Theory of the Stochastic Resonance Effect in Signal Detection: Part I—Fixed Detectors

Hao Chen, *Student Member, IEEE*, Pramod K. Varshney, *Fellow, IEEE*, Steven M. Kay, *Fellow, IEEE*, and James H. Michels, *Fellow, IEEE*

***Abstract*—This paper develops the mathematical framework to analyze the stochastic resonance (SR) effect in binary hypothesis testing problems. The mechanism for SR noise enhanced signal detection is explored. The detection performance of a noise modified detector is derived in terms of the probability of detection $P_{\rm D}$ and the probability of false alarm $P_{\rm FA}$. Furthermore, sufficient conditions are established to determine the improvability of a fixed detector using SR. The form of the optimal noise pdf is determined and the optimal stochastic resonance noise pdf which renders the maximum $P_{\rm D}$ without increasing $P_{\rm FA}$ is derived. Finally, an illustrative example is presented where performance comparisons are made between detectors where the optimal stochastic resonance noise, as well as Gaussian, uniform, and optimal symmetric noises are applied to enhance detection performance.**

***Index Terms*—Hypothesis testing, non-Gaussian noise, nonlinear systems, signal detection, stochastic resonance (SR).**

I. Introduction

STOCHASTIC RESONANCE (SR) is a nonlinear physical phenomenon in which the output signals of some nonlinear systems can be enhanced by adding suitable noise under certain conditions. Since its discovery by Benzi *et al.* in 1981 [1], the SR effect has been observed and applied in numerous nonlinear systems [2]. The classic SR signature is the signal-to-noise ratio (SNR) gain of certain nonlinear systems, i.e., the output SNR is higher than the input SNR when an appropriate amount of noise is added [3]–[17]. Some approaches have been proposed to tune the SR system by maximizing SNR. It has been shown that the SNR of a summing network of excitable units is optimum at a certain level of noise [3]. Later, for some SR systems, robustness enhancement using non-Gaussian noises was reported by Castro *et al.* [10]. For a fixed type of noise, Mitaim and Kosko [18] proposed an adaptive stochastic learning scheme performing a stochastic gradient ascent on the SNR to determine the optimal noise level based on the samples from the process. Rather than adjusting the input noise level, Xu *et al.* [19] proposed a numerical method for realizing SR by tuning system parameters to maximize SNR gain. Although SNR is a very important measure of system performance, SNR gain based SR approaches have several limitations. First, the definition of SNR is not uniform and it varies from one application to another. Second, to optimize the performance, the complete *a priori* knowledge of the signal is required. Finally, for detection problems where the noise is non-Gaussian, SNR is not always directly related to detection performance; i.e., optimizing output SNR does not guarantee optimizing probability of detection.

SR was also found to enhance the mutual information (MI) between input and output signals [20]–[25]. Similar to the SNR scenario, for a specified type of SR noise, Mitaim and Kosko [25] showed that almost all noise probability density functions produce some SR effect in threshold neurons and a new statistically robust learning law was proposed to find the optimal noise level. McDonnell *et al.* [26] pointed out that the capacity of a SR channel can not exceed the actual capacity at the input. Compared to SNR, MI is more directly correlated with the transferred input signal information.

In signal detection theory, SR also plays a very important role in improving the signal detectability. In [27] and [16], improvement of detection performance of a weak sinusoid signal is reported. To detect a dc signal in a Gaussian mixture noise background, Kay [28] showed that under certain conditions, performance of the sign detector can be enhanced by adding some white Gaussian noise. For another suboptimal detector, the locally optimal detector (LOD), Zozor and Amblard [17] pointed out that detection performance is optimum when the noise parameters and detector parameters are matched. A study of the stochastic resonance phenomenon in quantizers conducted by Saha and Anand showed that a better detection performance can be achieved by a proper choice of the quantizer thresholds [29]. Recently, Rousseau and Blondeau [30] pointed out that the detection performance can be further improved by using an optimal detector on the output signal. Despite the progress achieved by the above approaches, the study of the SR effect in signal detection systems is rather limited and does not fully consider the underlying theory. In this paper, we explore the underlying mechanism of the SR phenomenon for a more general two hypotheses detection problem which can be formulated as follows.

Manuscript received December 14, 2005; revised July 24, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. L. Collins. This work was supported by AFOSR under contract FA9550-05-C-0139. This paper was presented in part at the 2006 International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France, May 2006.

H. Chen and P. K. Varshney are with the Department of EECS, Syracuse University, Syracuse, NY 13244 USA (e-mail: hchen21@syr.edu; varshney@syr.edu).

S. M. Kay is with the Department of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI 02881 USA (e-mail: kay@ele.uri.edu).

J. H. Michels is with the JHM Technologies, Ithaca, NY 14852 USA (e-mail: jmichels@americu.net).

Color versions of one or more of the figures in this paper available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.893757

Consider a two hypotheses detection problem where given an N dimensional data vector $\mathbf{x} \in R^N$, we have to decide between two hypotheses H_1 or H_0

$$\begin{cases} H_0 : p_{\mathbf{x}}(\mathbf{x}; H_0) = p_0(\mathbf{x}) \\ H_1 : p_{\mathbf{x}}(\mathbf{x}; H_1) = p_1(\mathbf{x}) \end{cases} \tag{1}$$

where $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$ are the pdfs of $\mathbf{x}$ under H_0 and H_1, respectively. In order to make a decision, a test (random or nonrandom) is needed to choose between the two hypotheses. This test can be completely characterized by a *critical function* (*decision function*) ϕ [31] where $0 \leq \phi(\mathbf{x}) \leq 1$ for all $\mathbf{x}$. For any observation $\mathbf{x}$, this test chooses the hypothesis H_1 with probability $\phi(\mathbf{x})$. In many cases, $\phi(\mathbf{x})$ can be implicitly expressed by using a test statistic T which is a function of $\mathbf{x}$ and a threshold η such that

$$T(\mathbf{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2}$$

where its corresponding critical function is

$$\phi_T(\mathbf{x}) = \begin{cases} 1 : T(\mathbf{x}) > \eta \\ \alpha : T(\mathbf{x}) = \eta \\ 0 : T(\mathbf{x}) < \eta \end{cases} \tag{3}$$

where $0 \leq \alpha \leq 1$ is a suitable number. The probability of detection P_D is given by

$$P_D^x = \int_{R^N} \phi(\mathbf{x}) p_1(\mathbf{x}) d\mathbf{x} \tag{4}$$

and the probability of false alarm P_{FA} is given by

$$P_{\mathrm{FA}}^x = \int_{R^N} \phi(\mathbf{x}) p_0(\mathbf{x}) d\mathbf{x} \tag{5}$$

where the superscripts on P_D and P_{FA} in (4) and (5) indicate that the test in (2) is employed for the data vector $\mathbf{x}$. Although the critical function $\phi(\mathbf{x})$ and the test statistic T can take any form, we know that the optimum Neyman-Pearson detector involves a likelihood ratio test (LRT) where $T_{LRT}(\mathbf{x}) = p_1(\mathbf{x})/p_0(\mathbf{x})$. Although a Neyman-Pearson detector is optimum in the sense of maximizing P_D given a fixed P_{FA}, the associated LRT requires the complete knowledge of the pdfs $p_0(\cdot)$ and $p_1(\cdot)$ which is not always available in a practical application. Also, the input data statistics may vary with time or may change from one application to another. To make matters worse, there are many detection problems where the exact form of the LRT is too complicated to be implemented. Therefore, simpler and more robust suboptimal detectors are used in numerous applications [32]. To improve a suboptimal detector detection performance, two approaches are widely used. In the first approach, the detector parameters are varied [15]–[17], [29], [33]. Alternatively, when the detector itself cannot be altered or the optimum parameter values are difficult to obtain, adjusting the observed data becomes a viable approach. It is well known that the detection performance can be improved by adding additional noise that is statistically dependent on the existing noise and/or with pdf that depends upon which hypothesis is true [28]. However, adding a dependent noise is not always possible because pertinent prior information is usually not available. Therefore, in this paper, we constrain the additive noise to be independent noise. For some suboptimal detectors, as Kay pointed out in [28], detection performance can be improved by adding an independent noise to the data under certain conditions. For a given type of SR noise, the optimal amount of noise can be determined that maximizes the detection performance for a given suboptimal detector [34]. In an effort to explain this noise enhanced phenomenon, for some integrate-and-fire neuron models, Tougaard demonstrated that the detection performance gain is caused by the nonlinear properties of the spike-generation process itself [35]. However, despite the progress made in the literature, the underlying mechanism of this Stochastic Resonance phenomenon in detection problems has not fully been explored. For example, an interesting question is to determine the best 'noise' to be added in order to achieve the best achievable detection performance for the suboptimal detector and this question remains unsolved. In this case, the detection problem can be stated as: Given that the test is fixed; i.e., the critical function $\phi(\cdot)$ (for example, T and η) is fixed, can we improve the detection performance by adding SR noise? If the answer is yes, what kind of noise and how much noise (i.e., noise pdf) should we add to the observed data to maximize P_D without increasing P_{FA}? In this paper, a theoretical analysis is presented to gain further insight into the SR phenomenon and the detection performance of the noise modified observations is obtained. Furthermore, the optimum noise pdf, i.e., not only the noise level but also the noise type is determined. As an illustrative example, the optimum noise pdf and some suboptimal noise pdfs for the sign detector are derived. Compared to the earlier definitions of SR [1], [2], we further extend the concept of "SR" to a pure noise enhanced phenomenon, i.e., a phenomenon of some nonlinear systems in which the system performance is enhanced due to the addition of an independent noise at the input. In this paper, the terminologies "SR" and "noise enhanced" are used interchangeably. However, we point out that the latter is actually the generalization of the former.

The paper is organized as follows. In Section II, we formulate the noise modified detection problem and the conditions for the best SR noise pdf are derived. The exact form of the optimum SR noise pdf is derived in Section III. An illustrative example is presented in Section IV. Conclusions and further comments are given in Section V.

II. Problem Formulation

In order to study a possible enhancement of the detection performance, we add noise to the original data process $\mathbf{x}$ and obtain a new data process $\mathbf{y}$ given by

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \tag{6}$$

where $\mathbf{n}$ is either an independent random process with pdf $p_{\mathbf{n}}(\cdot)$ or a nonrandom signal. Notice that here we do not have any constraint on $\mathbf{n}$. For example, $\mathbf{n}$ can be white noise, colored noise, or even be a deterministic signal A, corresponding to $p_{\mathbf{n}}(\mathbf{n}) = \delta(\mathbf{n} - A)$. As will be shown later, depending on the detection problem, an improvement of detection performance may not always be possible. In that case, the optimal noise is equal to zero. The pdf of $\mathbf{y}$ is expressed by the convolutions of the pdfs such that

$$p_{\mathbf{y}}(\mathbf{y}) = p_{\mathbf{x}}(\mathbf{x}) * p_{\mathbf{n}}(\mathbf{x}) = \int_{R^N} p_{\mathbf{x}}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{x}. \quad (7)$$

The binary hypotheses testing problem for this new observed data $\mathbf{y}$ can be expressed as

$$\begin{cases} H_0 : p_{\mathbf{y}}(\mathbf{y}; H_0) = \int_{R^N} p_0(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{x} \\ H_1 : p_{\mathbf{y}}(\mathbf{y}; H_1) = \int_{R^N} p_1(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{x}. \end{cases} \quad (8)$$

Since the detector is fixed, i.e., the critical function ϕ of $\mathbf{y}$ is the same as that for $\mathbf{x}$, the probability of detection based on data $\mathbf{y}$ is given by

$$\begin{aligned} P_D^y &= \int_{R^N} \phi(\mathbf{y}) p_{\mathbf{y}}(\mathbf{y}; H_1) d\mathbf{y} \\ &= \int_{R^N} \phi(\mathbf{y}) \int_{R^N} p_1(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ &= \int_{R^N} p_1(\mathbf{x}) \left(\int_{R^N} \phi(\mathbf{y}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{y} \right) d\mathbf{x} \\ &= \int_{R^N} p_1(\mathbf{x}) C_{\mathbf{n},\phi}(\mathbf{x}) d\mathbf{x} = E_1\left(C_{\mathbf{n},\phi}(\mathbf{x})\right) \quad (9) \end{aligned}$$

where

$$C_{\mathbf{n},\phi}(\mathbf{x}) \equiv \int_{R^N} \phi(\mathbf{y}) p_{\mathbf{n}}(\mathbf{y}-\mathbf{x}) d\mathbf{y}. \quad (10)$$

Alternatively

$$\begin{aligned} P_D^y &= \int_{R^N} p_{\mathbf{n}}(\mathbf{x}) \left(\int_{R^N} \phi(\mathbf{y}) p_1(\mathbf{y}-\mathbf{x}) d\mathbf{y} \right) d\mathbf{x} \\ &= \int_{R^N} F_{1,\phi}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} = E_{\mathbf{n}}\left(F_{1,\phi}(\mathbf{x})\right). \quad (11) \end{aligned}$$

Similarly, we have

$$\begin{aligned} P_{\mathrm{FA}}^y &= \int_{R^N} p_0(\mathbf{x}) C_{\mathbf{n},\phi}(\mathbf{x}) d\mathbf{x} = E_0\left(C_{\mathbf{n},\phi}(\mathbf{x})\right) \quad (12) \\ &= \int_{R^N} F_{0,\phi}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} = E_{\mathbf{n}}\left(F_{0,\phi}(\mathbf{x})\right) \quad (13) \end{aligned}$$

where

$$F_{i,\phi}(\mathbf{x}) \equiv \int_{R^N} \phi(\mathbf{y}) p_i(\mathbf{y}-\mathbf{x}) d\mathbf{y} \quad (14)$$

corresponding to hypothesis H_i. $E_i(\cdot)$, $E_{\mathbf{n}}(\cdot)$ are the expected values based on distributions p_i and $p_{\mathbf{n}}$, respectively, and $P_{\mathrm{FA}}^x = F_{0,\phi}(0)$, $P_D^x = F_{1,\phi}(0)$. To simplify notation, we omit the subscript ϕ of F and C and denote them as F_1, F_0, and C_n, respectively. Further, from (14), $F_1(\mathbf{x}_0)$ and $F_0(\mathbf{x}_0)$ are actually the probability of detection and probability of false alarm, respectively, for this detection scheme with input $\mathbf{y} = \mathbf{x} + \mathbf{x}_0$. For example, $F_1(-2)$ is the P_D of this detection scheme with input $\mathbf{x} - 2$. Therefore, it is very convenient for us to obtain the F_1 and F_0 values by analytical computation if p_0, p_1 and ϕ are known. When they are not available, F_1 and F_0 can be obtained from the data itself by processing it through the detector and recording the detection performance.[1] From (11) and (13), we may formalize the optimal SR noise definition as follows.

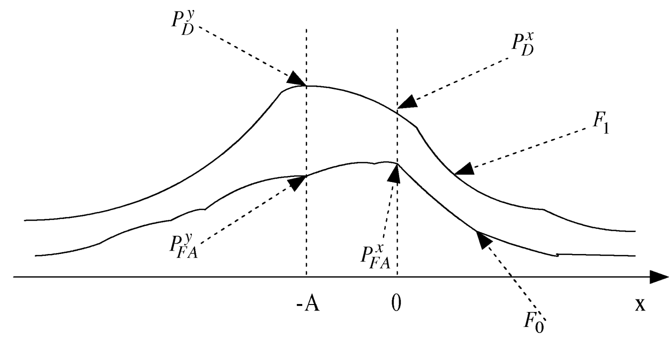


Fig. 1. An example of the relationship between P_D^x, P_{FA}^x, P_D^y, P_{FA}^y and p_n. The optimum noise pdf $p_n^{\mathrm{opt}}(n) = \delta(n + A)$.

Consider the two hypotheses detection problem as in (1). The pdf of optimum SR noise is given by

$$p_{\mathbf{n}}^{\mathrm{opt}} = \arg\max_{p_{\mathbf{n}}} \int_{R^N} F_1(\mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} \quad (15)$$

where

1) $p_{\mathbf{n}}(\mathbf{x}) \geq 0$, $\mathbf{x} \in R^N$;
2) $\int_{R^N} p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} = 1$;
3) $\int_{R^N} F_0(\mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} \leq F_0(0)$.

Conditions 1) and 2) are fundamental properties of a pdf function. Condition 3) ensures that $P_{\mathrm{FA}}^y \leq P_{\mathrm{FA}}^x$, i.e., the P_{FA} constraint specified under the Neyman-Pearson Criterion is satisfied. Further, if the inequality of condition 3) becomes equality, the constant false alarm rate (CFAR) property of the original detector is maintained.

A simple illustration of the effect of additive noise is shown in Fig. 1. In this example, $F_1(-A) = \max_{\mathbf{x}} F_1(\mathbf{x})$ and $F_0(-A) < F_0(0)$, hence $p_{\mathbf{n}}^{\mathrm{opt}} = \delta(\mathbf{x} + A)$ which means the optimal SR noise $\mathbf{n} \equiv -A$ is a dc signal with value $-A$. In practical applications, some additional restrictions on the noise may also be applied. For example, the type of noise may be restricted, (e.g., $\mathbf{n}$ may be specified as Gaussian noise), or we may require a noise with even symmetric pdf $p_{\mathbf{n}}(\mathbf{x}) = p_{\mathbf{n}}(-\mathbf{x})$ to ensure that the mean value of $\mathbf{y}$ is equal to the mean value of $\mathbf{x}$. However, regardless of the additional restrictions, the conditions 1), 2), and 3) are always valid and the optimum noise pdf can be determined for these conditions.

III. Optimum SR Noise for Neyman-Pearson Detection

In general, it is difficult to find the exact form of $p_{\mathbf{n}}(\cdot)$ directly because of condition 3). However, an alternative approach considers the relationship between $p_{\mathbf{n}}(\mathbf{x})$ and $F_i(\mathbf{x})$. From (14), for a given value f_0 of F_0, we have $\mathbf{x} = F_0^{-1}(f_0)$, where F_0^{-1} is the inverse function of F_0. When F_0 is a one-to-one mapping function, $\mathbf{x}$ is a unique vector. Otherwise, $F_0^{-1}(f_0)$ is a set of $\mathbf{x}$

[1]Thus, it is not necessary to have complete knowledge regarding $\phi(\cdot)$ and $p_i(\cdot)$.

for which $F_0(\mathbf{x}) = f_0$. Therefore, we can express a value or a set of values f_1 of F_1 as

$$f_1 = F_1(\mathbf{x}) = F_1\left(F_0^{-1}(f_0)\right). \tag{16}$$

Given the noise distribution of $p_{\mathbf{n}}(\cdot)$ in the original R^N domain, $p_{\mathbf{n},f_0}(\cdot)$, the noise distribution in the f_0 domain can also be uniquely determined. Further, the conditions on the optimum noise can be rewritten in terms of f_0 equivalently as

4) $p_{\mathbf{n},f_0}(f_0) \geq 0$;
5) $\int_0^1 p_{\mathbf{n},f_0}(f_0)df_0 = 1$;
6) $\int_0^1 f_0 p_{\mathbf{n},f_0}(f_0)df_0 \leq P_{\mathrm{FA}}^x$;

and

$$P_D^y = \int_0^1 f_1 p_{\mathbf{n},f_0}(f_0)df_0 \tag{17}$$

where $p_{\mathbf{n},f_0}(f_0)$ is the SR noise pdf in the f_0 domain.

Compared to the original conditions 1), 2), and 3), this equivalent form has some advantages. First, the problem complexity is dramatically reduced. Instead of searching for an optimal solution in R^N, we are now looking for an optimal solution in a single dimensional space. Second, by applying these new conditions, we avoid the direct use of the underlying pdfs $p_1(\cdot)$ and $p_0(\cdot)$ and replace them with f_1 and f_0. Note that, in some cases, it is not very easy to find the exact form of f_0 and f_1. However, recall that $F_1(\mathbf{x}_0)$ and $F_0(\mathbf{x}_0)$ are the Probability of Detection and Probability of False Alarm, respectively, of the original system with input $\mathbf{x}+\mathbf{x}_0$. In practical applications, we may learn the relationship by Monte Carlo simulation using importance sampling. In general, compared to p_1 and p_0, f_1 and f_0 are much easier to estimate and once the optimum $p_{\mathbf{n},f_0}$ is found, the optimum $p_n(\mathbf{x})$ is determined, as well by the inverse of the functions F_0 and F_1.

Let us now consider the function $J(t)$ such that $J(t) = \sup(f_1 : f_0 = t)$ is the maximum value of f_1 given f_0. Clearly, $J(P_{\mathrm{FA}}^x) \geq F_1(0) = P_D^x$. From (17), it follows that for any noise $p_{\mathbf{n}}$, we have

$$P_D^y(p_{\mathbf{n}}) \leq \int_0^1 J(f_0)p_{\mathbf{n},f_0}(f_0)df_0. \tag{18}$$

Therefore, the optimum P_D^y is attained when $f_1(f_0) = J(f_0)$ and $P_{\mathrm{D,opt}}^y = E_{\mathbf{n}}(J)$.

A. Determination of the Improvability of Detection via SR

Improvability of the given detector when SR noise is added can be determined by computing and comparing $P_{\mathrm{D,opt}}^y$ and P_D^x. When $P_{\mathrm{D,opt}}^y > P_D^x$, the given detector is improvable by adding SR noise. However, it requires the complete knowledge of $F_1(\cdot)$ and $F_0(\cdot)$ and significant computation. For a large class of detectors, however, depending on the specific properties of J, we may determine the sufficient conditions for improvability and nonimprovability more easily. These are given in the following theorems.

Theorem 1 (Improvability of Detection via SR): If $J(P_{\mathrm{FA}}^x) > P_D^x$ or $J''(P_{\mathrm{FA}}^x) > 0$ when $J(t)$ is second-order continuously differentiable around P_{FA}^x, then there exists at least one noise process $\mathbf{n}$ with pdf $p_{\mathbf{n}}(\cdot)$ that can improve the detection performance.

Proof: First, when $J(P_{\mathrm{FA}}^x) > P_D^x$, from the definition of J function, we know that there exist at one least one $\mathbf{n}_0$ such that $F_0(\mathbf{n}_0) = P_{\mathrm{FA}}^x$ and $F_1(\mathbf{n}_0) = J(P_{\mathrm{FA}}^x) > P_D^x$, therefore, the detection performance can be improved by choosing a SR noise pdf $P_{\mathbf{n}}(\mathbf{n}) = \delta(\mathbf{n}-\mathbf{n}_0)$. When $J''(P_{\mathrm{FA}}^x) > 0$ and is continuous around P_{FA}^x, there exists an $\epsilon > 0$ such that $J''(\cdot) > 0$ on $I = (P_{\mathrm{FA}}^x - \epsilon, P_{\mathrm{FA}}^x + \epsilon)$. Therefore, from Theorem A-1, J is convex on I.[2] Let us add a noise $\mathbf{n}$ with pdf $p_{\mathbf{n}}(\mathbf{x}) = (1/2)\delta(\mathbf{x}-\mathbf{x}_0) + (1/2)\delta(\mathbf{x}+\mathbf{x}_1)$ where $F_0(\mathbf{x}_0) = P_{\mathrm{FA}}^x + (\epsilon/2)$ and $F_0(\mathbf{x}_1) = P_{\mathrm{FA}}^x - (\epsilon/2)$. Due to the convexity of J, $P_D^y = (J(P_{\mathrm{FA}}^x - (\epsilon/2)) + J(P_{\mathrm{FA}}^x + (\epsilon/2)))/2 > J(P_{\mathrm{FA}}^x) \geq P_D^x$. Thus, detection performance can be improved via the addition of SR noise. ■

We will illustrate this result with an example in the next section.

Theorem 2 (Nonimprovability of Detection via SR): If there exists a nondecreasing concave function $\Psi(f_0)$ where $\Psi(P_{\mathrm{FA}}^x) = J(P_{\mathrm{FA}}^x) = F_1(0)$ and $\Psi(f_0) \geq J(f_0)$ for every f_0, then $P_D^y \leq P_D^x$ for any independent noise, i.e., the detection performance cannot be improved by adding noise.

Proof: For any noise $\mathbf{n}$ and corresponding $\mathbf{y}$, we have

$$\begin{aligned} P_D^y &\leq \int_0^1 J(f_0)p_{\mathbf{n},f_0}(f_0)df_0 \\ &\leq \int_0^1 \Psi(f_0)p_{\mathbf{n},f_0}(f_0)df_0 \\ &\leq \Psi\left(\int_0^1 f_0 p_{\mathbf{n},f_0}(f_0)df_0\right) \\ &\leq \Psi\left(P_{\mathrm{FA}}^x\right) = P_D^x. \end{aligned} \tag{19}$$

The third inequality of the Right Hand Side (RHS) of (19) is obtained using the concavity of the Ψ function. The detection performance cannot be improved via the addition of SR noise. ■

Again, we will illustrate this result in the next section.

B. Determination of the Form of Optimum SR Noise PDF

Before determining the exact pdf of $p_{\mathbf{n}}^{\mathrm{opt}}$, we first present the following result for the form of optimum SR noise.

Theorem 3 (Form of Optimum SR Noise): To maximize P_D^y, under the constraint that $P_{\mathrm{FA}}^y \leq P_{\mathrm{FA}}^x$, the optimum noise can be expressed as[3]

$$p_{\mathbf{n}}^{\mathrm{opt}}(\mathbf{n}) = \lambda\delta(\mathbf{n}-\mathbf{n}_1) + (1-\lambda)\delta(\mathbf{n}-\mathbf{n}_2) \tag{20}$$

where $0 \leq \lambda \leq 1$. In other words, to obtain the maximum achievable detection performance given the false alarm constraints, the optimum noise is a randomization of two discrete vectors added with probability λ and $1-\lambda$, respectively.

Proof: Let $U = \{(f_1, f_0)|f_1 = F_1(\mathbf{x}), f_0 = F_0(\mathbf{x}), \mathbf{x} \in R^N)$ be the set of all pairs of (f_1, f_0). Since $0 \leq f_1, f_0 \leq 1$, U is a subset of the linear space R^2. Furthermore, let V be the convex hull of U. Since $V \subset R^2$, its dimension $\mathrm{Dim}(V) \leq 2$.

[2]Please refer to [36] or the Appendix for the related definitions and Theorems.

[3]This form of optimum noise pdf is not necessarily unique. There may exist other forms of noise pdf that achieve the same detection performance.

Similarly, let the set of all possible $(P_D^y, P_{\rm FA}^y)$ be W. Since any convex combination of the elements of U, say $(\chi, \varnothing) = \sum_{i=1}^{M} \alpha_i (f_{1,i}, f_{0,i})$ can be obtained by setting the SR noise pdf such that $p_{\mathbf{n},f_0}(f_0) = \sum_{i=1}^{M} \alpha_i \delta(f_0 - f_{0,i})$, we have, $V \subseteq W$. It can also be shown that $W \subseteq V$. Otherwise, there would exist at least one element z such that $z \in W$, but $z \notin V$. In this case, there exists a small set S and a positive number τ such that $S = \{(x,y) | \, \|(x,y) - z\|_2^2 < \tau\}$ and $S \bigcap V =$ '{}', where '{}' denotes an empty set. However, since $0 \le f_1, f_0 \le 1$, by the well known property of integration, there always exists a finite set E with finite elements such that $E \subseteq U$ and (x_1, y_1), a convex combination of the elements of E, such that $\|(x_1,y_1) - z\|_2^2 < \tau$. Since $(x_1, y_1) \in V$, $(x_1, y_1) \in (V \bigcap S)$ which contradicts the definition of S. Therefore, $W \subseteq V$. Hence, $W = V$. From Theorem A-4, $(P_D^y, P_{\rm FA}^y)$ can be expressed as a convex combination of three elements. Also, since we are only interested in maximizing P_D under the constraint that $P_{\rm FA}^y \le P_{\rm FA}^x$, the optimum pair can only belong to B, the set of the boundary elements of V. To show this, let (f_1^*, f_0^*) be an arbitrary nonboundary point inside V. We know that there exists a $\tau > 0$ such that $(f_1^* + \tau, f_0^*) \in V$. Therefore, (f_1^*, f_0^*) is inadmissible as an optimum pair. Thus, the optimum pair can only exist on the boundary. Therefore, each z on the boundary of V can be expressed as the convex combination of only two elements in U. Hence,

$$\left(P_{D,\rm opt}^y, P_{\rm FA,opt}^y\right) = \lambda(f_{11}, f_{01}) + (1-\lambda)(f_{12}, f_{02}) \quad (21)$$

where $(f_{11}, f_{01}), (f_{12}, f_{02}) \in U$, $0 \le \lambda \le 1$. Therefore, we have

$$p_{\mathbf{n},f_0}^{\rm opt} = \lambda\delta(f_0 - f_{01}) + (1-\lambda)\delta(f_0 - f_{02}). \quad (22)$$

Equivalently, $p_{\mathbf{n}}^{\rm opt}(\mathbf{n}) = \lambda\delta(\mathbf{n} - \mathbf{n}_1) + (1-\lambda)\delta(\mathbf{n} - \mathbf{n}_2)$, where $\mathbf{n}_1$ and $\mathbf{n}_2$ are determined by the equations

$$\begin{cases} F_0(\mathbf{n}_1) = f_{01} \\ F_1(\mathbf{n}_1) = f_{11} \\ F_0(\mathbf{n}_2) = f_{02} \\ F_1(\mathbf{n}_2) = f_{12} \end{cases}. \quad (23)$$

Alternatively, the optimum SR noise can also be expressed in terms of $C_{\mathbf{n}}$, such that

$$C_{\mathbf{n}}^{\rm opt}(\mathbf{x}) = \lambda\phi(\mathbf{x} + \mathbf{n}_1) + (1-\lambda)\phi(\mathbf{x} + \mathbf{n}_2). \quad (24)$$

■

From (22), we have

$$P_{D,\rm opt}^y = \lambda J(f_{01}) + (1-\lambda)J(f_{02}) \quad (25)$$

and

$$P_{\rm FA,opt}^y = \lambda f_{01} + (1-\lambda) f_{02} \le P_{\rm FA}^x. \quad (26)$$

C. Determination of the pdf of Optimum SR Noise

Depending on the location of the maxima of $J(\cdot)$, we have the following theorem.

Theorem 4: Let $F_{1M} = \max(J(t))$ and $t_o = \arg\min_t (J(t) = F_{1M})$. It follows that

Case 1) If $t_o \le P_{\rm FA}^x$, then $P_{\rm FA,opt}^y = t_o$ and $P_{\rm D,opt}^y = F_{1M}$, i.e., the maximum achievable detection performance is obtained when the optimum noise is a dc signal with value $\mathbf{n}_o$, i.e.

$$p_{\mathbf{n}}^{\rm opt}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o) \quad (27)$$

where $F_0(\mathbf{n}_o) = t_o$ and $F_1(\mathbf{n}_o) = F_{1M}$.

Case 2) If $t_o > P_{\rm FA}^x$, then $P_{\rm FA,opt}^y = F_0(0) = P_{\rm FA}^x$, i.e., the inequality of (26) becomes equality. Furthermore

$$P_{\rm FA,opt}^y = \lambda f_{01} + (1-\lambda) f_{02} = P_{\rm FA}^x. \quad (28)$$

Proof: For Case 1, notice that $P_D^y = \int_0^1 J(f_0) p_{\mathbf{n},f_0}(f_0) df_0 \le \int_0^1 F_{1M} p_{\mathbf{n},f_0}(f_0) df_0 = F_{1M}$ and $F_1(\mathbf{n}_o) = F_{1M}$. Therefore, the optimum detection performance is obtained when the noise is a dc signal with value $\mathbf{n}_o$ with $P_{\rm FA}^y = t_o$.

We use the contradiction method here to prove Case 2. First, let us suppose that the optimum detection performance is obtained when $P_{\rm FA,opt}^y = \kappa < P_{\rm FA}^x$ with noise pdf $p_{\mathbf{n},f_0}^{\rm opt}(f_0)$. Let $p_{\mathbf{n}_1,f_0}(f_0) = ((P_{\rm FA}^x - \kappa)/(t_o - \kappa))\delta(f_0 - t_o) + ((t_o - P_{\rm FA}^x)/(t_o - \kappa))p_{\mathbf{n},f_0}^{(}f_0)$. It is easy to verify that $p_{\mathbf{n}_1,f_0}(f_0)$ is a valid pdf. Let $\mathbf{y}_1 = \mathbf{x} + \mathbf{n}_1$. We now have

$$P_{\rm FA}^{y_1} = \frac{P_{\rm FA}^x - \kappa}{t_o - \kappa} t_o + \frac{t_o - P_{\rm FA}^x}{t_o - \kappa}\kappa = P_{\rm FA}^x$$

and

$$P_D^{y_1} = \frac{P_{\rm FA}^x - \kappa}{t_o - \kappa} F_{1M} + \frac{t_o - P_{\rm FA}^x}{t_o - \kappa} P_{\rm D,opt}^y > P_{\rm D,opt}^y.$$

But, this contradicts (15), the definition of $p_{\mathbf{n}}^{\rm opt}$. Therefore, $P_{\rm FA,opt}^y = P_{\rm FA}^x$, i.e., the maximum achievable detection performance is obtained when the probability of false alarm remains the same for the SR noise modified observation $\mathbf{y}$. ■

For Case 2 of Theorem 4, i.e., when $t_o > P_{\rm FA}^x$,[4] let us consider the following construction to derive the form of the optimum noise pdf. From Theorem 4, we have the condition that $P_{\rm FA,opt}^y = F_0(0) = P_{\rm FA}^x$ is a constant. Let us define an auxiliary function G such that

$$G(f_0, k) = J(f_0) - k f_0 \quad (29)$$

where $k \in R$. We have $P_D^y = E_{\mathbf{n}}(J) = E_{\mathbf{n}}(G(f_0,k)) + kE_{\mathbf{n}}(f_0) = E_{\mathbf{n}}(G(f_0,k)) + kP_{\rm FA}^x$. Hence, $p_{\mathbf{n},f_0}^{\rm opt}$ also maximizes $E_{\mathbf{n}}(G(f_0,k))$ and vice versa. Therefore, under the condition that $P_{\rm FA}^y = P_{\rm FA}^x$, maximization of P_D^y is equivalent to maximization of $E_{\mathbf{n}}(G(f_0,k))$. Let us divide the domain of f_0 into two intervals $I_1 = [0, P_{\rm FA}^x]$ and $I_2 = [P_{\rm FA}^x, 1]$. Let $f_{01}(k)$ be the minimum value that maximizes $G(f_0,k)$ in I_1 and let $f_{02}(k)$ be the minimum value that maximizes $G(f_0,k)$ in I_2. Also, let $\nu_1(k) = G(f_{01},k)$ and $\nu_2(k) = G(f_{02},k)$ be the corresponding maximum values. Since for any f_0, $G(f_0,k)$ is monotonically decreasing when k is increasing, $\nu_1(k)$ and $\nu_2(k)$ are monotonically decreasing while $f_{01}(k)$ and $f_{02}(k)$ are monotonically nonincreasing when k is increasing. Since $G(f_0, 0) = J$, therefore, $\nu_2(0) = F_{1M} > \nu_1(0)$, furthermore, when k is very large, we have $\nu_1(k) = J(0) > \nu_2(k) = J(P_{\rm FA}^x) - kP_{\rm FA}^x$. Hence, there exists at least one $k_0 > 0$ such that $\nu_1(k_0) = \nu_2(k_0) \equiv \nu$. For illustration purposes, the plots of $G(f_0,k)$ for the detection problem discussed in Section IV are shown in Fig. 4. Let

[4]This case is usually true because, for a reasonable detector, a higher $P_{\rm FA}$ yields a higher P_D.

us divide the [0,1] interval into two nonoverlapping parts A, $\{f_{01}(k_0), f_{02}(k_0)\}$, such that $\{f_{01}(k_0), f_{02}(k_0)\} \bigcup A = [0,1]$ and $\{f_{01}(k_0), f_{02}(k_0)\} \bigcap A = \{\}$. Next, represent $p_{\mathbf{n},f_0}(f_0)$ as

$$p_{\mathbf{n},f_0}(f_0) = \alpha_1 \delta(f_0 - f_{01}(k_0)) + \alpha_2 \delta(f_0 - f_{02}(k_0)) + I_A(f_0) p_{\mathbf{n},f_0}(f_0) \quad (30)$$

where $I_A(f_0) = 1$ for $f_0 \in A$ and is zero otherwise (an indicator function). From (5), we must have

$$\alpha_1 + \alpha_2 + \int_A p_{\mathbf{n},f_0} df_0 = 1 \quad (31)$$

and

$$\begin{aligned} E_{\mathbf{n}}(G) &= (\alpha_1 + \alpha_2)\nu + \int_A G(f_0, k_0) p_{\mathbf{n},f_0} df_0 \\ &= \nu + \int_A \underbrace{(G(f_0, k_0) - \nu)}_{\leq 0} p_{\mathbf{n},f_0} df_0 \leq \nu. \end{aligned} \quad (32)$$

Note that $J(f_0) \leq \nu$ for all $f_0 \in A$. Clearly, the upper bound can be attained when $p_{\mathbf{n},f_0} = 0$ for all $f_0 \in A$, i.e., $\alpha_1 + \alpha_2 = 1$. Therefore, $P_{\mathrm{D,opt}}^y = E_{\mathbf{n}}(G) + k_0 P_{\mathrm{FA}}^x = \nu + k_0 P_{\mathrm{FA}}^x$. From (28), we have

$$\begin{aligned} p_{\mathbf{n},f_0}^{\mathrm{opt}}(f_0) = &\frac{f_{02}(k_0) - P_{\mathrm{FA}}^x}{f_{02}(k_0) - f_{01}(k_0)} \delta(f_0 - f_{01}(k_0)) \\ &+ \frac{P_{\mathrm{FA}}^x - f_{01}(k_0)}{f_{02}(k_0) - f_{01}(k_0)} \delta(f_0 - f_{02}(k_0)). \end{aligned} \quad (33)$$

Notice that by letting $\lambda = (f_{02}(k_0) - P_{\mathrm{FA}}^x)/(f_{02}(k_0) - f_{01}(k_0))$, (33) is equivalent to (22).

Equivalently, we have the expression of $p_{\mathbf{n}}^{\mathrm{opt}}(\mathbf{n})$ as

$$\begin{aligned} p_{\mathbf{n}}^{\mathrm{opt}}(\mathbf{n}) = &\frac{f_{02}(k_0) - P_{\mathrm{FA}}^x}{f_{02}(k_0) - f_{01}(k_0)} \delta(\mathbf{n} - \mathbf{n}_1) \\ &+ \frac{P_{\mathrm{FA}}^x - f_{01}(k_0)}{f_{02}(k_0) - f_{01}(k_0)} \delta(\mathbf{n} - \mathbf{n}_2). \end{aligned} \quad (34)$$

Further, in the special case where f_1 is continuously differentiable, G is also continuously differentiable. Since f_{01} and f_{02} are at least local maxima, we have $(\partial G/\partial f_0)(f_{01}, k_0) = (\partial G/\partial f_0)(f_{02}, k_0) = 0$. Therefore, from the derivative of (29), we have

$$\frac{dJ}{df_0}(f_{01}(k_0)) = \frac{dJ}{df_0}(f_{02}(k_0)) = k_0 \quad (35)$$

$$J(f_{02}(k_0)) - J(f_{01}(k_0)) = k_0 (f_{02}(k_0) - f_{01}(k_0)). \quad (36)$$

In other words, the line connecting $(J(f_{01}(k_0)), f_{01}(k_0))$ and $J(f_{02}(k_0)), f_{02}(k_0))$ is the bitangent line of $J(\cdot)$ and k_0 is its slope. Also,

$$P_{\mathrm{D,opt}}^y = \nu + k_0 P_{\mathrm{FA}}^x. \quad (37)$$

In this section, we have derived the condition under which SR noise can improve detection performance. Also, we have obtained the specific form of the optimum SR noise. Next we illustrate the ideas by applying the theory to a specific detection problem.

IV. A Detection Example

Here, we consider the same detection problem as considered by Kay [28]. The two hypotheses H_0 and H_1 are given as

$$\begin{cases} \mathrm{H}_0 : x[i] = w[i] \\ \mathrm{H}_1 : x[i] = A + w[i] \end{cases} \quad (38)$$

for $i = 0, 1, \cdots, N-1$, $A > 0$ is a known dc signal, and $w[i]$ are i.i.d noise samples with a symmetric Gaussian mixture noise pdf

$$p_w(w) = \frac{1}{2}\gamma(w; -\mu, \sigma_0^2) + \frac{1}{2}\gamma(w; \mu, \sigma_0^2) \quad (39)$$

where $\gamma(w; \mu, \sigma^2) = (1/\sqrt{2\pi\sigma^2}) \exp[-((w-\mu)^2/2\sigma^2)]$. Here, we set $\mu = 3$, $A = 1$ and $\sigma_0 = 1$. A suboptimal detector is considered with test statistic

$$T(x) = \frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{1}{2} + \frac{1}{2} \mathrm{sgn}(x[i]) \right) = \frac{1}{N} \sum_{i=0}^{N-1} \varpi_x[i] \quad (40)$$

where $\varpi_x[i] = (1/2) + (1/2)\mathrm{sgn}(x[i])$. From (40), this detector is essentially a fusion of the decision results of N i.i.d. sign detectors.

When $N = 1$, the detection problem reduces to a problem with the test statistic $T_1(x) = x$, threshold $\eta = 0$ (sign detector) and the probability of false alarm $P_{\mathrm{FA}}^x = 0.5$. The distribution of x under the H_0 and H_1 hypotheses can be expressed as

$$p_0(x) = \frac{1}{2}\gamma(x; -\mu, \sigma_0^2) + \frac{1}{2}\gamma(x; \mu, \sigma_0^2) \quad (41)$$

and

$$p_1(x) = \frac{1}{2}\gamma(x; -\mu + A, \sigma_0^2) + \frac{1}{2}\gamma(x; \mu + A, \sigma_0^2) \quad (42)$$

respectively. The critical function is given by

$$\phi(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}. \quad (43)$$

The problem of determining the optimal SR noise is to find the optimal $p(n)$ where for the new observation $y = x + n$, the probability of detection $P_D^y = p(y > 0; H_1)$ is maximum while the probability of false alarm $P_{\mathrm{FA}}^y = p(y > 0; H_0) \leq P_{\mathrm{FA}}^x = 1/2$.

When $N > 1$, the detector is equivalent to a fusion of N individual detectors and the detection performance monotonically increases with N. Like the $N = 1$ case, when the decision function is fixed, the optimum SR noise can be obtained by a similar procedure. Due to space limitations, here only the suboptimal case where the additive noise n is assumed to be an i.i.d noise is considered. Under this constraint, since the P_Ds and P_{FA}s of each and every detector are the same, it can be shown that the optimal noise for the case $N > 1$ is the same as $N = 1$ because again, we need to fix $P_{\mathrm{FA}} \leq 0.5$ for each individual detector while increasing its P_D. Hence, in the following discussion, we only consider the one sample case ($N = 1$). However, the performance of the $N > 1$ case can be derived similarly.

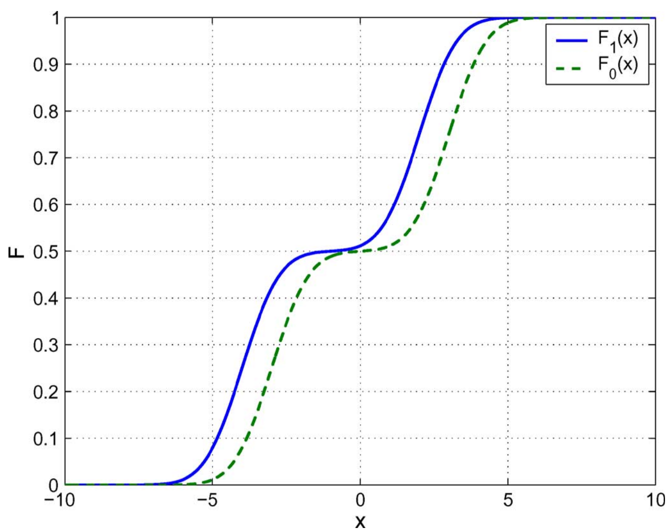


Fig. 2. $F_1(x)$ and $F_0(x)$ as a function of x as given in (44) and (45) for $\mu = 3$, $A = 1$, and $\sigma_0 = 1$, respectively.

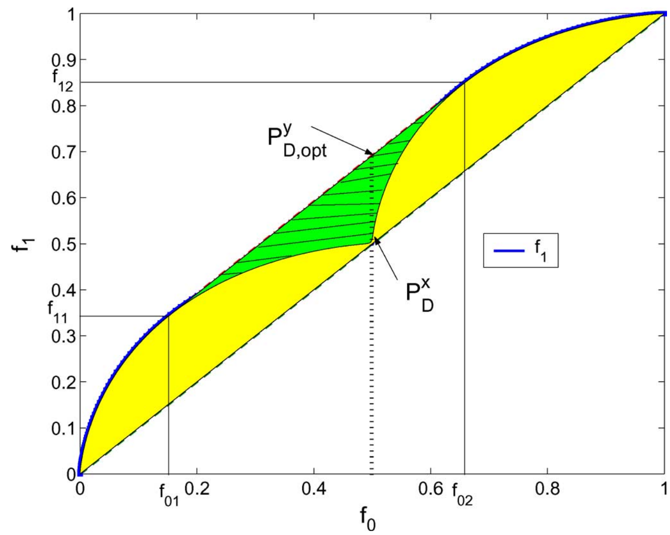


Fig. 3. Relationship between $F_1(x)$ and $F_0(x)$ as derived from (44) and (45), respectively. The shadowed region [including both yellow (light gray) and dashed green (dashed dark gray)] is the convex hull V of U. The green dashed region is the region of (f_1, f_0) where possible SR effect may take place.

A. Determination of the Optimal SR Noise pdf

From (11) and (13), it can be shown that in this case

$$\begin{aligned} F_1(x) &= \int_0^{+\infty} \phi(y)p_1(y-x)dy \\ &= \frac{1}{2}\left(\int_0^{+\infty} \left(\gamma\left(y-x;-\mu+A,\sigma_0^2\right)\right.\right. \\ &\qquad \left.\left.+\gamma\left(y-x;\mu+A,\sigma_0^2\right)dy\right)\right) \\ &= \frac{1}{2}Q\left(\frac{-x-\mu-A}{\sigma_0}\right)+\frac{1}{2}Q\left(\frac{-x+\mu-A}{\sigma_0}\right) \end{aligned} \tag{44}$$

and

$$\begin{aligned} F_0(x) &= \int_0^{+\infty} \phi(y)p_0(y-x)dy \\ &= \frac{1}{2}\left(\int_0^{+\infty} \left(\gamma\left(y-x;-\mu,\sigma_0^2\right)+\gamma\left(y-x;\mu,\sigma_0^2\right)dy\right)\right) \\ &= \frac{1}{2}Q\left(\frac{-x-\mu}{\sigma_0}\right)+\frac{1}{2}Q\left(\frac{-x+\mu}{\sigma_0}\right), \end{aligned} \tag{45}$$

where $Q(x) = \int_x^{\infty}(1/\sqrt{2\pi})\mathrm{e}^{-t^2/2}dt$. It is also easy to show that in this case, $F_1(x) > F_0(x)$ and both are monotonically increasing with x. Therefore, $J(f_0) = f_1(f_0) = f_1$, and $U = (f_1, f_0)$ is a single curve. Fig. 2 shows the values of f_1 and f_0 as a function of x while the relationship between F_1 and F_0 is shown in Fig. 3. V, the convex hull of all possible P_D and P_{FA} after n is added is shown as the light and dark shadowed regions, respectively, in Fig. 3. Note that a similar nonconcave ROC occurs in distributed detection systems [37] and dependent randomization is employed to improve system performance [38], [39].

Taking the derivative of f_1 w.r.t. f_0, we have

$$\frac{d(f_1)}{d(f_0)} = \frac{\frac{d(f_1)}{d(x)}}{\frac{d(f_0)}{d(x)}} = \frac{p_1(-x)}{p_0(-x)} \tag{46}$$

and,

$$\begin{aligned} \frac{d^2(f_1)}{d\left(f_0^2\right)} &= \frac{1}{p_0(-x)}\frac{d\left(\frac{p_1(-x)}{p_0(-x)}\right)}{dx} \\ &= \frac{-p_1'(-x)p_0(-x)+p_0'(-x)p_1(-x)}{p_0^3(-x)}, \end{aligned} \tag{47}$$

where $x = F_0^{-1}(f_0)$. Since $d\gamma(y-x;\mu,\sigma^2)/dx = (\mu - x/\sigma^2)\gamma(y-x;\mu,\sigma^2)$, we have $p_0'(-x)|_{x=0} = 0$ and

$$\begin{aligned} \left.\frac{d^2(f_1)}{d\left(f_0^2\right)}\right|_{f_0=f_0(0)} &= \left.\frac{-p_1'(-x)p_0(-x)+p_0'(-x)p_1(-x)}{p_0^3(-x)}\right|_{x=0} \\ &= \left.\frac{-p_1'(-x)}{p_0^2(-x)}\right|_{x=0} \\ &= \frac{(\mu-A)\exp\left(-\frac{(\mu-A)^2}{2\sigma_0^2}\right)}{\sqrt{2\pi}\sigma_0^3p_0^2(0)} \\ &\quad - \frac{(\mu+A)\exp\left(-\frac{(\mu+A)^2}{2\sigma_0^2}\right)}{\sqrt{2\pi}\sigma_0^3p_0^2(0)}. \end{aligned} \tag{48}$$

Next, let us discuss the improvability of this detector. First, when $A < \mu$, setting (48) equal to zero and solving the equation for σ_0, we have σ_1, the zero pole of (48)

$$\sigma_1 = \sqrt{2\frac{\mu A}{\ln\left(\frac{\mu+A}{\mu-A}\right)}}.$$

When $\sigma_0 < \sigma_1$, we have $d^2(f_1)/d(f_0^2)|_{f_0=F_0(0)} > 0$ and in this example, $\sigma_1^2 = 8.6562 > \sigma_0^2 = 1$. From Theorem 1, this detector is improvable by adding independent SR noise. When $A > \mu$, $d^2(f_1)/d(f_0^2)|_{f_0=f_0(0)} < 0$, and the improvability cannot be determined by Theorem 1. However, for this particular detector, as we discuss later, the detection performance can still be improved.

We now determine the two discrete values as well as the probability of their occurrence by solving equations (35) and (36). From (44) and (45), the relationship between f_1, f_0 and x, and (46), we have

$$\frac{p_1(-n_1)}{p_0(-n_1)} = \frac{p_1(-n_2)}{p_0(-n_2)}$$
$$\frac{F_1(n_1)-F_1(n_2)}{F_0(n_1)-F_0(n_2)} = \frac{p_1(-n_2)}{p_0(-n_2)}. \tag{49}$$

Although it is generally very difficult to solve the above equation analytically, fortunately, in this particular detection problem, we have

$$\begin{aligned} p_1\left(-\left(\mu-\frac{A}{2}\right)\right) &= 0.5\gamma\left(-\frac{A}{2};0,\sigma_0^2\right) \\ &\quad + 0.5\gamma\left(2\mu+\frac{A}{2};0,\sigma_0^2\right) \\ p_0\left(-\left(\mu-\frac{A}{2}\right)\right) &= 0.5\gamma\left(-\frac{A}{2};0,\sigma_0^2\right) \\ &\quad + 0.5\gamma\left(2\mu-\frac{A}{2};0,\sigma_0^2\right) \\ p_1\left(-\left(-\mu-\frac{A}{2}\right)\right) &= 0.5\gamma\left(-\frac{A}{2};0,\sigma_0^2\right) \\ &\quad + 0.5\gamma\left(2\mu-\frac{A}{2};0,\sigma_0^2\right) \\ p_0\left(-\left(-\mu-\frac{A}{2}\right)\right) &= 0.5\gamma\left(-\frac{A}{2};0,\sigma_0^2\right) \\ &\quad + 0.5\gamma\left(2\mu+\frac{A}{2};0,\sigma_0^2\right) \end{aligned}$$

so that $p_1(-(\mu-(A/2)))/p_0(-(\mu-(A/2))) \cong 1$, $p_1(-(-\mu-(A/2)))/p_0(-(-\mu-(A/2))) \cong 1$ and $F_1((\mu-(A/2))) - F_1((-\mu-(A/2))) = F_0((\mu-(A/2))) - F_0((-\mu-(A/2)))$, given $2\mu - A/2 > 3\sigma_0$. Thus, the roots n_1, n_2 of (49) can be approximately expressed as $n_1 = -\mu-(A/2)$ and $n_2 = \mu-(A/2)$. Correspondingly, $\lambda = (F_0(n_2)-F_0(0))/(F_0(n_2)-F_0(n_1))$ and $1-\lambda = (F_0(0)-F_0(n_1))/(F_0(n_2)-F_0(n_1))$.

Hence

$$\begin{aligned} p_n^{\text{opt}}(n) &= \lambda\delta(n-n_1) + (1-\lambda)\delta(n-n_2) \\ &= 0.3085\delta(n+3.5) + 0.6915\delta(n-2.5) \end{aligned} \tag{50}$$

and

$$P_{\text{D,opt}}^{y} = \lambda F_1(n_1) + (1-\lambda)F_1(n_2) = 0.6915. \tag{51}$$

B. Optimal Symmetrical Noises

In this subsection, we consider the special cases where the SR noise is constrained to be symmetric. These include symmetric noise with arbitrary pdf $p_s(x)$, white Gaussian noise $p_g(x) = \gamma(x;0,\sigma^2)$ and white uniform noise $p_u(x) = 1/a$, $a > 0$, $-(a/2) \le x \le a/2$. The noise modified data processes are denoted as y_s, y_g and y_u, respectively. Here, for illustration purposes, we find the pdfs of these suboptimal SR noises using the $C(x)$ functions. The same results can be obtained by applying the same approach as in the previous subsection using $F_1(\cdot)$ and $F_0(\cdot)$ functions.

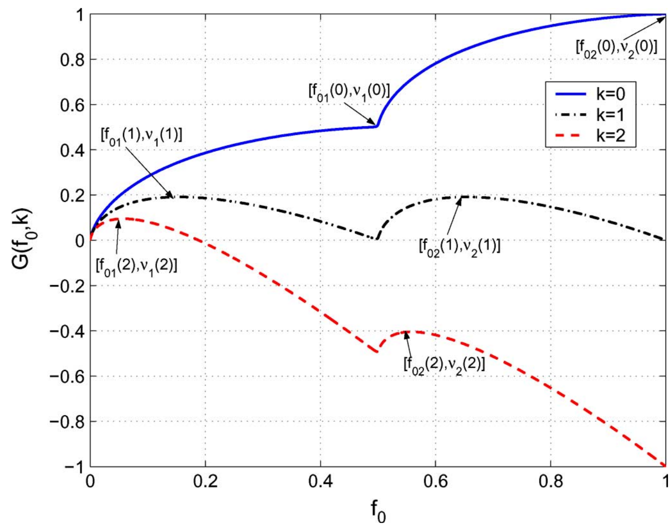


Fig. 4. An illustration of the relationship between $G(f_0,k)$, f_0, $f_{0i}(k)$, $\nu_i(k)$ with $i = 1, 2$ and different k value 0,1 and 2.

For the arbitrary symmetrical noise case, we have the condition

$$p_s(x) = p_s(-x). \tag{52}$$

Therefore, $p(y_s|H_0)$ is also a symmetric function, so that $P_{\text{FA}}^{y_s} = 1/2$. By (43) and (52), we have

$$\begin{aligned} C_s(x) &= \int_0^\infty p_s(t-x)dt \\ &= \int_{-x}^\infty p_s(t)dt \\ &= \int_{-\infty}^x p_s(t)dt \\ &= 1 - C_s(-x). \end{aligned} \tag{53}$$

Since $p_s(x) \ge 0$, we also have

$$C_s(x_1) \ge C_s(x_0) \quad \text{for any } x_1 \ge x_0, \tag{54}$$
$$C_s(0) = 1/2, \quad C_s(-\infty) = 0, \text{ and } C_s(\infty) = 1. \tag{55}$$

From (9) and (53), we have the P_D of y_s given by

$$\begin{aligned} P_D^{y_s} &= \int_{-\infty}^\infty p_1(x)C_s(x)dx \\ &= \int_{-\infty}^0 p_1(x)C_s(x)dx + \int_0^\infty (1-C_s(-x))\,p_1(x)dx \\ &= \int_{-\infty}^0 (p_1(x)-p_1(-x))\,C_s(x)dx + P_D^x \\ &= \int_{-\infty}^0 H(x)C_s(x)dx + P_D^x \end{aligned} \tag{56}$$

TABLE I
COMPARISON OF DETECTION PERFORMANCE FOR DIFFERENT SR NOISE ENHANCED DETECTORS

SR noise	p_n^{opt}	p_s^{opt}	p_u^{opt}	p_g^{opt}	No SR noise
P_D^y	0.6915	0.6707	0.6011	0.5807	0.5114

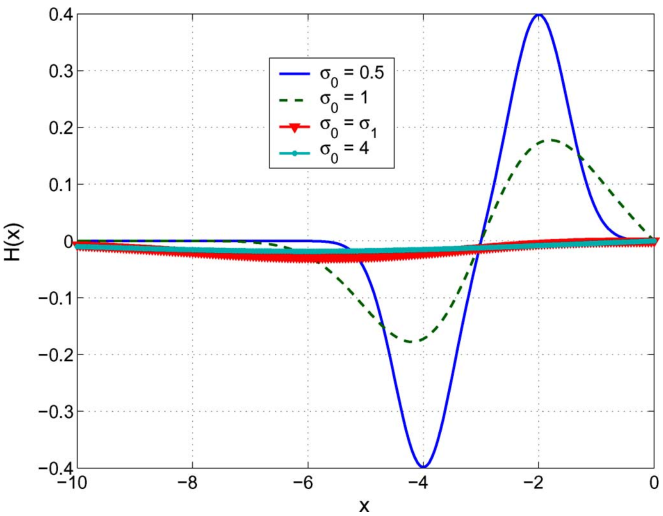


Fig. 5. Different $H(x)$ curves where $\mu = 3$, $A = 1$.

where $H(x) \equiv p_1(x) - p_1(-x)$. Fig. 5 shows a plot of $H(x)$ for several σ_0 values. Finally, from (42), we have

$$p_1(-x) = \frac{1}{2}\gamma\left(x; \mu - A, \sigma_0^2\right) + \frac{1}{2}\gamma\left(x; -\mu - A, \sigma_0^2\right).$$

When $A \geq \mu$, since $p_1(-x) \geq p_1(x)$ when $x < 0$, we have, $H(x) < 0$, $x < 0$. From (56), $P_D^y \leq P_D^x$ for any $H(x)$, i.e., in this case, the detection performance of this detector cannot be improved by adding symmetric noise. When $A < \mu$ and $\sigma_0 \geq \sigma_1$ we also have $H(x) < 0, \forall x < 0$. Therefore, adding symmetric noise will not improve the detection performance as well. However, when $\sigma_0 < \sigma_1$, $H(x)$ has only a single root x_0 for $x < 0$ and $H(x) < 0, \forall x < x_0, H(x) > 0, \forall x \in (x_0, 0)$ and detection performance can be improved by adding symmetric SR noise. From (56), we have

$$C_s^{\text{opt}}(x) = \begin{cases} 0, & x < x_0 \\ \frac{1}{2}, & x_0 \leq x \leq 0 \end{cases} \tag{57}$$

and

$$p_s^{\text{opt}} = \frac{1}{2}\delta(x - x_0) + \frac{1}{2}\delta(x + x_0).$$

Furthermore, since $\gamma(-\mu; -\mu - A, \sigma_0^2) = \gamma(-\mu; -\mu + A, \sigma_0^2)$ and $\gamma(-\mu; \mu + A, \sigma_0^2) \cong \gamma(-\mu; \mu - A, \sigma_0^2) \cong 0$ given $2\mu - A \gg \sigma_0$, we have $x_0 \cong -\mu$. Therefore

$$p_s^{\text{opt}} = \frac{1}{2}\delta(x - \mu) + \frac{1}{2}\delta(x + \mu). \tag{58}$$

The pdf of y for the H_1 hypothesis becomes

$$p_{1,y_s}^{\text{opt}}(y) = \frac{1}{2}\gamma\left(y; A, \sigma_0^2\right) + \frac{1}{4}\gamma\left(y; 2\mu + A, \sigma_0^2\right) + \frac{1}{4}\gamma\left(y; -2\mu + A, \sigma_0^2\right). \tag{59}$$

Hence, when μ is large enough, $P_{\text{D,opt}}^{y_s} = (1/2)Q(-(A/\sigma_0)) + (1/4) = 0.6707$. Note that, as σ_0 decreases, $P_{\text{D,opt}}^{y_s}$ increases, i.e., better detection performance can be achieved by adding the optimal symmetric noise.

Similarly, for the uniform noise case,

$$C_u(x) = \int_{-x}^{\infty} p_u(t)dt = \begin{cases} 0, & x < \frac{-a}{2} \\ \frac{x}{a} + \frac{1}{2}, & -\frac{a}{2} \leq x \leq 0 \end{cases}. \tag{60}$$

Substituting (60) for $C_s(x)$ in (56) and taking the derivative w.r.t a, we have

$$\frac{dP_D^{y_s}}{da} = -\frac{1}{a^2}\int_{-\frac{a}{2}}^{0} xH(x)dx. \tag{61}$$

Setting it equal to zero and solving, we have $a_{\text{opt}} = 8.4143$ in the pdf of uniform noise defined earlier. Additionally, we have $P_{\text{D,opt}}^{y_u} = 0.6011$.

For the Gaussian case, the optimal WGN level is readily determined since

$$P_D^{y_g} = \frac{1}{2}Q\left(\frac{-A-\mu}{\sqrt{\sigma_0^2 + \sigma^2}}\right) + \frac{1}{2}Q\left(\frac{-A+\mu}{\sqrt{\sigma_0^2 + \sigma^2}}\right). \tag{62}$$

Let $\sigma_2^2 = \sigma_0^2 + \sigma^2$. Taking the derivative w.r.t σ_2^2 in (62), setting it equal to zero and solving, we obtain

$$\sigma_2^2 = 2\frac{\mu A}{\ln\left(\frac{\mu+A}{\mu-A}\right)} = 8.6562 \tag{63}$$

and $\sigma_{\text{opt}}^2 = \sigma_2^2 - \sigma_0^2 = 7.6562$, and, correspondingly, $P_{\text{D,opt}}^{y_g} = 0.5807$. Therefore, when $\sigma_0^2 < \sigma_2^2$, adding WGN with variance σ_{opt}^2 can improve the detection performance to a constant level $P_{\text{D,opt}}^{y_g}$.

C. Detection Performance Results

Table I shows the values of $P_{\text{D,opt}}^y$ for these different types of SR noise. Compared to the original data process with $P_D^x = 0.5114$, the improvement of different detectors are 0.1811, 0.1593, 0.0897, and 0.0693 for optimum SR noise, optimum symmetric noise, optimum uniform noise, and optimum Gaussian noise enhanced detectors, respectively.

Fig. 6 shows P_D^x as well as the maximum achievable P_D^y with different values of A. The detection performance is significantly improved by adding optimal SR noise. When $A \leq \mu$, a certain degree of improvement is also observed by adding suboptimal SR noise. When A is small, $n_1 \approx -\mu$ and $n_2 \approx \mu$, the detection performance of the optimum SR noise enhanced detector is close to the optimum symmetric noise enhanced one. However, when $A > 0.6$, the difference is significant. When $A > \mu = 3$, $H(x) < 0, \forall x < 0$, so that $P_{\text{D,opt}}^{y_s} = P_{\text{D,opt}}^{y_u} = P_{\text{D,opt}}^{y_g} = P_D^x$, i.e, the optimal symmetric noise is zero (no SR noise). However, by adding optimal SR noise, $P_{\text{D,opt}}^y$ is still larger than P_D^x, i.e., the detection performance can still be improved. When

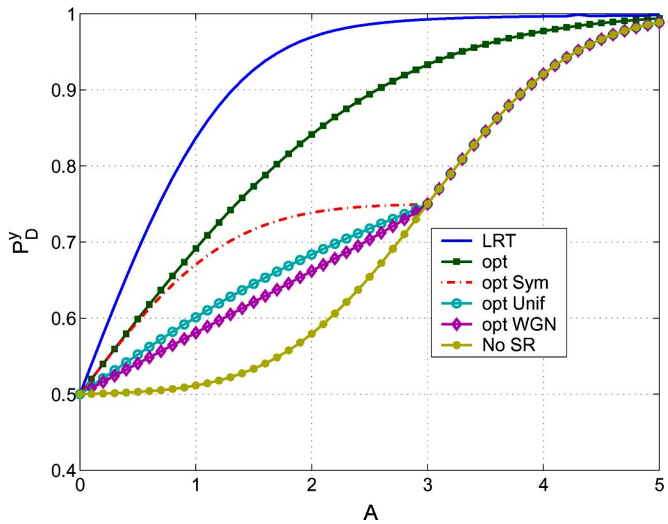


Fig. 6. P_D^y as a function of signal level A in Gaussian mixture noise when $\mu = 3$ and $\sigma_0 = 1$. "LRT" is the P_D^x obtained by applying the optimum LRT on the observed data x. "opt", "opt Sym", "opt Unif", "opt WGN" and "No SR" are the P_D^y of the optimum noise, optimum symmetric noise, optimum uniform noise, optimum white Gaussian noise and original data (no SR noise), respectively.

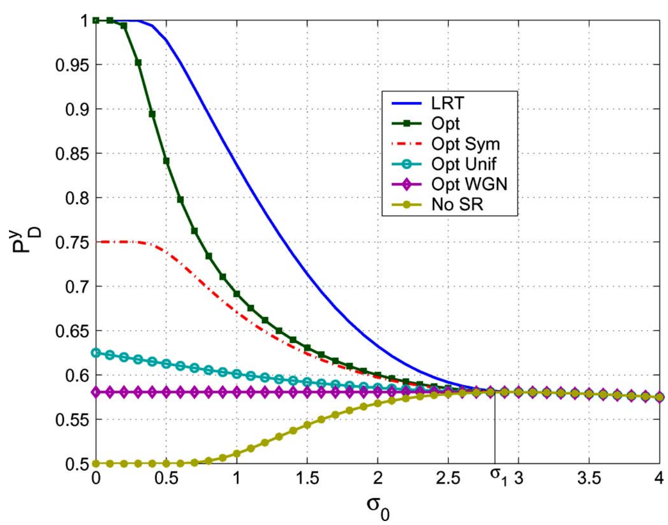


Fig. 7. P_D^y as a function of σ_0 for different types of noise enhanced detectors when $\mu = 3$ and $A = 1$.

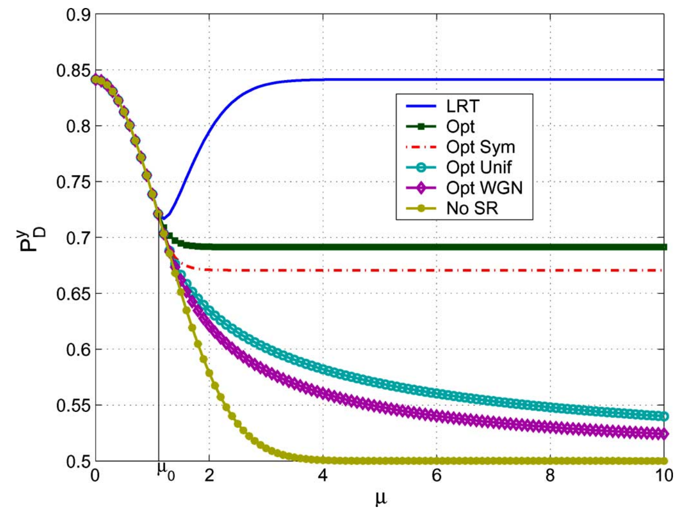


Fig. 8. P_D^y as a function of μ for different types of noise enhanced detectors when $\sigma_0 = 1$ and $A = 1$.

$A \geq 5$, the P_D improvement is not that significant because $P_D^x > 0.97 \approx 1$ which is already a very good detector.

The maximum achievable detection performance of different SR noise enhanced detectors with different background noise σ_0 is shown in Fig. 7. When σ_0 is small, for the optimum SR noise enhanced detectors $P_{\mathrm{D,opt}}^{y} \approx 1$, while for the symmetric SR noise case $P_{\mathrm{D,opt}}^{y_s} \approx 0.75$. When σ_0 increases, P_D^x increases and the detection performance of SR noise enhanced detectors degrades. When $\sigma_0 \geq \sigma_1$, $p_0(x)$ becomes a unimodal noise and the decision function ϕ is the same as the decision function decided by the optimum LRT test given the false alarm $P_{\mathrm{FA}} = 0.5$. Therefore, adding any SR noise will not improve P_D. Hence, all the detection results converge to P_D^x.

Fig. 8 compares the detection performance of different detectors w.r.t. μ when $A = 1$ and $\sigma_0 = 1$ is fixed. P_D^x, $P_{\mathrm{D,opt}}^{y_u}$ and $P_{\mathrm{D,opt}}^{y_g}$ monotonically decrease when μ increases. Also, there exist a unique μ value μ_0, such that when $\mu < \mu_0$ is small, p_0 is still a unimodal pdf, so that the decision function ϕ is the optimum one for $P_{\mathrm{FA}} = 0.5$. An interesting observation from Fig. 8 is that the P_D of the "optimum LRT", after the lowest value is reached, increases when μ increases. The explanation of this phenomenon is that when μ is sufficiently large, the separation of the two peaks of the Gaussian mixtures increases as μ increases so that the detectability is increased. When $\mu \to \infty$, the two peaks are sufficiently separated, so that the detection performance of "LRT" is equal to the P_D when $\mu = 0$.

Finally, Fig. 9 shows the ROC curves for this detection problem when $N = 30$ and the different types of i.i.d SR noise determined previously are added. Different degrees of improvement are observed for different SR noises pdfs. The optimum SR detector and the optimum symmetric SR detector performance levels are superior to those of the uniform and Gaussian SR detectors and more closely approximate the LRT curve.

V. Concluding Remarks

In this paper, we have established the mathematical theory for the SR noise modified detection problem. Several fundamental theorems on SR in detection theory are established. We analyzed the detection performance of a SR noise enhanced detector where, for any additive noise, the detection performance in terms of P_D and P_{FA} can be obtained by applying the expressions we have developed. Based on that, we have established the conditions of potential improvement of P_D via the SR effect. This leads to the sufficient condition for the improvability/non-improvability of most suboptimal detectors. The exact form of the optimal SR noise pdf has been proposed. The optimal SR noise is shown to be a proper randomization of no more than two discrete signals. Also, the upper limit of the SR enhanced detection performance is obtained. Given the distributions p_1 and p_0, a theoretical approach is proposed to determine the optimal SR consisting of the two discrete signals and their corresponding weights.

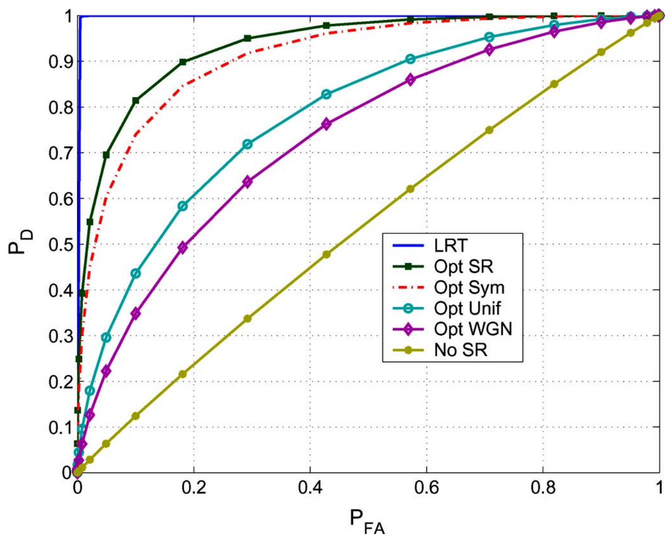


Fig. 9. ROC curves for different SR noise enhanced sign detectors, $N = 30$. For "LRT", its performance nearly perfect ($P_D \approx 1$ for all P_{FA}'s).

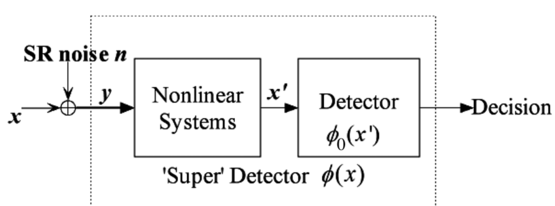


Fig. 10. An illustration of a SR detection system and the corresponding "Super" detector.

It should be pointed out that the results obtained in this paper are very general and are applicable to a variety of SR detectors considered in the literature, e.g., bistable systems. The SR detectors presented in [22], [30], [33]–[35] can be included in our framework as shown in Fig. 10. For example, the nonlinear system block of Fig. 10 can depict the bistable system [33]–[35]. Let $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$ be the input to the nonlinear system, and $\mathbf{x}' = [x'_1, x'_2, \cdots, x'_N]^T$ be the output of the system as shown, where $\mathbf{x}' = f(\mathbf{x})$ is the appropriate nonlinear function. The decision problem based on $\mathbf{x}'$ can be described by decision function $\phi_0(\cdot)$ as shown. It is easy to observe that the corresponding decision function $\phi(\cdot)$ for the "super" detector (nonlinear system plus detector) is $\phi(\mathbf{x}) = \phi_0(f(\mathbf{x}))$. Thus the SR detectors proposed in the literature can be incorporated in our framework and the theory developed in the paper is applicable to these general situations.

Based on our mathematical framework, for a particular detection problem, we have compared the detection performance of six different detectors, namely, the optimum LRT detector, optimum noise enhanced sign detector, optimum symmetric noise enhanced sign detector, optimum uniform noise enhanced sign detector, optimum Gaussian noise enhanced sign detector and the original sign detector. Compared to the traditional SR approach where the noise type is predetermined, much better detection performance is obtained by adding the proposed optimum SR noise to the observed data process.

This fundamental theory well explains the observed SR phenomenon in signal detection problems, and greatly advances our ability to determine the applicability of SR in signal detection. It can also be applied to many other signal processing problems such as distributed detection and fusion as well as pattern recognition applications.

APPENDIX
REVIEW OF CONVEX FUNCTIONS AND CONVEX SETS [36]

In this section, we put together some background information of convex functions and convex sets for reader's convenience. More details are available in [36].

A. *Convex Functions*

A function $f : I \to \mathrm{R}$ is called **convex** if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{64}$$

for all $x, y \in I$ and λ in the open interval (0,1). It is called **strictly convex** provided that the inequality (64) is strict for $x \neq y$. Similarly, if $-f : I \to \mathrm{R}$ is convex, then we say that $f : I \to \mathrm{R}$ is **concave**.

Theorem A-1: Suppose f'' exists on (a, b). Then f is convex if and only if $f''(x) \geq 0$. And if $f''(x) > 0$ on (a, b), then f is strictly convex on the interval.

B. *Convex Sets*

Let U be a subset of a linear space $\mathbf{L}$. We say that U is **convex** if $\mathbf{x}, \mathbf{y} \in U$ implies that $z = [\lambda x + (1-\lambda)y] \in U$ for all $\lambda \in [0, 1]$.

Theorem A-2: A set $U \subseteq L$ is convex if and only if every convex combination of points of U lies in U.

We call the intersection of all convex sets containing a given set U the *convex hull* of U, denoted by $H(U)$.

Theorem A-3: For any $U \subseteq L$, the convex hull of U consists precisely of all convex combinations of elements of U.

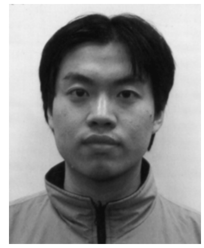
Furthermore, for the convex hull, we have Carathéodory's theorem for convex sets.

Theorem A-4 (Carathéodory's Theorem): If $U \subseteq L$ and its convex hull $H(U)$ has dimension m, then for each $z \in H(U)$, there exist $m+1$ point $x_0, \ldots, x_m$ of U such that z is a convex combination of these points.

REFERENCES

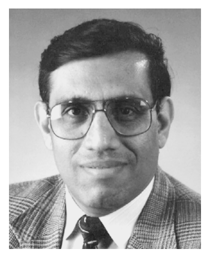
[1] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A: Math. General*, vol. 14, pp. L453–L457, 1981.
[2] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223–287, Jan. 1998.
[3] J. J. Collins, C. C. Chow, and T. T. Imhoff, "Stochastic resonance without tuning," *Nature*, vol. 376, pp. 236–238, 1995.
[4] 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.
[5] S. M. Bezrukov and I. Vodyanoy, "Stochastic resonance in non-dynamical systems without response thresholds," *Nature*, vol. 385, pp. 319–321, Jan. 1997.
[6] F. Chapeau-Blondeau, "Input-output gains for signal in noise in stochastic resonance," *Phys. Lett. A*, vol. 232, pp. 41–48, 1997.
[7] F. Chapeau-Blondeau and X. Godivier, "Theory of stochastic resonance in signal transmission by static nonlinear systems," *Phys. Rev. E*, vol. 55, no. 2, pp. 1478–1495, 1997.
[8] F. Chapeau-Blondeau, "Periodic and aperiodic stochastic resonance with output signal-to-noise ratio exceeding that at the input," *Int. J. Bifurc. Chaos*, vol. 9, pp. 267–272, 1999.

[9] P. Hänggi, M. E. Inchiosa, D. Fogliatti, and A. R. Bulsara, "Nonlinear stochastic resonance: The saga of anomalous output-input gain," *Phys. Rev. E*, vol. 62, no. 5, pp. 6155–6163, 2000.
[10] F. J. Castro, M. N. Kuperman, M. Fuentes, and H. S. Wio, "Experimental evidence of stochastic resonance without tuning due to non-Gaussian noises," *Phys. Rev. E*, vol. 64, no. 051105, 2001.
[11] Z. Gingl, P. Makra, and R. Vajtai, "High signal-to-noise ratio gain by stochastic resonance in a double well," *Fluctuat. Noise Lett.*, vol. 1, no. 3, pp. L181–L188, 2001.
[12] 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.
[13] 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.
[14] P. Makra, Z. Gingl, and T. Fulei, "Signal-to-noise ratio gain in stochastic resonators driven by coloured noises," *Phys. Lett. A*, vol. 317, pp. 228–232, 2003.
[15] S. Zozor and P.-O. Amblard, "Stochastic resonance in discrete time nonlinear ar(1) models," *IEEE Trans. Signal Process.*, vol. 47, pp. 108–122, Jan. 1999.
[16] S. Zozor and P.-O. Amblard, "On the use of stochastic resonance in sine detection," *Signal Process.*, vol. 7, pp. 353–367, Mar. 2002.
[17] S. Zozor and P.-O. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177–3181, Dec. 2003.
[18] S. Mitaim and B. Kosko, "Adaptive stochastic resonance," *Proc. IEEE*, vol. 86, pp. 2152–2183, 1998.
[19] B. Xu, J. Li, and J. Zheng, "How to tune the system parameters to realize stochastic resonance," *J. Phys. A: Math. General*, vol. 36, pp. 11 969–11 980, 2003.
[20] 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.
[21] I. Goychuk and P. Hänggi, "Stochastic resonance in ion channels characterized by information theory," *Phys. Rev. E*, vol. 61, no. 4, pp. 4272–4280, 2000.
[22] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Mar. 2000.
[23] B. Kosko and S. Mitaim, "Stochastic resonance in noisy threshold neurons," *Neural Netw.*, vol. 16, pp. 755–761, 2003.
[24] B. Kosko and S. Mitaim, "Robust stochastic resonance for simple threshold neurons," *Phys. Rev. E*, vol. 70, no. 031911, 2004.
[25] S. Mitaim 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.
[26] M. McDonnell, N. Stocks, C. Pearce, and D. Abbott, "Stochastic resonance and data processing inequality," *Electron. Lett.*, vol. 39, no. 17, pp. 1287–1288, 2003.
[27] A. Asdi and A. Tewfik, "Detection of weak signals using adaptive stochastic resonance," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP 95)*, May 1995, vol. 2, pp. 1332–1335.
[28] S. Kay, "Can detectability be improved by adding noise?," *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.
[29] A. A. Saha and G. Anand, "Design of detectors based on stochastic resonance," *Signal Process.*, vol. 83, pp. 1193–1212, 2003.
[30] D. Rousseau and F. Chapeau-Blondeau, "Stochastic resonance and improvement by noise in optimal detection strategies," *Digit. Signal Process.*, vol. 15, pp. 19–32, 2005.
[31] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Springer, 1986.
[32] J. Thomas, "Nonparametric detection," *Proc. IEEE*, vol. 58, no. 5, pp. 623–631, May 1970.
[33] V. Galdi, V. Pierro, and I. M. Pinto, "Evaluation of stochastic-resonance-based detectors of weak harmonic signals in additive white Gaussian noise," *Phys. Rev. E*, vol. 57, no. 6, pp. 6470–6479, Jun. 1998.
[34] M. E. Inchiosa and A. R. Bulsara, "Signal detection statistics of stochastic resonators," *Phys. Rev. E*, vol. 53, no. 3, pp. R2021–R2024, Mar. 1996.
[35] J. Tougaard, "Signal detection, detectability and stochastic resonance effects," *Biolog. Cybern.*, vol. 87, pp. 79–90, 2002.
[36] A. W. Roberts and D. E. Varberg, *Convex Functions*. New York: Academic, 1973.
[37] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer, 1997.
[38] P. Willet and D. Warren, "The suboptimality of randomized tests in distributed and quantized detection systems," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 355–361, 1992.
[39] F. Gini, F. Lombardini, and L. Verrazzani, "Decentralised detection strategies under communication constraints," *Inst. Elect. Eng. Proc. Radar, Sonar and Navigat.*, vol. 145, no. 4, pp. 199–208, Aug. 1998.



Hao Chen (S'06) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China (USTC), Hefei, in 1999 and 2002, respectively.

Since 2002, he has been pursuing the Ph.D. degree with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY. His research interests are in the areas of statistical signal processing and its applications, including detection, estimation, remote sensing, and image processing.



Pramod K. Varshney (S'72–M'77–SM'82–F'97) was born in Allahabad, India, on July 1, 1952. He received the B.S. degree in electrical engineering and computer science (with highest honors), and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, in 1972, 1974, and 1976, respectively.

During 1972–1976, he held teaching and research assistantships with the University of Illinois. Since 1976 he has been with Syracuse University, Syracuse, NY, where he is currently a Professor of electrical engineering and computer science and the Research Director of the New York State Center for Advanced Technology in Computer Applications and Software Engineering. He served as the Associate Chair of the department during 1993–1996. He is also an Adjunct Professor of Radiology at Upstate Medical University, Syracuse. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, and remote sensing. He has published extensively. He is the author of *Distributed Detection and Data Fusion* (New York: Springer-Verlag, 1997). He has served as a consultant to several major companies.

Dr. Varshney was a James Scholar, a Bronze Tablet Senior, and a Fellow, all while he was with the University of Illinois. He is a member of Tau Beta Pi and is the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. He was elected to the grade of Fellow of the IEEE in 1997 for his contributions in the area of distributed detection and data fusion. He was the Guest Editor of the Special Issue on Data Fusion of the PROCEEDINGS OF THE IEEE, January 1997. In 2000, he received the Third Millennium Medal from the IEEE and Chancellor's Citation for exceptional academic achievement at Syracuse University. He serves as a distinguished lecturer for the AES Society of the IEEE. He is on the editorial board of *Information Fusion*. He was the President of International Society of Information Fusion during 2001.

Steven M. Kay (M'75–S'76–M'78–SM'83–F'89) was born in Newark, NJ, on April 5, 1951. He received the B.E. degree from Stevens Institute of Technology, Hoboken, NJ, in 1972, the M.S. degree from Columbia University, New York, NY, in 1973, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, in 1980, all in electrical engineering.

From 1972 to 1975, he was with Bell Laboratories, Holmdel, NJ, where he was involved with transmission planning for speech communications and simulation and subjective testing of speech processing algorithms. From 1975 to 1977, he attended the Georgia Institute of Technology to study communication theory and digital signal processing. From 1977 to 1980, he was with the Submarine Signal Division, Portsmouth, RI, where he engaged in research on autoregressive spectral estimation and the design of sonar systems. He is presently a Professor of Electrical Engineering with the University of Rhode Island, Kingston, and a consultant to numerous industrial concerns, the Air Force, the Army, and the Navy. As a leading expert in statistical signal processing, he has been invited to teach short courses to scientists and engineers at government laboratories, including NASA and the CIA. He has written numerous journal and conference papers and is a contributor to several edited books. He is the author of the textbooks *Modern Spectral Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 1988), *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1993), *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1998), and *Intuitive Probability and Random Processes using MATLAB* (New York: Springer, 2005). His current interests are spectrum analysis, detection and estimation theory, and statistical signal processing.

Dr. Kay is a member of Tau Beta Pi and Sigma Xi. He has been a distinguished lecturer for the IEEE Signal Processing Society. He has served on the IEEE Acoustics, Speech, and Signal Processing Committee on Spectral Estimation and Modeling, on the IEEE Oceans committees, has been an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has received the IEEE Signal Processing Society Education Award "for outstanding contributions in education and in writing scholarly books and texts. . .." He has recently been included on a list of the 250 most-cited researchers in the world in engineering.

James H. Michels (M'82–SM'90–F'05) received the M.S. degrees from Syracuse University, Syracuse, NY, in both physics and electrical engineering, and the Ph.D. degree in electrical engineering from the same university in 1991. His Ph.D. thesis in multichannel signal processing provided the foundation for in-house and contractual research programs in his former position with the Air Force Research Laboratory (AFRL) in Rome, NY.

Currently, he is the founder and Senior Research Engineer of JHM Technologies, Ithaca, NY. He also serves as a volunteer Research Professor and Adjunct Faculty member with Syracuse University, where he assists a graduate student research group. His current research interests include detection, estimation, multichannel adaptive signal processing, multi- and hyperspectral image processing, time series analyses, space-time adaptive processing (STAP), change detection, particle filtering, and stochastic resonance. He has more than fifty papers and six patents in these technical areas.