

# Optimal Noise Benefits in Neyman–Pearson and Inequality-Constrained Statistical Signal Detection

Ashok Patel and Bart Kosko

**Abstract**—We present theorems and an algorithm to find optimal or near-optimal “stochastic resonance” (SR) noise benefits for Neyman–Pearson hypothesis testing and for more general inequality-constrained signal detection problems. The optimal SR noise distribution is just the randomization of two noise realizations when the optimal noise exists for a single inequality constraint on the average cost. The theorems give necessary and sufficient conditions for the existence of such optimal SR noise in inequality-constrained signal detectors. There exists a sequence of noise variables whose detection performance limit is optimal when such noise does not exist. Another theorem gives sufficient conditions for SR noise benefits in Neyman–Pearson and other signal detection problems with inequality cost constraints. An upper bound limits the number of iterations that the algorithm requires to find near-optimal noise. The appendix presents the proofs of the main results.

**Index Terms**—Inequality-constrained signal detection, Neyman–Pearson test, noise-finding algorithm, optimal noise, stochastic resonance.

## I. NOISE BENEFITS IN SIGNAL DETECTION

STOCHASTIC resonance (SR) occurs when noise benefits a nonlinear system [1]–[14]. SR noise benefits occur in a wide range of applications in physics, biology, and medicine [15]–[26]. The noise benefit can take the form of an increase in an entropy-based bit count [27]–[29], a signal-to-noise ratio [30]–[32], a cross-correlation [3], [30], or a detection probability for a preset level of false-alarm probability [33], [34], or a decrease in the error probability [35] or in the average sample number of sequential detection problems [36]. An SR noise benefit requires some form of nonlinear signal detection [10]. Its signature often takes the form of an inverted-U curve or a non-monotonic plot of a bit count or SNR against the variance or dispersion of the noise process.

We focus first on the special case of SR in signal detection that uses Neyman–Pearson (N-P) hypothesis testing [37] to decide between two simple alternatives. We define the noise as *N-P SR noise* if adding such noise to the received signal before making a decision increases the signal detection probability  $P_D$  while the false-alarm probability  $P_{FA}$  stays at or below a preset level  $\alpha$  for a given detection strategy. Fig. 1 shows this type of noise benefit for a suboptimal receiver and does not involve the typical

inverted-U curve of SR (but would if it used uniform noise and we plotted the detection probability against the noise variance).

An SR noise benefit does *not* occur in an optimal receiver if the noise is independent of the concurrent received signal and the hypotheses. This follows from the so-called irrelevance theorem of optimal detection [38], [39]. But Section V shows that SR noise benefits can occur even if the receiver is optimal when the noise depends on the received signal. Fig. 3 shows such an SR noise benefit in optimal antipodal signal detection when the average signal power constrains the signal transmission.

Sections II and III present three SR results for Neyman–Pearson signal detection. The first SR result gives necessary and sufficient conditions for the existence of optimal N-P SR noise. The existence of some N-P SR noise does not itself imply the existence of optimal noise. But there exists a sequence of noise variables whose detection performance limit is optimal when the optimal N-P SR noise does not exist. The second SR result is a sufficient condition for SR noise benefits in N-P signal detection. The third SR result is an algorithm that finds near-optimal N-P SR noise from a finite set  $\tilde{\mathcal{N}}$  of noise realizations. This noise is nearly optimal if the detection and false alarm probabilities in  $\tilde{\mathcal{N}}$  and in the actual noise space  $\mathcal{N} \supset \tilde{\mathcal{N}}$  are sufficiently close. An upper bound limits the number of iterations that the algorithm needs to find near-optimal noise. Section IV extends these results to more general statistical decision problems that have one inequality constraint on the average cost.

These SR results extend and correct prior work in “detector randomization” or adding noise for improving the performance of N-P signal detection. Tsitsiklis [40] explored the mechanism of detection-strategy randomization for a finite set of detection strategies (operating points) in decentralized detection. He showed that there exists a randomized detection strategy that uses a convex or random combination of at most two existing detection strategies and that gives the optimal N-P detection performance. Such optimal detection strategies lie on the upper boundary of the convex hull of the receiver operating characteristic (ROC) curve. Scott [41] later used the same optimization principle in classification systems while Appadwedula [42] used it for energy-efficient detection in sensor networks. Then Chen *et al.* [33] used a fixed detector structure: they injected independent noise in the received signal to obtain a proper random combination of operating points on the ROC curve for a given suboptimal detector. They showed that the optimal N-P SR noise for suboptimal detectors randomizes no more than two noise realizations.

But Chen *et al.* [33] assumed that the convex hull  $V$  of the set of ROC curve operating points  $U \subseteq \mathbf{R}^2$  always contains its boundary  $\partial V$  and thus that the convex hull  $V$  is closed. This is not true in general. The topological problem is that the convex hull  $V$  need not be closed if  $U$  is not compact [43]: the convex

Manuscript received June 10, 2008; accepted December 10, 2008. First published January 13, 2009; current version published April 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Andreas Jakobsson.

The authors are with the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: kosko@usc.edu; ashokpat@usc.edu).

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

Digital Object Identifier 10.1109/TSP.2009.2012893

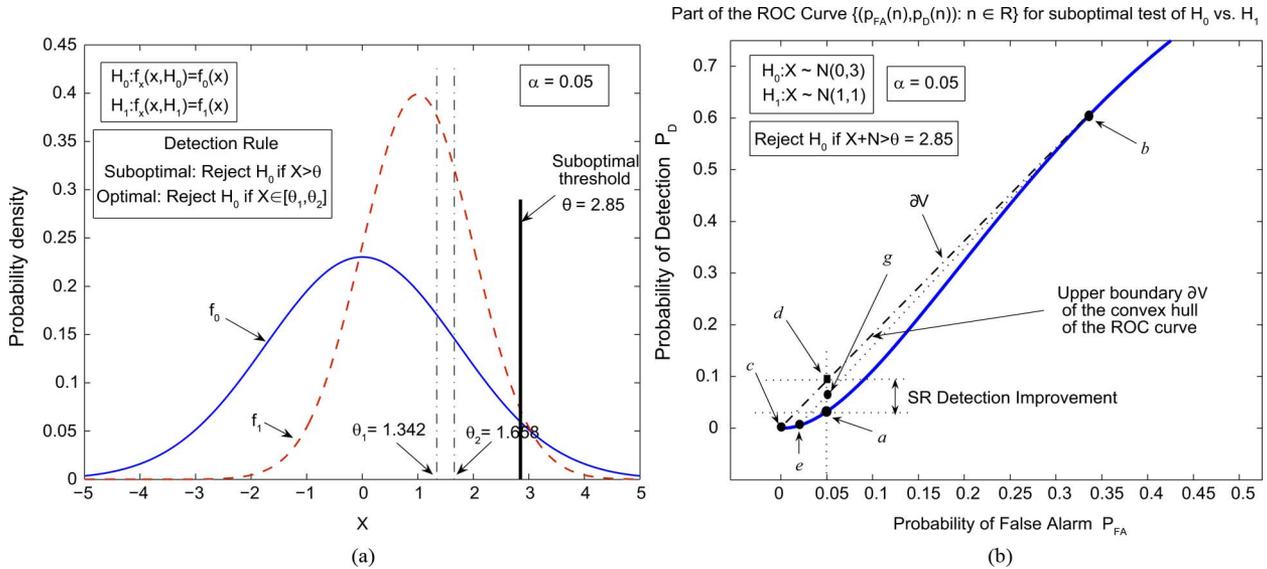


Fig. 1. SR noise benefits in suboptimal Neyman–Pearson signal detection. (a) The thin solid line shows the probability density function (pdf)  $f_0$  of signal  $X$  under the normal hypothesis  $H_0 : X \sim N(0, 3)$  while the dashed line shows the pdf  $f_1$  of  $X$  under the alternative normal hypothesis  $H_1 : X \sim N(1, 1)$ . The one-sample optimal detection scheme requires the two thresholds  $\theta_1 = 1.342$  and  $\theta_2 = 1.658$ . It rejects  $H_0$  if  $X \in [\theta_1, \theta_2]$  for the test significance or preset false-alarm probability  $\alpha = 0.05$ . The thick vertical solid line shows the best *single* threshold  $\theta = 2.85$  if the detector cannot use two thresholds due to resource limits or some design constraint. The respective suboptimal detector rejects  $H_0$  if  $X > \theta$ . (b) The solid line shows the monotonic but nonconcave ROC curve  $U = \{(p_{FA}(n), p_D(n)) : n \in \mathbf{R}\}$  of the suboptimal detector for different realizations  $n$  of the additive noise  $N$  in  $X$ . Here  $p_{FA}(n) = 1 - \Phi(\theta - n/\sqrt{3})$  and  $p_D(n) = 1 - \Phi(\theta - 1 - n)$  for the standard normal cumulative distribution function  $\Phi$ . The detector operates at point  $a = (p_{FA}(0), p_D(0)) = (0.05, 0.0322)$  on the ROC curve in the absence of noise. Nonconcavity of the ROC curve  $U$  between the points  $b = (p_{FA}(n_2), p_D(n_2))$  and  $c = (0, 0)$  allows the N-P SR effect to occur. A convex or random combination of two operating points  $b$  and  $e = (p_{FA}(n_1), p_D(n_1))$  gives a better detection performance (point  $g$ ) than point  $a$  at the same false-alarm level  $p_{FA}(0) = \alpha = 0.05$ . Such a random combination of operating points results from adding an independent discrete noise  $N$  with pdf  $f_N(n) = \lambda\delta(n - n_1) + (1 - \lambda)\delta(n - n_2)$  to the random sample  $X$  where  $\lambda = (p_{FA}(n_2) - \alpha)/(p_{FA}(n_2) - p_{FA}(n_1))$ . Point  $d$  is on the upper boundary  $\partial V$  of the ROC curve's convex hull (dashed tangent line between  $b$  and  $c$ ). So  $d$  is the supremum of detection performances that the random or convex combination of operating points on the ROC can achieve such that  $\alpha$  stays at 0.05. Point  $d$  is the convex combination of  $b$  and  $c$  but it is not realizable by adding only noise to the sample  $X$  because point  $c = (0, 0)$  is not on the ROC curve since there is no noise realization  $n \in \mathbf{R}$  such that  $1 - \Phi(\theta - n/\sqrt{3}) = 0 = 1 - \Phi(\theta - 1 - n)$ . So the N-P SR noise exists but the *optimal* N-P SR noise does not exist in the noise space  $\mathcal{N} = \mathbf{R}$ .

hull of  $U$  is open if  $U$  itself is open [44]. Chen *et al.* argued correctly in their proof of Theorem 3 in [33] when they concluded that the “optimum pair can only exist on the boundary.” But their later claim that “each  $z$  on the boundary can be expressed as the convex combination of only two elements of  $U$ ” is not true in general because  $V$  may not include *all* of its boundary points. The optimal N-P SR noise need not exist at all in a fixed detector [45].

Fig. 1 shows a case where the N-P SR noise exists but where the optimal N-P SR noise does not exist in the noise space  $\mathcal{N} = \mathbf{R}$ . Section V shows that the optimal SR noise does exist if we restrict the noise space to a compact interval such as  $[-3, 3]$ . The algorithm finds nearly optimal N-P SR noise realizations from a discretized set of noise realizations  $\hat{\mathcal{N}} = [-3 : 0.0001 : 3]$  in 17 iterations. Section V also shows that the detection performance of the maximum *a posteriori* (MAP) receiver can sometimes benefit from signal-power randomization in an average-power-constrained antipodal signal transmitter if the channel noise pdf is multimodal. We assume that the transmitter transmits equiprobable antipodal signals  $\{-S, S\}$  with  $S \in \mathcal{S} = [0.5, 3.75]$  and that the additive channel noise has a symmetric bimodal Gaussian-mixture probability density. Then the respective error probability of the optimal MAP receiver is nonconvex. So the transmitter can improve the detection performance by time-sharing or randomizing between two power levels for some values of the constraining maximum average power. The algorithm finds a near-optimal signal power randomization from a discretized subset of signal-strength realizations  $\hat{\mathcal{S}} = [0.5 : 0.0001 : 3.75]$ . The algorithm finds this

signal-power distribution in just 13 iterations. The next four sections present and illustrate the formal SR results. All examples use single samples in statistical decision making but the results still hold for multiple samples.

## II. OPTIMAL NOISE DENSITIES FOR NEYMAN–PEARSON SIGNAL DETECTION

We now derive two theorems that fully characterize the optimal noise probability densities for Neyman–Pearson signal detection. Then we give a sufficient condition for SR noise benefits in N-P signal detection.

Consider a binary hypothesis test where we decide between  $H_0 : f_{\mathbf{X}}(\mathbf{x}; H_0) = f_0(\mathbf{x})$  and  $H_1 : f_{\mathbf{X}}(\mathbf{x}; H_1) = f_1(\mathbf{x})$  using an  $m$ -dimensional noisy observation vector  $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ . Here  $\mathbf{X} \in \mathbf{R}^m$  is the received signal vector,  $\mathbf{N} \in \mathcal{N} \subseteq \mathbf{R}^m$  is a noise vector with pdf  $f_N$ , and  $\mathcal{N}$  is the noise space. The noise vector  $\mathbf{N}$  can be random or a deterministic constant such as when  $f_N(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$ . Here  $f_0$  and  $f_1$  are the pdfs of the signal  $\mathbf{X}$  under the hypotheses  $H_0$  and  $H_1$ . We need not know the prior probabilities  $P(H_i)$  of the hypotheses  $H_i$ . Then we want to determine when the optimal additive noise  $\mathbf{N}_{\text{opt}}$  exists and gives the best achievable detection performance at the significance level  $\alpha$  for the given detection strategy.

Define  $P_D(\mathbf{n})$  and  $P_{FA}(\mathbf{n})$  as the respective probabilities of correct detection and false detection (alarm) when the noise realization is  $\mathbf{n}$ . Define  $P_D(f_N) = \int_{\mathcal{N}} P_D(\mathbf{n})f_N(\mathbf{n})d\mathbf{n}$  and  $P_{FA}(f_N) = \int_{\mathcal{N}} P_{FA}(\mathbf{n})f_N(\mathbf{n})d\mathbf{n}$  as the respective probabilities of detection and false alarm when the noise pdf is  $f_N$ . Let  $f_{N_{\text{opt}}}$  be the pdf of the optimal SR noise  $\mathbf{N}_{\text{opt}}$  that we add to the

received signal  $\mathbf{X}$  to maximize the probability of detection  $P_D$  while keeping  $P_{FA} \leq \alpha$ . Let  $\mathcal{F}$  denote the set of all probability density functions on  $\mathbf{R}$ . So we need to find

$$f_{N_{opt}} = \arg \max_{f_N \in \mathcal{F}} \int_{\mathcal{N}} P_D(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} \quad (1)$$

such that

$$f_{N_{opt}}(\mathbf{n}) \geq 0 \quad \text{for all } \mathbf{n} \quad (2)$$

$$\int_{\mathcal{N}} f_{N_{opt}}(\mathbf{n}) d\mathbf{n} = 1 \quad \text{and} \quad (3)$$

$$P_{FA}(f_{N_{opt}}) = \int_{\mathcal{N}} P_{FA}(\mathbf{n}) f_{N_{opt}}(\mathbf{n}) d\mathbf{n} \leq \alpha. \quad (4)$$

Conditions (2) and (3) are general pdf properties while (1) and (4) state the Neyman–Pearson criteria for the optimal SR noise pdf  $f_{N_{opt}}$ . The conditional noise pdf  $f(\mathbf{n}|\mathbf{x}, H_i)$  obeys  $f(\mathbf{n}|\mathbf{x}, H_i) = f_N(\mathbf{n})$  if the noise random variable  $\mathbf{N}$  is independent of the concurrent received signal random variable  $\mathbf{X}$  and the hypotheses  $H_i$ . Then the irrelevance theorem [38], [39] implies that the optimal likelihood-ratio test based on the received signal realization  $\mathbf{x}$  and noise realization  $\mathbf{n}$  is the same as the optimal likelihood-ratio test based on only  $\mathbf{x}$ . So the optimal detector can always ignore the noise realization  $\mathbf{n}$  without affecting its detection performance. This implies that an N-P SR noise benefit will not occur for the noise  $\mathbf{N}$  if the receiver uses the optimal likelihood-ratio test based on  $\mathbf{x}$ . But computing optimal likelihood-ratio thresholds is not simple for many non-Gaussian noise types [46], [47]. We may also need multiple thresholds to partition the test-statistic sample space into acceptance and rejection regions if the likelihood ratio is not a monotone function of the test statistic. So some detection systems use suboptimal tests if they have special hardware resource limits or if they constrain the number of detection thresholds [48], [49]. Then SR noise benefits may occur and so we may need to compute the optimal or near-optimal SR noise pdf.

The primal-dual method [50], [51] directly solves the above optimization problem (1)–(4) in the noise domain  $\mathbf{R}^m$ . This approach gives both the conditions for the existence of optimal SR noise and the exact form of the optimal noise pdf. It also leads to an algorithm that can find near-optimal SR noise if the detection probability  $P_D$  on the noise space  $\mathcal{N}$  is sufficiently close to its restriction to the discrete noise realizations  $\tilde{\mathcal{N}} \subset \mathcal{N}$ .

The Lagrangian of the inequality-constrained optimization problem (1)–(4) is

$$\begin{aligned} L(f_N, k) &= \int_{\mathcal{N}} P_D(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} \\ &\quad - k \left( \int_{\mathcal{N}} P_{FA}(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} - \alpha \right) \quad (5) \\ &= \int_{\mathcal{N}} (P_D(\mathbf{n}) - k(P_{FA}(\mathbf{n}) - \alpha)) f_N(\mathbf{n}) d\mathbf{n}. \quad (6) \end{aligned}$$

The Lagrange duality [50], [51] implies that

$$\sup_{f_N \in \mathcal{F}} \int_{\mathcal{N}} P_D(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} = \min_{k \geq 0} \sup_{f_N \in \mathcal{F}} L(f_N, k). \quad (7)$$

So solving the optimization problem equates to finding  $k^* \geq 0$  and the pdf  $f_{N_{opt}}$  such that

$$\min_{k \geq 0} \sup_{f_N \in \mathcal{F}} L(f_N, k) = L(f_{N_{opt}}, k^*). \quad (8)$$

The next two theorems give necessary and sufficient conditions for the existence of the optimal N-P SR noise and for the form of its pdf if it exists. Define first the sets

$$D^+ = \{\mathbf{n} \in \mathcal{N} : (P_{FA}(\mathbf{n}) - \alpha) \geq 0\} \quad (9)$$

and

$$D^- = \{\mathbf{n} \in \mathcal{N} : (P_{FA}(\mathbf{n}) - \alpha) \leq 0\}. \quad (10)$$

Assume  $D^- \neq \emptyset$  so that  $N_{opt}$  always exists. Let  $P_{D^+_{sup}}$ ,  $P_{D^-_{sup}}$ , and  $P_{D_{sup}}$  be the respective suprema of  $P_D(\mathbf{n})$  over the sets  $D^+$ ,  $D^-$ , and  $\mathcal{N}$ :

$$P_{D^+_{sup}} = \sup_{\mathbf{n}} \{P_D(\mathbf{n}) : \mathbf{n} \in D^+\} \quad (11)$$

$$P_{D^-_{sup}} = \sup_{\mathbf{n}} \{P_D(\mathbf{n}) : \mathbf{n} \in D^-\} \quad (12)$$

and

$$P_{D_{sup}} = \sup_{\mathbf{n}} \{P_D(\mathbf{n}) : \mathbf{n} \in \mathcal{N}\}. \quad (13)$$

Define

$$g(\mathbf{n}, k) = P_D(\mathbf{n}) - k(P_{FA}(\mathbf{n}) - \alpha) \quad (14)$$

and let  $d^+(k)$ ,  $d^-(k)$ , and  $d(k)$  be the respective suprema of  $g$  over the sets  $D^+$ ,  $D^-$ , and  $\mathcal{N}$ :

$$d^+(k) = \sup_{\mathbf{n}} \{g(\mathbf{n}, k) : \mathbf{n} \in D^+\} \quad (15)$$

$$d^-(k) = \sup_{\mathbf{n}} \{g(\mathbf{n}, k) : \mathbf{n} \in D^-\} \quad (16)$$

and

$$d(k) = \sup_{\mathbf{n}} \{g(\mathbf{n}, k) : \mathbf{n} \in \mathcal{N}\}. \quad (17)$$

Define

$$G^+ = \{\mathbf{n} \in D^+ : P_D(\mathbf{n}) = P_{D^+_{sup}}\} \quad (18)$$

and

$$G^- = \{\mathbf{n} \in D^- : P_D(\mathbf{n}) = P_{D^-_{sup}}\}. \quad (19)$$

Rewrite the Lagrangian (6) as

$$L(f_N, k) = \int_{\mathcal{N}} g(\mathbf{n}, k) f_N(\mathbf{n}) d\mathbf{n}. \quad (20)$$

Then (8) becomes

$$\min_{k \geq 0} \sup_{f_N \in \mathcal{F}} L(f_N, k) = \min_{k \geq 0} \sup_{f_N \in \mathcal{F}} \int_{\mathcal{N}} g(\mathbf{n}, k) f_N(\mathbf{n}) d\mathbf{n} \quad (21)$$

$$= \min_{k \geq 0} d(k). \quad (22)$$

Equality (22) follows because

$$\sup_{f_N \in \mathcal{F}} \int_{\mathcal{N}} g(\mathbf{n}, k) f_N(\mathbf{n}) d\mathbf{n} \leq \sup_{f_N \in \mathcal{F}} \int_{\mathcal{N}} d(k) f_N(\mathbf{n}) d\mathbf{n} \quad (23)$$

$$= d(k) \quad (24)$$

$$= \sup_{\mathbf{n}} \{g(\mathbf{n}, k) : \mathbf{n} \in \mathcal{N}\} \quad (25)$$

and because strict inequality in (23) implies that there exists an  $\mathbf{n}_1 \in \mathcal{N}$  such that

$$\begin{aligned} \sup_{f_N \in \mathcal{F}} \int_{\mathcal{N}} g(\mathbf{n}, k) f_N(\mathbf{n}) d\mathbf{n} &< g(\mathbf{n}_1, k) \\ &\text{because of (24)–(25)} \\ &= \int_{\mathcal{N}} g(\mathbf{n}, k) \delta(\mathbf{n} - \mathbf{n}_1) d\mathbf{n}. \end{aligned} \quad (26)$$

But this is a contradiction because the supremum of a set of numbers cannot be less than any one of those numbers. The definition (17) implies that  $d(k) = \max\{d^-(k), d^+(k)\}$ . So (21)–(22) reduces to

$$\min_{k \geq 0} \sup_{f_N \in \mathcal{F}} L(f_N, k) = \min_{k \geq 0} \max\{d^-(k), d^+(k)\}. \quad (27)$$

Theorem 1a) below gives necessary and sufficient conditions for the existence of optimal SR noise. It also gives the exact form of the optimal N-P SR noise pdf  $f_{N_{\text{opt}}}$  if it exists when  $P_{D^{\text{sup}}} \geq P_{D^{\text{inf}}}$ . Theorem 2 likewise gives necessary and sufficient conditions for the existence of  $f_{N_{\text{opt}}}$  when  $P_{D^{\text{sup}}} < P_{D^{\text{inf}}}$ . The proofs are in the Appendix.

*Theorem 1:*

**a)** Suppose that  $P_{D^{\text{sup}}} \geq P_{D^{\text{inf}}}$  and that  $G^- \neq \emptyset$ . Then

$$f_{N_{\text{opt}}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o) \quad (28)$$

is an optimal SR noise pdf for Neyman–Pearson detection for some  $\mathbf{n}_o \in G^-$  and  $P_{\text{FA}}(f_{N_{\text{opt}}}) \leq \alpha$ . The Neyman–Pearson optimal SR noise does not exist for the given test level  $\alpha$  if  $G^- = \emptyset$ . But there exists a noise pdf sequence  $\{f_{N_r}\}_{r=1}^{\infty}$  of the form (28) such that  $P_{\text{FA}}(f_{N_r}) \leq \alpha$  for all  $r$  and such that

$$\lim_{r \rightarrow \infty} P_D(f_{N_r}) = P_{D^{\text{sup}}}. \quad (29)$$

**b)** Suppose that  $P_{D^{\text{sup}}} < P_{D^{\text{inf}}}$ . Then  $P_{\text{FA}}(f_{N_{\text{opt}}}) = \alpha$  if the Neyman–Pearson optimal SR noise pdf  $f_{N_{\text{opt}}}(\mathbf{n})$  exists. The optimal noise pdf in Theorem 1a) is not unique if  $G^-$  contains more than one noise realization.

*Theorem 2:*

Suppose that  $P_{D^{\text{sup}}} < P_{D^{\text{inf}}}$ . Then a)–d) hold:

**a)** There exists  $k^* \geq 0$  such that  $d^+(k^*) = d^-(k^*) = d(k^*)$  and  $\min\{d^+(k), d^-(k)\} \leq d(k^*) \leq \max\{d^+(k), d^-(k)\}$  for any  $k \geq 0$ .

**b)** Suppose the noise pdf  $f_N$  satisfies  $P_D(f_N) = d(k^*) > P_D(0)$  and  $P_{\text{FA}}(f_N) = \alpha$ . Then  $f_N$  is a Neyman–Pearson optimal noise pdf. So  $d(k^*)$  is the optimal N-P SR detection probability  $P_{D_{\text{opt}}}$ .

**c)** Suppose that there exist  $\mathbf{n}_1 \in D^-$  and  $\mathbf{n}_2 \in D^+$  such that  $g(\mathbf{n}_1, k^*) = d^-(k^*) = d(k^*) = g(\mathbf{n}_2, k^*) = d^+(k^*)$ . Then

$$f_{N_{\text{opt}}}(\mathbf{n}) = \lambda \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda) \delta(\mathbf{n} - \mathbf{n}_2) \quad (30)$$

is the optimal Neyman–Pearson SR noise pdf if  $d(k^*) > P_D(0)$  and if

$$\lambda = \frac{P_{\text{FA}}(\mathbf{n}_2) - \alpha}{P_{\text{FA}}(\mathbf{n}_2) - P_{\text{FA}}(\mathbf{n}_1)}. \quad (31)$$

**d)** Neyman–Pearson optimal SR noise does not exist if c) does not hold. But there does exist a noise pdf sequence  $\{f_{N_r}\}_{r=1}^{\infty}$  of the form (30)–(31) such that

$$\lim_{r \rightarrow \infty} P_D(f_{N_r}) = d(k^*). \quad (32)$$

The optimal noise pdf is not unique if more than one pair of noise realizations satisfy Theorem 2c).

The optimal randomization in (30)–(31) of two noise realizations resembles the optimal “minimax” randomization of pure strategies in finite zero-sum games [52]–[54]. But the noise result differs because the minimax optimization does not impose inequality constraints on the expected payoff.

Theorem 2 also implies the following necessary conditions for the optimal N-P SR noise.

*Corollary 1:*

Suppose that  $P_{D^{\text{sup}}} < P_{D^{\text{inf}}}$  and that  $P_D$  and  $P_{\text{FA}}$  are differentiable in the interior of the noise space  $\mathcal{N}$ .

**a)** Suppose that  $f_{N_{\text{opt}}}$  is an optimal N-P SR noise pdf of the form (30)–(31) in Theorem 2c) and that  $\mathbf{n}_1$  and  $\mathbf{n}_2$  of (30)–(31) are the interior points of  $\mathcal{N}$ . Then  $\mathbf{n}_1$  and  $\mathbf{n}_2$  satisfy

$$P_D(\mathbf{n}_1) - k P_{\text{FA}}(\mathbf{n}_1) = P_D(\mathbf{n}_2) - k P_{\text{FA}}(\mathbf{n}_2) \quad (33)$$

$$\nabla P_D(\mathbf{n}_1) - k \nabla P_{\text{FA}}(\mathbf{n}_1) = \mathbf{0} \quad (34)$$

$$\nabla P_D(\mathbf{n}_2) - k \nabla P_{\text{FA}}(\mathbf{n}_2) = \mathbf{0} \quad (35)$$

for some  $k \geq 0$ .

**b)** Suppose further that for each  $k \geq 0$  at most one solution of  $\nabla P_D(\mathbf{n}) - k \nabla P_{\text{FA}}(\mathbf{n}) = \mathbf{0}$  in  $\mathbf{R}^m$  is a global maximum of  $P_D(\mathbf{n}) - k(P_{\text{FA}}(\mathbf{n}) - \alpha)$ . Then  $f_{N_{\text{opt}}} = \delta(\mathbf{n} - \mathbf{n}_\alpha)$  is the optimal N-P SR noise pdf if such a solution  $\mathbf{n}_\alpha$  exists in  $D^0$ . There is otherwise no optimal N-P SR noise in the interior of  $\mathcal{N}$ .

Equalities (33)–(35) are necessary but not sufficient because the noise realizations  $n_1$  and  $n_2$  that satisfy (34)–(35) need not be global maxima. They need not be in  $D^+$  and  $D^-$  even if they are global maxima. So (33)–(35) may not help find the optimal SR noise. But Corollary 1b) shows that these necessary conditions can suggest when optimal SR noise does not exist. Section V-A applies Corollary 1b) to a hypothesis test between two Gaussian densities.

Theorem 3 gives two sufficient conditions to detect an N-P SR noise benefit in detectors that use a *single* noisy observation  $Y \in \mathbf{R}$  to decide between  $H_0$  and  $H_1$ .

*Theorem 3:*

Let the detection and false-alarm probabilities  $P_D$  and  $P_{\text{FA}}$  be real-valued functions that are differentiable in a neighborhood of 0. Suppose that  $P_D''(0)$  and  $P_{\text{FA}}''(0)$  exist and that  $P_{\text{FA}}(0) = \alpha$ . Suppose also that  $P_{\text{FA}}$  does not have a local minimum at 0

and that  $P_D$  does not have a local maximum at 0. Then an N-P SR noise exists if

$$P''_D(0)|P'_{FA}(0)| > P''_{FA}(0)|P'_D(0)| \quad (36)$$

or if  $\text{sgn}(P'_{FA}(0))\text{sgn}(P'_D(0)) \leq 0$ .

Theorem 3 implies the following corollary. It gives a sufficient condition for a SR noise benefit in N-P signal detectors if they partition the real line  $\mathbf{R}$  into acceptance and rejection intervals and if they use a single noisy observation  $Y \in \mathbf{R}$  to decide between  $H_0$  and  $H_1$ .

*Corollary 2:*

Suppose that the thresholds  $\Theta = \{\theta_1, \dots, \theta_k\}$  partition the real line  $\mathbf{R}$  into acceptance and rejection regions and that  $P_{FA}(0) = \alpha$ . Suppose also that the hypothesized pdfs  $f_i$  are differentiable at all the thresholds in  $\Theta$ . Define  $s(j) = 1$  if  $\theta_j$  is a left endpoint of any rejection interval. Else let  $s(j) = -1$ . Then additive noise can improve the N-P detection of such a detector if  $P_{FA}$  does not have a local minimum at 0, if  $P_D$  does not have local maximum at 0, and if

$$\left( \sum_j s(j)f'_0(\theta_j) \right) \left| \left( \sum_j s(j)f_1(\theta_j) \right) \right| > \left( \sum_j s(j)f'_1(\theta_j) \right) \left| \left( \sum_j s(j)f_0(\theta_j) \right) \right| \quad (37)$$

or  $\text{sgn}(\sum_j s(j)f_0(\theta_j))\text{sgn}(\sum_j s(j)f_1(\theta_j)) \leq 0$ .

The inequality (37) holds because the hypotheses of Corollary 2 imply that

$$P''_D(0)|P'_{FA}(0)| = - \left( \sum_j s(j)f'_1(\theta_j) \right) \left| \left( \sum_j s(j)f_0(\theta_j) \right) \right|$$

and

$$P''_{FA}(0)|P'_D(0)| = - \left( \sum_j s(j)f'_0(\theta_j) \right) \left| \left( \sum_j s(j)f_1(\theta_j) \right) \right|$$

in (36).

### III. N-P SR NOISE FINDING ALGORITHM

This section presents an algorithm for finding a near-optimal SR noise density. Theorems 1 and 2 give the exact form for the optimal SR noise pdf. But such a noise pdf may not be easy to find in a given noise space  $\mathcal{N}$ . So we present an algorithm that uses Theorems 1 and 2 and successive approximations to find a near-optimal SR noise from a finite set of noise realizations  $\tilde{\mathcal{N}} \subset \mathcal{N}$ .

The algorithm takes as input  $\epsilon$ ,  $\alpha$ ,  $\tilde{\mathcal{N}}$  in (9)–(19), and the respective detection and false alarm probabilities  $P_D$  and  $P_{FA}$  in  $\tilde{\mathcal{N}}$ . The algorithm first searches for a constant noise from the set  $G^-$  if the inequality  $P_{D_{\text{sup}}} \geq P_{D_{\text{sup}}}$  holds. The algorithm otherwise finds a number  $k(i)$  at each iteration  $i$  such that  $|d^-(k(i)) - d(k^*)| \leq 2^{-i+1}$  and this  $k(i)$  gives  $|d^+(k(i)) - d^-(k(i))| < \epsilon$  in at most  $i_{\text{max}} = \lceil \log_2(2/\epsilon) \rceil + 1$  iterations. Then the algorithm defines the noise  $\tilde{N}'$  as the random combination of  $\tilde{n}_1 \in D^-$  and  $\tilde{n}_2 \in D^+$  so that  $g(\tilde{n}_1, k(i_{\text{max}})) =$

$d^-(k(i_{\text{max}})), g(\tilde{n}_2, k(i_{\text{max}})) = d^+(k(i_{\text{max}}))$ , and  $P_{FA}(f_{\tilde{N}'}) = \alpha$ .

Theorem 4a) below shows that for all  $\epsilon > 0$  the algorithm finds an SR noise  $\tilde{N}'$  from  $\tilde{\mathcal{N}}$  in at most  $i_{\text{max}} = \lceil \log_2(2/\epsilon) \rceil + 1$  iterations such that  $0 \leq P_D(f_{\tilde{N}_{\text{opt}}}) - P_D(f_{\tilde{N}'}) \leq \epsilon$ . Here  $\tilde{N}_{\text{opt}}$  is the optimal N-P SR noise in  $\tilde{\mathcal{N}}$  and  $f_{\tilde{N}_{\text{opt}}}$  is the pdf of  $\tilde{N}_{\text{opt}}$ . Theorem 4b) shows that  $0 \leq P_D(f_{N_{\text{opt}}}) - P_D(f_{\tilde{N}'}) \leq (\tau + \epsilon)$  if for each  $n \in \mathcal{N}$  there exists an  $\tilde{n} \in \tilde{\mathcal{N}}$  so that

$$|P_D(n) - P_D(\tilde{n})| \leq \tau \quad (38)$$

and

$$P_{FA}(\tilde{n}) \leq P_{FA}(n) \quad (39)$$

and if  $N_{\text{opt}}$  is the optimal N-P SR noise in  $\mathcal{N}$  with pdf  $f_{N_{\text{opt}}}$ . Thus the algorithm will find a near-optimal noise  $\tilde{N}'$  for any small positive  $\epsilon$  if we choose  $\tilde{\mathcal{N}}$  such that  $\tau$  is sufficiently small.

#### SR Noise-Finding Algorithm

Let  $D^+ = \{\tilde{n} \in \tilde{\mathcal{N}} : (P_{FA}(\tilde{n}) - \alpha) \geq 0\}$

Let  $D^- = \{\tilde{n} \in \tilde{\mathcal{N}} : (P_{FA}(\tilde{n}) - \alpha) \leq 0\}$

Let  $P_{D_{\text{sup}}} = \max\{P_D(\tilde{n}) : \tilde{n} \in D^+\}$

Let  $P_{D_{\text{sup}}} = \max\{P_D(\tilde{n}) : \tilde{n} \in D^-\}$

Let  $G^- = \{\tilde{n} \in D^- : P_D(\tilde{n}) = P_{D_{\text{sup}}}\}$

If  $P_{D_{\text{sup}}} \geq P_{D_{\text{sup}}}$   
 $f_{\tilde{N}_{\text{opt}}}(n) = \delta(n - \tilde{n}_0)$  for any  $\tilde{n}_0 \in G^-$

**Else**

Let  $D^0 = \{\tilde{n} \in \tilde{\mathcal{N}} : (P_{FA}(\tilde{n}) - \alpha) = 0\}$  and  $k(0) = 1$

Let  $d^-(k(0)) = \max\{P_D(\tilde{n}) - (P_{FA}(\tilde{n}) - \alpha) : \tilde{n} \in D^-\}$

Let  $d^+(k(0)) = \max\{P_D(\tilde{n}) - (P_{FA}(\tilde{n}) - \alpha) : \tilde{n} \in D^+\}$

Let  $ds(1) = d^-(k(0))$  and  $df(1) = d^+(k(0))$

Let  $i = 1$  and  $i_{\text{stop}} = \lceil \log_2(\frac{2}{\epsilon}) \rceil$

**While**  $|d^-(k(i)) - d^+(k(i))| > \epsilon$  and  $i \leq i_{\text{stop}}$

Let  $dr(i) = (ds(i) + df(i))/2$

Let  $k(i) = \min\{(P_D(\tilde{n}) - dr(i))/(P_{FA}(\tilde{n}) - \alpha) : \tilde{n} \in D^- \setminus D^0\}$

Let  $d^+(k(i)) = \max\{P_D(\tilde{n}) - k(i)(P_{FA}(\tilde{n}) - \alpha) : \tilde{n} \in D^+\}$

Let  $d^-(k(i)) = dr(i)$  and  $ds(i+1) = dr(i)$

**If**  $d^+(k(i)) > d^-(k(i))$

Let  $df(i+1) = \min\{D^+(k(i), \max\{ds(i), df(i)\})\}$

**Else**

Let  $df(i+1) = \max\{D^+(k(i), \min\{ds(i), df(i)\})\}$

**End If**

Let  $i = i + 1$

**End While**

**If**  $|d^+(k(i-1)) - d^-(k(i-1))| > \epsilon$

Let  $t = \text{sgn}[d^+(k(i-1)) - d^-(k(i-1))]$

Let  $k(i) = \max\left\{\frac{(P_D(\tilde{n}) - (d^-(k(i-1)) + t\epsilon))}{(P_{FA}(\tilde{n}) - \alpha)} : \tilde{n} \in d^+ \setminus D^0\right\}$

Let  $d^+(k(i)) = d^-(k(i-1)) + t\epsilon$

Let  $d^-(k(i)) = \max\{P_D(\tilde{n}) - k(i)(P_{FA}(\tilde{n}) - \alpha) : \tilde{n} \in D^-\}$

**Else**

Let  $k(i) = k(i-1)$

**End If**

$f_{\tilde{N}'}(n) = \lambda\delta(n - \tilde{n}_1) + (1 - \lambda)\delta(n - \tilde{n}_2)$  (\*)

where  $\tilde{n}_1 \in D^- : P_D(\tilde{n}_1) - k(i)(P_{FA}(\tilde{n}_1) - \alpha) = D^-(k(i))$ ,

$\tilde{n}_2 \in D^+ : P_D(\tilde{n}_2) - k(i)(P_{FA}(\tilde{n}_2) - \alpha) = D^+(k(i))$ ,

and  $\lambda = P_{FA}(\tilde{n}_2) - \alpha / P_{FA}(\tilde{n}_2) - P_{FA}(\tilde{n}_1)$

**End If**

*Theorem 4:*

**a)** Pick any  $\epsilon > 0$ . Then the above algorithm finds an N-P SR noise  $\tilde{N}'$  from  $\tilde{N}$  in at most  $i_{\max} = \lceil \log_2(2/\epsilon) \rceil + 1$  iterations so that

$$P_D(f_{\tilde{N}'_{\text{opt}}}) \geq P_D(f_{\tilde{N}'}) \geq P_D(f_{\tilde{N}'_{\text{opt}}}) - \epsilon \quad (40)$$

and

$$P_{\text{FA}}(f_{\tilde{N}'}) \leq \alpha. \quad (41)$$

**b)** The suboptimal detection performance with noise  $\tilde{N}'$  is at most  $\tau + \epsilon$  less than the optimal SR detection with noise  $N_{\text{opt}}$  if  $\tilde{N}$  satisfies (38)–(39)

$$P_D(f_{N_{\text{opt}}}) \geq P_D(f_{\tilde{N}'}) \geq P_D(f_{N_{\text{opt}}}) - (\tau + \epsilon). \quad (42)$$

We next show that the noise-finding algorithm is much faster than exhaustive search if  $P_{D-\text{sup}} < P_{D+\text{sup}}$  for large number of noise realizations. Suppose that the number of elements in the discretized sets  $D^-$ ,  $D^+$ , and  $D^0$  are respectively  $M$ ,  $T$ , and  $V$ . Then Theorem 2c) implies that exhaustive search needs to consider all  $M \times T$  noise realization pairs  $(\tilde{n}^-, \tilde{n}^+) \in D^- \times D^+$  and find the pair that corresponds to the maximal value of  $\lambda P_D(\tilde{n}^-) + (1 - \lambda) P_D(\tilde{n}^+)$  where  $\lambda = (P_{\text{FA}}(\tilde{n}^+) - \alpha) / (P_{\text{FA}}(\tilde{n}^+) - P_{\text{FA}}(\tilde{n}^-))$  to find the optimal noise  $\tilde{N}_{\text{opt}}$  in  $\tilde{N}$ . Thus exhaustive search needs to find the maximum from the set of  $M \times T$  elements. But the algorithm first finds the maximum from the smaller set of only  $M$  elements and from the smaller set of only  $T$  elements to find the respective values of  $d^-(k(0))$  and  $d^+(k(0))$ . The algorithm next finds the minimum from the set  $\{(P_D(\tilde{n}) - dr(i)) / (P_{\text{FA}}(\tilde{n}) - \alpha) : \tilde{n} \in D^- \setminus D^0\}$  of only  $M - V$  elements and finds the maximum from the set  $\{P_D(\tilde{n}) - k(i) / (P_{\text{FA}}(\tilde{n}) - \alpha) : \tilde{n} \in D^+\}$  of only  $T$  elements to get the respective values of  $k(i)$  and  $d^+(k(i))$  at each iteration  $i = 1, \dots, \lceil \log_2(2/\epsilon) \rceil$ . Then it finds the maximum from the set of  $T - V$  elements and from the set of  $M$  elements during the last iteration if needed. Theorem 4a) shows that the algorithm finds the near-optimal noise  $\tilde{N}'$  or randomization in just  $i_{\max} = \lceil \log_2(2/\epsilon) \rceil + 1$  iterations such that  $P_D(f_{\tilde{N}'_{\text{opt}}}) - P_D(f_{\tilde{N}'}) \leq \epsilon$ . So the algorithm is faster than exhaustive search even for very small values of  $\epsilon$  if both  $M$  and  $T$  are large. The two applications in Section V confirm that the noise-finding algorithm requires fewer iterations than the upper-bound number of iterations  $i_{\max}$ .

#### IV. NOISE BENEFITS IN INEQUALITY-CONSTRAINED STATISTICAL DECISION PROBLEMS

We show that randomization and noise benefits extend beyond Neyman–Pearson signal detection. Researchers have found randomization benefits in inequality-constrained statistical decisions such as average-power-constrained signal transmission and jamming strategies [55] as well as in pricing and scheduling for a network access point that maximizes the time-average profit subject to an average-transmission-rate constraint [56]. Azizoglu [55] showed that the optimal channel switching strategy for an average-power-constrained transmitter time-shares between at most two channels and power

levels to minimize the receiver-end error probability in multiple additive noise channels. Huang and Neely [56] studied a pricing and transmission scheduling problem for a network access point to maximize the time-average profit. They showed that randomization of at most two business-price tuples suffices for the access point to achieve its optimal time-average profit under the average-transmission-rate constraint. These examples are all special cases of noise benefits in expected payoff maximization in statistical decision problems that have one inequality constraint on the expected cost.

We now extend the previous results to a broad class of expected payoff maximization problems subject to an inequality-constrained expected cost such that the payoff and cost are real-valued and bounded nonnegative Borel-measurable functions on the noise space  $\mathcal{N} \subseteq \mathbf{R}^m$ . Let  $h$  be the payoff function and  $c$  be the cost function. We want to maximize the average payoff  $E_{f_N}(h(\mathbf{N})) = \int_{\mathcal{N}} h(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n}$  subject to the average cost constraint  $E_{f_N}(c(\mathbf{N})) = \int_{\mathcal{N}} c(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} \leq \gamma$ . Define the noise  $\mathbf{N}$  as SR noise if its addition improves the expected payoff  $E_{f_N}(h(\mathbf{N}))$  while the expected cost  $E_{f_N}(c(\mathbf{N}))$  stays at or below a preset maximum expected cost  $\gamma$ . Suppose that the respective cost and payoff in the absence of noise are  $c(\mathbf{0}) \leq \gamma$  and  $h(\mathbf{0})$ . Then  $E_{f_N}(h(\mathbf{N})) > h(\mathbf{0})$  and  $E_{f_N}(c(\mathbf{N})) \leq \gamma$  hold if  $\mathbf{N}$  is SR noise. So we want to find the optimal SR noise  $\mathbf{N}_{\text{opt}}$  such that  $E_{f_N}(h(\mathbf{N})) \leq E_{f_{N_{\text{opt}}}}(h(\mathbf{N}_{\text{opt}}))$  for any other SR noise  $\mathbf{N}$  and such that  $E_{f_{N_{\text{opt}}}}(c(\mathbf{N}_{\text{opt}})) \leq \gamma$ .

Theorems 1–4 hold if we replace  $P_D$ ,  $P_{\text{FA}}$ , and  $\alpha$  in (9)–(42) with the respective real-valued bounded nonnegative payoff function  $h$ , cost function  $c$ , and the preset maximum expected cost  $\gamma$ . Theorem 4 holds with  $i_{\max} = \lceil \log_2(\xi/\epsilon) \rceil + 1$  if  $\xi/2$  bounds both  $h$  and  $c$ . So the SR noise-finding algorithm can find a near-optimal SR noise  $\tilde{N}'$  that improves the expected payoff in statistical decisions with one inequality constraint on the expected cost if we choose a small enough  $\epsilon$  and a set of noise realizations  $\tilde{N}$  such that  $\tau$  is sufficiently small in (38). We omit the proofs of the above statements because they are substantially the same as the proofs of Theorems 1–4 in the Appendix but with minor notational changes. The next section applies the algorithm to find a near-optimal signal power randomization for a power-constrained antipodal signal transmitter that improves the MAP detection in Gaussian-mixture noise.

#### V. APPLICATIONS OF THE SR NOISE-FINDING ALGORITHM

This section presents two applications of the SR noise-finding algorithm. The first application finds a near-optimal SR noise for a suboptimal one-sample Neyman–Pearson hypothesis test between two Gaussian distributions. The second application gives a near-optimal signal power randomization for a power-constrained antipodal signal transmitter in additive Gaussian-mixture channel noise where the receiver uses optimal MAP signal detection.

##### A. Near-Optimal SR Noise for a Suboptimal One-Sample Neyman–Pearson Hypothesis Test of Variance

Consider a hypothesis test between two Gaussian densities  $H_0 : f_0(x) = 1/\sqrt{6\pi} e^{-x^2/6}$  versus  $H_1 : f_1(x) = 1/\sqrt{2\pi} e^{-(x-1)^2/2}$  where we want to decide between  $H_0$  and  $H_1$  using only a *single* observation of  $X$  at the significance

level  $\alpha = 0.05$ . Fig. 1(a) shows both  $f_0$  and  $f_1$ . Note that the likelihood ratio  $f_1(x)/f_0(x) = \sqrt{3}e^{-2x^2-6x+3/6}$  is not monotonic. So the optimal Neyman–Pearson test function requires the two thresholds  $\theta_1 = 1.342$  and  $\theta_2 = 1.658$  for the optimal decision. It rejects  $H_0$  if  $X \in [\theta_1, \theta_2]$  and if the test significance or preset false-alarm probability is  $\alpha = 0.05$ . The optimal N-P detection probability is 0.11.

Suppose that the detection system can use only one threshold  $\theta$  due to resource limits or some design constraint. Such a suboptimal detector must reject  $H_0$  when  $X > \theta$ . Then  $\theta = 2.85$  for  $\alpha = P_{FA}(0) = 0.05$ . The noiseless detection probability is  $P_D(0) = 0.0322$ . Here  $P_D(n) = 1 - \Phi(\theta - 1 - n)$  and  $P_{FA}(n) = 1 - \Phi(\theta - n/\sqrt{3})$  for the standard normal cumulative distribution function  $\Phi(z) = 1/\sqrt{2\pi} \int_{-\infty}^z e^{-w^2/2} dw$ . Then inequality (37) of Corollary 2 becomes  $(\theta - 1) > \theta/3$ . So  $\theta = 2.85$  satisfies this sufficient condition. So there exists additive SR noise  $N$  that improves the detection performance of the suboptimal detector. But the existence of this SR noise  $N$  does not itself imply the existence of *optimal* SR noise  $N_{opt}$  if the noise space is  $\mathbf{R}$ . Note that  $P_D(n)$  and  $P_{FA}(n)$  are monotonic increasing on  $\mathbf{R}$  and  $P_{D-sup} = P_D(0) < P_{D+sup} = 1$ . Then Theorem 1b) implies that the corresponding false-alarm probability is  $\alpha = 0.05$  if the optimal N-P SR noise exists. Suppose that the optimal N-P SR noise exists. Then the necessary condition of Corollary 1a) becomes  $P'_D(n) - kP'_{FA}(n) = f_1(\theta - n) - kf_0(\theta - n) = 0$  where  $f_0$  and  $f_1$  are the pdfs defined above. Then there exists  $u > 0$  such that  $f_1(n) > kf_0(n)$  for all  $n \in (n_i, n_i + u)$  and  $f_1(n) < kf_0(n)$  for all  $n \in (n_i - u, n_i)$  if  $n_i$  maximizes  $P_D(n) - k(P_{FA}(n) - \alpha)$ . Then at most one solution of  $P'_D(n) - kP'_{FA}(n) = 0$  is a global maximum of  $P_D(n) - k(P_{FA}(n) - \alpha)$  for each  $k \geq 0$ . So Corollary 1b) implies that optimal N-P SR noise does not exist if the noise space is  $\mathbf{R}$ . But the hypothesis of Theorem 2c) does hold if we restrict the noise space to a compact interval (say  $[-3, 3]$ ) because  $P_D(n)$  and  $P_{FA}(n)$  are continuous functions of  $n$ .

We next apply the algorithm to find near optimal noise in  $\mathcal{N} = [-3, 3]$  for  $\epsilon = 2^{-20}$ . Consider the discretized set  $\tilde{\mathcal{N}}$  of noise realizations starting from  $-3$  up to  $3$  with an increment of  $0.0001$ :  $\tilde{\mathcal{N}} = [-3 : 0.0001 : 3]$ .  $\tilde{\mathcal{N}}$  satisfies (38)–(39) for  $\tau = 0.00004$  because  $0.4$  bounds  $f_0$  and  $f_1$ . Fig. 2 plots  $g(\tilde{n}, k(i)) = P_D(\tilde{n}) - k(i)(P_{FA}(\tilde{n}) - \alpha)$  before the first iteration ( $i = 0$ ) where  $k(0) = 1$  and after the 17th iteration ( $i = 17$ ). The noise-finding algorithm finds the value  $k(17) = 1.8031$  in just  $17 (< i_{max} = 22)$  iterations such that  $|d^+(k(17)) - d^-(k(17))| < \epsilon = 2^{-20}$ . Note that  $g(\tilde{n}_1, k(17)) = d^-(k(17))$  at  $\tilde{n}_1 = -3 \in D^-$  and that  $g(\tilde{n}_2, k(17)) = d^+(k(17))$  at  $\tilde{n}_2 = 2.1433 \in D^+$ . Then

$$f_{\tilde{N}'}(n) = \lambda\delta(n + 3) + (1 - \lambda)\delta(n - 2.1433) \quad (43)$$

$$\text{with } \lambda = \frac{(P_{FA}(2.1433) - 0.05)}{(P_{FA}(2.1433) - P_{FA}(-3))} = 0.8547 \quad (44)$$

is the pdf of a near-optimal N-P additive SR noise  $\tilde{N}'$  because  $P_D(f_{\tilde{N}'}) = 0.0894$  and because Theorem 4b) implies that the detection probability  $P_D(f_{N_{opt}})$  for the optimal N-P SR noise  $N_{opt}$  in  $\mathcal{N} = [-3, 3]$  will be at most  $0.00004 + 2^{-20}$  more than  $P_D(f_{\tilde{N}'})$ . So the algorithm finds a near-optimal N-P SR noise that gives a 177% increase in the detection probability of

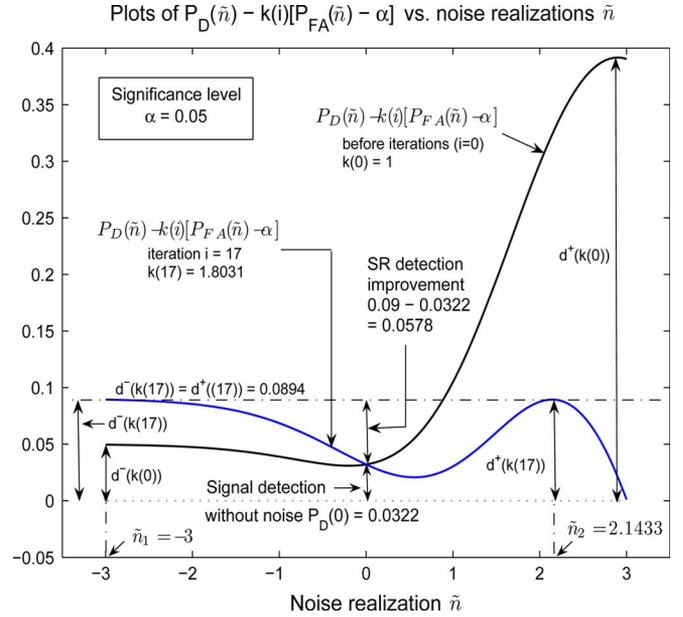


Fig. 2. Finding near-optimal Neyman–Pearson SR noise. The two curves plot  $g(\tilde{n}, k(i)) = P_D(\tilde{n}) - k(i)(P_{FA}(\tilde{n}) - \alpha)$  before the first iteration ( $i = 0$ ) and after the 17th iteration ( $i = 9$ ) where  $k(0) = 1$ . The detection probability is  $P_D(0) = 0.0322$  in the absence of additive noise. The noise-finding algorithm finds a value of  $k(17) = 1.8031$  in 17 iterations such that  $|d^+(k(17)) - d^-(k(17))| < \epsilon = 2^{-20}$ . Note that  $g(\tilde{n}_1, k(17)) = d^-(k(17))$  at  $\tilde{n}_1 = -3 \in D^-$  and  $g(\tilde{n}_2, k(17)) = d^+(k(17))$  at  $\tilde{n}_2 = 2.1433 \in D^+$ . Then (43)–(44) give the pdf of a near-optimal N-P SR noise  $\tilde{N}'$  and  $P_D(f_{\tilde{N}'}) = 0.0894$ . So the N-P SR noise  $\tilde{N}'$  increases the detection probability  $P_D$  177% from 0.0322 to 0.0894.

the single-threshold suboptimal test from 0.0322 to 0.0894. The noise-enhanced detection probability 0.0894 is fairly close to the optimal N-P detection probability 0.11.

### B. Near-Optimal Signal Power Randomization for a Power-Constrained Signal Transmitter

The detection performance of a MAP receiver can sometimes benefit from signal power randomization or time-sharing in an average-power-constrained antipodal signal transmitter if the channel noise pdf is multimodal. The noise-finding algorithm finds a near-optimal signal power distribution or randomization in an average-power-constrained transmitter that improves the MAP receiver's detection performance.

Consider a signal detection hypothesis test where the transmitter transmits antipodal signals  $X \in \{-S, S\}$  where  $S \in \mathcal{S} = [0.5, 3.75]$  and both signal values are equally likely:  $H_0 : X = -S$  versus  $H_1 : X = S$  and  $P(H_0) = P(-S) = P(S) = P(H_1)$ . Suppose that the transmitter can use at most 4.75 units of expected power  $E(S^2)$  and that the receiver decides between  $H_0$  and  $H_1$  using a single noisy observation  $Y = X + N$ . Here  $N$  is an additive symmetric Gaussian-mixture channel noise where the signal probability density is  $f_0(y) = 1/2\sqrt{2\pi}e^{-(y-2-S)^2/2} + 1/2\sqrt{2\pi}e^{-(y+2-S)^2/2}$  at the receiver under the hypothesis  $H_0$  and  $f_1(y) = 1/2\sqrt{2\pi}e^{-(y-2+S)^2/2} + 1/2\sqrt{2\pi}e^{-(y+2+S)^2/2}$  under the hypothesis  $H_1$ . Such Gaussian-mixture channel noise can occur due to co-channel interference in communication systems [57]–[59]. The receiver is optimal and hence it uses

maximum *a posteriori* (MAP) signal detection to maximize the probability of correct decision. Then the receiver rejects  $H_0$  if the likelihood ratio obeys  $f_1(y)/f_0(y) > P(H_0)/P(H_1) = 1$ . We assume that the transmitter can time-share or randomize between the signal power levels so that the receiver knows the signal power value  $S^2$  but does not know whether the signal is  $+S$  or  $-S$ .

Let  $f_S$  be the probability density of the transmitter's signal strength  $S$ . Define  $P_{CD}(f_S)$  as the probability of correct decision and let  $E_{f_S}(S^2)$  be the average signal power when the signal-strength pdf is  $f_S$ . Then we want to find  $f_{S_{opt}}$  such that  $P_{CD}(f_S) \geq P_{CD}(f_{S_{opt}})$  for any  $f_S$  such that  $E_{f_{S_{opt}}}(S^2) \leq 4.75$ . We can view  $P_{CD}(f_S)$  as the average payoff  $E_{f_S}(h(S))$  and view  $E_{f_S}(S^2)$  as the average cost  $E_{f_S}(c(S))$  with the maximum average cost  $\gamma = 4.75$ . So we can apply the SR noise finding algorithm to find a near-optimal signal-power pdf  $f_{\hat{S}}$  from a discretized set  $\hat{\mathcal{S}}$  of signal-power realizations. We use  $\hat{\mathcal{S}} = [0.5 : 0.0001 : 3.75]$  and  $\epsilon = 2^{-20}$ . Then  $\hat{\mathcal{S}}$  satisfies conditions (38)–(39) for  $\tau = 0.0002$  because 0.2 bounds  $f_0$  and  $f_1$ . Both  $h$  and  $c$  have upper bound  $\xi/2 = 10/2$ . So the iteration upper bound is  $i_{max} = \lceil \log_2(\xi/\epsilon) \rceil + 1 = 24$ . The generalization of Theorem 3 to this example gives the sufficient condition

$$P''_{CD}(s)2s > 2|P'_{CD}(s)| \quad (45)$$

for the SR noise benefits if the maximal average signal power is  $s^2$ .

Fig. 3 plots the correct-decision probability  $P_{CD}$  versus the signal power  $S^2$ . The randomized signal power is optimal because the graph is nonconcave. Azizoglu [55] proved that the plot of  $P_{CD}$  versus  $S^2$  is concave if the channel noise has finite variance and if it has a unimodal pdf that is continuously differentiable at every point except the mode. The nonconcavity of the graph in Fig. 3 arises from the bimodal Gaussian-mixture pdf even though the bimodality of the channel noise pdf does not itself give a nonconcave plot of  $P_{CD}$  versus  $S^2$ . The probability of correct decision  $P_{CD}$  is 0.7855 (point *a*) if the transmitter uses a constant power  $s^2 = 4.75$  (a constant signal strength  $s = 2.1794$ ). The dashed tangent line shows that we can achieve a better detection performance using the same average signal power  $E(S^2) = 4.75$ . The probability of correct decision  $P_{CD}$  is 0.8421 (point *d*) if the transmitter time-shares or randomizes appropriately between the signal power levels  $s_1^2 = 1.4908$  (point *b*) and  $s_2^2 = 9.3697$  (point *c*). The sufficient condition (45) does not hold for  $s^2 = 4.75$  because the plot of  $P_{CD}$  versus  $S^2$  is locally concave at the constant-power operating point *a*:  $P''_{CD}(s) < 0 < P'_{CD}(s)$  at  $s = \sqrt{4.75}$ . So the noise benefit occurs even when the sufficient condition (45) does not hold.

The algorithm finds the SR noise or signal-strength randomization  $\hat{S}'$  with pdf

$$f_{\hat{S}'}(s) = \lambda\delta(s - 1.221) + (1 - \lambda)\delta(s - 3.061) \quad (46)$$

$$\text{where } \lambda = \frac{(3.061^2 - 4.74)}{(3.061^2 - 1.221^2)} = 0.5876 \quad (47)$$

in just 13 ( $< i_{max} = 24$ ) iterations. So the transmitter should time-share or randomize between the antipodal signals

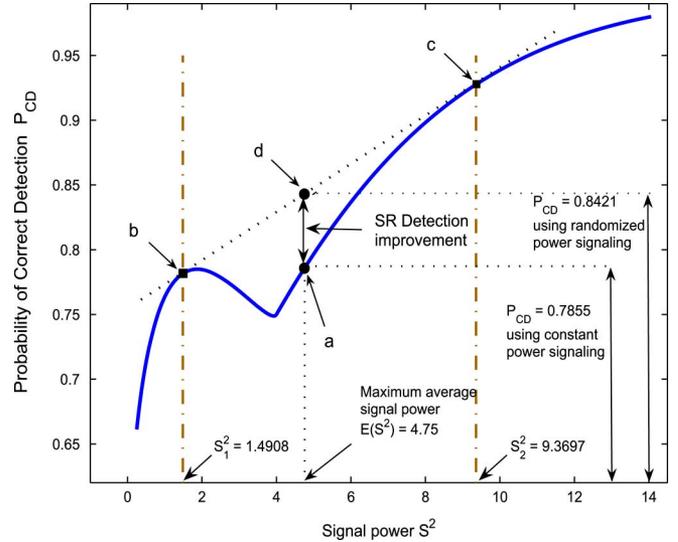


Fig. 3. SR noise (signal-strength randomization) benefits in optimal antipodal signal detection. Signal power randomization in an average-power-constrained antipodal transmitter improves the detection performance of the optimal receiver. The transmitter transmits antipodal signals  $X \in \{-S, S\}$  such that  $E(S^2) \leq \gamma = 4.75$  and such that both signals values are equally likely:  $H_0 : X = -S$  versus  $H_1 : X = S$  and  $P(H_0) = P(-S) = P(S) = P(H_1)$ . The receiver receives the noisy observation  $Y = X + N$ .  $N$  is symmetric Gaussian-mixture channel noise. The signal probability density is  $f_0(y) = 1/2\sqrt{2\pi}e^{-(y-2-S)^2/2} + 1/2\sqrt{2\pi}e^{-(y+2-S)^2/2}$  at the receiver under the hypothesis  $H_0$  and is  $f_1(y) = 1/2\sqrt{2\pi}e^{-(y-2+S)^2/2} + 1/2\sqrt{2\pi}e^{-(y+2+S)^2/2}$  under the hypothesis  $H_1$ . The receiver uses the single noisy observation  $Y$  and the optimal MAP decision rule to decide between  $H_0$  and  $H_1$ . The solid line shows the nonmonotonic and nonconcave plot of the probability of correct decision  $P_{CD}$  versus the signal power  $S^2$ . Nonconcavity of the plot between the points *b* and *c* allows the SR effect to occur. The respective probability of correct decision  $P_{CD}$  is 0.7855 (point *a*) if the transmitter uses a constant power  $s^2 = 4.75$  (a constant signal strength  $s = 2.1794$ ). The dashed tangent line shows that we can achieve a better probability of correct decision (0.8421 at point *d*) at the same average signal power  $E(S^2) = 4.75$  if the transmitter time-shares or randomizes appropriately between the signal power levels  $s_1^2 = 1.4908$  (point *b*) and  $s_2^2 = 9.3697$  (point *c*).

$\{-1.221, 1.221\}$  and  $\{-3.061, 3.061\}$  with respective probabilities  $\lambda = 0.5876$  and  $1 - \lambda = 0.4124$ . This signal-strength randomization pdf  $f_{\hat{S}'}$  is nearly optimal because Theorem 4b) implies that  $P_{CD}(f_{S_{opt}})$  will be at most  $0.0002 + 2^{-20}$  more than  $P_{CD}(f_{\hat{S}'}) = 0.8421$  for the optimal signal-strength randomization or for the optimal SR noise  $S_{opt}$  in  $\mathcal{S} = [0.5, 3.75]$ . Thus the SR noise algorithm can find a near-optimal signal power randomization that improves the average probability of correct decision (from 0.7855 to 0.8421) over the constant power signaling. Chapeau-Blondeau and Rousseau showed related SR noise benefits in the optimal Bayesian detection of a periodic square-wave signal in bimodal Gaussian-mixture phase noise [60] and of a constant signal in additive bimodal noise [61]. But they did not find either the optimal or near-optimal noise pdf as in (46)–(47) for inequality-constrained optimal signal detection.

## VI. CONCLUSION

Adding noise can sometimes benefit Neyman–Pearson statistical signal detection. Theorems 1–3 give necessary and sufficient conditions for optimal noise densities in such signal detection. The noise-finding algorithm finds near-optimal SR noise from a finite set of noise realizations. An upper bound limits

the number of iterations that the algorithm requires to find this near-optimal noise. These results directly extend to the general case of maximizing the expected payoff in statistical decision making with an inequality constraint on the average cost. Nor are the noise benefits limited to suboptimal signal detection. Noise or randomization can sometimes improve the performance of even optimal MAP detectors. Adding such optimal or near-optimal noise should benefit many other problems of signal processing and communication. The general optimization structure should also lead to adaptive noise-finding algorithms.

APPENDIX

PROOFS OF THEOREMS AND COROLLARY 1

*Theorem 1:*

**a)** Suppose that  $P_{D-\text{sup}} \geq P_{D+\text{sup}}$  and that  $G^- \neq \emptyset$ . Then

$$f_{N_{\text{opt}}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o) \tag{48}$$

is an optimal SR noise pdf for Neyman–Pearson detection for some  $\mathbf{n}_o \in G^-$  and  $P_{\text{FA}}(f_{N_{\text{opt}}}) \leq \alpha$ . The Neyman–Pearson optimal SR noise does not exist for the given test level  $\alpha$  if  $G^- = \emptyset$ . But there exists a noise pdf sequence  $\{f_{N_r}\}_{r=1}^\infty$  of the form (48) such that  $P_{\text{FA}}(f_{N_r}) \leq \alpha$  for all  $r$  and such that

$$\lim_{r \rightarrow \infty} P_D(f_{N_r}) = P_{D_{\text{sup}}}. \tag{49}$$

**b)** Suppose that  $P_{D-\text{sup}} < P_{D+\text{sup}}$ . Then  $P_{\text{FA}}(f_{N_{\text{opt}}}) = \alpha$  if the Neyman–Pearson optimal SR noise pdf  $f_{N_{\text{opt}}}(\mathbf{n})$  exists.

*Proof:*

**Part a):** Suppose that  $P_{D-\text{sup}} \geq P_{D+\text{sup}}$  and that  $G^- \neq \emptyset$ . Then for all  $k \geq 0$

$$d^-(k) = \sup_{\mathbf{n}} \{P_D(\mathbf{n}) - k(P_{\text{FA}}(\mathbf{n}) - \alpha) : \mathbf{n} \in D^-\} \tag{50}$$

$$\geq \sup_{\mathbf{n}} \{P_D(\mathbf{n}) : \mathbf{n} \in D^-\} = P_{D-\text{sup}} \tag{51}$$

and similarly  $d^+(k) \leq P_{D+\text{sup}}$ . Then  $d^-(k) \geq d^+(k)$  for all  $k \geq 0$  if  $P_{D-\text{sup}} \geq P_{D+\text{sup}}$ . The left-hand side of (27) becomes

$$\min_{k \geq 0} \sup_{f_N \in \mathcal{F}} L(f_N, k) = \min_{k \geq 0} d^-(k) \tag{52}$$

$$\begin{aligned} &= d^-(0) \text{ because } d^-(k) \text{ is a} \\ &\quad \text{nondecreasing function of } k \\ &= P_{D-\text{sup}} = P_{D_{\text{sup}}}. \end{aligned} \tag{53}$$

Thus (6), (8), and (53) imply that  $k^* = 0$ . So we need to find the pdf  $f_{N_{\text{opt}}}$  such that

$$L(f_{N_{\text{opt}}}, 0) = \int_{\mathcal{N}} P_D(\mathbf{n}) f_{N_{\text{opt}}}(\mathbf{n}) d\mathbf{n} = P_{D-\text{sup}}. \tag{54}$$

The definition of  $G^-$  implies that  $P_D(\mathbf{n}_o) = P_{D-\text{sup}} = P_{D_{\text{sup}}}$  for any  $\mathbf{n}_o \in G^-$ . Choose  $f_N(\mathbf{n})$  as the unit impulse at  $\mathbf{n}_o$ :  $f_N(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$ . Then

$$\int_{\mathcal{N}} P_D(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} = P_D(\mathbf{n}_o) = P_{D-\text{sup}} \tag{55}$$

$$\int_{\mathcal{N}} P_{\text{FA}}(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} = P_{\text{FA}}(\mathbf{n}_o) \leq \alpha. \tag{56}$$

So  $f_N$  is the optimal SR noise pdf and hence

$$f_{N_{\text{opt}}} = \delta(\mathbf{n} - \mathbf{n}_o) \text{ for any } \mathbf{n}_o \in G^-. \tag{57}$$

Suppose now that  $G^- = \emptyset$ . Then  $P_D(\mathbf{n}) < P_{D_{\text{sup}}} = P_{D-\text{sup}}$  for all  $\mathbf{n} \in \mathcal{N}$ . Suppose that  $f'_N$  is the optimal SR noise pdf. Then

$$\begin{aligned} &P_D(f'_N) \\ &= \int_{\mathcal{N}} P_D(\mathbf{n}) f'_N(\mathbf{n}) d\mathbf{n} \end{aligned} \tag{58}$$

$$< \int_{\mathcal{N}} P_{D_{\text{sup}}} f'_N(\mathbf{n}) d\mathbf{n} \tag{59}$$

because  $P_D(\mathbf{n}) < P_{D_{\text{sup}}}$  and

$$\text{because } \int_{\mathcal{N}} [P_{D_{\text{sup}}} - P_D(\mathbf{n})] f'_N(\mathbf{n}) d\mathbf{n} = 0$$

$$\begin{aligned} &\text{iff } [P_{D_{\text{sup}}} - P_D(\mathbf{n})] f'_N(\mathbf{n}) = 0 \text{ a.e. on } \mathcal{N} \\ &= P_{D_{\text{sup}}}. \end{aligned} \tag{60}$$

$G^- = \emptyset$  and the supremum definition of  $P_{D-\text{sup}}$  further imply [62] that there exists a sequence of noise realizations  $\{\mathbf{n}_r\}_{r=1}^\infty$  in  $G^-$  such that

$$\lim_{r \rightarrow \infty} P_D(\mathbf{n}_r) = P_{D_{\text{sup}}}. \tag{61}$$

So there exists  $\mathbf{n}_s \in \{\mathbf{n}_r\}_{r=1}^\infty \subset G^-$  such that  $P_D(\mathbf{n}_s) > P_D(f'_N)$  and  $P_{\text{FA}}(\mathbf{n}_s) \leq \alpha$ . Define a sequence of noise pdfs  $f_{N_r}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_r)$ . Then  $f_{N_s}$  contradicts the optimality of  $f'_N$  while  $P_{\text{FA}}(f_{N_r}) \leq \alpha$  for all  $r$  and

$$\lim_{r \rightarrow \infty} P_D(f_{N_r}) = \lim_{r \rightarrow \infty} P_D(\mathbf{n}_r) = P_{D_{\text{sup}}}. \tag{62}$$

**Part b):** Suppose that  $P_{D-\text{sup}} < P_{D+\text{sup}}$  and that  $G^+ \neq \emptyset$ . Suppose also that  $f'_N$  is the optimal SR noise pdf such that  $P_{\text{FA}}(f'_N) = v < \alpha$  and that  $P_D(f'_N) \geq P_D(f_N)$  for any other noise pdf  $f_N$ . The definition of  $G^+$  implies that  $P_{\text{FA}}(\mathbf{n}_1) \geq \alpha > v$  if  $\mathbf{n}_1 \in G^+$ . Then  $P_D(\mathbf{n}_1) = P_{D_{\text{sup}}} > P_D(f'_N)$  because  $P_{\text{FA}}(f'_N) = v < \alpha$  and  $P_{D-\text{sup}} < P_{D+\text{sup}} = P_{D_{\text{sup}}}$ . Define

$$f_N(\mathbf{n}) = \frac{\alpha - v}{P_{\text{FA}}(\mathbf{n}_1) - v} \delta(\mathbf{n} - \mathbf{n}_1) + \frac{P_{\text{FA}}(\mathbf{n}_1) - \alpha}{P_{\text{FA}}(\mathbf{n}_1) - v} f'_N(\mathbf{n}). \tag{63}$$

Then  $f_N$  is a pdf because  $f_N(\mathbf{n}) \geq 0$  for all  $\mathbf{n}$  and  $\int_{\mathcal{N}} f_N(\mathbf{n}) d\mathbf{n} = 1$ . So

$$P_{\text{FA}}(f_N) = \int_{\mathcal{N}} P_{\text{FA}}(\mathbf{n}) f_N(\mathbf{n}) d\mathbf{n} \tag{64}$$

$$\begin{aligned} &= \frac{\alpha - v}{P_{\text{FA}}(\mathbf{n}_1) - v} P_{\text{FA}}(\mathbf{n}_1) \\ &\quad + \frac{P_{\text{FA}}(\mathbf{n}_1) - \alpha}{P_{\text{FA}}(\mathbf{n}_1) - v} \int_{\mathcal{N}} P_{\text{FA}}(\mathbf{n}) f'_N(\mathbf{n}) d\mathbf{n} \end{aligned} \tag{65}$$

$$= \frac{\alpha - v}{P_{\text{FA}}(\mathbf{n}_1) - v} P_{\text{FA}}(\mathbf{n}_1) + \frac{P_{\text{FA}}(\mathbf{n}_1) - \alpha}{P_{\text{FA}}(\mathbf{n}_1) - v} v \quad (66)$$

$$= \alpha \quad (67)$$

and

$$P_{\text{D}}(f_{\text{N}}) = \int_{\mathcal{N}} P_{\text{D}}(\mathbf{n}) f_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (68)$$

$$= \frac{\alpha - v}{P_{\text{FA}}(\mathbf{n}_1) - v} P_{\text{D}}(\mathbf{n}_1) + \frac{P_{\text{FA}}(\mathbf{n}_1) - \alpha}{P_{\text{FA}}(\mathbf{n}_1) - v} \int_{\mathcal{N}} P_{\text{D}}(\mathbf{n}) f'_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (69)$$

$$= \frac{\alpha - v}{P_{\text{FA}}(\mathbf{n}_1) - v} P_{\text{D}_{\text{sup}}} + \frac{P_{\text{FA}}(\mathbf{n}_1) - \alpha}{P_{\text{FA}}(\mathbf{n}_1) - v} P_{\text{D}}(f'_{\text{N}}) \quad (70)$$

$$> P_{\text{D}}(f'_{\text{N}}) \quad (71)$$

because  $P_{\text{D}_{\text{sup}}} > P_{\text{D}}(f'_{\text{N}})$ . But  $P_{\text{D}}(f_{\text{N}}) > P_{\text{D}}(f'_{\text{N}})$  contradicts the optimality of  $f'_{\text{N}}$ . So  $P_{\text{FA}}(f'_{\text{N}}) = \alpha$  if  $f'_{\text{N}}$  is the optimal SR noise pdf and if  $P_{\text{D}_{\text{sup}}} < P_{\text{D}_{\text{sup}}}$ .

Suppose now that  $P_{\text{D}_{\text{sup}}} < P_{\text{D}_{\text{sup}}}$  but  $G^+ = \emptyset$ . The definitions of  $P_{\text{D}_{\text{sup}}}$  and  $G^+$  imply that there exists an  $\mathbf{n}_1 \in D^+$  such that  $P_{\text{D}}(\mathbf{n}_1) > P_{\text{D}}(f'_{\text{N}})$  because  $P_{\text{FA}}(f'_{\text{N}}) = v < \alpha$ . Then  $f_{\text{N}}$  again contradicts the optimality of  $f'_{\text{N}}$  if we define  $f_{\text{N}}$  as in (63).  $\square$

**Theorem 2:**

Suppose that  $P_{\text{D}_{\text{sup}}} < P_{\text{D}_{\text{sup}}}$ . Then a)–d) hold:

**a)** There exists  $k^* \geq 0$  such that  $d^+(k^*) = d^-(k^*) = d(k^*)$  and  $\min\{d^+(k), d^-(k)\} \leq d(k^*) \leq \max\{d^+(k), d^-(k)\}$  for any  $k \geq 0$ .

**b)** Suppose the noise pdf  $f_{\text{N}}$  satisfies  $P_{\text{D}}(f_{\text{N}}) = d(k^*) > P_{\text{D}}(0)$  and  $P_{\text{FA}}(f_{\text{N}}) = \alpha$ . Then  $f_{\text{N}}$  is a Neyman–Pearson optimal noise pdf. So  $d(k^*)$  is the optimal N-P SR detection probability  $P_{\text{D}_{\text{opt}}}$ .

**c)** Suppose that there exist  $\mathbf{n}_1 \in D^-$  and  $\mathbf{n}_2 \in D^+$  such that  $g(\mathbf{n}_1, k^*) = d^-(k^*) = d(k^*) = g(\mathbf{n}_2, k^*) = d^+(k^*)$ . Then

$$f_{\text{N}_{\text{opt}}}(\mathbf{n}) = \lambda \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda) \delta(\mathbf{n} - \mathbf{n}_2) \quad (72)$$

is the optimal Neyman–Pearson SR noise pdf if  $d(k^*) > P_{\text{D}}(0)$  and if

$$\lambda = \frac{P_{\text{FA}}(\mathbf{n}_2) - \alpha}{P_{\text{FA}}(\mathbf{n}_2) - P_{\text{FA}}(\mathbf{n}_1)}. \quad (73)$$

**d)** Neyman–Pearson optimal SR noise does not exist if c) does not hold. But there does exist a noise pdf sequence  $\{f_{\text{N}_r}\}_{r=1}^{\infty}$  of the form (72)–(73) such that

$$\lim_{r \rightarrow \infty} P_{\text{D}}(f_{\text{N}_r}) = d(k^*). \quad (74)$$

*Proof:*

**Part a):** Definitions (14)–(15)  $d^+(k)$  and  $d^-(k)$  are continuous functions of  $k$ .  $d^+(k)$  is further an unbounded and decreasing function of  $k$ . But  $d^-(k)$  is a nondecreasing function of  $k$  because  $P_{\text{D}_{\text{sup}}} < P_{\text{D}_{\text{sup}}}$ . So there exists  $k^* > 1$  such

that  $d^+(k^*) = d^-(k^*)$  if  $d^+(1) > d^-(1)$ . There likewise exists  $0 \leq k^* < 1$  such that  $d^+(k^*) = d^-(k^*)$  if  $d^+(1) < d^-(1)$  because  $d^+(0) = P_{\text{D}_{\text{sup}}} > P_{\text{D}_{\text{sup}}} = d^-(0)$ . Thus there exists  $k^* \geq 0$  such that  $d^+(k^*) = d^-(k^*) = d(k^*)$  and  $\min\{d^+(k), d^-(k)\} \leq d(k^*) \leq \max\{d^+(k), d^-(k)\}$  for any  $k \geq 0$ .

**Part b):** Part a) above and (27) imply that

$$\min_{k \geq 0} \sup_{f_{\text{N}} \in \mathcal{F}} L(f_{\text{N}}, k) = \min_{k \geq 0} \max\{d^-(k), d^+(k)\} = d(k^*). \quad (75)$$

Let  $f_{\text{N}}$  be a noise pdf such that  $P_{\text{D}}(f_{\text{N}}) = d(k^*) > P_{\text{D}}(0)$  and  $P_{\text{FA}}(f_{\text{N}}) = \alpha$ . Then

$$L(f_{\text{N}}, k^*) = \int_{\mathcal{N}} (P_{\text{D}}(\mathbf{n}) - k^*(P_{\text{FA}}(\mathbf{n}) - \alpha)) f_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (76)$$

$$= \int_{\mathcal{N}} P_{\text{D}}(\mathbf{n}) f_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (77)$$

$$= P_{\text{D}}(f_{\text{N}}) = d(k^*) \quad (78)$$

$$= \min_{k \geq 0} \sup_{f_{\text{N}} \in \mathcal{F}} L(f_{\text{N}}, k) \quad \text{by (75)}. \quad (79)$$

So  $f_{\text{N}}$  is the optimal SR noise pdf.

**Part c):** Suppose that there exist  $\mathbf{n}_1 \in D^+$  and  $\mathbf{n}_2 \in D^-$  such that  $g(\mathbf{n}_1, k^*) = d(k^*) = g(\mathbf{n}_2, k^*)$ . Define

$$f_{\text{N}}(\mathbf{n}) = \lambda \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda) \delta(\mathbf{n} - \mathbf{n}_2) \quad (80)$$

$$\text{where } \lambda = \frac{P_{\text{FA}}(\mathbf{n}_2) - \alpha}{P_{\text{FA}}(\mathbf{n}_2) - P_{\text{FA}}(\mathbf{n}_1)}. \quad (81)$$

Then

$$P_{\text{D}}(f_{\text{N}}) = \int_{\mathcal{N}} P_{\text{D}}(\mathbf{n}) f_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (82)$$

$$= \int_{\mathcal{N}} P_{\text{D}}(\mathbf{n}) [\lambda \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda) \delta(\mathbf{n} - \mathbf{n}_2)] d\mathbf{n} \quad (83)$$

$$= \lambda P_{\text{D}}(\mathbf{n}_1) + (1 - \lambda) P_{\text{D}}(\mathbf{n}_2) \quad (84)$$

$$= \lambda d(k^*) + (1 - \lambda) d(k^*) = d(k^*) \quad (85)$$

and

$$P_{\text{FA}}(f_{\text{N}}) = \int_{\mathcal{N}} P_{\text{FA}}(\mathbf{n}) f_{\text{N}}(\mathbf{n}) d\mathbf{n} \quad (86)$$

$$= \int_{\mathcal{N}} P_{\text{FA}}(\mathbf{n}) [\lambda \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda) \delta(\mathbf{n} - \mathbf{n}_2)] d\mathbf{n} \quad (87)$$

$$= \lambda P_{\text{FA}}(\mathbf{n}_1) + (1 - \lambda) P_{\text{FA}}(\mathbf{n}_2) \quad (88)$$

$$= \frac{P_{\text{FA}}(\mathbf{n}_2) - \alpha}{P_{\text{FA}}(\mathbf{n}_2) - P_{\text{FA}}(\mathbf{n}_1)} P_{\text{FA}}(\mathbf{n}_1) + \left(1 - \frac{P_{\text{FA}}(\mathbf{n}_2) - \alpha}{P_{\text{FA}}(\mathbf{n}_2) - P_{\text{FA}}(\mathbf{n}_1)}\right) \times P_{\text{FA}}(\mathbf{n}_2) \quad (89)$$

$$= \alpha. \quad (90)$$

Then (85), (90), and the result of Part b) imply that  $f_{\text{N}}(\mathbf{n})$  is an optimal SR noise pdf. This optimal noise pdf is not unique if

there exist more than one pair of noise realizations that satisfy the hypothesis of Theorem 2c).

**Part d):** Define

$$H^+ = \{\mathbf{n} \in D^+ : g(\mathbf{n}, k^*) = d(k^*)\} \quad (91)$$

and

$$H^- = \{\mathbf{n} \in D^- : g(\mathbf{n}, k^*) = d(k^*)\}. \quad (92)$$

Suppose that  $H^+ \neq \emptyset$  but  $H^- = \emptyset$ . So there exists  $\mathbf{n}_1 \in D^+$  such that  $g(\mathbf{n}_1, k^*) = d(k^*)$  but there does not exist  $\mathbf{n} \in D^-$  such that  $g(\mathbf{n}, k^*) = d(k^*)$ . Then  $g(\mathbf{n}, k^*) < d^-(k^*) = d(k^*)$  for all  $\mathbf{n} \in D^-$  by the definition of  $d^-(k^*)$ . Suppose  $f'_N$  is the optimal SR noise pdf with  $P_{\text{FA}}(f'_N) = \alpha$ . Then

$$P_D(f'_N) = \int_{\mathcal{N}} g(\mathbf{n}, k^*) f'_N(\mathbf{n}) d(\mathbf{n}) \quad (93)$$

$$\leq \int_{D^+} g(\mathbf{n}, k^*) f'_N(\mathbf{n}) d(\mathbf{n}) + \int_{D^- \setminus D^0} g(\mathbf{n}, k^*) f'_N(\mathbf{n}) d(\mathbf{n}) \quad (94)$$

$$\leq \int_{D^+} d(k^*) f'_N(\mathbf{n}) d(\mathbf{n}) + \int_{D^- \setminus D^0} g(\mathbf{n}, k^*) f'_N(\mathbf{n}) d(\mathbf{n}) \quad (95)$$

$$< \int_{D^+} d(k^*) f'_N(\mathbf{n}) d(\mathbf{n}) + \int_{D^- \setminus D^0} d(k^*) f'_N(\mathbf{n}) d(\mathbf{n}) \quad (96)$$

because  $g(\mathbf{n}, k^*) < d(k^*)$  for all  $\mathbf{n} \in D^-$  and because

$$\int_{D^- \setminus D^0} [d(k^*) - g(\mathbf{n}, k^*)] f'_N(\mathbf{n}) d(\mathbf{n}) = 0$$

iff  $[d(k^*) - g(\mathbf{n}, k^*)] f'_N(\mathbf{n}) = 0$  a.e. on  $D^- \setminus D^0$   
so

$$P_D(f'_N) < d(k^*). \quad (97)$$

The supremum definition of  $d^-(k^*)$  and  $d(k^*) = d^-(k^*)$  imply [62] that there exists a sequence of noise realizations  $\{\mathbf{n}_r\}_{r=1}^{\infty}$  in  $D^-$  such that

$$\lim_{r \rightarrow \infty} g(\mathbf{n}_r, k^*) = d(k^*). \quad (98)$$

Now define a sequence of pdfs

$$f_{N_r}(\mathbf{n}) = \lambda_r \delta(\mathbf{n} - \mathbf{n}_1) + (1 - \lambda_r) \delta(\mathbf{n} - \mathbf{n}_r) \quad (99)$$

$$\text{where } \lambda_r = \frac{P_{\text{FA}}(\mathbf{n}_r) - \alpha}{P_{\text{FA}}(\mathbf{n}_r) - P_{\text{FA}}(\mathbf{n}_1)}. \quad (100)$$

Then  $P_{\text{FA}}(f_{N_r}) = \alpha$  for all  $r$  and

$$\lim_{r \rightarrow \infty} P_D(f_{N_r}) = \lim_{r \rightarrow \infty} \int_{\mathcal{N}} g(\mathbf{n}, k^*) f_{N_r}(\mathbf{n}) d(\mathbf{n}) \quad (101)$$

$$= \lim_{r \rightarrow \infty} [\lambda_r g(\mathbf{n}_1, k^*) + (1 - \lambda_r) g(\mathbf{n}_r, k^*)] \quad (102)$$

$$= \lim_{r \rightarrow \infty} [\lambda_r d(k^*) + (1 - \lambda_r) g(\mathbf{n}_r, k^*)] \quad (103)$$

$$= d(k^*) \lim_{r \rightarrow \infty} \lambda_r + \lim_{r \rightarrow \infty} g(\mathbf{n}_r, k^*) \lim_{r \rightarrow \infty} (1 - \lambda_r) \quad (104)$$

$$= d(k^*) \lim_{r \rightarrow \infty} \lambda_r + d(k^*) \lim_{r \rightarrow \infty} (1 - \lambda_r) \quad (105)$$

$$= d(k^*). \quad (106)$$

Hence  $P_{\text{FA}}(f_{N_r}) = \alpha$  and there exists a positive integer  $l$  such that  $P_D(f_{N_r}) > P_D(f'_N)$  for all  $r \geq l$ . This contradicts the optimality of  $f'_N$ . So optimal SR noise does not exist if  $H^- = \emptyset$  and  $H^+ \neq \emptyset$ . Similar arguments also prove the nonexistence of optimal SR noise in the more general case when either  $H^+ = \emptyset$  or  $H^- = \emptyset$ .  $\square$

*Corollary 1:*

Suppose that  $P_{D-\text{sup}} < P_{D^+\text{sup}}$  and that  $P_D$  and  $P_{\text{FA}}$  are differentiable in the interior of the noise space  $\mathcal{N}$ .

**a)** Suppose that  $f_{N_{\text{opt}}}$  is an optimal N-P SR noise pdf of the form (72)–(73) in Theorem 2c) and that  $\mathbf{n}_1$  and  $\mathbf{n}_2$  of (72)–(73) are the interior points of  $\mathcal{N}$ . Then  $\mathbf{n}_1$  and  $\mathbf{n}_2$  satisfy

$$P_D(\mathbf{n}_1) - k P_{\text{FA}}(\mathbf{n}_1) = P_D(\mathbf{n}_2) - k P_{\text{FA}}(\mathbf{n}_2) \quad (107)$$

$$\nabla P_D(\mathbf{n}_1) - k \nabla P_{\text{FA}}(\mathbf{n}_1) = \mathbf{0} \quad (108)$$

$$\nabla P_D(\mathbf{n}_2) - k \nabla P_{\text{FA}}(\mathbf{n}_2) = \mathbf{0} \quad (109)$$

for some  $k \geq 0$ .

**b)** Suppose further that for each  $k \geq 0$  at most one solution of  $\nabla P_D(\mathbf{n}) - k \nabla P_{\text{FA}}(\mathbf{n}) = \mathbf{0}$  in  $\mathbf{R}^m$  is a global maximum of  $P_D(\mathbf{n}) - k(P_{\text{FA}}(\mathbf{n}) - \alpha)$ . Then  $f_{N_{\text{opt}}} = \delta(\mathbf{n} - \mathbf{n}_\alpha)$  is the optimal N-P SR noise pdf if such a solution  $\mathbf{n}_\alpha$  exists in  $D^0$ . There is otherwise no optimal N-P SR noise in the interior of  $\mathcal{N}$ .

*Proof:*

**Part a):** Say that  $f_{N_{\text{opt}}}$  is an optimal SR noise pdf of the form (72)–(73) in Theorem 2c). Then  $g(\mathbf{n}_1, k^*) = d(k^*) = g(\mathbf{n}_2, k^*)$  for some  $k \geq 0$  and so (107) follows. The definition of  $d(k^*)$  implies that  $\mathbf{n}_1$  and  $\mathbf{n}_2$  are maximal points of  $g(\mathbf{n}, k)$  for  $k = k^*$ . So  $\mathbf{n}_1$  and  $\mathbf{n}_2$  satisfy (108)–(109) for  $k = k^*$  if  $P_D$  and  $P_{\text{FA}}$  are differentiable in the interior of  $\mathcal{N}$  and if  $\mathbf{n}_1$  and  $\mathbf{n}_2$  of (72)–(73) are the interior points of  $\mathcal{N}$ .

**Part b):** Theorem 2a) implies that there exists  $k^* \geq 0$  such that  $d^-(k^*) = d(k^*) = d^+(k^*)$ . Now either there exist  $\mathbf{n}_1 \in D^+$  and  $\mathbf{n}_2 \in D^-$  such that  $g(\mathbf{n}_1, k^*) = d^+(k^*) = d(k^*) = g(\mathbf{n}_2, k^*) = d^-(k^*)$  or else the optimal SR noise does not exist by Theorem 2d). The former case implies that  $\mathbf{n}_1$  and  $\mathbf{n}_2$  are solutions of  $\nabla P_D(\mathbf{n}) - k^* \nabla P_{\text{FA}}(\mathbf{n}) = \mathbf{0}$ . But the hypothesis implies that at most one solution of  $\nabla P_D(\mathbf{n}) - k \nabla P_{\text{FA}}(\mathbf{n}) = \mathbf{0}$

in  $\mathbf{R}^m$  is a global maximum of  $P_D(\mathbf{n}) - k(P_{FA}(\mathbf{n}) - \alpha)$  for each  $k \geq 0$ . So  $\mathbf{n}_1 = \mathbf{n}_2 =$  (say)  $\mathbf{n}_\alpha \in D^0 = D^- \cap D^+$ . Then  $P_{FA}(\mathbf{n}_\alpha) - \alpha = 0$  and  $f_{N_{opt}} = \delta(\mathbf{n} - \mathbf{n}_\alpha)$  is the optimal noise pdf by Theorem 2c).  $\square$

*Theorem 3:*

Let the detection and false-alarm probabilities  $P_D$  and  $P_{FA}$  be real-valued functions that are differentiable in a neighborhood of 0. Suppose that  $P_D''(0)$  and  $P_{FA}''(0)$  exist and that  $P_{FA}(0) = \alpha$ . Suppose also that  $P_{FA}$  does not have a local minimum at 0 and that  $P_D$  does not have a local maximum at 0. Then an N-P SR noise exists if

$$P_D''(0) |P_{FA}'(0)| > P_{FA}''(0) |P_D'(0)| \quad (110)$$

or if  $\text{sgn}(P_{FA}'(0))\text{sgn}(P_D'(0)) \leq 0$ .

*Proof:*

The local version of Taylor's theorem [63] gives

$$P_D(r) - P_D(0) = P_D'(0)r + P_D''(0)\frac{r^2}{2} + \eta_1(r) \quad (111)$$

where  $\eta_1(r)/r^2 \rightarrow 0$  as  $r \rightarrow 0$ . Likewise

$$\begin{aligned} P_D(0) - P_D(-r) \\ = P_D'(0)r - P_D''(0)\frac{r^2}{2} - \eta_2(r) \end{aligned} \quad (112)$$

$$\begin{aligned} P_{FA}(r) - P_{FA}(0) \\ = P_{FA}'(0)r + P_{FA}''(0)\frac{r^2}{2} + \eta_3(r) \end{aligned} \quad (113)$$

$$\begin{aligned} P_{FA}(0) - P_{FA}(-r) \\ = P_{FA}'(0)r - P_{FA}''(0)\frac{r^2}{2} - \eta_4(r) \end{aligned} \quad (114)$$

where  $\eta_i(r)/r^2 \rightarrow 0$  as  $r \rightarrow 0$  for  $i = 2, 3$ , and 4.

Now define

$$\eta_{\max}(r) = \max\{\eta_1(r), \eta_2(r), \eta_3(r), \eta_4(r)\}. \quad (115)$$

Suppose first that both  $P_{FA}$  and  $P_D$  increase at 0. Hence  $0 < P_D''(0)P_{FA}'(0) - P_{FA}''(0)P_D'(0)$  because of (110). Then (111)–(115) imply that there exists  $\zeta > 0$  such that for all  $0 < r < \zeta$

$$\frac{2\eta_{\max}(r)[P_D'(0) + P_{FA}'(0)]}{r^2} < P_D''(0)P_{FA}'(0) - P_{FA}''(0)P_D'(0) \quad (116)$$

and hence

$$\begin{aligned} \eta_{\max}(r)[P_D'(0)r + P_{FA}'(0)r] \\ < P_D''(0)\frac{r^2}{2}P_{FA}'(0)r - P_{FA}''(0)\frac{r^2}{2}P_D'(0)r. \end{aligned} \quad (117)$$

Rewrite the above inequality as

$$2P_D'(0)rP_{FA}''(0)\frac{r^2}{2} + 2P_D'(0)r\eta_{\max}(r)$$

$$< 2P_{FA}'(0)rP_D''(0)\frac{r^2}{2} - 2P_{FA}'(0)r\eta_{\max}(r). \quad (118)$$

This is equivalent to

$$\begin{aligned} [P_{FA}'(0)r + P_{FA}''(0)\frac{r^2}{2} + \eta_{\max}(r)] \\ \times [P_D'(0)r - P_D''(0)\frac{r^2}{2} + \eta_{\max}(r)] \\ < [P_{FA}'(0)r - P_{FA}''(0)\frac{r^2}{2} - \eta_{\max}(r)] \\ \times [P_D'(0)r + P_D''(0)\frac{r^2}{2} - \eta_{\max}(r)]. \end{aligned} \quad (119)$$

Then

$$\begin{aligned} [P_{FA}(r) - P_{FA}(0)][P_D(0) - P_D(-r)] \\ < [P_{FA}(0) - P_{FA}(-r)][P_D(r) - P_D(0)] \end{aligned} \quad (120)$$

for all  $0 < r < \zeta$  because of (111)–(115). Rewrite (120) as

$$\begin{aligned} P_D(0) < \frac{[P_{FA}(r) - P_{FA}(0)]}{[P_{FA}(r) - P_{FA}(-r)]}P_D(-r) \\ + \frac{[P_{FA}(0) - P_{FA}(-r)]}{[P_{FA}(r) - P_{FA}(-r)]}P_D(r). \end{aligned} \quad (121)$$

This inequality implies that

$$f_N(n) = \lambda\delta(n+r) + (1-\lambda)\delta(n-r) \quad (122)$$

$$\text{with } \lambda = \frac{[P_{FA}(r) - P_{FA}(0)]}{[P_{FA}(r) - P_{FA}(-r)]} \quad (123)$$

is an SR noise pdf.

Suppose now that both  $P_{FA}$  and  $P_D$  decrease at 0. Hence  $0 < P_{FA}''(0)P_D'(0) - P_D''(0)P_{FA}'(0)$  and there exists  $\zeta > 0$  such that for all  $0 < r < \zeta$

$$\begin{aligned} \frac{-2\eta_{\max}(r)[P_D'(0) + P_{FA}'(0)]}{r^2} \\ < P_{FA}''(0)P_D'(0) - P_D''(0)P_{FA}'(0). \end{aligned} \quad (124)$$

Then similar arguments show that (122)–(123) again give an SR noise pdf.

All other cases obey either  $P_D(0) < P_D(-r)$  and  $P_{FA}(0) \geq P_{FA}(-r)$  or  $P_D(0) < P_D(r)$  and  $P_{FA}(0) \geq P_{FA}(r)$  because  $P_{FA}$  does not have a local minimum at 0 and because  $P_D$  does not have a local maximum at 0 by hypothesis. So either  $f_N(n) = \delta(n+r)$  or  $f_N(n) = \delta(n-r)$  gives an SR noise pdf.  $\square$

*Theorem 4:*

**a)** Pick any  $\epsilon > 0$ . Then the above algorithm finds an N-P SR noise  $\tilde{N}'$  from  $\tilde{N}$  in at most  $i_{\max} = \lceil \log_2(2/\epsilon) \rceil + 1$  iterations so that

$$P_D(f_{\tilde{N}'_{opt}}) \geq P_D(f_{\tilde{N}'}) \geq P_D(f_{\tilde{N}'_{opt}}) - \epsilon \quad \text{and} \quad (125)$$

$$P_{FA}(f_{\tilde{N}'}) \leq \alpha. \quad (126)$$

**b)** The suboptimal detection performance with noise  $\tilde{N}'$  is at most  $\tau + \epsilon$  less than the optimal SR detection with noise  $N_{\text{opt}}$  if  $\tilde{N}'$  satisfies (38)–(39):

$$P_{\text{D}}(f_{N_{\text{opt}}}) \geq P_{\text{D}}(f_{\tilde{N}'}) \geq P_{\text{D}}(f_{N_{\text{opt}}}) - (\tau + \epsilon). \quad (127)$$

*Proof:*

**Part a):** The set  $G^-$  is nonempty and finite because the noise realization vector  $\tilde{N}$  is finite. Suppose that  $P_{\text{D}^-} \geq P_{\text{D}^+}$ . Then Theorem 1a) implies that the optimal noise  $\tilde{N}_{\text{opt}}$  in  $\tilde{N}$  has the form  $f_{\tilde{N}_{\text{opt}}}(n) = \delta(n - \tilde{n}_0)$  for any  $\tilde{n}_0 \in G^-$ . The algorithm finds such noise if  $P_{\text{D}^-} \geq P_{\text{D}^+}$ .

Suppose now that  $P_{\text{D}^-} < P_{\text{D}^+}$ . Then there exists  $k^* \in 0$  such that  $d^+(k^*) = d^-(k^*) = d(k^*)$  by Theorem 2a). Also there exist  $\tilde{n}_1^* \in D^-$  and  $\tilde{n}_2^* \in D^+$  such that  $g(\tilde{n}_1^*, k^*) = d^-(k^*) = d(k^*) = d^+(k^*) = g(\tilde{n}_2^*, k^*)$  because  $\tilde{N}$  is finite. Then Theorem 2c) implies that the optimal SR noise  $\tilde{N}_{\text{opt}}$  in  $\tilde{N}$  has the pdf  $f_{\tilde{N}_{\text{opt}}}$  of the form (72).  $P_{\text{D}}(f_{\tilde{N}_{\text{opt}}}) = \alpha$  by Theorem 1b) while Theorem 2b) implies that  $P_{\text{D}}(f_{\tilde{N}_{\text{opt}}}) = d(k^*)$ . The algorithm finds an SR noise  $\tilde{N}'$  from  $\tilde{N}$  and its pdf  $f_{\tilde{N}'}$  is of the form (\*) in the algorithm. This pdf  $f_{\tilde{N}'}$  satisfies (126) with equality. So we need to show only that  $f_{\tilde{N}'}$  satisfies (125).

We first show that the SR noise pdf  $f'_{\tilde{N}}$  satisfies (125) if  $|d^-(k(i)) - d(k^*)| \leq \epsilon$  and  $|d^+(k(i)) - d(k^*)| \leq \epsilon$  for some  $i$ . Theorem 2a) implies that  $d^+(k(i)) \geq d(k^*)$  if  $d^-(k(i)) \leq d(k^*)$  and  $d^+(k(i)) \leq d(k^*)$  if  $d^-(k(i)) \geq d(k^*)$ . So suppose first that  $d(k^*) - d^-(k(i)) \leq \epsilon$  and  $d^+(k(i)) \geq d(k^*)$ . Let  $\tilde{n}_1 \in D^-$  and  $\tilde{n}_2 \in D^+$  such that  $g(\tilde{n}_1, k(i)) = d^-(k(i))$  and  $g(\tilde{n}_2, k^*) = d^+(k(i))$ . Note that  $P_{\text{FA}}(f'_{\tilde{N}}) = \int_{\tilde{N}} P_{\text{FA}}(n) f'_{\tilde{N}}(n) d(n) = \alpha$ . Then

$$P_{\text{D}}(f'_{\tilde{N}}) = \int_{\tilde{N}} (P_{\text{D}}(n) - k(P_{\text{FA}}(n) - \alpha)) f'_{\tilde{N}}(n) d(n) \quad (128)$$

$$= \lambda g(\tilde{n}_1, k(i)) + (1 - \lambda) g(\tilde{n}_2, k(i)) \quad (129)$$

$$= \lambda d^-(k(i)) + (1 - \lambda) d^+(k(i)) \quad (130)$$

$$\geq \lambda(d(k^*) - \epsilon) + (1 - \lambda)d(k^*) \quad (131)$$

$$\text{because } d(k^*) - d^-(k(i)) \leq \epsilon$$

$$\text{and } d^+(k(i)) \geq d(k^*)$$

$$= d(k^*) - \lambda\epsilon \geq d(k^*) - \epsilon \quad (132)$$

$$\text{because } 0 \leq \lambda \leq 1$$

$$= P_{\text{D}}(f_{\tilde{N}_{\text{opt}}}) - \epsilon. \quad (133)$$

Similar arguments show that  $P_{\text{D}}(f'_{\tilde{N}}) \geq d(k^*) - \epsilon$  if  $d(k^*) - d^+(k(i)) \leq \epsilon$  and  $d^-(k(i)) \geq d(k^*)$ . So we need show only that  $|d^-(k(i)) - d(k^*)| \leq \epsilon$  and  $|d^+(k(i)) - d(k^*)| \leq \epsilon$  for some  $i$ .

We now show that  $|d^-(k(i)) - d(k^*)| \leq 2/2^i$  for all  $i$ . This implies that  $|d^-(k(i)) - d(k^*)| \leq \epsilon$  for  $i = i_{\text{stop}}$  where  $i_{\text{stop}} = \lceil \log_2(2/\epsilon) \rceil$ . We then prove that  $|d^+(k(i)) - d(k^*)| \leq \epsilon$  for  $i = i_{\text{stop}} + 1$ .

We use mathematical induction in  $i$  to prove that  $|d^-(k(i)) - d(k^*)| \leq 2/2^i$  for all  $i$ .

**Basis Step ( $i = 1$ ):** The definition of  $dr(i)$  gives  $dr(1) = ds(1) + df(1)/2$  where  $ds(1) = d^-(k(0))$ ,  $df(1) = d^+(k(0))$ ,

and  $k(0) = 1$ . Both  $ds(1)$  and  $df(1)$  are between 0 and 2 by hypothesis. So  $|ds(1) - df(1)| \leq 2$ . Then  $|dr(1) - ds(1)| = |dr(1) - df(1)| \leq 2/2$ . Theorem 2a) implies that  $d(k^*)$  is between  $d^-(k(0)) (= ds(1))$  and  $d^+(k(0)) (= df(1))$ . Then  $|dr(1) - d(k^*)| \leq |dr(1) - ds(1)| \leq 2/2$ . And  $d^-(k(1)) = dr(1)$  holds in the algorithm because  $k(1) = \min\{(P_{\text{D}}(\tilde{n}) - dr(1))/(P_{\text{FA}}(\tilde{n}) - \alpha) : \tilde{n} \in D^- \setminus D^0\}$ . Hence  $|d^-(k(1)) - d(k^*)| \leq 2/2$ .

**Induction Hypothesis ( $i = m$ ):** Suppose that  $|ds(m) - df(m)| \leq 2/2^{m-1}$ . Then the definition of  $dr(i)$  implies that  $|dr(m) - ds(m)| = |dr(m) - df(m)| \leq 2/2^m$ . Suppose also that  $d(k^*)$  lies between  $ds(m)$  and  $df(m)$ . Then  $|dr(m) - d(k^*)| \leq 2/2^m$ . Note that  $d^-(k(m)) = dr(m)$  where  $k(m) = \min\{(P_{\text{D}}(\tilde{n}) - dr(m))/(P_{\text{FA}}(\tilde{n}) - \alpha) : \tilde{n} \in D^- \setminus D^0\}$  in the algorithm. So  $|d^-(k(m)) - d(k^*)| \leq 2/2^m$ .

**Induction Step ( $i = m + 1$ ):**  $d(k^*)$  lies between  $d^-(k(m))$  and  $d^+(k(m))$  by Theorem 2a). Suppose that  $d^+(k(m)) > d^-(k(m))$ .  $d(k^*)$  also lies between  $d^-(k(m))$  and  $\max\{ds(m), df(m)\}$  because  $d^-(k(m)) = dr(m) = ds(m) + df(m)/2$  by definition. So  $d(k^*)$  lies between  $\min\{d^+(k(m)), \max\{ds(m), df(m)\}\}$  and  $d^-(k(m))$  and where  $d^-(k(m)) \leq \min\{d^+(k(m)), \max\{ds(m), df(m)\}\}$ . The algorithm defines  $ds(m+1) = d^-(k(m)) = dr(m)$  and  $df(m+1) = \min\{d^+(k(m)), \max\{ds(m), df(m)\}\}$ . Then  $dr(m) < df(m+1)$  and  $d(k^*)$  lies between  $ds(m+1)$  and  $df(m+1)$ .

Write  $|ds(m+1) - df(m+1)| = |dr(m) - df(m+1)| \leq |dr(m) - \max\{ds(m), df(m)\}| \leq 2/2^m$ . The last inequality follows from the induction hypothesis while the first inequality follows from the definition of  $df(m+1)$  and the fact that  $dr(m) < df(m+1)$ . Then  $|d^-(k(m+1)) - df(m+1)| = |dr(m+1) - df(m+1)| = |dr(m+1) - ds(m+1)| = |ds(m+1) - df(m+1)|/2 \leq 2/2^{m+1}$  because  $d^-(k(m+1)) = dr(m+1) = ds(m+1) + df(m+1)/2$  and  $|ds(m+1) - df(m+1)| \leq 2/2^m$ . This proves the induction claim  $|d^-(k(m+1)) - d(k^*)| \leq 2/2^{m+1}$  because  $d(k^*)$  lies between  $ds(m+1)$  and  $df(m+1)$ . Similar arguments prove the induction claim when  $d^+(k(m+1)) < d^-(k(m+1))$ .

We have so far proved that  $|d^-(k(i)) - d(k^*)| \leq \epsilon$  for  $i = i_{\text{stop}}$ . But this need not imply that  $|d^+(k(i_{\text{stop}})) - d^-(k(i_{\text{stop}}))| \leq \epsilon$ . The algorithm finds  $k(i_{\text{stop}} + 1)$  such that  $d^+(k(i_{\text{stop}} + 1)) = d^-(k(i_{\text{stop}})) + t\epsilon$  where  $t = \text{sgn}[d^+(k(i_{\text{stop}})) - d^-(k(i_{\text{stop}}))]$  if  $|d^+(k(i_{\text{stop}})) - d^-(k(i_{\text{stop}}))| > \epsilon$ . This implies that  $|d^-(k(i_{\text{stop}} + 1)) - d^+(k(i_{\text{stop}} + 1))| \leq \epsilon$  because  $d(k^*)$  lies between  $d^-(k(i_{\text{stop}} + 1))$  and  $d^+(k(i_{\text{stop}} + 1))$  and because  $|d^-(k(i_{\text{stop}} + 1)) - d(k^*)| \leq \epsilon$ .

**Part b):** The optimal SR noise pdf  $f_{N_{\text{opt}}}$  is of the form (72)–(73) and so  $P_{\text{FA}}(f_{N_{\text{opt}}}) = \lambda P_{\text{FA}}(n_1) + (1 - \lambda) P_{\text{FA}}(n_2) = \alpha$ . There exist  $\tilde{n}_1$  and  $\tilde{n}_2$  in  $\tilde{N}$  that satisfy (38)–(39) by hypothesis. Define  $\tilde{N}''$  as a noise restricted to  $\tilde{N}$  with pdf

$$f_{\tilde{N}''}(n) = \lambda \delta(n - \tilde{n}_1) + (1 - \lambda) \delta(n - \tilde{n}_2). \quad (134)$$

Then

$$\alpha = P_{\text{FA}}(f_{N_{\text{opt}}}) \geq \lambda P_{\text{FA}}(\tilde{n}_1) + (1 - \lambda) P_{\text{FA}}(\tilde{n}_2) \quad (135)$$

because of (39)

$$= P_{\text{FA}}(f_{\tilde{N}''}) \text{ because of (134)} \quad (136)$$

and

$$P_{\text{D}}(f_{N_{\text{opt}}}) = \lambda P_{\text{D}}(n_1) + (1 - \lambda) P_{\text{D}}(n_2) \quad (137)$$

$$\leq \lambda (P_{\text{D}}(\tilde{n}_1) + \tau) + (1 - \lambda) \times (P_{\text{D}}(\tilde{n}_2) + \tau) \quad (138)$$

because of (38)

$$= P_{\text{D}}(f_{\tilde{N}''}) + \tau \quad \text{because of (134).} \quad (139)$$

Inequalities (136) and (139) and the fact that  $\tilde{N}''$  is restricted to  $\tilde{N}$  imply that

$$P_{\text{D}}(f_{\tilde{N}_{\text{opt}}}) \geq P_{\text{D}}(f_{\tilde{N}''}) \geq P_{\text{D}}(f_{N_{\text{opt}}}) - \tau \quad (140)$$

if  $\tilde{N}_{\text{opt}}$  is the optimal SR noise restricted to  $\tilde{N}$  such that  $P_{\text{FA}}(f_{\tilde{N}_{\text{opt}}}) \leq \alpha$ . Then inequality (127) follows from (140) and the result of Part a).  $\square$

#### REFERENCES

- [1] S. M. Bezrukov and I. Vodyanov, "Noise-induced enhancement of signal transduction across voltage-dependent ion channels," *Nature*, vol. 378, pp. 362–364, Nov. 1995.
- [2] A. R. Bulsara, A. J. Maren, and G. Schnera, "Single effective neuron: Dendritic coupling effect and stochastic resonance," *Biol. Cybern.*, vol. 70, no. 2, pp. 145–156, Dec. 1993.
- [3] J. J. Collins, T. T. Imhoff, and P. Grigg, "Noise enhanced information transmission in rat sa1 cutaneous mechanoreceptors via aperiodic stochastic resonance," *J. Neurophys.*, vol. 76, no. 1, pp. 642–645, Jul. 1996.
- [4] T. Ditzinger, M. Stadler, D. Struber, and J. A. Kelso, "Noise improves three-dimensional perception: Stochastic resonance and other impacts of noise to the perception of autostereograms," *Phys. Rev. E*, vol. 62, no. 2, pp. 2566–2575, Aug. 2000.
- [5] J. K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, "Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance," *Nature*, vol. 365, pp. 337–340, Sep. 1993.
- [6] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," *Phys. Rev. E*, vol. 52, no. 5, pp. 4691–4698, Nov. 1995.
- [7] P. Jung, A. Cornell-Bell, F. Moss, S. Kadar, J. Wang, and K. Showalter, "Noise sustained waves in subexcitable media: From chemical waves to brain waves," *Chaos*, vol. 8, pp. 567–575, Sep. 1998.
- [8] K. Kitajo, D. Nozaki, L. M. Ward, and Y. Yamamoto, "Behavioral stochastic resonance within the human brain," *Phys. Rev. Lett.*, vol. 90, no. 21, pp. 218 103-1–218 103-4, May 2003.
- [9] J. E. Levin and J. P. Miller, "Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance," *Nature*, vol. 380, pp. 165–168, Mar. 1996.
- [10] F. Moss, L. M. Ward, and W. G. Sannita, "Stochastic resonance and sensory information processing: A tutorial and review of applications," *Clin. Neurophys.*, vol. 115, no. 2, pp. 267–281, Feb. 2004.
- [11] A. Patel and B. Kosko, "Stochastic resonance in continuous and spiking neuron models with Levy noise," *IEEE Trans. Neural Netw.*, vol. 19, pp. 1993–2008, Dec. 2008.
- [12] W. C. Stacey and D. M. Durand, "Stochastic resonance improves signal detection in hippocampal CA1 neurons," *J. Neurophys.*, vol. 83, no. 3, pp. 1394–1402, Mar. 2000.
- [13] R. A. Wannamaker, S. P. Lipshitz, and J. Vanderkooy, "Stochastic resonance as dithering," *Phys. Rev. E*, vol. 61, no. 1, pp. 233–236, Jan. 2000.
- [14] K. Wiesenfeld and F. Moss, "Stochastic resonance and the benefits of noise: From ice ages to crayfish and squids," *Nature*, vol. 373, pp. 33–36, Jan. 1995.
- [15] I. Goychuk and P. Hänggi, "Quantum stochastic resonance in parallel," *New J. Phys.*, vol. 1, pp. 14.1–14.14, Aug. 1999.
- [16] P. Hänggi, "Stochastic resonance in biology: How noise can enhance detection of weak signals and help improve biological information processing," *Chemphyschem*, vol. 3, pp. 285–290, Mar. 2002.
- [17] G. P. Harmer, B. R. Davis, and D. Abbott, "A review of stochastic resonance: Circuits and measurement," *IEEE Trans. Instrum. Meas.*, vol. 51, pp. 299–309, Apr. 2002.
- [18] F. Jaramillo and K. Wiesenfeld, "Mechano-electrical transduction assisted by Brownian motion: A role for noise in the auditory system," *Nature Neurosci.*, vol. 1, pp. 384–388, Sep. 1998.
- [19] G. G. Karapetyan, "Application of stochastic resonance in gravitational-wave interferometer," *Phys. Rev. D*, vol. 73, pp. 122 003-1–122 003-9, Jun. 2006.
- [20] B. Kosko, *Noise*. New York: Viking/Penguin, 2006.
- [21] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [22] I. Y. Lee, X. Liu, C. Zhou, and B. Kosko, "Noise-enhanced detection of subthreshold signals with carbon nanotubes," *IEEE Trans. Nanotechnol.*, vol. 5, pp. 613–627, Nov. 2006.
- [23] W. B. Levy and R. A. Baxter, "Energy-efficient neural computation via quantal synaptic failures," *J. Neurosci.*, vol. 22, no. 11, pp. 4746–4755, Jun. 2002.
- [24] Y. Sakumura and K. Aihara, "Stochastic resonance and coincidence detection in a single neuron," *Neural Process. Lett.*, vol. 16, no. 3, pp. 235–242, Dec. 2002.
- [25] N. G. Stocks, D. Appligam, and R. P. Morse, "The application of suprathreshold stochastic resonance to cochlear implant coding," *Fluctuat. Noise Lett.*, vol. 2, no. 3, pp. 169–181, Sep. 2002.
- [26] F. G. Zeng, Q. J. Fu, and R. Morse, "Human hearing enhanced by noise," *Brain Res.*, vol. 869, no. 2, pp. 251–255, Jun. 2000.
- [27] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, "Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance," *Phys. Lett. A*, vol. 352, pp. 183–189, Mar. 2006.
- [28] S. Mitaïm and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Trans. Neural Netw.*, vol. 15, pp. 1526–1540, Nov. 2004.
- [29] A. Patel and B. Kosko, "Stochastic resonance in noisy spiking retinal and sensory neuron models," *Neural Netw.*, vol. 18, pp. 467–478, Aug. 2005.
- [30] S. Mitaïm and B. Kosko, "Adaptive stochastic resonance," *Proc. IEEE*, vol. 86, pp. 2152–2183, Nov. 1998.
- [31] A. A. Saha and G. V. Anand, "Design of detectors based on stochastic resonance," *Signal Process.*, vol. 86, no. 3, pp. 1193–1212, Jun. 2003.
- [32] S. Zozor and P.-O. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, pp. 3177–3181, Dec. 2003.
- [33] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of stochastic resonance effects in signal detection: Part I—fixed detectors," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3172–3184, Jul. 2007.
- [34] S. Kay, "Can detectability be improved by adding noise?," *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.
- [35] D. Rousseau, G. V. 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, Nov. 2006.
- [36] M. Guerriero, P. Willett, S. Marano, and V. Matta, "Speedier sequential tests via stochastic resonance," in *Proc. 33rd Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2008, pp. 3901–3904.
- [37] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer, 2008.
- [38] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965, p. 220.
- [39] N. Benvenuto and G. Cherubini, *Algorithms for Communications Systems and their Applications*. New York: Wiley, 2002, p. 442.
- [40] J. N. Tsitsiklis, *Decentralized Detection, Advances in Statistical Signal Processing*, H. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI Press, 1993, pp. 297–344, Proposition 3.7.
- [41] M. J. J. Scott, M. Niranjana, and R. W. Prager, "Realisable classifiers: Improving operating performance on variable cost problems," in *Proc. 9th Brit. Machine Vision Conf.*, Sep. 1998, vol. 1, pp. 305–315.
- [42] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 693–702, Apr. 2005.

[43] J.-B. Hiriart-Urruty and C. Lemarchal, *Fundamentals of Convex Analysis*. New York: Springer, 2004, p. 31.

[44] K. C. Border, *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1989, p. 10.

[45] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson signal detection," in *Proc. 33rd Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Apr. 2008, pp. 3889-3892.

[46] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer-Verlag, 1988.

[47] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations*. Boston, MA: Birkhäuser, 2001.

[48] Y. Ching and L. Kurz, "Nonparametric detectors based on m-interval partitioning," *IEEE Trans. Inf. Theory*, vol. 18, pp. 251-257, 1972.

[49] J. D. Gibson and J. L. Melsa, *Introduction to Nonparametric Detection With Applications*. Piscataway, NJ: IEEE Press, 1996.

[50] D. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley-Interscience, 1969, pp. 224-225.

[51] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[52] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton Univ. Press, 1944.

[53] R. D. Luce and H. Raiffa, *Games and Decisions*. New York: Wiley, 1957.

[54] G. Owen, *Game Theory*, 2nd ed. Norwell, MA: Academic, 1982.

[55] M. Azizoglu, "Convexity property in binary detection," *IEEE Trans. Inf. Theory*, vol. 42, no. 4, pp. 1316-1321, Jul. 1996.

[56] L. Huang and M. Neely, "The optimality of two prices: Maximizing revenue in a stochastic network," in *Proc. 45th Annu. Allerton Conf. Communication, Control, Computing*, Sep. 2007.

[57] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Process.*, vol. 87, pp. 2569-2586, Nov. 2007.

[58] Y. Gong, Z. Ding, and C. F. N. Cowan, "MMSE turbo equalizer for channels with cochannel interference," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, May 2006, vol. 4, pp. 437-440.

[59] N. Veselinovic, T. Matsumoto, and M. Juntti, "Iterative pdf estimation-based multiuser diversity detection and channel estimation with unknown interference," *EURASIP J. Appl. Signal Process.*, vol. 6, pp. 872-882, Jan. 2005.

[60] F. Chapeau-Blondeau and D. Rousseau, "Noise improvements in stochastic resonance: From signal amplification to optimal detection," *Fluctuat. Noise Lett.*, vol. 2, no. 3, pp. L221-L233, Sep. 2002.

[61] F. Chapeau-Blondeau and D. Rousseau, "Constructive action of additive noise in optimal detection," *Int. J. Bifurcat. Chaos*, vol. 15, no. 9, pp. 2985-2994, Sep. 2005.

[62] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976, p. 35.

[63] T. W. Körner, *A Companion to Analysis: A Second First and First Second Course in Analysis*. Providence, RI: Amer. Math. Soc., 2004, pp. 135-138.



**Ashok Patel** received the B.E. degree and M.E. degrees in electrical engineering from Gujarat University, Ahmedabad, India, and the M.A. degree in applied mathematics and M.S. degree in electrical engineering both from the University of Southern California (USC), Los Angeles. He is completing the Ph.D. degree in the Department of Electrical Engineering at USC.

His current research interests are in the areas of nonlinear and statistical signal processing.



**Bart Kosko** received degrees in philosophy, economics, mathematics, electrical engineering, and law.

He is a Professor of electrical engineering at the University of Southern California (USC), Los Angeles, a past Director of USC's Signal and Image Processing Institute, and a licensed attorney. He has published the textbooks *Neural Networks and Fuzzy Systems* and *Fuzzy Engineering*, the trade books *Fuzzy Thinking* and *Heaven in a Chip*, and the novel *Nanotime*. He edited the volume *Neural Networks*

*for Signal Processing* and co-edited the volume *Intelligent Signal Processing*. His most recent book is *Noise* (Viking, 2006).

Dr. Kosko is an elected governor of the International Neural Networks Society. He has chaired or cochaired several international conferences on neural and fuzzy systems and sits on the editorial board of several scientific journals.