$  is the true hypothesis.

Since the prior distribution of  $\Theta$  is known (see (2)), the Bayesian approach is employed for the estimation part; that is, the Bayes estimation risk is used as the performance criterion. The Bayes estimation risk is given by

$$r(\hat{\boldsymbol{\theta}}) = E\{C[\Theta, \hat{\boldsymbol{\theta}}(\mathbf{Y})]\} , \quad (5)$$

which is the expectation of the cost function  $C[\Theta, \hat{\boldsymbol{\theta}}(\mathbf{Y})]$  over the joint distribution of noise modified observation  $\mathbf{Y}$  and parameter  $\Theta$ . Squared error, absolute error, and uniform cost functions are three most commonly used cost functions in the literature [38]; the choice may depend on the application. For the binary hypothesis-testing problem in (2) and the system model in Fig. 1, the Bayes risk in (5) can be expressed as

$$r(\hat{\boldsymbol{\theta}}) = \sum_{i=0}^1 \sum_{j=0}^1 P(\mathcal{H}_i) P_i(\hat{\mathcal{H}}_j) E\{C(\Theta, \hat{\boldsymbol{\theta}}(\mathbf{Y})|\mathcal{H}_i, \hat{\mathcal{H}}_j)\} . \quad (6)$$

In the joint detection and estimation system, the estimation is dependent on the detection result; hence, the overall Bayes estimation risk is not independent of the detection performance. Due to this dependency, the calculation of the Bayes estimation risk requires the evaluation of the conditional risks for different true hypothesis and decided hypothesis pairs. As it is clear from (6), it is not possible to analytically evaluate the overall Bayes estimation risk function  $r(\hat{\boldsymbol{\theta}})$  in the NP framework since the prior probabilities of the hypotheses,  $P(\mathcal{H}_i)$ , are unknown. To avoid this complication, the conditional Bayes estimation risk  $J(\phi, \hat{\boldsymbol{\theta}})$ , which is presented in [37] as the Bayes estimation risk under the true hypothesis  $\mathcal{H}_1$  and decision  $\hat{\mathcal{H}}_j$ , is adopted in the following. Furthermore, it should be noted that if the decision is not

correct, it is expected the estimation error is relatively higher and may be regarded as useless for specific applications. Therefore, taking into consideration only the estimation error when the decision is correct could be justified as a rational argument. Since a probability distribution for unknown parameter  $\Theta$  is not defined under true hypothesis  $\mathcal{H}_0$ , the estimation error conditioned on the true hypothesis testing event is equivalent to the estimation error given true hypothesis  $\mathcal{H}_1$  and decision  $\hat{\mathcal{H}}_j$ . The conditional Bayes estimation risk is defined as [37]

$$J(\phi, \hat{\theta}) = E\{C(\Theta, \hat{\theta}(\mathbf{Y}))|\mathcal{H}_1, \hat{\mathcal{H}}_1\} \quad (7)$$

which can be expressed based on (7) in [37], the hypotheses in (2), and the system model in Fig. 1, as

$$J(\phi, \hat{\theta}) = \frac{\int_{\mathbb{R}^K} f^N(\mathbf{n}) \int_{\Lambda} \int_{\mathbb{R}^K} C(\theta, \hat{\theta}(\mathbf{y})) \phi(\mathbf{y}) \pi(\theta) f_1^X(\mathbf{y}-\mathbf{n}|\theta) d\mathbf{y} d\theta d\mathbf{n}}{P_1(\hat{\mathcal{H}}_1)} . \quad (8)$$

After some manipulation, (3), (4) and (8) can be expressed as the expectations of certain auxiliary functions with respect to additive noise distribution:

$$P_0(\hat{\mathcal{H}}_1) = E\{T(\mathbf{n})\}, \quad \text{where } T(\mathbf{n}) \triangleq \int_{\mathbb{R}^K} \phi(\mathbf{y}) f_0^X(\mathbf{y}-\mathbf{n}) d\mathbf{y} \quad (9)$$

$$P_1(\hat{\mathcal{H}}_1) = E\{R(\mathbf{n})\}, \quad \text{where } R(\mathbf{n}) \triangleq \int_{\Lambda} \int_{\mathbb{R}^K} \phi(\mathbf{y}) \pi(\theta) f_1^X(\mathbf{y}-\mathbf{n}|\theta) d\mathbf{y} d\theta \quad (10)$$

$$J(\phi, \hat{\theta}) = \frac{E\{G_{11}(\mathbf{n})\}}{E\{R(\mathbf{n})\}}, \quad \text{where } G_{11}(\mathbf{n}) \triangleq \int_{\Lambda} \int_{\mathbb{R}^K} C(\theta, \hat{\theta}(\mathbf{y})) \phi(\mathbf{y}) \pi(\theta) f_1^X(\mathbf{y}-\mathbf{n}|\theta) d\mathbf{y} d\theta . \quad (11)$$

It is noted that  $R(\mathbf{n}_0)$ ,  $T(\mathbf{n}_0)$ , and  $G_{11}(\mathbf{n}_0)/R(\mathbf{n}_0)$  correspond respectively to the detection probability, false alarm probability, and estimation risks of the system when the additive noise is equal to  $\mathbf{n}_0$ .

In the NP framework, the aim is to obtain the optimal probability distribution of the additive noise that minimizes the conditional estimation risk in (8) subject to the constraints on the probability of false alarm in (3) and the probability of detection in (4). The constraints are selected as the probabilities of false alarm and detection of the original system without any additive noise, which correspond to  $R(\mathbf{0})$  and  $T(\mathbf{0})$ , respectively (see (9) and (10)). In other words, while minimizing the conditional estimation risk, no degradation is allowed in the detection part of the system. The proposed optimization problem

can be expressed as

$$\underset{f^N(\mathbf{n})}{\text{minimize}} \frac{\mathbb{E}\{G_{11}(\mathbf{n})\}}{\mathbb{E}\{R(\mathbf{n})\}} \text{ subject to } \mathbb{E}\{T(\mathbf{n})\} \leq T(\mathbf{0}) \text{ and } \mathbb{E}\{R(\mathbf{n})\} \geq R(\mathbf{0}). \quad (12)$$

It is important to emphasize that a generic and fixed decision rule and estimator structure is considered in (12), and the aim is to improve the performance of the given system via the optimal additive noise as shown in Fig 1. Since the optimal decision rule and estimator in [37] can be quite complex in some cases, the use of a practical (low-complexity) decision rule and estimator structure together with noise enhancement can be considered for motivating the problem in (12).

### B. Bayesian Hypothesis-Testing Framework

When the prior probabilities of the hypotheses,  $P(\mathcal{H}_i)$ , are known, the Bayesian detection framework can be employed. In this framework, the Bayes detection risk,  $r(\phi)$ , is the main objective, which is defined as the average of the conditional risks as follows:

$$r(\phi) = P(\mathcal{H}_0) \sum_{j=0}^1 C_{j0} P_0(\hat{\mathcal{H}}_j) + P(\mathcal{H}_1) \sum_{j=0}^1 C_{j1} P_1(\hat{\mathcal{H}}_j) \quad (13)$$

where  $C_{ji}$  is the cost of choosing  $\mathcal{H}_j$  (i.e., the decision is  $\hat{\mathcal{H}}_j$ ) when  $\mathcal{H}_i$  is the true hypothesis, and  $\sum_{j=0}^1 C_{ji} P_i(\hat{\mathcal{H}}_j)$  is the conditional risk when hypothesis  $\mathcal{H}_i$  is true [38].

Determining the values of the costs  $C_{ji}$  generally depends on the application. As a reasonable choice,  $C_{ji}$  can be set to zero when  $j = i$  and to one when  $j \neq i$ , which is called uniform cost assignment (UCA) [38]. In that case, the Bayes detection risk is calculated as

$$r(\phi) = P(\mathcal{H}_0) P_0(\hat{\mathcal{H}}_1) + P(\mathcal{H}_1) P_1(\hat{\mathcal{H}}_0). \quad (14)$$

Based on the expressions in (9) and (10), (14) can be stated as

$$\begin{aligned} r(\phi) &= P(\mathcal{H}_0) \mathbb{E}\{T(\mathbf{n})\} + P(\mathcal{H}_1) (1 - \mathbb{E}\{R(\mathbf{n})\}) \\ &= P(\mathcal{H}_1) + \mathbb{E}\{P(\mathcal{H}_0)T(\mathbf{n}) - P(\mathcal{H}_1)R(\mathbf{n})\}. \end{aligned} \quad (15)$$

Since the prior probabilities of the hypotheses are known, the overall Bayes estimation risk function given in (6) can be evaluated in this case. As discussed previously, under  $\mathcal{H}_0$ , parameter  $\Theta$  is assumed to



have a deterministic value, which is equal to  $\theta_0$ . If the decision is  $\hat{\mathcal{H}}_1$ , the estimate  $\hat{\theta}(\mathbf{y})$  is produced for given observation  $\mathbf{y}$ . If the decision is  $\hat{\mathcal{H}}_0$ , the trivial estimation result is  $\theta_0$ . Notice that if the decision is correct when the true hypothesis is  $\mathcal{H}_0$ , the conditional estimation risk for this case is equal to zero. With this remark, the Bayes estimation risk in (6) becomes

$$\begin{aligned} r(\hat{\theta}) &= P(\mathcal{H}_0)P_0(\hat{\mathcal{H}}_1)\mathbb{E}\{C(\Theta, \hat{\theta}(\mathbf{Y})|\mathcal{H}_0, \hat{\mathcal{H}}_1)\} \\ &+ P(\mathcal{H}_1)\left(P_1(\hat{\mathcal{H}}_0)\mathbb{E}\{C(\Theta, \hat{\theta}(\mathbf{Y})|\mathcal{H}_1, \hat{\mathcal{H}}_0)\} + P_1(\hat{\mathcal{H}}_1)\mathbb{E}\{C(\Theta, \hat{\theta}(\mathbf{Y})|\mathcal{H}_1, \hat{\mathcal{H}}_1)\}\right) \end{aligned} \quad (16)$$

Similar to (7) and (8), (16) can be calculated as

$$\begin{aligned} r(\hat{\theta}) &= \int_{\mathbb{R}^K} f^N(\mathbf{n}) \left[ P(\mathcal{H}_0) \int_{\mathbb{R}^K} C(\theta_0, \hat{\theta}(\mathbf{y}))\phi(\mathbf{y})f_0^X(\mathbf{y}-\mathbf{n}) d\mathbf{y} \right. \\ &+ P(\mathcal{H}_1) \left( \int_{\Lambda} \int_{\mathbb{R}^K} C(\theta, \theta_0)(1-\phi(\mathbf{y}))f_1^X(\mathbf{y}-\mathbf{n}|\theta)\pi(\theta) d\mathbf{y} d\theta \right. \\ &\left. \left. + \int_{\Lambda} \int_{\mathbb{R}^K} C(\theta, \hat{\theta}(\mathbf{y}))\phi(\mathbf{y})f_1^X(\mathbf{y}-\mathbf{n}|\theta)\pi(\theta) d\mathbf{y} d\theta \right) \right] d\mathbf{n}. \end{aligned} \quad (17)$$

With the introduction of new auxiliary functions  $G_{01}(\mathbf{n})$  and  $G_{10}(\mathbf{n})$ , in addition to  $G_{11}(\mathbf{n})$  in (11), the Bayes estimation risk in (17) can be expressed as

$$r(\hat{\theta}) = \mathbb{E}\{P(\mathcal{H}_0)G_{01}(\mathbf{n}) + P(\mathcal{H}_1)[G_{11}(\mathbf{n}) + G_{10}(\mathbf{n})]\} \quad (18)$$

where

$$G_{01}(\mathbf{n}) \triangleq \int_{\mathbb{R}^K} C(\theta_0, \hat{\theta}(\mathbf{y}))\phi(\mathbf{y})f_0^X(\mathbf{y}-\mathbf{n}) d\mathbf{y} \quad (19)$$

$$G_{10}(\mathbf{n}) \triangleq \int_{\Lambda} \int_{\mathbb{R}^K} C(\theta, \theta_0)(1-\phi(\mathbf{y}))f_1^X(\mathbf{y}-\mathbf{n}|\theta)\pi(\theta) d\mathbf{y} d\theta. \quad (20)$$

In the Bayesian framework, the aim is the minimization of the Bayes estimation risk under a constraint on the Bayes detection risk. The Bayes detection risk constraint for the noise modified system is specified as the Bayes detection risk of the original system, which is  $P(\mathcal{H}_1) + P(\mathcal{H}_0)T(\mathbf{0}) - P(\mathcal{H}_1)R(\mathbf{0})$ . Then,

the proposed optimization problem is given by

$$\begin{aligned} & \underset{f^N(\mathbf{n})}{\text{minimize}} \quad \mathbb{E}\{P(\mathcal{H}_0)G_{01}(\mathbf{n}) + P(\mathcal{H}_1)[G_{11}(\mathbf{n}) + G_{10}(\mathbf{n})]\} \\ & \text{subject to} \quad \mathbb{E}\{P(\mathcal{H}_0)T(\mathbf{n}) - P(\mathcal{H}_1)R(\mathbf{n})\} \leq P(\mathcal{H}_0)T(\mathbf{0}) - P(\mathcal{H}_1)R(\mathbf{0}). \end{aligned} \quad (21)$$

### III. OPTIMUM NOISE DISTRIBUTIONS

The optimization problems in (12) and (21) require searches over all possible probability density functions (PDFs). These complex problems can be simplified by specifying the structure of the optimum noise probability distribution. Similar approaches are employed in various studies related to noise enhanced detection such as [15], which utilizes Caratheodory's theorem for noise enhanced binary hypothesis testing problems. It is proved that the optimum additive noise is characterized by a probability mass function (PMF) with at most two point masses under certain conditions in the binary hypothesis testing problem, where the objective function is the detection probability and the constraint function is the false alarm probability [15]. Using the primal-dual concept, [23] reaches PMFs with at most two point masses under certain conditions for binary hypothesis testing problems. In [29] and [25], the proof given in [15] is extended to hypothesis testing problems with  $(M - 1)$  constraint functions and the optimum noise distribution is found to have  $M$  point masses.

In this study, the objective function is the Bayes estimation risk in both of the proposed optimization problems in (12) and (21), and the constraint functions are defined in terms of the probability of detection and the probability of false alarm. The structures of the proposed problems are similar to those in [15], [23], [25], [29]. The same principles can be applied to both of the optimization problems in (12) and (21) and the optimum noise distribution structure can be specified under certain conditions as follows:

**Theorem 3.1:** Define set  $Z$  as  $Z = \{\mathbf{z} = (z_0, z_1, \dots, z_{K-1}) : z_i \in [a_i, b_i], i = 1, 2, \dots, K\}$ , where  $a_i$  and  $b_i$  are finite numbers, and define set  $U$  as  $U = \{\mathbf{u} = (u_0, u_1, u_2) : u_0 = R(\mathbf{n}), u_1 = T(\mathbf{n}), u_2 = G_{11}(\mathbf{n}), \text{ for } \mathbf{n} \in Z\}$ . Assume that the support set of the additive noise random variable is set  $Z$ . If  $U$  is a compact set in  $\mathbb{R}^K$ , the optimal solution of (12) can be represented by a discrete probability distribution with at most three point masses; that is,

$$f_{\text{opt}}^N(\mathbf{n}) = \sum_{i=1}^3 \kappa_i \delta(\mathbf{n} - \mathbf{n}_i). \quad (22)$$

*Proof:*  $U$  is the set of all possible detection probability, false alarm probability and conditional estimation risk triples for a given additive noise value  $\mathbf{n}$ , where  $\mathbf{n} \in Z$ . By the assumption in the theorem;  $U$  is a compact set; hence, it is a closed and bounded set. (A subset of  $\mathbb{R}^K$  is a closed and bounded set if and only if it is a compact set by Heine-Borel theorem.) Define set  $V$  as the convex hull of set  $U$ , and define  $W$  as the set of all possible values of  $E\{R(\mathbf{n})\}$ ,  $E\{T(\mathbf{n})\}$  and  $E\{G_{11}(\mathbf{n})\}$  triples as follows:

$$W = \{(w_0, w_1, w_2) : w_0 = E\{R(\mathbf{n})\}, w_1 = E\{T(\mathbf{n})\}, w_2 = E\{G_{11}(\mathbf{n})\}; \forall f^N(\mathbf{n}), \mathbf{n} \in Z\}. \quad (23)$$

It is already shown in the literature that set  $W$  and set  $V$  are equal [39]; that is,  $W = V$ . By the corollary of Carathéodory's theorem,  $V$  is also a compact set [40]. Due to the structure of the minimization problem in (12), the optimal solution must lie on the boundary of  $W$ ; equivalently,  $V$  [15], [23], [25], [29]. Then, from Carathéodory's theorem [40], it can be concluded that any point on the boundary of  $V$  can be expressed as the convex combination of at most three different points in  $U$ . The convex combination of three elements of  $U$  is equivalent to an expectation operation over additive noise  $N$ , where its distribution is a probability mass function with three point masses. ■

The same approach can be adopted to obtain the optimal solution of the problem (21) and it is stated without a proof. Define  $U$  as the set of all possible Bayes detection risk (14) and Bayes estimation risk (18) pairs for a given additive noise value  $\mathbf{n} \in Z$ , where  $Z$  is  $Z = \{\mathbf{z} = (z_0, z_1, \dots, z_{K-1}) : z_i \in [a_i, b_i], i = 1, 2, \dots, K\}$ , with  $a_i$  and  $b_i$  being finite numbers. Assume that the support set of the additive noise random variable is set  $Z$ . If  $U$  is a compact set in  $\mathbb{R}^K$ , the optimal solution of the (21) is given by a probability mass function with at most two point masses; that is,

$$f_{\text{opt}}^N(\mathbf{n}) = \sum_{i=1}^2 \kappa_i \delta(\mathbf{n} - \mathbf{n}_i). \quad (24)$$

The results in Theorem 3.1 and (24) state that when additive noise values are confined to some finite intervals (which always holds for practical systems), the optimal additive noise can be represented by a discrete random variable with three (two) point masses in the NP (Bayesian) detection framework. Then, from (22) and (24), the optimization problems in (12) and (21) can be restated as follows:

For the NP detection framework:

$$\begin{aligned}
& \underset{\kappa_1, \kappa_2, \kappa_3, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3}{\text{minimize}} && \frac{\sum_{i=1}^3 \kappa_i G_{11}(\mathbf{n}_i)}{\sum_{i=1}^3 \kappa_i R(\mathbf{n}_i)} && (25) \\
& \text{subject to} && \sum_{i=1}^3 \kappa_i T(\mathbf{n}_i) \leq T(\mathbf{0}), \quad \sum_{i=1}^3 \kappa_i R(\mathbf{n}_i) \geq R(\mathbf{0}) \\
& && \kappa_1, \kappa_2, \kappa_3 \geq 0 \text{ and } \kappa_1 + \kappa_2 + \kappa_3 = 1.
\end{aligned}$$

For the Bayes detection framework:

$$\begin{aligned}
& \underset{\kappa_1, \kappa_2, \mathbf{n}_1, \mathbf{n}_2}{\text{minimize}} && \sum_{i=1}^2 \kappa_i \left[ P(\mathcal{H}_0) G_{01}(\mathbf{n}_i) + P(\mathcal{H}_1) [G_{11}(\mathbf{n}_i) + G_{10}(\mathbf{n}_i)] \right] && (26) \\
& \text{subject to} && \sum_{i=1}^2 \kappa_i \left[ P(\mathcal{H}_0) T(\mathbf{n}_i) - P(\mathcal{H}_1) R(\mathbf{n}_i) \right] \leq P(\mathcal{H}_0) T(\mathbf{0}) - P(\mathcal{H}_1) R(\mathbf{0}) \\
& && \kappa_1, \kappa_2 \geq 0 \text{ and } \kappa_1 + \kappa_2 = 1.
\end{aligned}$$

Compared to the optimization problems in (12) and (21), which require searches over all possible PDFs, the formulations in (25) and (26) provide significant reductions in the computational complexity. However, the computational complexity of (25) and (26) can still be quite high in some cases since the problems are not convex in general. (The non-convexity of (25) and (26) is mainly due to the generality of the auxiliary functions, and the multiplication and division of functions involving the optimization variables.) Hence, a practical approach is considered in the next section.

#### IV. LINEAR PROGRAMMING (LP) APPROACH

The characteristics of the optimization problems in (25) and (26) are related to the given joint detection and estimation mechanism and the statistics of observation signal  $\mathbf{X}$  and parameter  $\Theta$ . The problems may not be convex in general. Therefore, the application of global optimization techniques can be necessary to obtain the solutions [41], [42]. As an alternative method, the optimization problems in (12) and (21) can be approximated as linear programming (LP) problems. LP problems are a special case of convex problems and they have lower computational load (solvable in polynomial time) than the possible global optimization techniques [43].

In order to achieve the LP approximation of the problem in (12), the support of the additive noise is

restricted to a finite set  $\mathbb{S} = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_M\}$ . In real life applications, it is not possible to generate an additive noise component which can take infinitely many different values in an interval; hence, it is a reasonable assumption that additive noise component can only have finite precision. With this approach, the possible values of  $R(\mathbf{n})$ ,  $T(\mathbf{n})$  and  $G_{11}(\mathbf{n})$  can be expressed as  $M$  dimensional column vectors and the expectation operation reduces to a convex combination of the elements of these column vectors with weights  $\lambda_1, \lambda_2, \dots, \lambda_M$ . The optimal values of the LP approximated problems are worse than or equal to the optimal values of the original optimization problems in (12) and (21) (equivalently, in (25) and (26)), and the gap between these results is dependent upon the number of noise samples, which is denoted by  $M$  in this formulation. For notational convenience, the following column vectors are defined:

$$\begin{aligned}\mathbf{t}^\top &= [T(\mathbf{n}_1) \quad T(\mathbf{n}_2) \quad \dots \quad T(\mathbf{n}_M)] \\ \mathbf{r}^\top &= [R(\mathbf{n}_1) \quad R(\mathbf{n}_2) \quad \dots \quad R(\mathbf{n}_M)] \\ \mathbf{g}^\top &= [G_{11}(\mathbf{n}_1) \quad G_{11}(\mathbf{n}_2) \quad \dots \quad G_{11}(\mathbf{n}_M)]\end{aligned}$$

Then, the optimization problem in (12), which considers the minimization of the conditional Bayes estimation risk, can be approximated as the following linear fractional programming (LFP) problem:

$$\begin{aligned}\text{minimize}_{\boldsymbol{\lambda}} \quad & \frac{\mathbf{g}^\top \boldsymbol{\lambda}}{\mathbf{r}^\top \boldsymbol{\lambda}} \\ \text{subject to} \quad & \mathbf{r}^\top \boldsymbol{\lambda} \geq R(\mathbf{0}) \\ & \mathbf{t}^\top \boldsymbol{\lambda} \leq T(\mathbf{0}) \\ & \mathbf{1}^\top \boldsymbol{\lambda} = 1 \\ & \boldsymbol{\lambda} \succeq 0.\end{aligned}\tag{27}$$

An example of transformation from an LFP problem to an LP problem is presented in [43]. The same approach can be adopted to obtain an LP problem as explained in the following. The optimization variable  $\mathbf{l}$  in the LP problem, to be employed in (29), is expressed as

$$\mathbf{l} = \frac{\boldsymbol{\lambda}}{\mathbf{r}^\top \boldsymbol{\lambda}}.\tag{28}$$

Notice that  $\mathbf{r}$  and  $\boldsymbol{\lambda}$  have non-negative components, and  $\mathbf{r}^\top \boldsymbol{\lambda}$  represents the detection probability of the noise modified mechanism. Therefore, it can be assumed that  $\mathbf{r}^\top \boldsymbol{\lambda}$  is positive valued and less than or equal

to 1. With this assumption, it is straightforward to prove the equivalence of the LP and LFP problems by showing that if  $\lambda$  is feasible in (27), then  $\mathbf{l}$  is also feasible in (29) with the same objective value, and vice versa. Hence, the following problem is obtained:

$$\begin{aligned}
& \underset{\mathbf{l}}{\text{minimize}} && \mathbf{g}^\top \mathbf{l} \\
& \text{subject to} && \mathbf{t}^\top \mathbf{l} \leq T(\mathbf{0})(\mathbf{1}^\top \mathbf{l}) \\
& && \mathbf{1}^\top \mathbf{l} \leq 1/R(\mathbf{0}) \\
& && \mathbf{r}^\top \mathbf{l} = 1 \\
& && \mathbf{l} \succeq 0.
\end{aligned} \tag{29}$$

The LP approximation of the optimization problem (21) is also obtained by limiting the possible additive noise values to a finite set  $\mathbb{S}' = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{M'}\}$ . With that restriction, the LP problem is given by

$$\begin{aligned}
& \underset{\lambda}{\text{minimize}} && \mathbf{q}^\top \lambda \\
& \text{subject to} && \mathbf{p}^\top \lambda \leq P(\mathcal{H}_0)T(\mathbf{0}) - P(\mathcal{H}_1)R(\mathbf{0}) \\
& && \mathbf{1}^\top \lambda = 1 \\
& && \lambda \succeq 0.
\end{aligned} \tag{30}$$

where

$$\begin{aligned}
\mathbf{p}^\top &= [p_1 \ p_2 \ \dots \ p_{M'}], \quad p_i = P(\mathcal{H}_0)T(\mathbf{n}_i) - P(\mathcal{H}_1)R(\mathbf{n}_i) \\
\mathbf{q}^\top &= [q_1 \ q_2 \ \dots \ q_{M'}], \quad q_i = P(\mathcal{H}_0)G_{01}(\mathbf{n}_i) + P(\mathcal{H}_1)[G_{11}(\mathbf{n}_i) + G_{10}(\mathbf{n}_i)].
\end{aligned}$$

Compared to (25) and (26), the problems in (29) and (30) can have significantly lower computational complexity in general since they are in the form of linear programs. In addition, as the number of possible noise values increases (i.e., as  $M$  or  $M'$  increases), the solution obtained from the LP approach gets closer to the optimal solution. Therefore, the LP approach can be more preferable in practical applications.

## V. IMPROVABILITY AND NON-IMPROVABILITY CONDITIONS

Before attempting to solve the optimization problems in (25) and (26), or the LP problems in (29) and (30), it is worthwhile to investigate the improvability of the given system via additive noise. The

joint detection and estimation system in the NP framework is called *improvable* if there exists a PDF  $f^N(\mathbf{n})$  for the additive noise  $\mathbf{N}$  such that  $J(\phi, \hat{\boldsymbol{\theta}}) < G_{11}(\mathbf{0})/R(\mathbf{0})$ ,  $P_1(\hat{\mathcal{H}}_1) \geq R(\mathbf{0})$  and  $P_0(\hat{\mathcal{H}}_1) \leq T(\mathbf{0})$ , and *non-improvable* if there does not exist such a PDF (cf. (9)-(12)). Similarly, the joint system in the Bayes detection framework is called *improvable* if there exists a PDF  $f^N(\mathbf{n})$  such that  $r(\hat{\boldsymbol{\theta}}) < P(\mathcal{H}_0)G_{01}(\mathbf{0}) + P(\mathcal{H}_1)[G_{11}(\mathbf{0}) + G_{10}(\mathbf{0})]$  and  $r(\phi) \leq P(\mathcal{H}_0)T(\mathbf{0}) - P(\mathcal{H}_1)R(\mathbf{0})$ , and *non-improvable* otherwise (cf. (15), (18), and (21)). Improvable and non-improvable joint detection and estimation systems under the LP approximation can also be defined in a similar fashion for both detection frameworks.

In the following, necessary and sufficient conditions are presented for the non-improvability (improvability) of given detection and estimation systems under the NP and Bayesian detection frameworks for the LP formulations.

**Theorem 5.1:** Consider the LFP problem in (27), where the aim is to optimize the system performance in the NP detection framework via additive noise, which is restricted to a finite set  $\mathbb{S} = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_M\}$ . Then, the joint detection and estimation system is non-improvable if and only if there exist  $\gamma_1, \gamma_2, \nu \in \mathbb{R}$ , with  $\gamma_1, \gamma_2 \geq 0$ , and  $\nu \leq -[G_{11}(\mathbf{0}) + \gamma_2]/R(\mathbf{0})$ , that satisfy the following set of inequalities:

$$G_{11}(\mathbf{n}_i) + \gamma_1(T(\mathbf{n}_i) - T(\mathbf{0})) + \gamma_2 + \nu R(\mathbf{n}_i) \geq 0, \quad \forall i \in \{1, 2, \dots, M\}. \quad (31)$$

*Proof:* In equation (29), the equivalent LP problem of the LFP problem in (27) is given. The dual problem of the LP problem is found as the following:

$$\begin{aligned} & \underset{\nu, \gamma_1, \gamma_2, \mathbf{u}}{\text{maximize}} && -\nu - \gamma_2/R(\mathbf{0}) \\ & \text{subject to} && G_{11}(\mathbf{n}_i) + \gamma_1(T(\mathbf{n}_i) - T(\mathbf{0})) + \gamma_2 + \nu R(\mathbf{n}_i) = u_i, \quad \forall i \in \{1, 2, \dots, M\} \\ & && \gamma_1, \gamma_2, u_1, u_2, \dots, u_M \geq 0 \end{aligned} \quad (32)$$

where  $\mathbf{u}^\top = [u_1 \ u_2 \ \dots \ u_M]$ . Let  $\mathbb{P}$  and  $\mathbb{D}$  be the feasible sets of the primal (27) and dual (32) problems, respectively. The objective functions of the primal and dual problems are denoted respectively as  $f_{\text{obj}}^P(p)$  and  $f_{\text{obj}}^D(d)$ , where  $p \in \mathbb{P}$  and  $d \in \mathbb{D}$ . Also,  $p^*$  and  $d^*$  represent the optimal solutions of the primal and dual problems, respectively. By the strong duality property of the LP problems,  $p^* = d^*$  [43].

Sufficient condition for non-improvability: Assume that  $\exists \gamma_1, \gamma_2, \nu \in \mathbb{R}$ ,  $\mathbf{u} \in \mathbb{R}^K$  such that  $\gamma_1, \gamma_2 \geq 0$ ,  $\mathbf{u} \succeq 0$ ,  $\nu \leq -[G_{11}(\mathbf{0}) + \gamma_2]/R(\mathbf{0})$ , and  $\gamma_1, \gamma_2, \nu, \mathbf{u}$  satisfy the following set of equations:  $G_{11}(\mathbf{n}_i) +$

$\gamma_1(T(\mathbf{n}_i) - T(\mathbf{0})) + \gamma_2 + \nu R(\mathbf{n}_i) = u_i \geq 0, \forall i \in \{1, 2, \dots, M\}$ . These variables describe an element of the dual feasible set  $d^\circ = (\gamma_1, \gamma_2, \nu, \mathbf{u}) \in \mathbb{D}$ .  $f_{\text{obj}}^D(d^\circ) = -\nu - \gamma_2/R(\mathbf{0}) \geq G_{11}(\mathbf{0})/R(\mathbf{0})$  by the assumption. This implies that  $G_{11}(\mathbf{0})/R(\mathbf{0}) \leq f_{\text{obj}}^D(d^\circ) \leq d^* = p^*$ ; hence, the conditional Bayes risk of the system in the NP framework cannot be reduced from its original value.

Necessary condition for non-improvability: To prove the necessary condition, it is equivalent to show that the system performance can be improved if  $\forall \gamma_1, \gamma_2, \nu \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^K$  such that  $\gamma_1, \gamma_2 \geq 0, \mathbf{u} \succeq 0, \nu \geq -[G_{11}(\mathbf{0}) + \gamma_2]/R(\mathbf{0})$ , the following set of equations are satisfied:  $G_{11}(\mathbf{n}_i) + \gamma_1(T(\mathbf{n}_i) - T(\mathbf{0})) + \gamma_2 + \nu R(\mathbf{n}_i) = u_i \geq 0, \forall i \in \{1, 2, \dots, M\}$ . Observe that  $\gamma_2$  or  $\nu$  can always be picked arbitrarily large to satisfy the equality constraints given in (32), since  $1 \geq R(\mathbf{n}_i) \geq 0, 1 \geq T(\mathbf{n}_i) \geq 0$  and  $G_{11}(\mathbf{n}_i) \geq 0$ . Therefore, the feasible set of the dual problem cannot be empty, i.e.,  $\mathbb{D} \neq \emptyset$ . Notice that the assumption implies  $\forall d \in \mathbb{D}, f_{\text{obj}}^D(d) < G_{11}(\mathbf{0})/R(\mathbf{0})$ . For this reason and with the strong duality property it can be asserted that  $d^* = p^* < G_{11}(\mathbf{0})/R(\mathbf{0})$  since  $d^* = f_{\text{obj}}^D(d^{\text{opt}}), d^{\text{opt}} \in \mathbb{D}$ . ■

**Theorem 5.2:** Consider the LP problem in (30), where the aim is to optimize the system performance in the Bayes detection framework via additive noise, which is restricted to a finite set  $\mathbb{S}' = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{M'}\}$ . Then, the joint detection and estimation system is non-improvable if and only if there exist  $\gamma, \nu \in \mathbb{R}$  with  $\gamma \geq 0$  that satisfy the following set of inequalities:

$$P(\mathcal{H}_0)[\gamma T(\mathbf{0}) + G_{01}(\mathbf{0})] + P(\mathcal{H}_1)[G_{11}(\mathbf{0}) + G_{10}(\mathbf{0}) - \gamma R(\mathbf{0})] + \nu \leq 0 \quad (33)$$

$$P(\mathcal{H}_0)[\gamma T(\mathbf{n}_i) + G_{01}(\mathbf{n}_i)] + P(\mathcal{H}_1)[G_{11}(\mathbf{n}_i) + G_{10}(\mathbf{n}_i) - \gamma R(\mathbf{n}_i)] + \nu \geq 0 \quad (34)$$

$\forall i \in \{1, 2, \dots, M'\}$ . Note that if  $\mathbf{0} \in \mathbb{S}' = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{M'}\}$ , then the inequality in (33) must be satisfied with equality. With this, the necessary and sufficient conditions in (33) and (34) are expressed as

$$\begin{aligned} & P(\mathcal{H}_1)[G_{11}(\mathbf{n}_i) + G_{10}(\mathbf{n}_i) - \gamma R(\mathbf{n}_i) - G_{11}(\mathbf{0}) - G_{10}(\mathbf{0}) + \gamma R(\mathbf{0})] \\ & + P(\mathcal{H}_0)[G_{01}(\mathbf{n}_i) + \gamma T(\mathbf{n}_i) - G_{01}(\mathbf{0}) - \gamma T(\mathbf{0})] \geq 0. \end{aligned} \quad (35)$$

The proof of Theorem 5.2 is not presented since it can be obtained based on an approach similar to that employed in the proof of Theorem 5.1. It should be emphasized that the conditions in these theorems specify whether the system can be improved via additive noise or not. In other words, they are



both improvability and non-improvability conditions: If the conditions in the theorems are satisfied, the system is non-improvable (performance cannot be enhanced via additive noise); otherwise, the system is improvable (performance can be enhanced via additive noise).

Notice that the LP approach is based on sampling the objective and constraint functions. Therefore, the presented sufficient and necessary conditions in Theorems 5.1 and 5.2 demonstrate the convex geometry of the optimization problems in (12) and (21). For similar problem formulations, different necessary or sufficient improvability or nonimprovability conditions are stated in the literature [15], [23]–[25], [28]. In [28], firstly, a necessary and sufficient condition is presented for a similar single inequality constrained problem with a continuous support set. It should be noted that (21) is a single inequality constrained problem and its necessary and sufficient non-improvability condition for the LP approach in (35) share the same structure with the inequality (10) in [28] under a certain condition. Theorem 5.1 extends this result to the problems with multiple inequality constraints and finite noise random variable support set from a completely different perspective. The merit of this approach, which is presented in the proof of Theorem 5.1, is that it is generic and can easily be adapted to different problems. In this study, the main focus is on the justification of the LP approach for noise enhancement problems in joint detection and estimation systems. A natural extension of Theorem 5.1 which is the formulation for a continuous support set is omitted.

## VI. ANALYSIS OF A GIVEN JOINT DETECTION ESTIMATION SYSTEM

In this section, a binary hypothesis testing example is analyzed to demonstrate the noise enhancement effects on the described joint detection and estimation system.

The hypothesis testing problem is specified as follows:

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{X} = \boldsymbol{\epsilon} \\ \mathcal{H}_1 &: \mathbf{X} = \boldsymbol{\epsilon} + \Theta\end{aligned}\tag{36}$$

where  $\mathbf{X}$  is the observation with  $\mathbf{X} = [X_1 \ X_2 \ \cdots \ X_K]^\top$ ,  $\Theta$  is the parameter with  $\Theta = \Theta \mathbf{1}$  ( $\mathbf{1} = [1 \ 1 \ \cdots \ 1]^\top$ ), and  $\boldsymbol{\epsilon} = [\epsilon_1 \ \epsilon_2 \ \cdots \ \epsilon_K]^\top$  is the system noise. In the example,  $\Theta$  is taken to be Gaussian distributed random variable and its value is to be estimated. Also,  $\epsilon_k$ 's are identically and independent dis-

tributed according to a known Gaussian mixture distribution. It is assumed that both of these distributions are known.

More specifically, the parameter  $\Theta$  is taken as Gaussian distributed with  $\Theta \sim \mathcal{N}(a, b^2)$ ; that is,

$$\pi(\theta) = \frac{1}{\sqrt{2\pi b^2}} \exp \left\{ -\frac{(\theta - a)^2}{2b^2} \right\}. \quad (37)$$

In addition, the components of the system noise  $\epsilon$  are identical, independent and Gaussian mixture distributed as follows:

$$f_{\epsilon_k}(\epsilon) = \sum_{i=1}^{N_m} \frac{\nu_i}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(\epsilon - \mu_i)^2}{2\sigma^2} \right\} \quad (38)$$

Notice that each element of the Gaussian mixture has different mean  $\mu_i$  and weight  $\nu_i$  with the same standard deviation  $\sigma$ . The mixture background noise is encountered in a variety of contexts [44] (and references therein) such as co-channel interference [45], ultra-wideband synthetic aperture radar (UWB SAR) imaging [46], and underwater noise modeling [47]. The standard deviation values are taken to be equal for all the mixture components just to simplify the analytical evaluation of this problem for  $K > 1$ . The standard deviation values can also be taken to be different for each mixture component.

Noise  $\mathbf{N}$  is added to observation  $\mathbf{X}$  as shown in Fig. 1 for the purpose of noise enhancement, and the noise modified observation  $\mathbf{Y}$  is obtained as in (1). The decision rule is a threshold detector, and it outputs the probability of deciding in favor of  $\mathcal{H}_1$  as follows:

$$\phi(\mathbf{y}) = \begin{cases} 1 & \text{if } \frac{1}{K} \sum_{i=1}^K y_i > \tau_{PF} \\ 0 & \text{if } \frac{1}{K} \sum_{i=1}^K y_i \leq \tau_{PF} \end{cases} \quad (39)$$

where the subscript  $PF$  is used for the threshold  $\tau$  to emphasize that threshold  $\tau_{PF}$  is determined according to the predetermined probability of false alarm. The decision rule in (39) is a simple and reasonable one which compares the sample mean of the observations against the threshold. In addition, the estimation

cost function in (5) is a uniform cost function specified by

$$C(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{y})) = \begin{cases} 1, & \text{if } |\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}| > \Delta \\ 0, & \text{otherwise} \end{cases} \quad (40)$$

where  $\Delta > 0$  is the threshold of the uniform cost function [38]. More specifically, according to this cost function, estimation errors up to  $\Delta$  are tolerable and they incur no cost, whereas estimation errors higher than  $\Delta$  are associated with a unit cost. For the estimator in Fig. 1, a sample mean estimator is considered, which is expressed as

$$\hat{\boldsymbol{\theta}}(\mathbf{y}) = \frac{1}{K} \sum_{i=1}^K y_i. \quad (41)$$

Notice that the considered detector and estimator structures are not optimal due to the presence of Gaussian mixture noise  $\epsilon$ . However, they are practical ones with low-computational complexity. [The optimal detector and estimator in \[37\] would have significantly higher computational complexity for the considered scenario with Gaussian mixture noise.](#) Therefore, the aim is to employ the low-complexity structures in (39) and (41) and to improve their performance via additive noise.

#### A. Scalar Case, $K = 1$

In this case, functions  $T(n)$ ,  $R(n)$ ,  $G_{11}(n)$ ,  $G_{01}(n)$ , and  $G_{10}(n)$  defined respectively in (9), (10), (11), (19), and (20) are derived as follows:<sup>2</sup>

$$T(n) = \sum_{i=1}^{N_m} \nu_i \mathcal{Q} \left( \frac{\tau_{PF} - \mu_i - n}{\sigma_i} \right) \quad (42)$$

$$R(n) = \sum_{i=1}^{N_m} \nu_i \mathcal{Q} \left( \frac{\tau_{PF} - \mu_i - a - n}{\sqrt{\sigma_i^2 + b^2}} \right) \quad (43)$$

<sup>2</sup>The details of the derivations are not presented.

$$G_{11}(n) = \sum_{i=1}^{N_m} \nu_i \left( Q\left(\frac{\tau_{PF} - \mu_i - n - a}{\sqrt{b^2 + \sigma_i^2}}\right) - \int_{\tau_{PF} - \Delta}^{\tau_{PF} + \Delta} \pi(\theta) Q\left(\frac{\tau_{PF} - \mu_i - n - \theta}{\sigma_i}\right) d\theta \right. \\ \left. + Q\left(\frac{\Delta - \mu_i - n}{\sigma_i}\right) Q\left(\frac{\tau_{PF} - \Delta - a}{b}\right) + Q\left(\frac{-\Delta - \mu_i - n}{\sigma_i}\right) Q\left(\frac{\tau_{PF} + \Delta - a}{b}\right) \right) \quad (44)$$

$$G_{01}(n) = \begin{cases} \sum_{i=1}^{N_m} \nu_i Q\left(\frac{\tau_{PF} - \mu_i - n}{\sigma_i}\right), & \text{if } \tau_{PF} > \Delta \\ \sum_{i=1}^{N_m} \nu_i Q\left(\frac{\Delta - \mu_i - n}{\sigma_i}\right), & \text{if } \Delta \geq \tau_{PF} > -\Delta \\ \sum_{i=1}^{N_m} \nu_i \left( Q\left(\frac{\tau_{PF} - \mu_i - n}{\sigma_i}\right) + Q\left(\frac{\Delta - \mu_i - n}{\sigma_i}\right) - Q\left(\frac{-\Delta - \mu_i - n}{\sigma_i}\right) \right), & \text{if } -\Delta \geq \tau_{PF} \end{cases} \quad (45)$$

$$G_{10}(n) = \sum_{i=1}^{N_m} \nu_i \left( \Phi\left(\frac{\tau_{PF} - \mu_i - n - a}{\sqrt{b^2 + \sigma_i^2}}\right) - \int_{-\Delta}^{\Delta} \pi(\theta) \Phi\left(\frac{\tau_{PF} - \mu_i - n - \theta}{\sigma_i}\right) d\theta \right) \quad (46)$$

where  $Q(\cdot)$  and  $\Phi(\cdot)$  are respectively the tail probability function and the cumulative distribution function of the standard Gaussian random variable.

### B. Vector Case, $K > 1$

To evaluate the performance of this system (with and without noise enhancement), the statistics of  $\frac{1}{K} \sum_{i=1}^K x_i$  need to be revealed. Additive noise and noise modified observation are represented as  $\mathbf{N} = [N_1 \ N_2 \ \dots \ N_K]^\top$  and  $\mathbf{Y} = [Y_1 \ Y_2 \ \dots \ Y_K]^\top$ . Denote  $\frac{1}{K} \sum_{i=1}^K N_i$  with  $\tilde{N}$  and  $\frac{1}{K} \sum_{i=1}^K \epsilon_i$  with  $\tilde{\epsilon}_K$ . Under  $\mathcal{H}_1$  and with additive noise, this vector joint detection and estimation problem can be reexpressed as a scalar problem as follows:

$$\text{Under } \mathcal{H}_1 : \frac{1}{K} \sum_{i=1}^K Y_i = \frac{1}{K} \sum_{i=1}^K (\Theta + N_i + \epsilon_i) = \Theta + \tilde{N} + \tilde{\epsilon}_K. \quad (47)$$

It can be shown that  $\tilde{\epsilon}$  has the following Gaussian mixture distribution [48]:

$$f_{\tilde{\epsilon}_K}(\epsilon) = \sum_{j=1}^{\tilde{N}_m} \frac{\tilde{\nu}_j}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left\{-\frac{(\epsilon - \tilde{\mu}_j)^2}{2\tilde{\sigma}^2}\right\}, \quad (48)$$

where

$$\tilde{N}_m = \binom{K + N_m - 1}{N_m - 1}, \quad \tilde{\sigma}^2 = \frac{\sigma^2}{K}, \quad \tilde{\nu}_j = \left( \frac{K!}{l_1! \ l_2! \ \dots \ l_{N_m}!} \right) \left( \prod_{i=1}^{N_m} \nu_i^{l_i} \right), \quad \tilde{\mu}_j = \frac{1}{K} \sum_{i=1}^{N_m} \mu_i l_i$$

for each distinct  $\{l_1, l_2, \dots, l_{N_m}\}$  set satisfying  $l_1 + l_2 + \dots + l_{N_m} = K$ ,  $l_i \in \{1, 2, \dots, K\}$ ,  $i \in$

$\{1, 2, \dots, N_m\}$ . With this result, the vector case reduces to the scalar case. The derived expressions in the  $K = 1$  case for function  $T(n)$ ,  $R(n)$ ,  $G_{11}(n)$ ,  $G_{10}(n)$ , and  $G_{01}(n)$  (see (42)-(46)) do also apply to the  $K > 1$  case, where the only necessary modification is the usage of new mean  $\tilde{\mu}_j$ , weight  $\tilde{\nu}_j$ , and standard deviation  $\tilde{\sigma}$  values. With this approach, the optimal statistics for the design of random variable  $\tilde{N}$  is revealed. The mapping from  $\tilde{N}$  to  $\mathbf{N}$  is left to the designer. A very straightforward choice can be  $\mathbf{N} = [\tilde{N}K \ 0 \ \dots \ 0]^\top$ .

In this joint detection and estimation problem, the components of the system noise  $\epsilon$  are independent and identically distributed Gaussian mixture random variables. A similar analysis can also be carried out for a system noise with components being generalized Gaussian distributed. However, in general, it is not possible to express the density of the sum of the independent and identically distributed generalized Gaussian random variables with an exact analytical expression. The distribution of the sum is not generalized Gaussian (only exception is the Gaussian distribution) [49]. However, functions  $T$ ,  $R$ ,  $G_{11}$ ,  $G_{10}$ , and  $G_{01}$  can be evaluated numerically and the LP approximation can be applied.

### C. Asymptotic Behaviour of the System, Large $K$ Values

As  $K$  goes to infinity ( $K \rightarrow \infty$ ), by Lindeberg Lévy Central Limit Theorem,  $\sqrt{K} \left( \left( \frac{1}{K} \sum_{i=1}^K \epsilon_i \right) - \mu_\epsilon \right)$  converges in distribution to a Gaussian random variable  $\mathcal{N}(0, \sigma_\epsilon^2)$  given that  $\{\epsilon_1, \epsilon_2, \dots, \epsilon_K\}$  is a sequence of independent and identically distributed random variables with  $E\{\epsilon_i\} = \mu_\epsilon$ ,  $\text{Var}\{\epsilon_i\} = \sigma_\epsilon^2 < \infty$ . This general result applies to the analysis of the given joint detection and estimation problem in this section. For large  $K$  values, the probability density function of  $\tilde{\epsilon}_K = \frac{1}{K} \sum_{i=1}^K \epsilon_i$  can be approximated by the distribution of a Gaussian random variable  $\mathcal{N}(\mu_\epsilon, \sigma_\epsilon^2/K)$ .

## VII. NUMERICAL RESULTS FOR THE JOINT DETECTION AND ESTIMATION SYSTEM

For the numerical examples, the joint detection and estimation system in Section VI is considered, and the parameter values are set as follows: For each element  $\epsilon_k$  of the Gaussian mixture noise specified by the PDF in (38), the weights and the means of the Gaussian components are set to  $\boldsymbol{\nu} = [0.40 \ 0.15 \ 0.45]$  and  $\boldsymbol{\mu} = [5.25 \ -0.5 \ -4.5]$ , respectively. Also, the standard deviation of the mixture components, denoted by  $\sigma$  in (38), is considered as a variable to evaluate the performance of noise enhancement for various

signal-to-noise ratio (SNR) values. The decision rule is as specified in (39), where  $\tau_{PF}$  is set in such a way that the probability of false alarm of the given system is equal to 0.15 for all standard deviation values. Regarding the prior PDF of the unknown parameter in (37), the mean parameter  $a$  is set to 4.5 and the standard deviation  $b$  is equal to 1.25. It can be shown that the estimator in (41) is unbiased for the considered scenario in the absence of additive noise. In addition, for the uniform estimation cost function in (40), parameter  $\Delta$  is taken as 0.75. Furthermore, the support of the additive noise is considered as  $[-10, 10]$ .

First, the NP detection framework is considered and the optimal additive noise distributions are obtained from (25) for the exact (global) solution, and from (29) for the LP based solution. The conditional estimation risks are plotted versus  $\sigma$  in Fig. 2 for  $K = 1$  and in Fig. 3 for  $K = 4$ , in the absence of (original) and in the presence of additive noise. It is observed that the performance improvement via additive noise is reduced as the standard deviation  $\sigma$  increases. In other words, the noise enhancement is more significant in the high SNR region. The improvement is mainly caused by the multimodal nature of the observation statistics and increasing the standard deviation  $\sigma$  reduces this effect. In both of the figures, the performances of the LP approximations are also illustrated in comparison with the global (exact) solution. In obtaining the LP based solutions, the additive noise samples are taken uniformly from the support of the additive noise with the specified step size values in the figures. It is observed that the LP based solution achieves improved performance as the step size decreases; i.e., as more additive noise values are considered. It is also noted that the LP based approach achieves very close performance to the global solution for reasonable step sizes. **Another observation is that the conditional estimation risk is not always monotone with respect to the standard deviation for the original system and the LP approach with step size 0.5, which is mainly due to the suboptimality of the employed decision rule and the estimator (see, e.g., [15] and [29] for similar observations).** As it is clear from Fig. 2 and Fig. 3, the performance of the given joint detection system is superior for  $K = 4$  in comparison to the scalar case,  $K = 1$ . In this numerical example, the vector case corresponds to taking more samples and an increase in the SNR. Some numerical values of the conditional estimation risk, the detection probability, and the false alarm probability of this noise enhanced system are presented in Table I, together with the original values in the absence of additive noise. Also, the values of the detector threshold,  $\tau_{PF}$ , are shown in the table. It

is observed that the noise enhanced systems have the same detection and false alarm probabilities as the original system (i.e., they satisfy the constraints in (25) and (29) with equality), and they achieve a lower conditional estimation risk than the original system.

In Figs. 4–6 the solutions of the optimization problem in (25) are presented together with the solutions of the LP problem in (29) for various step sizes, where the standard deviations for the components of the Gaussian mixture system noise are equal to 0.3, 0.3, and 0.4, and  $K$  is set to 1, 4, and 4 for Fig. 4, Fig. 5, and Fig. 6, respectively. In the figures, the locations and weights of the point masses are presented. According to Theorem 3.1, the optimal solution of the optimization problem in (12) is a probability mass function with at most three point masses as shown in the figures. It should also be emphasized that Theorem 3.1 is valid only for the exact (global) solution; hence, the LP approach can in general have a solution with more than three point masses although it is not the case in this specific example.

Next, for the same system noise PDF  $f_{\epsilon_k}(\varepsilon)$ , the problem in the Bayes detection framework is studied for  $P(\mathcal{H}_0) = 0.5$  and  $\tau_{PF} = a/2$  (see (37)). The optimization problems in (26) and (30) are considered for obtaining the exact (global) and LP based solutions. In Fig. 7 ( $K = 1$ ) and Fig. 8 ( $K = 4$ ), the Bayes estimation risks are plotted versus the standard deviation in the absence (original) and presence of additive noise, where both the global and LP based solutions are shown for noise enhancement. For the LP based solution, various step sizes are considered. The behaviors of the curves are very similar to those in the NP detection framework. In particular, the noise enhancement effects are again more significant in the high SNR region, and the LP based approach gets very close to the global solution for reasonable step sizes. Some numerical values of the Bayes estimation risk and the Bayes detection risk of both the original and the noise enhanced systems are presented in Table II.

In practice, the step size for the LP based approach can be adapted as follows: Starting from a reasonably large step size, the step size is decreased (say by a certain fraction) and the difference between the estimated values is monitored. This operation of step size reduction can continue until the difference between consecutive estimated values gets smaller than a certain threshold. (If there is a difference between the decisions, i.e., selection of different hypotheses, then this is considered as a significant difference and the step size reduction continues.)

As discussed in Section III, the optimal additive noise in the Bayes detection framework is specified

by a probability mass function with one or two point masses. In Figs. 9–11 the optimal solutions of the problem in (26) and the solutions of the LP problem in (30) are illustrated for  $K = 1$  and  $\sigma = 0.5$ ,  $K = 4$  and  $\sigma = 0.5$ , and  $K = 4$  and  $\sigma = 0.75$ , respectively. It is observed from Fig. 9 and Fig. 10 that the optimal solution is a single point mass in the first two scenarios whereas it involves two point masses in the third scenario (Fig. 11). Notice that the LP solution with step size 1 corresponds to a single point mass at location 0 in the third scenario, which means that the LP solution becomes the same as the original system that involves no additive noise. This observation can also be verified based on the results in Table II.

### VIII. CONCLUSION

In this study, noise benefits have been investigated for joint detection and estimation systems under the NP and Bayesian detection frameworks and according to the Bayesian estimation criterion. It has been shown that the optimal additive noise can be represented by a discrete random variable with three and two point masses under the NP and Bayesian detection frameworks, respectively. Also, the proposed optimization problems have been approximately modeled by LP problems and conditions under which the performance of the system can or cannot be improved via additive noise have been derived. Numerical examples have been presented to provide performance comparison between the noise enhanced system and the original system and to support the theoretical analysis.

### IX. ACKNOWLEDGMENTS

This research was supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013).

### REFERENCES

- [1] M. D. McDonnell, "Is electrical noise useful?" *Proc. IEEE*, vol. 99, no. 2, pp. 242–246, Feb. 2011.
- [2] M. E. Inchiosa and A. R. Bulsara, "Signal detection statistics of stochastic resonators," *Phys. Rev. E*, vol. 53, pp. R2021–R2024, Mar. 1996.
- [3] F. Chapeau-Blondeau, "Noise-aided nonlinear Bayesian estimation," *Phys. Rev. E*, vol. 66, p. 032101, Sep. 2002.
- [4] F. Chapeau-Blondeau and D. Rousseau, "Noise-enhanced performance for an optimal Bayesian estimator," *IEEE Trans. Sig. Process.*, vol. 52, no. 5, pp. 1327–1334, May 2004.



- [5] A. A. Saha and G. Anand, "Design of detectors based on stochastic resonance," *Signal Process.*, vol. 83, no. 6, pp. 1193–1212, 2003.
- [6] D. Rousseau and F. Chapeau-Blondeau, "Stochastic resonance and improvement by noise in optimal detection strategies," *Digit. Signal Process.*, vol. 15, no. 1, pp. 19–32, Jan. 2005.
- [7] F. Chapeau-Blondeau and D. Rousseau, "Raising the noise to improve performance in optimal processing," *J. Stat. Mech.*, vol. 2009, no. 01, p. P01003, 2009.
- [8] F. Duan and B. Xu, "Parameter-induced stochastic resonance and baseband binary PAM signals transmission over an AWGN channel," *Int. J. of Bifurcation and Chaos*, vol. 13, no. 02, pp. 411–425, 2003.
- [9] B. Xu, F. Duan, and F. Chapeau-Blondeau, "Comparison of aperiodic stochastic resonance in a bistable system realized by adding noise and by tuning system parameters," *Phys. Rev. E*, vol. 69, p. 061110, June 2004.
- [10] S. Jiang, F. Guo, Y. Zhou, and T. Gu, "Parameter-induced stochastic resonance in an over-damped linear system," *Physica A*, vol. 375, no. 2, pp. 483 – 491, 2007.
- [11] Q. He and J. Wang, "Effects of multiscale noise tuning on stochastic resonance for weak signal detection," *Digit. Signal Process.*, vol. 22, no. 4, pp. 614–621, 2012.
- [12] Y. Tadokoro and A. Ichiki, "A simple optimum nonlinear filter for stochastic-resonance-based signal detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2013, pp. 5760–5764.
- [13] S. Kay, "Can detectability be improved by adding noise?" *IEEE Sig. Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.
- [14] S. Kay, J. Michels, H. Chen, and P. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Sig. Process. Lett.*, vol. 13, no. 11, pp. 695–698, Nov. 2006.
- [15] H. Chen, P. K. Varshney, J. H. Michels, and S. M. Kay, "Theory of the stochastic resonance effect in signal detection: Part I — Fixed detectors," *IEEE Trans. Sig. Process.*, vol. 55, no. 7, pp. 3172–3184, July 2007.
- [16] F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Fisher information as a metric of locally optimal processing and stochastic resonance," *PLoS ONE*, vol. 7, no. 4 (e34282), Apr. 2012.
- [17] —, "Fisher-information condition for enhanced signal detection via stochastic resonance," *Phys. Rev. E*, vol. 84, Nov. 2011.
- [18] S. Kay, "Noise enhanced detection as a special case of randomization," *IEEE Sig. Process. Lett.*, vol. 15, pp. 709–712, Nov. 2008.
- [19] H. Chen, P. K. Varshney, and J. H. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Sig. Process.*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.
- [20] H. Chen and P. Varshney, "Theory of the stochastic resonance effect in signal detection: Part II — Variable detectors," *IEEE Trans. Sig. Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2008.
- [21] H. Chen, P. Varshney, and J. Michels, "Improving sequential detection performance via stochastic resonance," *IEEE Sig. Process. Lett.*, vol. 15, pp. 685–688, 2008.
- [22] H. Chen, P. Varshney, S. Kay, and J. Michels, "Noise enhanced nonparametric detection," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 499–506, Feb. 2009.
- [23] A. Patel and B. Kosko, "Optimal noise benefits in Neyman Pearson and inequality-constrained statistical signal detection," *IEEE Trans. Sig. Process.*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [24] S. Bayram and S. Gezici, "On the improvability and nonimprovability of detection via additional independent noise," *IEEE Sig. Process. Lett.*, vol. 16, no. 11, pp. 1001–1004, Nov. 2009.

- [25] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Process.*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.
- [26] W. Chen, J. Wang, H. Li, and S. Li, "Stochastic resonance noise enhanced spectrum sensing in cognitive radio networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec 2010, pp. 1–6.
- [27] G. Balkan and S. Gezici, "CRLB based optimal noise enhanced parameter estimation using quantized observations," *IEEE Sig. Process. Lett.*, vol. 17, no. 5, pp. 477–480, May 2010.
- [28] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, "Optimum power allocation for average power constrained jammers in the presence of non-Gaussian noise," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1153–1156, Aug. 2012.
- [29] S. Bayram and S. Gezici, "Stochastic resonance in binary composite hypothesis-testing problems in the Neyman-Pearson framework," *Digit. Signal Process.*, vol. 22, no. 3, pp. 391–406, May 2012.
- [30] J.-Y. Liu and Y.-T. Su, "Noise-enhanced blind multiple error rate estimators in wireless relay networks," *IEEE Trans. Veh. Tech.*, vol. 61, no. 3, pp. 1145–1161, Mar. 2012.
- [31] A. Vempaty, V. Nadendla, and P. Varshney, "Further results on noise-enhanced distributed inference in the presence of byzantines," in *Proc. 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, June 2013, pp. 1–5.
- [32] G. Guo, X. Yu, Y. Jing, and M. Mandal, "Optimal design of noise-enhanced binary threshold detector under AUC measure," *IEEE Sig. Process. Lett.*, vol. 20, no. 2, pp. 161–164, Feb. 2013.
- [33] S. Zhang, J. Wang, and S. Li, "Adaptive bistable stochastic resonance aided spectrum sensing," in *Proc IEEE Int. Conf. on Commun. (ICC)*, June 2013, pp. 2622–2626.
- [34] S. Bayram, S. Gultekin, and S. Gezici, "Noise enhanced hypothesis-testing according to restricted Neyman-Pearson criterion," *Digit. Signal Process.*, vol. 25, no. 0, pp. 17–27, Feb. 2014.
- [35] D. Rousseau, G. Anand, and F. Chapeau-Blondeau, "Noise-enhanced nonlinear detector to improve signal detection in non-Gaussian noise," *Signal Process.*, vol. 86, no. 11, pp. 3456 – 3465, 2006.
- [36] S. Bayram and S. Gezici, "On the performance of single-threshold detectors for binary communications in the presence of Gaussian mixture noise," *IEEE Trans. Comm.*, vol. 58, no. 11, pp. 3047–3053, Nov. 2010.
- [37] G. V. Moustakides, G. H. Jajamovich, A. Tajer, and X. Wang, "Joint detection and estimation: Optimum tests and applications," *IEEE Trans. Inform. Theory*, vol. 58, no. 7, pp. 4215–4229, July 2012.
- [38] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [39] P. Whittle, *Probability Via Expectation*. New York: Springer-Verlag, 1992.
- [40] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*, ser. Athena Scientific optimization and computation series. Boston, MA: Athena Scientific, 2003.
- [41] 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*.
- [42] 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.
- [43] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [44] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Process.*, vol. 87, no. 11, pp. 2569–2586, 2007.

- [45] C. Luschy and B. Mulgrew, "Nonparametric trellis equalization in the presence of non-Gaussian interference," *IEEE Trans. Comm.*, vol. 51, no. 2, pp. 229–239, Feb. 2003.
- [46] A. Banerjee, P. Burlina, and R. Chellappa, "Adaptive target detection in foliage-penetrating sar images using alpha-stable models," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1823–1831, Dec. 1999.
- [47] M. Bouvet and S. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 621–626, May 1989.
- [48] A. B. Akbay, "Noise benefits in joint detection and estimation systems," *M.S. Thesis*, Bilkent University, Aug. 2014.
- [49] Q. Zhao, H. wei Li, and Y. tong Shen, "On the sum of generalized Gaussian random signals," in *Proc. Int. Conf. Signal Proc. (ICSP)*, vol. 1, Aug. 2004, pp. 50–53.

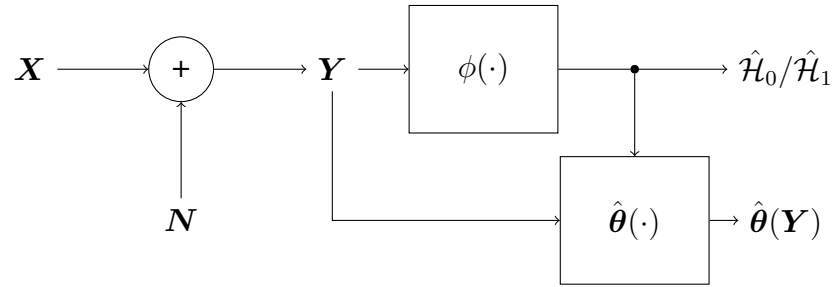


Fig. 1: Joint detection and estimation scheme with noise enhancement: The only modification on the original system is the introduction of the additive noise  $\mathbf{N}$ .

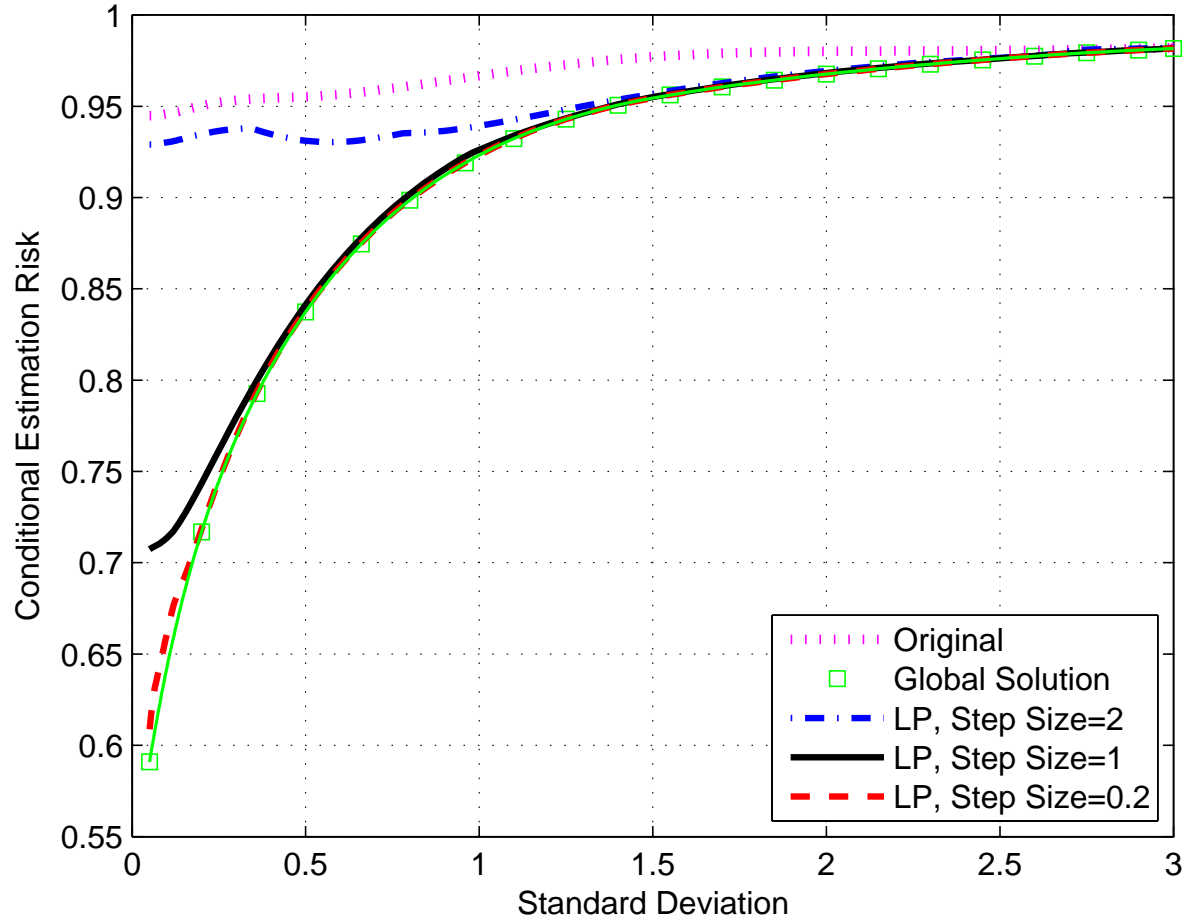


Fig. 2: Noise enhancement effects in the NP framework for  $K = 1$ .

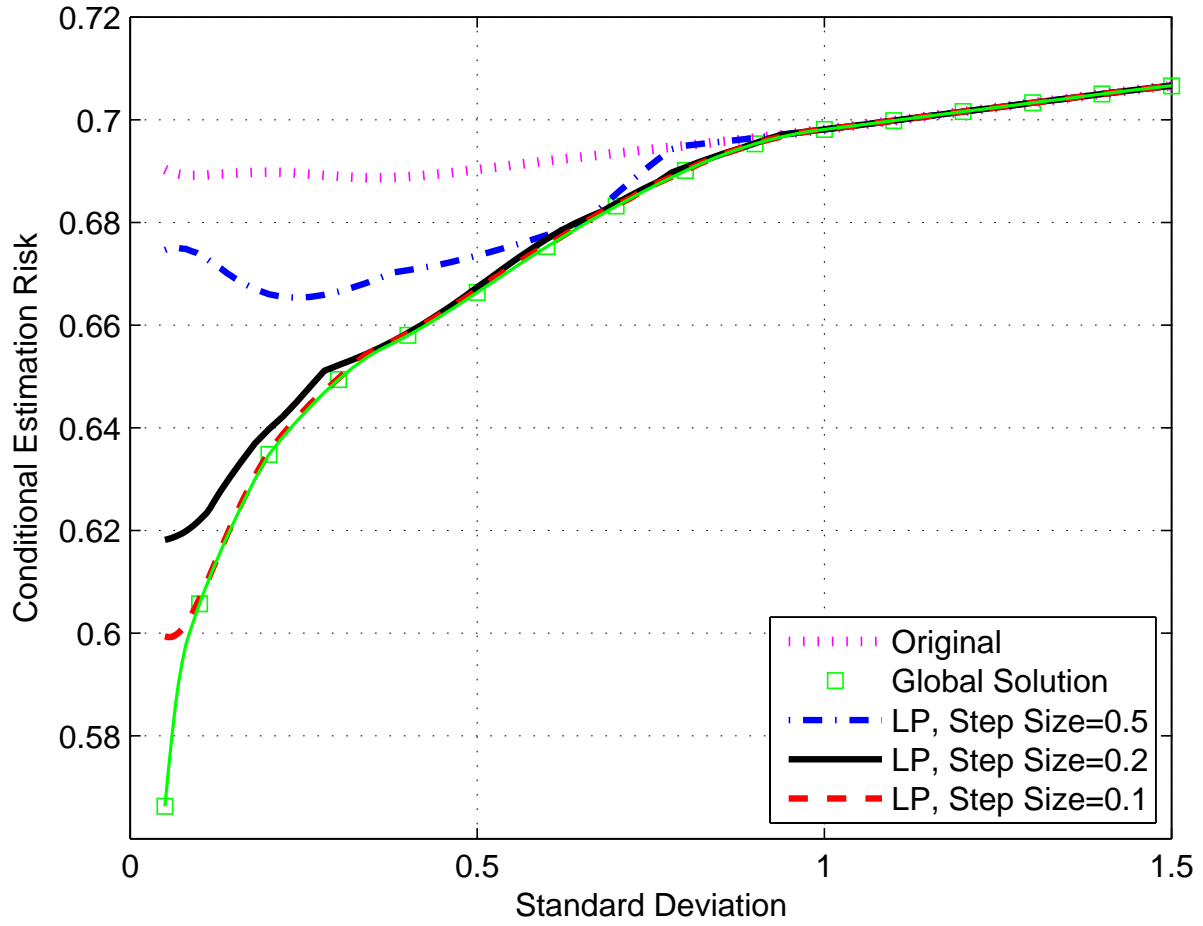


Fig. 3: Noise enhancement effects in the NP framework for  $K = 4$ .

$\sigma=0.3, K=1$	$\tau_{PF}$	$E\{T(0)\}$	$E\{R(0)\}$	$\frac{E\{G_{11}(0)\}}{E\{R(0)\}}$	$E\{T(\mathbf{n})\}$	$E\{R(\mathbf{n})\}$	$\frac{E\{G_{11}(\mathbf{n})\}}{E\{R(\mathbf{n})\}}$
LP (2.0)	5.3456	0.1500	0.4220	0.9533	0.1500	0.4220	0.9376
LP (1.0)	5.3456	0.1500	0.4220	0.9533	0.1500	0.4220	0.7796
LP (0.2)	5.3456	0.1500	0.4220	0.9533	0.1500	0.4220	0.7694
Opt. Sol.	5.3456	0.1500	0.4220	0.9533	0.1500	0.4220	0.7684
$\sigma=0.3, K=4$	$\tau_{PF}$	$E\{T(0)\}$	$E\{R(0)\}$	$\frac{E\{G_{11}(0)\}}{E\{R(0)\}}$	$E\{T(\mathbf{n})\}$	$E\{R(\mathbf{n})\}$	$\frac{E\{G_{11}(\mathbf{n})\}}{E\{R(\mathbf{n})\}}$
LP (0.5)	2.7140	0.1500	0.7474	0.6890	0.1500	0.7474	0.6651
LP (0.2)	2.7140	0.1500	0.7474	0.6890	0.1500	0.7474	0.6522
LP (0.1)	2.7140	0.1500	0.7474	0.6890	0.1500	0.7474	0.6496
Opt. Sol.	2.7140	0.1500	0.7474	0.6890	0.1500	0.7474	0.6494
$\sigma=0.4, K=4$	$\tau_{PF}$	$E\{T(0)\}$	$E\{R(0)\}$	$\frac{E\{G_{11}(0)\}}{E\{R(0)\}}$	$E\{T(\mathbf{n})\}$	$E\{R(\mathbf{n})\}$	$\frac{E\{G_{11}(\mathbf{n})\}}{E\{R(\mathbf{n})\}}$
LP (0.5)	2.6867	0.1500	0.7505	0.6889	0.1500	0.7505	0.6707
LP (0.2)	2.6867	0.1500	0.7505	0.6889	0.1500	0.7505	0.6584
LP (0.1)	2.6867	0.1500	0.7505	0.6889	0.1500	0.7505	0.6584
Opt. Sol.	2.6867	0.1500	0.7505	0.6889	0.1500	0.7505	0.6579

TABLE I: The conditional estimation risk, the detection probability, and the false alarm probability for the original system (i.e., no additive noise), for the optimal solution of the problem in (25) and for the solution of the LP problem in (29).

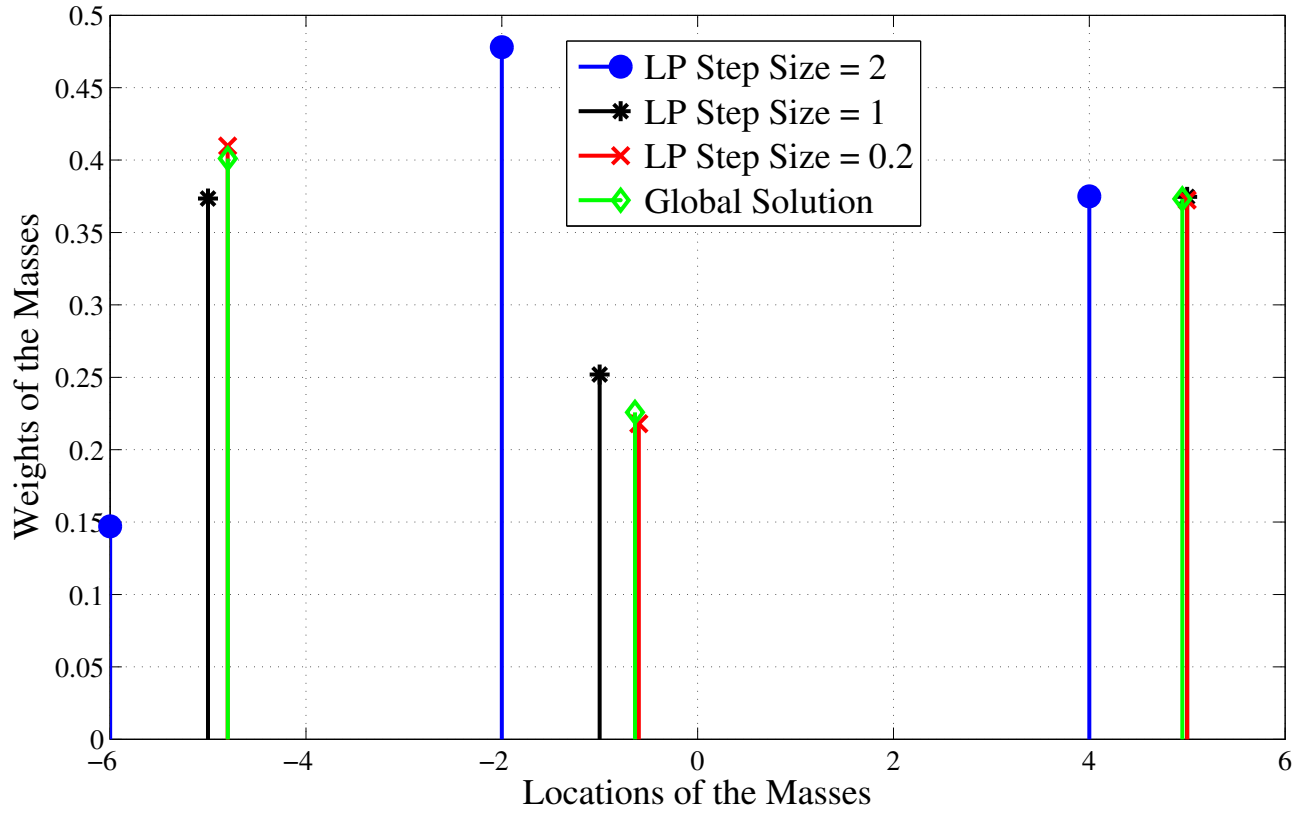


Fig. 4: Optimal solutions of the problem in (25) and the solutions of the LP problem defined in (29) for  $K = 1$  and  $\sigma = 0.3$ .



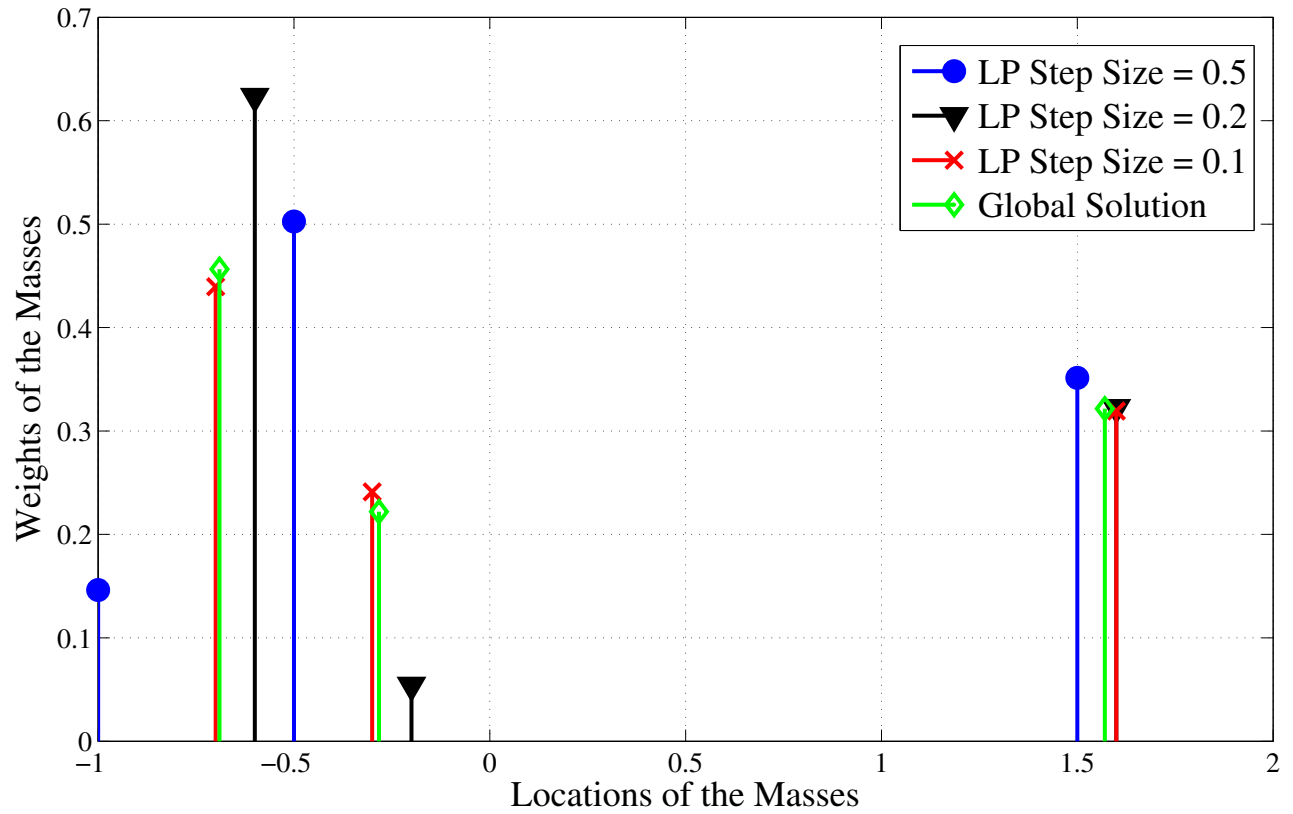


Fig. 5: Optimal solutions of the problem in (25) and the solutions of the LP problem defined in (29) for  $K = 4$  and  $\sigma = 0.3$ .

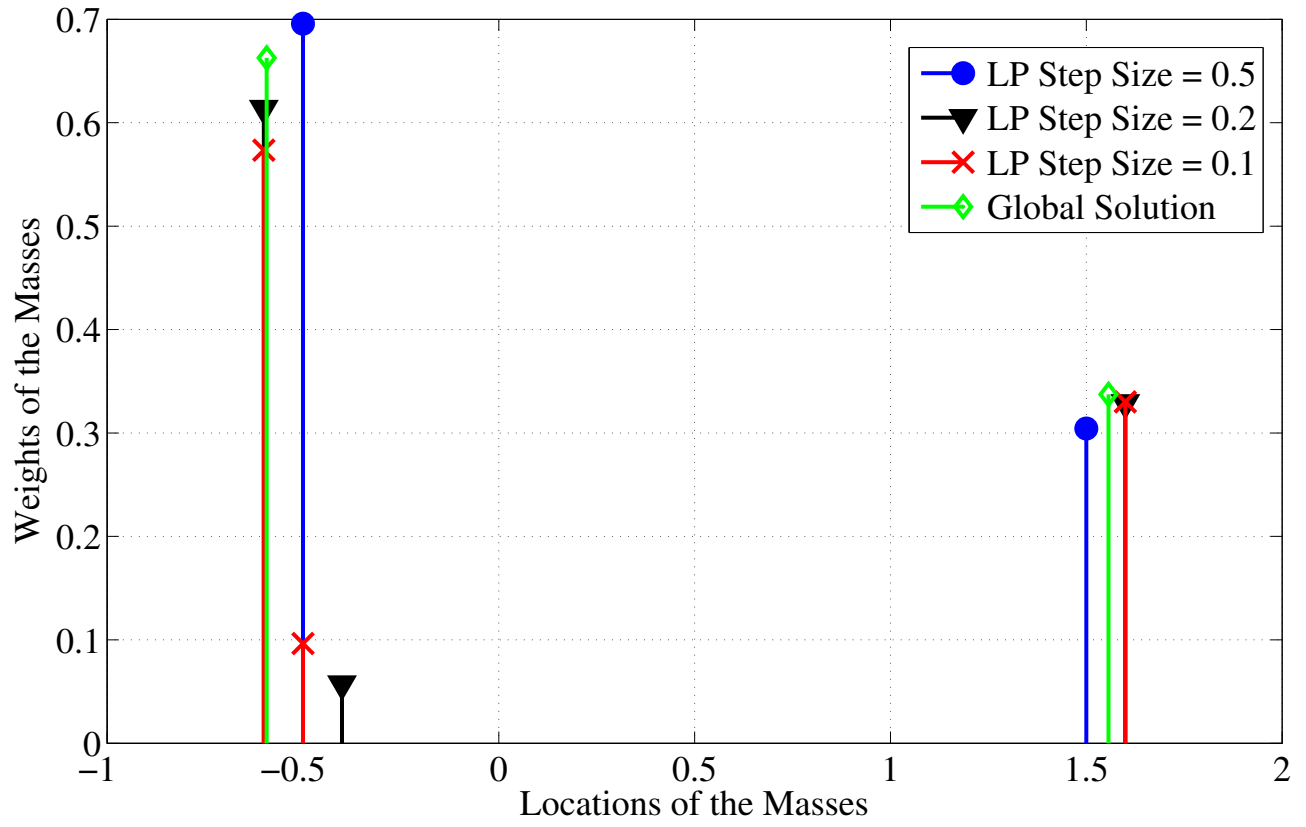


Fig. 6: Optimal solutions of the problem in (25) and the solutions of the LP problem defined in (29) for  $K = 4$  and  $\sigma = 0.4$ .

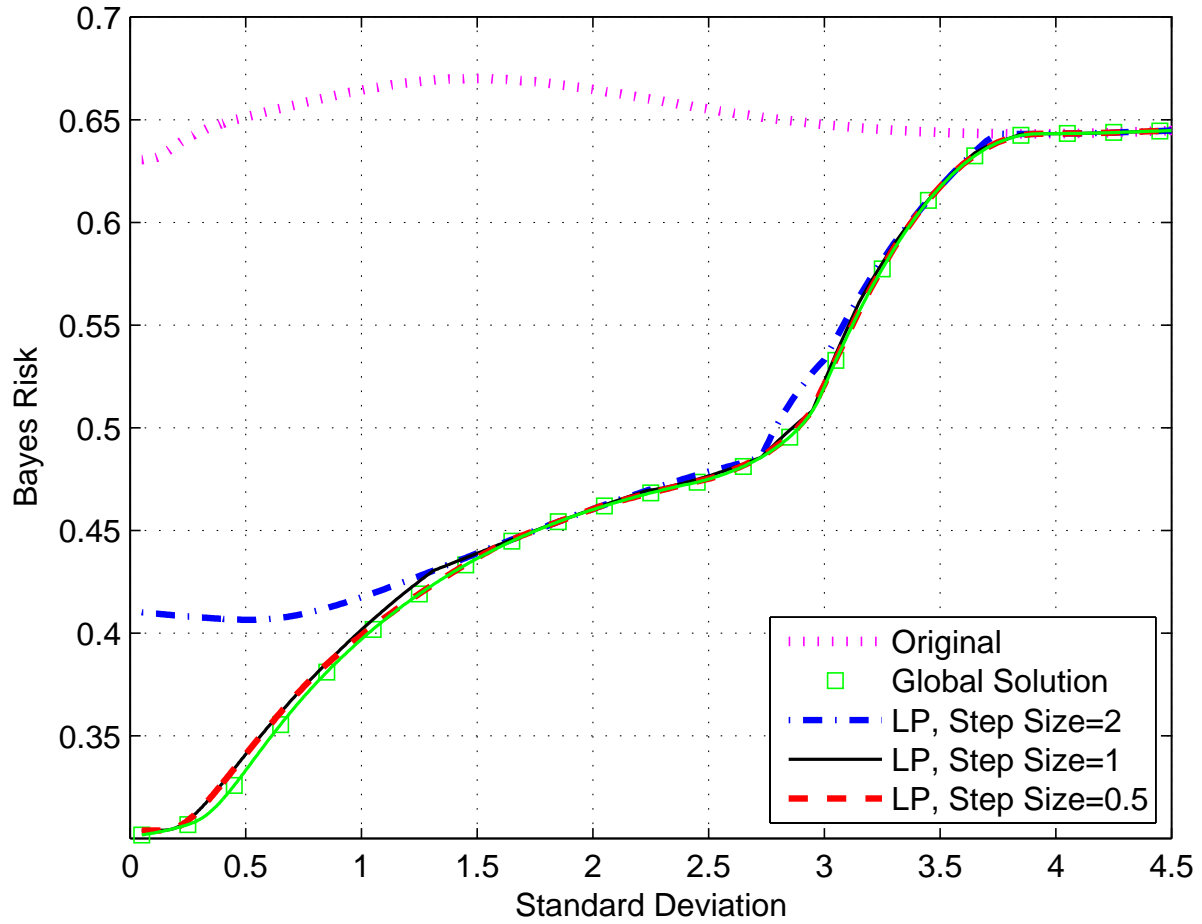


Fig. 7: Noise enhancement effects in the Bayes detection framework for  $K = 1$ .

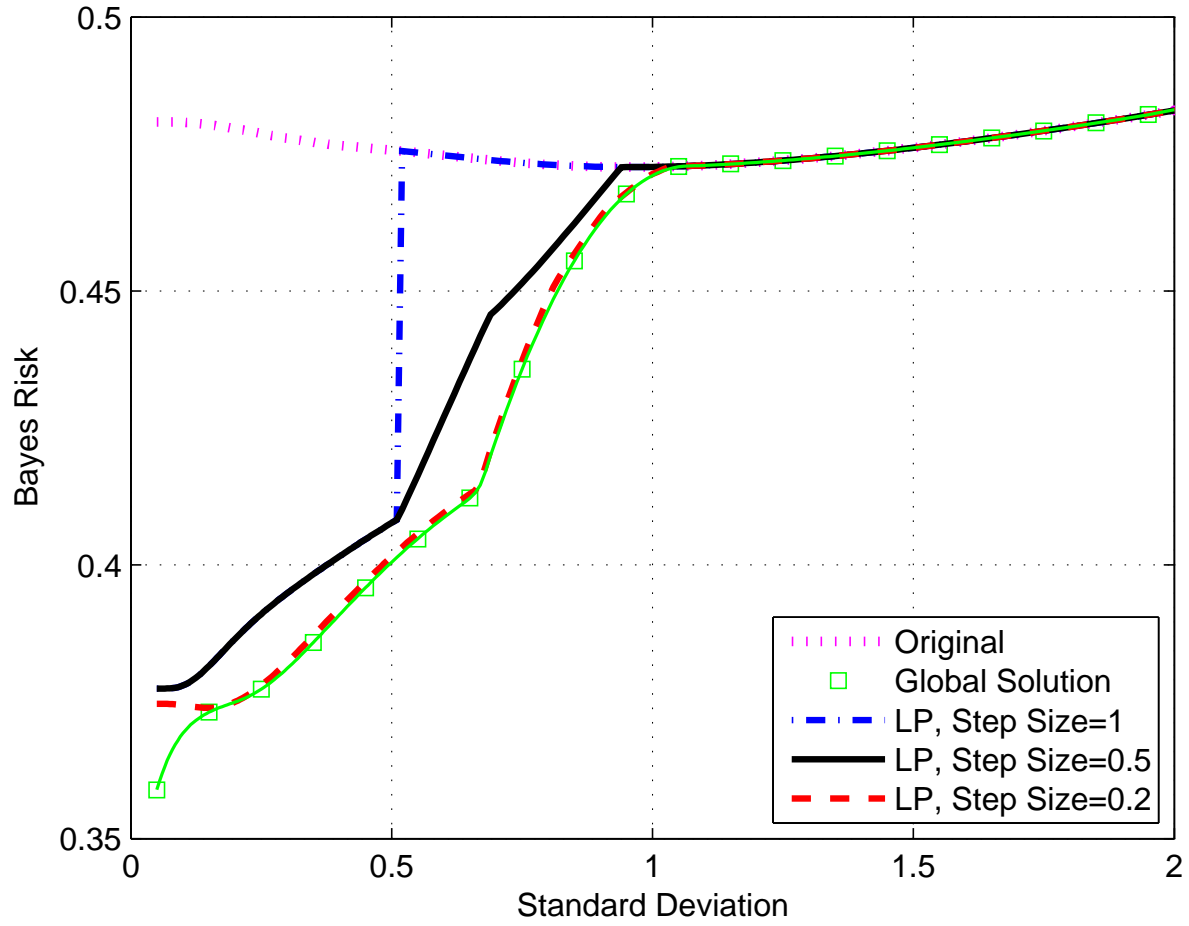


Fig. 8: Noise enhancement effects in the Bayes detection framework for  $K = 4$ .

$\sigma=0.5, K=1$	$r(\phi), n=0$	$r(\hat{\theta}), n=0$	$r(\phi)$	$r(\hat{\theta})$
LP (1.0)	0.1959	0.4757	0.3266	0.4061
LP (0.5)	0.4216	0.6514	0.3057	0.3411
LP (0.2)	0.4216	0.6514	0.3057	0.3411
Opt. Sol.	0.4216	0.6514	0.3088	0.3333
$\sigma=0.5, K=4$	$r(\phi), n=0$	$r(\hat{\theta}), n=0$	$r(\phi)$	$r(\hat{\theta})$
LP (1.0)	0.1959	0.4757	0.1956	0.4076
LP (0.5)	0.1959	0.4757	0.1956	0.4076
LP (0.2)	0.1959	0.4757	0.1899	0.4015
Opt. Sol.	0.1959	0.4757	0.1906	0.4005
$\sigma=0.75, K=4$	$r(\phi), n=0$	$r(\hat{\theta}), n=0$	$r(\phi)$	$r(\hat{\theta})$
LP (1.0)	0.1933	0.4734	0.1933	0.4734
LP (0.5)	0.1933	0.4734	0.1933	0.4514
LP (0.2)	0.1933	0.4734	0.1933	0.4372
Opt. Sol.	0.1933	0.4734	0.1933	0.4357

TABLE II: The Bayes estimation risk and the Bayes detection risk for the original system (i.e.,  $n=0$ ), for the optimal solution of the problem in (26), and for the solution of the LP problem in (30).

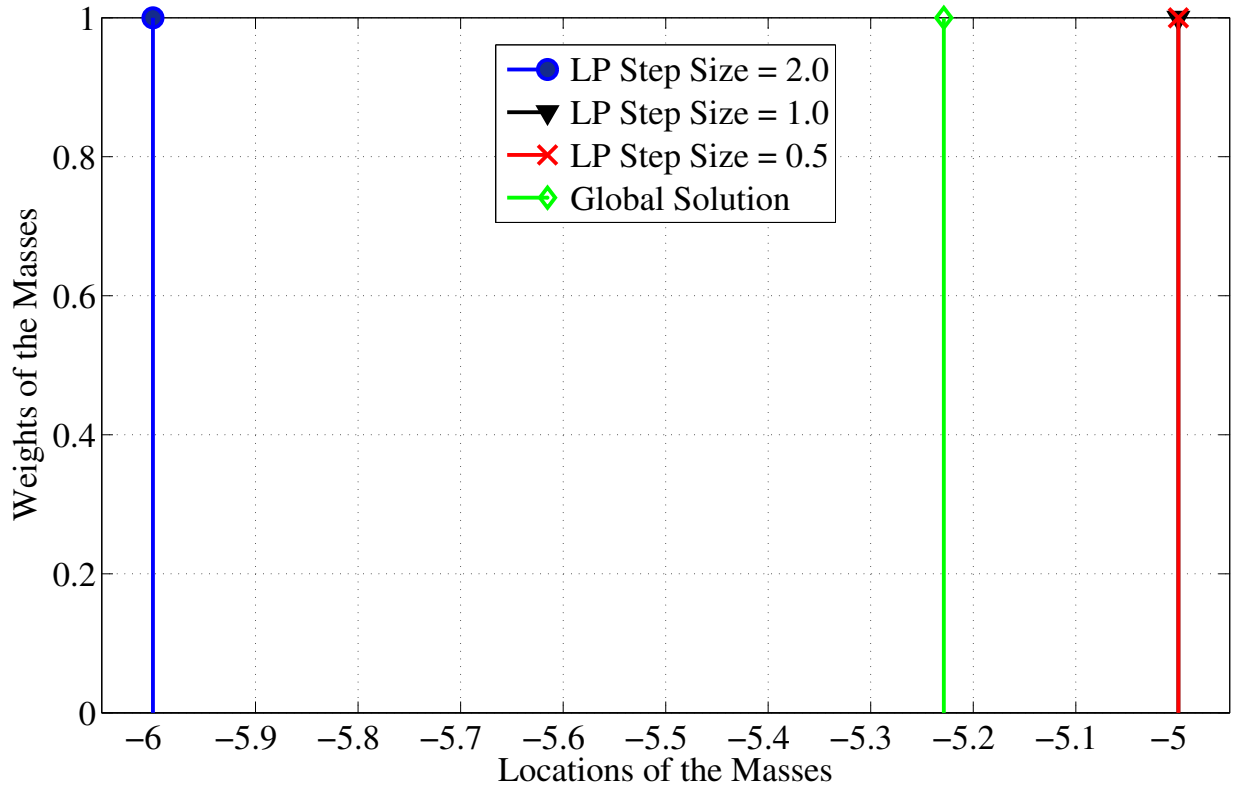


Fig. 9: Optimal solutions of the problem in (26) and the solutions of the LP problem in (30) for  $K = 1$  and  $\sigma = 0.5$ .

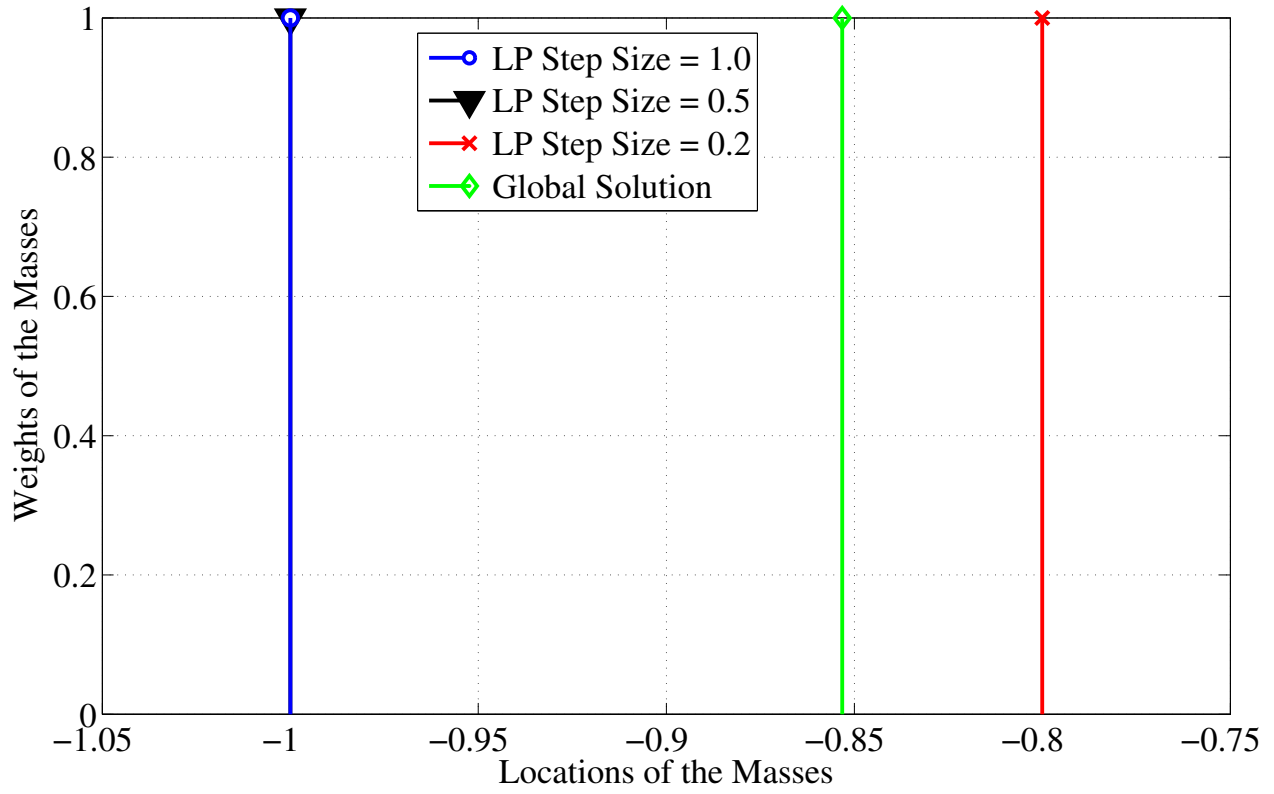


Fig. 10: Optimal solutions of the problem in (26) and the solutions of the LP problem in (30) for  $K = 4$  and  $\sigma = 0.5$ .

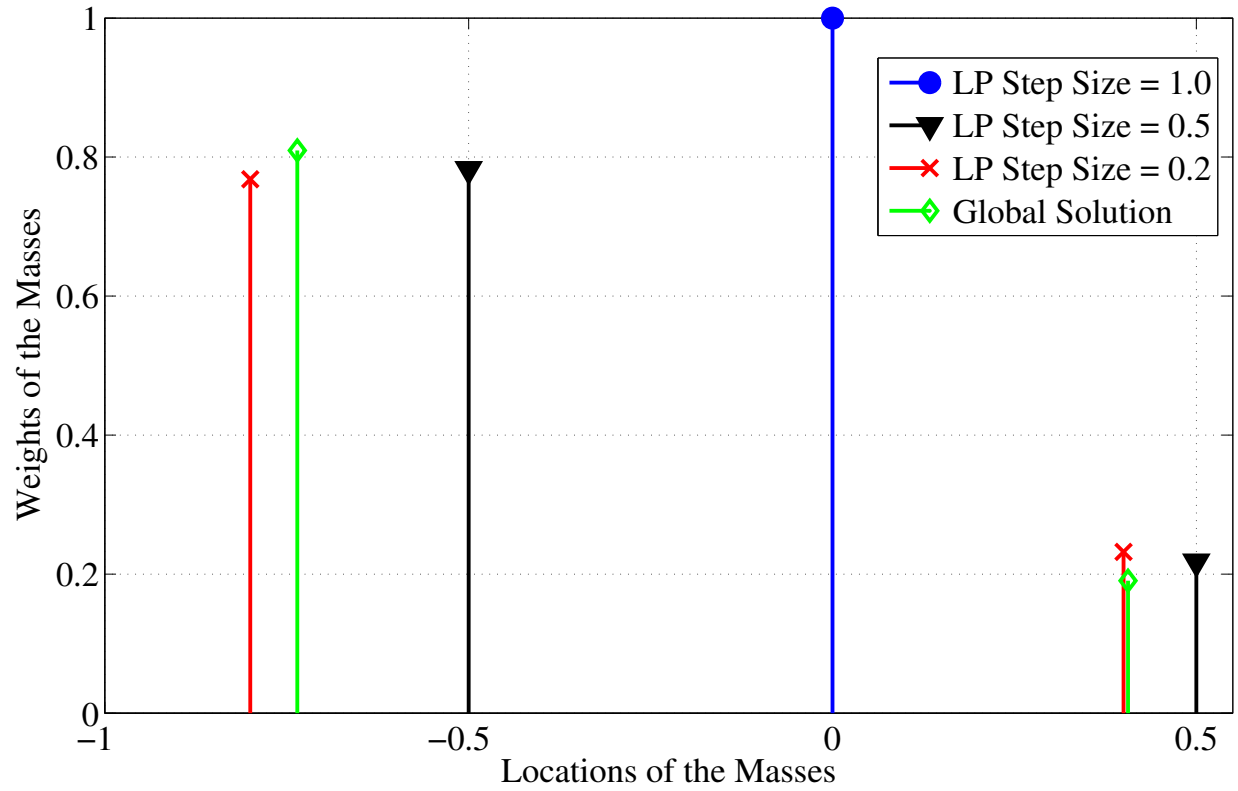


Fig. 11: Optimal solutions of the problem in (26) and the solutions of the LP problem in (30) for  $K = 4$  and  $\sigma = 0.75$ .