Noise Benefits in Quantizer-Array Correlation Detection and Watermark Decoding

Ashok Patel, *Member, IEEE*, and Bart Kosko, *Fellow, IEEE*

***Abstract*—Quantizer noise can improve statistical signal detection in array-based nonlinear correlators in Neyman-Pearson and maximum-likelihood (ML) detection. This holds even for infinite-variance symmetric alpha-stable channel noise and for generalized-Gaussian channel noise. Noise-enhanced correlation detection leads to noise-enhanced watermark extraction based on such nonlinear detection at the pixel or bit level. This yields a noise-based algorithm for digital watermark decoding using two new noise-benefit theorems. The first theorem gives a necessary and sufficient condition for quantizer noise to increase the detection probability of a constant signal for a fixed false-alarm probability if the channel noise is symmetric and if the sample size is large. The second theorem shows that the array must contain more than one quantizer for such a stochastic-resonance noise benefit if the symmetric channel noise is unimodal. It also shows that the noise-benefit rate improves in the small-quantizer noise limit as the number of array quantizers increases. The second theorem further shows that symmetric uniform quantizer noise gives the optimal rate for an initial noise benefit among all finite-variance symmetric scale-family noise. Two corollaries give similar results for stochastic-resonance noise benefits in ML detection of a signal sequence with known shape but unknown amplitude.**

***Index Terms*—Maximum-likelihood (ML) detection, Neyman-Pearson (NP) detection, noise benefits, nonlinear correlation detectors, quantizer arrays, quantizer noise, scale-family noise, stochastic resonance, watermark decoding.**

I. Noise Benefits in Nonlinear Signal Detection

NOISE can sometimes improve nonlinear signal processing [1]–[10]. This noise-benefit *stochastic resonance* (SR) effect in signal detection occurs when small amounts of noise improves detection performance while too much noise degrades it [11]–[15]. Such SR noise benefits arise in many physical and biological signal systems from carbon nanotubes to neurons [16]–[23]. We focus here on the special case of SR for quantizer-array-based nonlinear correlators in Neyman-Pearson (NP) and maximum-likelihood (ML) signal detection in non-Gaussian channel noise. This channel noise includes symmetric α-stable ($S\alpha S$) noise and generalized-Gaussian noise. Such detection problems occur in sonar, radar, and watermark detection [24]–[26]. The two theorems below prove that injecting small amounts of quantizer noise in these types of nonlinear detectors can improve their detection performance. These new SR theorems extend the growing list of formal proofs of SR noise benefits [1]–[3], [8]–[14], [27]–[31].

Fig. 1 shows an SR noise benefit in the ML watermark extraction of the "yin-yang" image embedded in the discrete-cosine transform (DCT-2) coefficients of the "Lena" image [32]. The yin-yang image of Fig. 1(a) is the 64×64 binary watermark message embedded in the midfrequency DCT-2 coefficients of the 512×512 gray-scale Lena image using direct-sequence spread spectrum [33]. Fig. 1(b) shows the result when the yin-yang figure watermarks the Lena image. Figs. 1(c)–1(g) shows that small amounts of additive uniform quantizer noise improve the watermark-extraction performance of the noisy quantizer-array ML detector while too much noise degrades the performance. Uniform quantizer noise with standard deviation $\sigma = 1$ reduces more than 33% of the pixel-detection errors in the extracted watermark image. Section VI gives the details of such noise-enhanced watermark decoding.

The quantizer-array detector consists of two parts. It consists of a nonlinear preprocessor that precedes a correlator and a likelihood-ratio test of the correlator's output. This nonlinear detector takes K samples of a noise-corrupted signal and then sends each sample to the nonlinear preprocessor array of Q noisy quantizers connected in parallel. Each quantizer in the array adds its independent quantizer noise to the noisy input sample and then quantizes this doubly noisy data sample into a binary value. The quantizer array output for each sample is just the sum of all Q quantizer outputs. The correlator then correlates these preprocessed K samples with the signal. The detector's final stage applies either the NP likelihood-ratio test in Section II or the ML-ratio test in Section V.

Section III presents two SR noise-benefit theorems that apply to broad classes of channel and quantizer noises for the quantizer-array NP and ML detectors. Theorem 1 gives a necessary and sufficient condition for an SR noise benefit in NP detection of a constant ("dc") signal. The condition characterizes when the detection probability P_D will have a positive noise-based derivative—when $dP_D/d\sigma_N > 0$. Theorem 1 applies to all symmetric channel noise and to all symmetric quantizer noise so long as the number K of data samples is large. Corollary 1 in Section V gives a similar condition for the ML detection of a known sequence of unknown amplitude. It gives a simple method to find a near-optimal quantizer noise intensity for the ML detection and does not need the error probability. Section VI uses this method for watermark decoding.

Theorem 2 of Section III contains three SR results for quantizer-array NP detectors when the quantizer noise comes from a symmetric scale-family probability density function (pdf) with finite variance. The first result shows that $Q > 1$ is necessary

Manuscript received December 24, 2009; revised April 26, 2010; accepted October 21, 2010. Date of publication November 11, 2010; date of current version January 12, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Maciej Niedzwiecki.

The authors are with the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: ashokpat@usc.edu; kosko@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.2010.2091409

1053-587X/$26.00 © 2010 IEEE

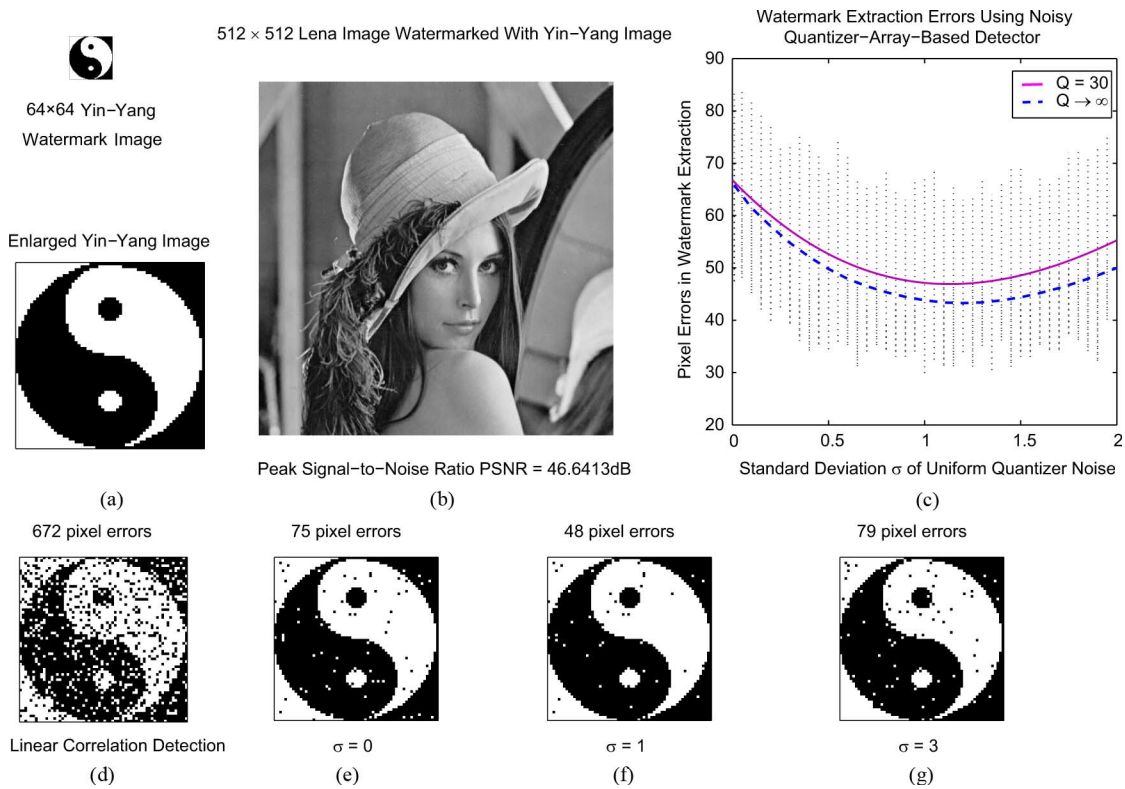


Fig. 1. Noise-enhanced digital watermark extraction using a noise-based algorithm: SR noise benefits in quantizer-array ML watermark decoding: (a) Binary 64×64 watermark yin-yang image. (b) Watermarked 512×512 Lena image. Direct-sequence spread spectrum embeds each message bit of the yin-yang image in a set of mid-frequency discrete cosine transform (DCT-2) coefficients of the gray-scale Lena image (c) Nonmonotonic quantizer-noise-enhanced watermark-detection performance plot of the array-based ML detectors. The noisy array detector had $Q = 30$ quantizers. Uniform quantizer noise decreased the pixel-detection error by more than 33%. The solid U-shaped line shows the average pixel-detection errors of 200 simulation trials. The dashed vertical lines show the total min-max deviations of pixel-detection errors in these simulation trials. The dashed U-shaped line shows the average pixel-detection errors of the limiting-array $(Q \rightarrow \infty)$ correlation detector. This dashed U-shaped line gives the lower bound on the pixel-detection error for any finite $(Q < \infty)$ quantizer-array detector with symmetric uniform quantizer noise. (d) Retrieved yin-yang image using the ML linear correlation detector. (e) Retrieved yin-yang image using the ML noiseless quantizer-array detector. This nonlinear detector outperforms the linear correlation detector. (f) Retrieved yin-yang image using the ML noisy quantizer-array detector. Additive uniform quantizer noise improves the detection of the quantizer-array detector by more than 33% as the uniform quantizer noise standard deviation σ increases from $\sigma = 0$ to $\sigma = 1$. (g) Too much quantizer noise degrades the watermark detection. The SR effect is robust against the quantizer noise intensity since the pixel-detection error in (g) is still less than the pixel-detection errors in (d).

for an initial SR effect if the symmetric channel noise is unimodal. The second result is that the rate of the initial SR effect in the small quantizer noise limit $(\lim_{\sigma_N \rightarrow 0} dP_D/d\sigma_N)$ improves if the number Q of quantizers in the array increases. This result implies that we should replace the noisy quantizer-array nonlinearity with its deterministic limit $(Q \rightarrow \infty)$ to achieve the upper-bound detection performance if the respective quantizer-noise cumulative distribution function has a simple closed form. The third result is that symmetric *uniform* quantizer noise gives the best initial SR effect rate among all symmetric scale-family noise types. Corollary 2 in Section V extends Theorem 2 to the ML detection of a known sequence of unknown amplitude. All these results hold for any symmetric unimodal channel noise even though we focus on $S\alpha S$ noise and symmetric generalized-Gaussian channel noise. The scope of these new theorems extends well beyond watermark decoding and detection. They show how quantizer noise can enhance a wide range of array-based NP and ML detection problems in non-Gaussian channel noise. Applications include radar, sonar, and telecommunications [24]–[26], [34]–[36] when optimal detectors do not have a closed form or when we cannot easily estimate channel noise parameters.

Array-based noise benefits have only a recent history. Stocks [10] first showed that adding quantizer noise in an array of parallel-connected quantizers improves the mutual information between the array's input and output. This produced a type of suprathreshold SR effect (or SSR as Stocks calls it [37]) because it did not require subthreshold signals [38]. Then Rousseau and Chapeau-Blondeau [9], [13] used such a quantizer array for signal detection. They first showed the SR effect for NP detection of time-varying signals and for Bayesian detection of both constant and time-varying signals in different types of non-Gaussian but finite-variance channel noise. We proved in [12] that noise in parallel arrays of threshold neurons can improve the ML detection of a constant signal in symmetric channel noise. Theorem 5 in [12] showed that collective noise benefits can occur in a large number of parallel arrays of threshold units even when an individual threshold unit does not itself produce a noise benefit.

II. NP Binary Signal Detection in α-Stable Noise

This section develops the NP hypothesis-testing framework for the two noise-benefit theorems that follow. The problem is to

detect a known deterministic signal s_k with amplitude A in additive white symmetric α-stable $(S\alpha S)$ channel noise V_k given K random samples $X_1, \ldots, X_K$:

$$\begin{aligned} H_0 &: X_k = V_k \\ H_1 &: X_k = As_k + V_k \end{aligned} \tag{1}$$

such that the signal detection probability $P_D = P(\text{Decide } H_1 | H_1 \text{ is true})$ is maximal while the false-alarm probability $P_{FA} = P(\text{Decide } H_1 | H_0 \text{ is true})$ stays at a preset level τ. The V_k are independent and identically distributed (i.i.d.) zero-location $S\alpha S$ random variables. We consider only constant (dc) signals so that $s_k = 1$ for all k. So the null hypothesis H_0 states that the signal s_k is not present in the noisy sample X_k while the alternative hypothesis H_1 states that s_k is present.

The characteristic function φ of the $S\alpha S$ noise random variable V_k has the exponential form [39], [40]

$$\varphi(\omega) = \exp(j\delta\omega - \gamma|\omega|^\alpha) \tag{2}$$

where real δ is the location parameter, $\alpha \in (0, 2]$ is the characteristic exponent that controls the density's tail thickness, $\gamma = \sigma^\alpha > 0$ is the dispersion that controls the width of the bell curve, and σ is the scale parameter. The bell curve's tails get thicker as α falls from 2 to near zero. So energetic impulses become more frequent for smaller values of α.

$S\alpha S$ pdfs can model heavy-tailed or impulsive noise in applications that include underwater acoustic signals, telephone noise, clutter returns in radar, internet traffic, financial data, and transform domain image or audio signals [24], [26], [40]–[44]. The only known closed-form $S\alpha S$ pdfs are the thick-tailed Cauchy with $\alpha = 1$ and the thin-tailed Gaussian with $\alpha = 2$. The Gaussian pdf alone among $S\alpha S$ pdfs has a finite variance and finite higher-order moments. The mth lower-order moments of an α-stable pdf with $\alpha < 2$ exist if and only if $m < \alpha$. The location parameter δ serves as a proxy for the mean if $1 < \alpha \leq 2$ and as a proxy for the median if $0 < \alpha \leq 1$. The uniformly most powerful detector for the hypotheses in (1) is a NP log-likelihood ratio test [45], [46]

$$\Lambda_{NP}(\mathbf{X}) = \sum_{k=1}^{K} \log(f_\alpha(X_k - s_k)) - \log(f_\alpha(X_k)) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_\tau \tag{3}$$

because the random K samples $X = \{X_1, \ldots, X_K\}$ are i.i.d. We choose λ_τ so that it has a preset false-alarm probability P_{FA} of $P_{FA} = \tau$. This $S\alpha S$ NP detector (3) is hard to implement because again the $S\alpha S$ pdf f_α has no closed form except when $\alpha = 1$ or $\alpha = 2$. The NP detector (3) does reduce to the simpler test

$$\Lambda_{Lin}(\mathbf{X}) = \sum_{k=1}^{K} s_k X_k \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_\tau \tag{4}$$

if the additive channel noise V_k is Gaussian $(\alpha = 2)$ [45]. But this linear correlation detector is suboptimal when the channel noise is non-Gaussian. Its detection performance degrades severely as the channel noise pdf departs further from Gaussianity [45], [47] and thus when $\alpha < 2$ holds.

An important special case is the NP detector for Cauchy $(\alpha = 1)$ channel noise. The zero-location Cauchy random variable V_k $(\alpha = 1)$ has the closed-form pdf

$$f_{V_k}(v_k) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + v_k^2} \tag{5}$$

for real v_k and positive dispersion γ. The NP detector is nonlinear for such Cauchy channel noise and has the form

$$\Lambda_{C_{NP}}(\mathbf{X}) = \sum_{k=1}^{K} \log\left(\frac{\gamma^2 + (X_k)^2}{\gamma^2 + (X_k - As_k)^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_\tau. \tag{6}$$

So the NP Cauchy detector (6) does not have a simple correlation structure. It is also more computationally complex than the NP linear correlation detector (4). But the NP Cauchy detector performs well for highly impulsive $S\alpha S$ noise cases [26], [40]. Section III shows that three other simple nonlinear correlation detectors can perform as well or even better than the Cauchy detector does when the $S\alpha S$ noise is mildly impulsive (when $\alpha \geq 1.6$).

The *locally optimal* detector has the familiar correlation structure [40], [48]

$$\begin{aligned} \Lambda_{LO}(\mathbf{X}) &= \sum_{k=1}^{K} s_k \frac{-f'_\alpha(X_k)}{f_\alpha(X_k)} \\ &= \sum_{k=1}^{K} s_k g_{LO}(X_k) \underset{H_0}{\overset{H_1}{\gtrless}} \tilde{\lambda}_\tau \end{aligned} \tag{7}$$

and coincides with the linear correlator (4) for Gaussian $(\alpha = 2)$ channel noise V_k. The score function g_{LO} is nonlinear for $\alpha < 2$. The locally optimal detector (7) performs well when the signal amplitude A is small. But this test is not practical when f_α does not have a closed form because g_{LO} requires both f_α and f'_α. So researchers have suggested other suboptimal detectors that preserve the correlation structure but that replace g_{LO} with different zero-memory nonlinear functions g[46], [49]–[52]. These nonlinearities range from simple *ad-hoc* soft-limiters

$$g_{SL}(X_k) = \begin{cases} X_k & \text{if } |X_k| \leq c \\ 1 & \text{if } X_k > c \\ -1 & \text{if } X_k < -c \end{cases} \tag{8}$$

and hole-puncher functions

$$g_{HP}(X_k) = \begin{cases} X_k & \text{if } |X_k| \leq c \\ 0 & \text{else} \end{cases} \tag{9}$$

to more complex nonlinearities that better approximate g_{LO}. The latter may use a scale-mixture approach [51] or a simplified Cauchy-Gaussian mixture model [52].

The next section presents the two main SR theorems for the nonlinear correlation detectors that replace the deterministic nonlinearity g_{LO} with a noisy quantizer-array-based *random* nonlinearity g_{NQ} or with its deterministic limit $g_{N\infty}$. We show that these detectors enjoy SR noise benefits. We then compare their detection performances with the Cauchy detector (6) and with the nonlinear correlation detectors based on the simple soft-limiter and hole-puncher nonlinearities (8)–(9).

III. Quantizer Noise Benefits in Nonlinear-Correlation-Detector-Based NP Detection

This section presents the two main SR noise-benefit theorems for NP detectors. The Appendix gives the proof of Theorem 2. We start with the nonlinear correlation detector:

$$\Lambda_{NQ}(\mathbf{X}) = \sum_{k=1}^{K} s_k g_{NQ}(X_k) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \tag{10}$$

$$\text{where} \quad g_{NQ}(X_k) = \frac{1}{Q}\sum_{q=1}^{Q} sign(X_k + N_q - \theta). \tag{11}$$

Here λ is the detection threshold, θ is the quantization threshold, and $sign(X_k + N_q - \theta) = \pm 1$ for $q = 1, \ldots, Q$. We chose $\theta = A/2$ because both the channel noise V_k and the quantizer noise N_q are symmetric.

We assume that the additive quantizer noise N_q has a symmetric scale-family [53] noise pdf $f_N(\sigma_N, n) = 1/\sigma_N f_{\tilde{N}}(n/\sigma_N)$. Here σ_N is the noise standard deviation and $f_{\tilde{N}}$ is the standard pdf for the whole family [53]. Then the noise cumulative distribution function (CDF) is $F_N(\sigma_N, n) = F_{\tilde{N}}(n/\sigma_N)$ where $F_{\tilde{N}}$ is the standard CDF for the whole family. Scale-family densities include many common densities such as the Gaussian and uniform but not the Poisson. We assume that the quantizer noise random variables N_q have finite variance and are independent and come from a symmetric scale family noise. The quantizer noise can arise from electronic noise such as thermal noise or from avalanche noise in analog circuits [54], [55]. The noisy quantizer-array detector (10)–(11) is easy to use and requires only one bit to represent each quantizer's output. This favors sensor networks and distributed systems that have limited energy or that allow only limited data handling and storage [56], [57].

Define next $\mu_i(\sigma_N)$ and $\sigma_i^2(\sigma_N)$ as the respective population mean and population variance of Λ_{NQ} under the hypothesis H_i ($i = 0$ or $i = 1$) when σ_N is the quantizer noise intensity: $\mu_i(\sigma_N) = E(\Lambda_{NQ}|H_i)$ and $\sigma_i^2(\sigma_N) = V\text{ar}(\Lambda_{NQ}|H_i)$. Then $\mu_0(\sigma_N) = -\mu_1(\sigma_N)$ and $\sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N)$ for all σ_N because both the additive channel noise V and the quantizer noise N are symmetric. The mean μ_i and variance σ_i^2 of the test statistic Λ_{NQ} depend on both V and N. So μ_i and σ_i^2 depend on the noise intensities σ_V and σ_N. We write these two terms as $\mu_i(\sigma_N)$ and $\sigma_i^2(\sigma_N)$ because we control only the quantizer noise intensity σ_N and not the channel noise intensity σ_V. The Appendix derives the complete form of $\mu_i(\sigma_N)$ and $\sigma_i^2(\sigma_N)$ in the respective (72) and (92) as part of the proof of Theorem 2.

The additive structure of Λ_{NQ} in (10) gives rise to a key simplification. The pdf of Λ_{NQ} is *approximately Gaussian* for both hypotheses because the central limit theorem [53] applies to (10) if the sample size K is large since the random variables $X_1, \ldots, X_K$ have finite variance and are independent and identically distributed (i.i.d.). Then Theorem 1 gives a necessary and sufficient inequality condition for the SR effect in the quantizer-array detector (10)–(11). This SR condition depends only on $\mu_1(\sigma_N)$ and $\sigma_1^2(\sigma_N)$ and on their first derivatives. It is equivalent to $d/d\sigma_N(\ln\mu_1(\sigma_N)/\sigma_1(\sigma_N)) > 0$ since $\mu_1 > 0$ in (72).

Theorem 1: Suppose that the detection statistic Λ_{NQ} in (10) has sufficiently large sample size K so that it is approximately conditionally normal: $\Lambda_{NQ}|H_0 \approx \mathbf{N}(\mu_0(\sigma_N), \sigma_0^2(\sigma_N))$ and $\Lambda_{NQ}|H_1 \approx \mathbf{N}(\mu_1(\sigma_N), \sigma_1^2(\sigma_N))$ where $\mu_0(\sigma_N) = -\mu_1(\sigma_N)$ and $\sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N)$. Then the inequality

$$\sigma_1(\sigma_N)\mu_1'(\sigma_N) > \mu_1(\sigma_N)\sigma_1'(\sigma_N) \tag{12}$$

is necessary and sufficient for the SR noise benefit $dP_D/d\sigma_N > 0$ in NP signal detection based on the nonlinear test statistic Λ_{NQ}.

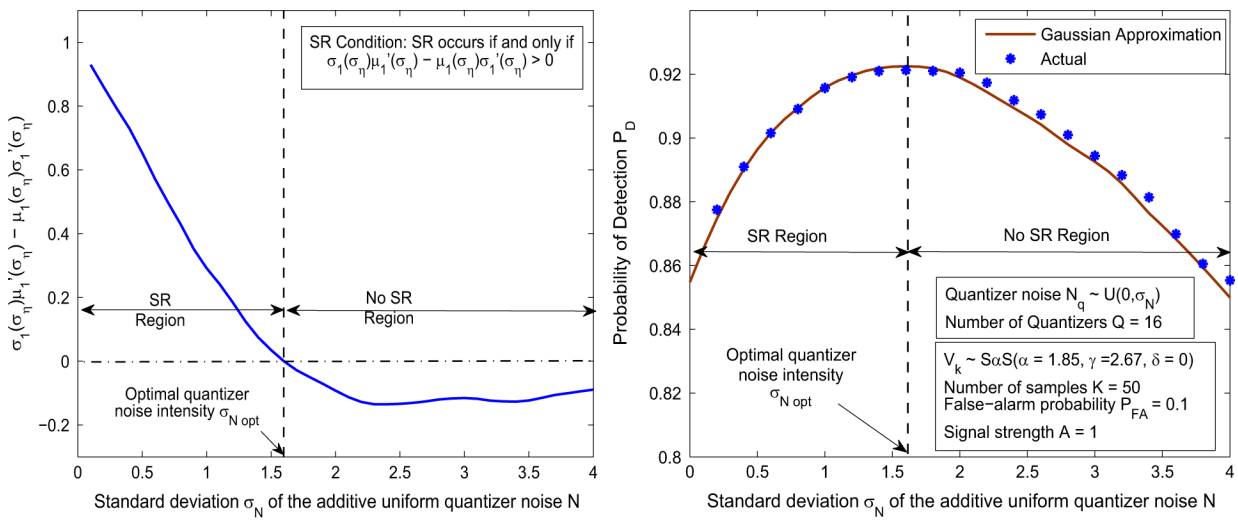
Proof: We first derive an approximate linear form for the detection threshold λ_τ: $\lambda_\tau \approx z_\tau\sigma_0(\sigma_N) + \mu_0(\sigma_N)$. The NP detection rule based on Λ_{NQ} rejects H_0 if $\Lambda_{NQ} > \lambda_\tau$ because we choose the detection threshold λ_τ such that $P(\Lambda_{NQ} > \lambda_\tau|H_0) = \tau$. We also need to define the constant z_τ so that $1 - \Phi(z_\tau) = \tau$ where $\Phi(z) = \int_{-\infty}^{z}(1/\sqrt{2\pi})e^{-x^2/2}dx$. Then standardizing Λ_{NQ} under the assumption that the null hypothesis H_0 is true gives

$$\begin{aligned} &P(\Lambda_{NQ} > \lambda_\tau \,|\, H_0) \\ &\quad = P\left(\frac{\Lambda_{NQ} - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} > \frac{\lambda_\tau - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right) && (13)\\ &\quad \approx P\left(Z > \frac{\lambda_\tau - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right) && (14)\\ &\quad \text{for } Z \sim \mathrm{N}(0,1) \text{ by the central limit theorem} \\ &\quad = 1 - \Phi\left(\frac{\lambda_\tau - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right). && (15) \end{aligned}$$

So $1 - \Phi(z_\tau) \approx 1 - \Phi\left(\lambda_\tau - \mu_0(\sigma_N)/\sigma_0(\sigma_N)\right)$ and thus $z_\tau \approx \lambda_\tau - \mu_0(\sigma_N)/\sigma_0(\sigma_N)$. So the detection threshold λ_τ has the approximate linear form $\lambda_\tau \approx z_\tau\sigma_0(\sigma_N) + \mu_0(\sigma_N)$.

Standardizing Λ_{NQ} under the assumption that the alternative hypothesis H_1 is true likewise gives the detection probability P_D as

$$\begin{aligned} P_D(\sigma_N) &= P(\Lambda_{NQ} > \lambda_\tau \,|\, H_1) \quad \text{from (10)} && (16)\\ &= P\left(\frac{\Lambda_{NQ} - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} > \frac{\lambda_\tau - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (17)\\ &\approx P\left(Z > \frac{\lambda_\tau - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (18)\\ &\quad \text{for } Z \sim \mathrm{N}(0,1) \text{ by the central limit theorem} \\ &= 1 - P\left(Z \le \frac{\lambda_\tau - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (19)\\ &= 1 - \Phi\left(\frac{\lambda_\tau - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (20)\\ &\approx 1 - \Phi\left(\frac{z_\tau\sigma_0(\sigma_N) + \mu_0(\sigma_N) - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (21)\\ &\quad \text{because } \lambda_\tau \approx z_\tau\sigma_0(\sigma_N) + \mu_0(\sigma_N) \\ &= 1 - \Phi\left(\frac{z_\tau\sigma_1(\sigma_N) - \mu_1(\sigma_N) - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) && (22)\\ &\quad \text{because } \mu_0(\sigma_N) = -\mu_1(\sigma_N) \text{ and } \sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N) \\ &= 1 - \Phi\left(z_\tau - \frac{2\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right). && (23) \end{aligned}$$



(a) (b)

Fig. 2. SR noise benefits based on inequality (12) for constant (dc) signal detection in α-stable channel noise (a) The plot of $\sigma_1(\sigma_N)\mu_1'(\sigma_N) - \mu_1(\sigma_N)\sigma_1'(\sigma_N)$ versus the standard deviation σ_N of zero-mean additive uniform quantizer noise. The zero crossing occurs at the quantizer noise standard deviation $\sigma_{N_{\text{opt}}}$. (b) The solid line and circle markers show the respective plots of the detection probabilities P_D with and without the Gaussian approximation of the Λ_{NQ} pdf. Adding small amounts of quantizer noise N improved the detection probability P_D by more than 7.5%. This SR effect occurred until inequality (12) held. So $\sigma_{N_{\text{opt}}}$ maximized the detection probability.

Then the normal pdf $\phi(z) = d\Phi(z)/dz$ and the chain rule of differential calculus give

$$\frac{dP_D}{d\sigma_N} \approx 2\phi\left(z_\tau - \frac{2\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \times \frac{\sigma_1(\sigma_N)\mu_1'(\sigma_N) - \mu_1(\sigma_N)\sigma_1{}'(\sigma_N)}{\sigma_1^2(\sigma_N)} \quad (24)$$

because z_τ is a constant. So $\sigma_1(\sigma_N)\mu_1'(\sigma_N) > \mu_1(\sigma_N)\sigma_1'(\sigma_N)$ is necessary and sufficient for the SR effect $(dP_D/d\sigma_N > 0)$ because ϕ is a pdf and thus $\phi \geq 0$. □

Fig. 2 shows a simulation instance of the SR inequality condition in Theorem 1 for constant (dc) signal detection in impulsive infinite-variance channel noise. The signal has magnitude $A = 0.5$ and we set the false-alarm probability P_{FA} to $P_{FA} = 0.1$. The channel noise is $S\alpha S$ with parameters $\alpha = 1.85$, $\gamma = 1.7^{1.85} = 2.67$, and $\delta = 0$. The detector preprocesses each of the $K = 50$ noisy samples X_k with $Q = 16$ quantizers in the array. Each quantizer has quantization threshold $\theta = A/2$ and adds the independent uniform quantizer noise N to the noisy sample X_k before quantization. Fig. 3(a) plots the smoothed difference $\sigma_1(\sigma_N)\mu_1'(\sigma_N) - \mu_1(\sigma_N)\sigma_1'(\sigma_N)$ versus the standard deviation σ_N of the additive uniform quantizer noise. We used 10^5 simulation trials to estimate $\mu_1(\sigma_N)$ and $\sigma_1(\sigma_N)$ and then used the difference quotients $\mu_1(\sigma_{N_j}) - \mu_1(\sigma_{N_{j-1}})/\sigma_{N_j} - \sigma_{N_{j-1}}$ and $\sigma_1(\sigma_{N_j}) - \sigma_1(\sigma_{N_{j-1}})/\sigma_{N_j} - \sigma_{N_{j-1}}$ to estimate their first derivatives. Fig. 3(b) shows that adding small amounts of quantizer noise N improved the detection probability P_D by more than 7.5%. This SR effect occurs until (12) holds in Fig. 3(a). Fig. 3(b) also shows the accuracy of the Gaussian (central limit theorem) approximation of the detection statistic Λ_{NQ}'s pdf. Circle marks show the detection probabilities computed from the 10^5 Monte Carlo simulations. The solid line plots the detection probability P_D in (23).

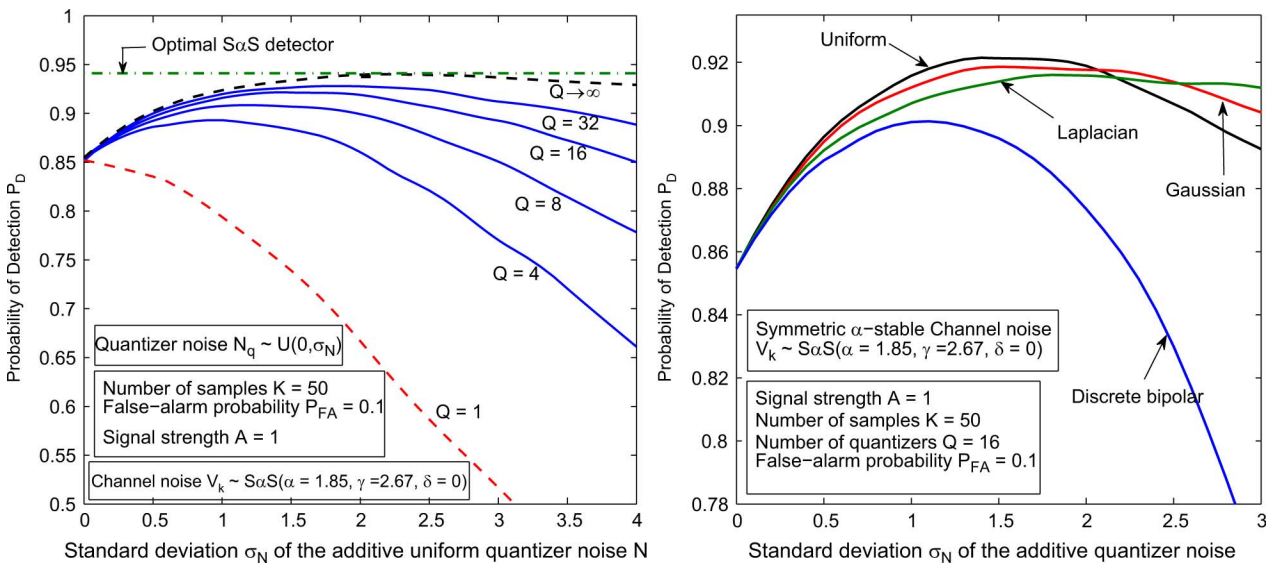
Theorem 2 states that it takes more than one quantizer to produce the initial SR effect and that the rate of the initial SR effect increases as the number Q of quantizers increases. It further states that uniform quantizer noise gives the maximal initial SR effect among all possible finite-variance symmetric scale-family quantizer noise. Theorem 2 and Corollary 2 involve an initial SR effect that either increases the detection probability P_D or decreases the error probability P_e for small amounts of noise. We define the SR effect as an *initial SR effect* if there exists some $b > 0$ such that $P_D(\sigma_N) > P_D(0)$ or that $P_e(\sigma_N) < P_e(0)$ for all $\sigma_N \in (0, b)$. Theorem 2 follows from Theorem 1 if we substitute the expressions that we derive in the Appendix for $\mu_1(\sigma_N)$, $\mu_1'(\sigma_N)$, $\sigma_1(\sigma_N)$, and $\sigma_1'(\sigma_N)$ and then pass to the limit $\sigma_N \longrightarrow 0$. The complete proof is in the Appendix because it is lengthy and uses real analysis.

Theorem 2: Suppose that the channel noise pdf is uniformly bounded and continuously differentiable at $-A/2$. Then

(a) $Q > 1$ is necessary for the initial SR effect in the NP detection of a constant signal in symmetric unimodal channel noise V if the test statistic is the nonlinear test statistic Λ_{NQ}.

(b) Suppose that the initial SR effect occurs in the quantizer-array detector (10)–(11) with Q_1 quantizers and with some symmetric quantizer noise. Then the rate of the initial SR effect in the quantizer-array detector (10)–(11) with Q_2 quantizers is larger than the rate of the initial SR effect with Q_1 quantizers if $Q_2 > Q_1$.

(c) Zero-mean uniform noise is the optimal finite-variance symmetric scale-family quantizer noise in that it gives the maximal rate of the initial SR effect among all possible finite-variance quantizer noise in the NP quantizer-array detector (10)–(11).



(a) (b)

Fig. 3. Initial SR effects in quantizer-array correlation detectors for constant (dc) signal detection in impulsive symmetric α-stable $(S\alpha S)$ channel noise $(\alpha = 1.85)$. (a) Initial SR effects for zero-mean uniform quantizer noise. The solid lines show that the detection probability P_D improves at first as the quantizer noise intensity σ_N increases. The dashed line shows that the SR effect does not occur if $Q = 1$ as Theorem 2(a) predicts. The solid lines also show that the rate of the initial SR effect increases as the number Q of quantizers increases as Theorem 2(b) predicts. The thick horizontal dash-dot line shows the detection probability of the optimal $S\alpha S$ NP detector (3). The limiting-array $(Q \to \infty)$ detector gave almost the same detection performance as the optimal $S\alpha S$ NP detector gave (b) comparison of initial SR effects in the quantizer-array correlation detector for different types of symmetric quantizer noise. Symmetric uniform noise gave the maximal rate of the initial SR effect as Theorem 2(c) predicts. Symmetric discrete bipolar noise gave the smallest SR effect and was the least robust.

Fig. 3 shows simulation instances of Theorem 2. The thin dashed line in Fig. 3(a) shows that the SR effect does not occur if $Q = 1$ as Theorem 2(a) predicts. The solid lines show that the initial SR effect increases as the number Q of quantizers increases as Theorem 2(b) predicts. $Q = 32$ quantizers gave a 0.925 maximal detection probability and thus gave an 8% improvement over the noiseless detection probability of 0.856. The thick dashed line in Fig. 3(a) shows the upper bound on the detection probability P_D that any noisy quantizer-array detector (10)–(11) with symmetric uniform quantizer noise can achieve if it increases the number Q of quantizers in its array. The thick horizontal dash-dot line shows the detection probability of the optimal $S\alpha S$ NP detector (3). This line does not depend on the quantizer-noise standard deviation σ_N and so is flat because the optimal $S\alpha S$ NP detector (3) does not use the quantizer noise N. The limiting-array $(Q \to \infty)$ detector gave almost the same detection performance as the optimal $S\alpha S$ NP detector gave.

Theorem 2(b) implies that the limit $\lim_{Q \to \infty} P_D(g_{NQ})$ gives the upper bound on the detection probability of the noisy quantizer-array detector (10)–(11). The right-hand side (RHS) of (11) is the conditional sample mean of the bipolar random variable $Y_{k,q} = sign(X_k + N_q - \theta)$ given X_k because the quantizer noise random variables N_q are i.i.d. Then the conditional expected value has the form

$$E[Y_{k,q}|X_k = x_k] = 1 - 2F_{\tilde{N}}(\sigma_N, \theta - x_k) \tag{25}$$

$$= 1 - 2F_{\tilde{N}}\left(\frac{\theta - X_k}{\sigma_N}\right) \tag{26}$$

where $\theta = A/2$. The CDF $F_N(\sigma_N, \cdot)$ is the CDF of the symmetric scale-family quantizer noise N_q that has standard deviation σ_N and that has standard CDF $F_{\tilde{N}}$ for the entire family. So the strong law of large numbers [53] implies that the sample mean $g_{NQ}(X_k)$ in (11) converges with probability one to its population mean in (26):

$$\lim_{Q \to \infty} g_{NQ}(X_k) = 1 - 2F_{\tilde{N}}\left(\frac{\theta - X_k}{\sigma_N}\right) \tag{27}$$

$$= 2F_{\tilde{N}}\left(\frac{X_k - \theta}{\sigma_N}\right) - 1. \tag{28}$$

Equality (28) follows because the quantizer noise is symmetric. So Theorem 2(b) implies that the detection probability $P_D(g_{N\infty})$ of the limiting nonlinear correlation detector

$$\Lambda_{N\infty}(\mathbf{X}) = \sum_{k=1}^{K} s_k g_{N\infty}(X_k) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \tag{29}$$

$$\text{with} \quad g_{N\infty}(X_k) = 2F_{\tilde{N}}\left(\frac{X_k - \theta}{\sigma_N}\right) - 1 \tag{30}$$

gives the upper bound on the detection performance of the quantizer-array detector (10)–(11) for $\theta = A/2$ when the quantizer noise N has standard deviation σ_N and when the scale-family CDF is $F_{\tilde{N}}$.

We can use this limiting *non-noisy* nonlinear correlation detector (29)–(30) if the quantizer noise CDF $F_{\tilde{N}}(x_k - A/2/\sigma_N)$ has a closed form. Simulations show that the detection performance of the noisy quantizer-array detector quickly approaches the detection performance of the limiting quantizer-array $(Q \to \infty)$ detectors. So we can often use the noisy quantizer-array detector (10)–(11) with Q near 100 to get a detection performance close to that of the limiting quantizer-array $(Q \to \infty)$ detector.

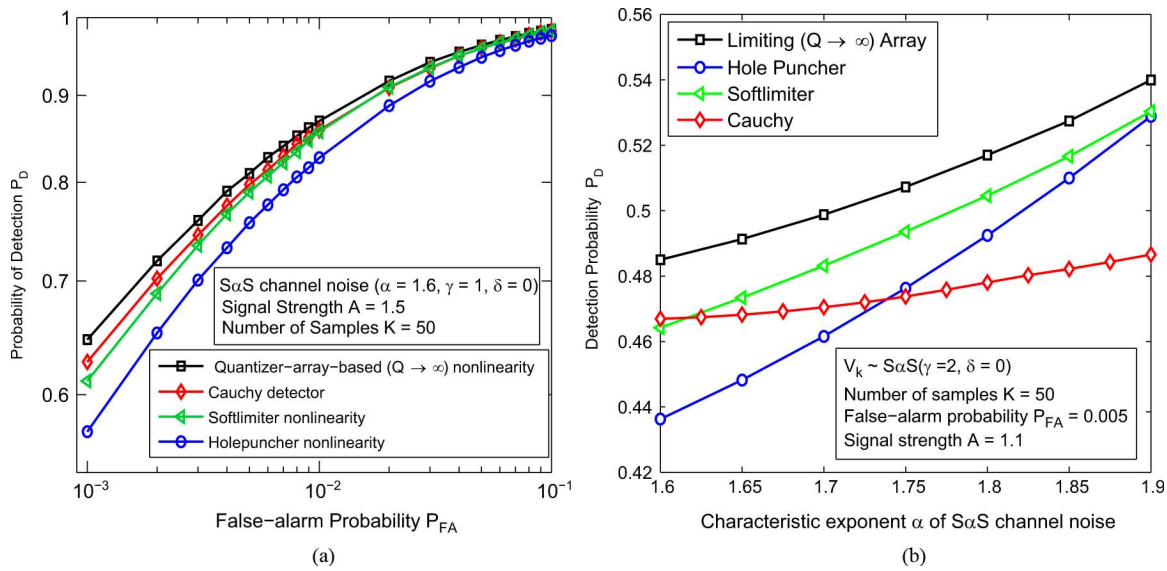


Fig. 4. Comparison of NP detection performance of four nonlinear detectors for different values of (a) false-alarm probabilities P_{FA}. (b) Characteristic exponent or tail-thickness parameter α of $S\alpha S$ channel noise dispersion γ when the signal strength A was $A = 1.1$ and the number K of samples was $K = 50$. The limiting $(Q \longrightarrow \infty)$ quantizer-array detector performed better than the Cauchy detector, the soft-limiter detector, and the hole-puncher detector for medium-to-low impulsive $S\alpha S$ noise cases $(1.6 \leq \alpha \leq 1.9)$ and for small false-alarm probabilities $(P_{FA} \leq 0.1)$.

The limiting nonlinearity $g_{N\infty}(X_k)$ is easy to use for symmetric uniform quantizer noise because it is a shifted soft-limiter with shift $\theta = A/2$:

$$g_{N\infty}(X_k) = \begin{cases} X_k - \frac{A}{2} & \text{if } |X_k - \frac{A}{2}| \leq c \\ 1 & \text{if } \left(X_k - \frac{A}{2}\right) > c \\ -1 & \text{if } \left(X_k - \frac{A}{2}\right) < -c \end{cases} \tag{31}$$

where $c = \sqrt{3}\sigma_N$. Fig. 3(a) shows that the limiting nonlinear correlation detector (29)–(30) with the shifted soft-limiter nonlinearity (31) gives almost the same detection performance as the optimal $S\alpha S$ detector (3). We used the numerical method of [58] to compute the $S\alpha S$ pdf f_α for $\alpha = 1.85$. Fig. 4 shows that the limiting $(Q \longrightarrow \infty)$ quantizer-array detector performed better than the Cauchy detector, the soft-limiter detector, and the hole-puncher detector for medium-to-low impulsive $S\alpha S$ noise cases $(1.6 \leq \alpha \leq 1.9)$ and for small false-alarm probabilities $(P_{FA} \leq 0.1)$.

The limiting-array nonlinearity (30) is monotone non-decreasing while the asymptotic behavior of the locally optimal nonlinearity in (7) is $g_{LO}(X_k) \approx (\alpha + 1)/X_k$. So a small signal strength A implies that quantizer-array detectors cannot perform better than nonlinear correlation detectors with non-monotonic nonlinearities such as [49], [50]

$$g_{LSO-P}(X_k) = \begin{cases} \frac{(\alpha+1)X_k}{c^2} & \text{if } |X_k| \leq c \\ \frac{(\alpha+1)}{X_k} & \text{else} \end{cases} \tag{32}$$

or such as [46]

$$g(X_k) = \frac{aAs_kX_k}{1 + bAs_kX_k^2}. \tag{33}$$

Fig. 3(b) shows simulation instances of Theorem 2(c). It compares the initial SR noise benefits for different types of simple zero-mean symmetric quantizer noises such as Laplacian, Gaussian, uniform, and discrete bipolar noise when there are $Q = 16$ quantizers in the array. Symmetric uniform noise gave the maximal rate of the initial SR effect as Theorem 2(c) predicts. It also gave the maximal SR effect (maximal increase in the detection probability) compared to Laplacian, Gaussian, and discrete bipolar noise. Theorem 2(c) guarantees only a maximal rate for the initial SR effect. It does not guarantee a maximal SR effect for symmetric uniform noise. So some other type of symmetric quantizer noise may give the maximal SR effect in other detection problems. Fig. 3(b) also shows that symmetric discrete bipolar noise gave the smallest SR effect and was the least robust. The SR effect was most robust against Laplacian quantizer noise.

IV. Maximum-Likelihood Binary Signal Detection in Symmetric α-Stable Channel Noise Or Generalized-Gaussian Channel Noise

Consider next the ML detection of a deterministic signal sequence of known shape s_k but unknown amplitude A in either additive i.i.d. $S\alpha S$ channel noise V_k or generalized-Gaussian channel noise V_k. We assume that the noise pdf has unknown parameters. The ML detector uses K random samples $X_1, \ldots, X_K$ to decide between the equally likely null hypothesis H_0 and alternative hypothesis H_1:

$$\begin{aligned} H_0 &: X_k = -As_k + V_k \\ H_1 &: X_k = As_k + V_k. \end{aligned} \tag{34}$$

The ML decision rule minimizes the average decision-error probability $P_e = p_0 P(\text{Decide } H_1 | H_0 \text{ is true}) + p_1 P(\text{Decide } H_0 | H_1 \text{ is true})$ [59]. The prior probabilities p_0 and p_1 are equal: $p_0 = P(H_0) = P(H_1) = p_1 = 1/2$. The ML detector for (34) is a log-likelihood ratio test [45], [46]:

$$\Lambda_{ML}(\mathbf{X}) = \sum_{k=1}^{K} \log\left(\frac{f(X_k - As_k)}{f(X_k + As_k)}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{35}$$

The optimal ML detector (35) again does not have a closed form when the channel noise is $S\alpha S$ except when $\alpha = 1$ and $\alpha = 2$. Nor does the optimal ML detector have a correlation structure if $\alpha < 2$ and thus if the $S\alpha S$ channel noise V_k is not Gaussian.

The optimal ML detector for the hypothesis test (34) has the different form

$$\Lambda_{GG_{ML}}(\mathbf{X}) = \sum_{k=1}^{K} [|X_k + As_k|^r - |X_k - As_k|^r] \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (36)$$

if the symmetric channel noise variables V_k are i.i.d. generalized Gaussian random variables with pdf

$$f_{V_k}(v_k) = De^{-B|v_k|^r}. \quad (37)$$

Here r is a positive shape parameter, B is an intensity parameter, and D is a normalizing constant. The generalized Gaussian family [60] is a two-parameter family of symmetric continuous pdfs. Its scale-family pdf $f(\sigma, v)$ has the form

$$f(\sigma, v) = \frac{1}{\sigma} f_{gg}\left(\frac{v}{\sigma}\right) = \frac{1}{\sigma}\frac{r}{2}\frac{\left[\Gamma\left(\frac{3}{r}\right)\right]^{1/2}}{\left[\Gamma\left(\frac{1}{r}\right)\right]^{3/2}} e^{-B|v/\sigma|^r} \quad (38)$$

where f_{gg} is the standard pdf of the family, σ is the standard deviation, and Γ is the gamma function. This family of pdfs includes all normal $(r = 2)$ and Laplace $(r = 1)$ pdfs. It includes in the limit $(r \longrightarrow \infty)$ all continuous uniform pdfs on bounded real intervals. It can also model symmetric platykurtic densities whose tails are heavier than normal $(r < 2)$ or symmetric leptokurtic densities whose tails are lighter than normal $(r > 2)$. Applications include noise modeling in image, speech, and multimedia processing [61]–[64].

Generalized Gaussian noise can apply to watermark detection or extraction [33], [65]–[67]. The generalized-Gaussian ML detector (36) does not use the scale parameter σ but we still need joint estimates of σ and r. The generalized-Gaussian ML detector (36) applies to watermark extraction in images when generalized Gaussian random variables V_k model mid-frequency discrete cosine transform (DCT-2) coefficients [33], [68] or subband discrete wavelet transform coefficients [65], [69]. But the mid-frequency DCT-coefficients of many images may have thicker tails than generalized Gaussian pdfs have. And using the generalized-Gaussian ML detector (36) may be difficult for non-Gaussian $(r \neq 2)$ noise because (36) requires joint estimation of the signal amplitude A and noise parameters and because (36) also requires exponentiation with floating point numbers when $r \neq 2$. So Briassouli and Tsakalides [26] have proposed using instead the Cauchy pdf to model the DCT coefficients.

The ML Cauchy detector has the form [40]

$$\Lambda_{C_{ML}}(\mathbf{X}) = \sum_{k=1}^{K} \log\left(\frac{\gamma^2 + (X_k + As_k)^2}{\gamma^2 + (X_k - As_k)^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0. \quad (39)$$

It does not use exponentiation with floating point numbers. But the nonlinear detectors (36) and (39) require that we jointly estimate the signal amplitude and the parameters of the channel-noise pdf. This joint estimation is not easy in the ML case.

We next analyze a noisy quantizer-array correlation statistic $\bar{g}_{NQ}$ and its limit $\bar{g}_{N\infty}$. Neither uses the value of A for the ML detection of (34). These nonlinearities are versions of (11) and (30) with $\theta = 0$. We show next that the results of Theorems 1 and 2 also hold for the ML correlation detectors based on $\bar{g}_{NQ}$ and $\bar{g}_{N\infty}$.

V. Quantizer Noise Benefits in ML Detection

The next two corollaries apply Theorems 1 and 2 to the ML detection problem in (34). We first restate the noisy quantizer-array correlation statistic $\bar{g}_{NQ}(11)$ and its limit $\bar{g}_{N\infty}(30)$ with $\theta = 0$ for the ML detection of the signal s_k in (34):

$$\Lambda_{NQ}(\mathbf{X}) = \sum_{k=1}^{K} s_k \bar{g}_{NQ}(X_k) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (40)$$

$$\text{where} \quad \bar{g}_{NQ}(X_k) = \frac{1}{Q}\sum_{q=1}^{Q} sign(X_k + N_q) \quad (41)$$

and

$$\Lambda_{N\infty}(\mathbf{X}) = \sum_{k=1}^{K} s_k \bar{g}_{N\infty}(X_k) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (42)$$

$$\text{where} \quad \bar{g}_{N\infty}(X_k) = 2F_{\tilde{N}}\left(\frac{X_k}{\sigma_N}\right) - 1. \quad (43)$$

We use $\theta = 0$ because the pdfs of the random samples X_k are symmetric about $-As_k$ and As_k given the hypotheses H_0 and H_1 of (34) and because both hypotheses are equally likely. The two ML detectors (40)–(43) require that we know the quantizer noise N_q and its intensity σ_N. But they do not require that we know the signal amplitude A or the channel noise pdf parameters r or γ. The generalized-Gaussian ML detector (36) and the Cauchy ML detector (39) do require such knowledge.

Corollary 1 requires that the mean and variance of the detection statistics Λ_{NQ} and $\Lambda_{N\infty}$ in (40) and (42) obey $\mu_0(\sigma_N) = -\mu_1(\sigma_N)$ and $\sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N)$ for all σ_N. These equalities hold because (40)–(43) imply that $\Lambda_{NQ}|H_0 = -\Lambda_{NQ}|H_1$ and $\Lambda_{N\infty}|H_0 = -\Lambda_{N\infty}|H_1$. The pdf of Λ_{NQ} is approximately Gaussian for both hypotheses because the central limit theorem applies to (40) and (42) if the sample size K is large since the random variables $s_k\bar{g}_{NQ}(X_k)$ and $s_k\bar{g}_{N\infty}(X_k)$ are independent even though they are *not* identically distributed. This holds for uniformly bounded random variables that satisfy the Lindeberg condition [70]: $\lim_{K \longrightarrow \infty} Var(\Lambda_{NQ}) = \infty$ and $\lim_{K \longrightarrow \infty} Var(\Lambda_{N\infty}) = \infty$. The variables $s_k\bar{g}_{NQ}(X_k)$ and $s_k\bar{g}_{N\infty}(X_k)$ are uniformly bounded so long as the sequence s_k is bounded. The Lindeberg condition then holds because the noise pdfs have infinite support since the noise is generalized Gaussian or $S\alpha S$. Then the SR noise-benefit conditions of Theorems 1 and 2 also hold for the quantizer-array ML detector (40)–(41) and for its limiting-array $(Q \longrightarrow \infty)$ ML detector (42)–(43). The proof of Corollary 1 mirrors that of Theorem 1 but uses the error probability as the performance measure. We state it for completeness because it is brief.

Corollary 1: Suppose that the detection statistic Λ_{NQ} in (40) has sufficiently large sample size K so that it is approximately conditionally normal: $\Lambda_{NQ}|H_0 \approx \mathbf{N}(\mu_0(\sigma_N), \sigma_0^2(\sigma_N))$ and $\Lambda_{NQ}|H_1 \approx \mathbf{N}(\mu_1(\sigma_N), \sigma_1^2(\sigma_N))$ where $\mu_0(\sigma_N) = -\mu_1(\sigma_N)$ and $\sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N)$. Then the inequality

$$\sigma_1(\sigma_N)\mu_1'(\sigma_N) > \mu_1(\sigma_N)\sigma_1'(\sigma_N) \quad (44)$$

is necessary and sufficient for the SR effect $(dP_e(\sigma_N)/d\sigma_N < 0)$ in the ML detection of (34) using the quantizer-array detector (40)–(41).

Proof: The ML detection rule (40) for Λ_{NQ} rejects H_0 if $\Lambda_{NQ} > 0$. The null hypothesis H_0 and alternative hypothesis H_1 partition the sample space. So the probability P_e of average decision error [59] is

$$P_e(\sigma_N) = p_0 P(\Lambda_{NQ} > 0 \,|\, H_0) + p_0 P(\Lambda_{NQ} < 0 \,|\, H_1) \quad (45)$$

$$= \frac{1}{2} P\left(\frac{\Lambda_{NQ} - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} > \frac{-\mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right) + \frac{1}{2} P\left(\frac{\Lambda_{NQ} - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} < \frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \text{ since } p_0 = p_1 \quad (46)$$

$$\approx \frac{1}{2} P\left(Z > \frac{-\mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right) + \frac{1}{2} P\left(Z < \frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \quad (47)$$

$$\text{for } Z \sim \mathrm{N}(0,1)$$

$$= \frac{1}{2}\left(1 - \Phi\left(\frac{-\mu_0(\sigma_N)}{\sigma_0(\sigma_N)}\right)\right) + \frac{1}{2}\Phi\left(\frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \quad (48)$$

$$= \frac{1}{2}\Phi\left(\frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) + \frac{1}{2}\Phi\left(\frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \quad (49)$$

$$\text{since } \mu_0(\sigma_N) = -\mu_1(\sigma_N) \text{ and } \sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N)$$

$$= \Phi\left(\frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right). \quad (50)$$

Then the normal pdf $\phi(z) = d\Phi(z)/dz$ and the chain rule of differential calculus give

$$\frac{dP_e(\sigma_N)}{d\sigma_N} \approx -\phi\left(\frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)}\right) \times \frac{\sigma_1(\sigma_N)\mu_1'(\sigma_N) - \mu_1(\sigma_N)\sigma_1{}'(\sigma_N)}{\sigma_1^2(\sigma_N)}. \quad (51)$$

So $\sigma_1(\sigma_N)\mu_1'(\sigma_N) > \mu_1(\sigma_N)\sigma_1'(\sigma_N)$ is necessary and sufficient for the SR effect $(dP_e/d\sigma_N < 0)$ because ϕ is a pdf and thus $\phi \geq 0$. □

The SR condition (44) in Corollary 1 allows us to find a near-optimal quantizer-noise standard deviation $\hat{\sigma}_{Nopt}$ for the ML detector if we have enough samples of the received signal vector $\mathbf{X}$ under both hypotheses H_0 and H_1 in (34). Then the pdf of Λ_{NQ} is a mixture of two equally likely Gaussian pdfs: $1/2N(-\mu_1(\sigma_N), \sigma_1^2(\sigma_N)) + 1/2N(\mu_1(\sigma_N), \sigma_1^2(\sigma_N))$. So $\Lambda_{NQ}^2/\sigma_1^2(\sigma_N)$ is a non-central chi-square random variable [53] with noncentrality parameter $\mu_1^2(\sigma_N)$ and 1 degree of freedom. Then

$$Var(\Lambda_{NQ}) = \sigma_1^2(\sigma_N) + \mu_1^2(\sigma_N) \quad (52)$$

$$Var(\Lambda_{NQ}^2) = 2\sigma_1^2(\sigma_N) + 4\sigma_1^2(\sigma_N)\mu_1^2(\sigma_N). \quad (53)$$

Putting $\mu_1^2(\sigma_N)$ from (52) in (53) gives a quadratic equation in $\sigma_1^2(\sigma_N)$ with solution

$$\sigma_1^2(\sigma_N) = Var(\Lambda_{NQ}) - \left[Var^2(\Lambda_{NQ}) - \frac{Var(\Lambda_{NQ}^2)}{2}\right]^{1/2}. \quad (54)$$

So the real part of (54) gives the consistent estimator $\hat{\sigma}_1^2(\sigma_N)$ of $\sigma_1^2(\sigma_N)$ for large sample size if we replace the population variances of Λ_{NQ} and Λ_{NQ}^2 with their sample variances because then the RHS of (54) is a continuous function of consistent estimators. Then $\hat{\sigma}_1^2(\sigma_N)$ can replace $\sigma_1^2(\sigma_N)$ in (52) to give the consistent estimator $\hat{\mu}_1^2(\sigma_N)$ of $\mu_1^2(\sigma_N)$. The same received vector $\mathbf{X}$ allows us to compute $\hat{\sigma}_1^2(\sigma_N)$ and $\hat{\mu}_1^2(\sigma_N)$ for all values of σ_N by (40)–(43). Then a zero-crossing of $\ln\hat{\mu}_1(\sigma_{N_j})/\hat{\sigma}_1(\sigma_{N_j}) - \ln\hat{\mu}_1(\sigma_{N_{j-1}})/\hat{\sigma}_1(\sigma_{N_{j-1}})$ estimates the optimal quantizer-noise standard deviation $\hat{\sigma}_{Nopt}$ for a small step-size of $\Delta\sigma_N = \sigma_{N_j} - \sigma_{N_{j-1}}$ because the SR condition (44) is equivalent to $d/d\sigma_N(\ln\mu_1(\sigma_N)/\sigma_1(\sigma_N)) > 0$ since $\mu_1 > 0$.Section VI uses this zero-crossing method to find a near-optimal standard deviation of uniform quantizer noise for the limiting array detector (42)–(43) in watermark decoding.

The lengthy proof of Corollary 2 below is nearly the same as the proof of Theorem 2 in the Appendix. It replaces the detection probability P_D with the error probability P_e and uses a zero threshold. So we omit it for reasons of space.

Corollary 2: Suppose that the channel noise pdf is uniformly bounded and continuously differentiable. Then

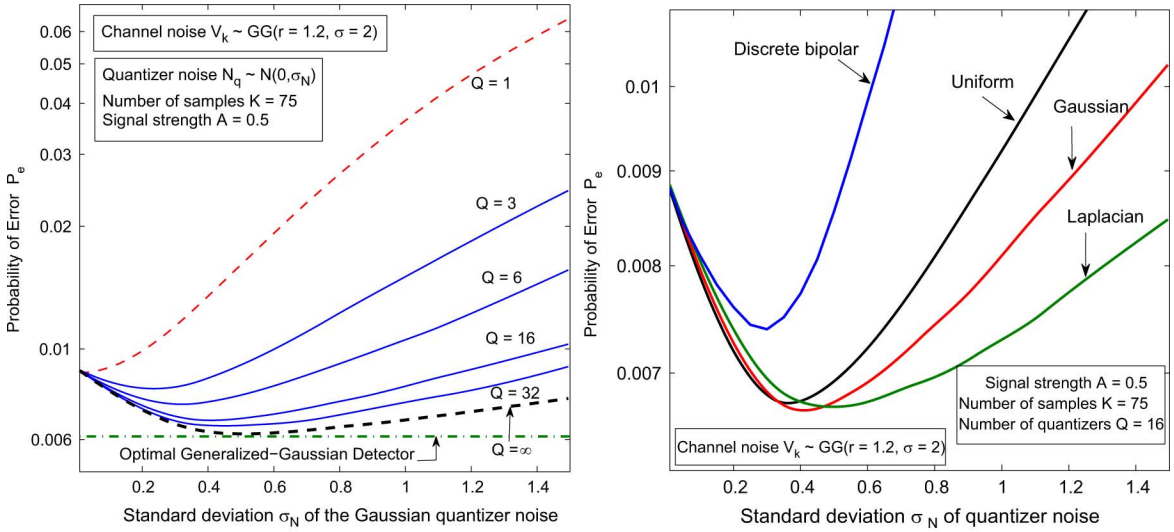
(a) $Q > 1$ is necessary for the initial SR effect in the quantizer-array detector (40)–(41) for the ML detection of (34) in any symmetric unimodal channel noise.

(b) Suppose that the initial SR effect occurs with Q_1 quantizers and with some symmetric quantizer noise in the quantizer-array detector (40)–(41). Then the rate of the initial SR effect in the quantizer-array detector (40)–(41) with Q_2 quantizers is larger than the rate of the initial SR effect with Q_1 quantizers if $Q_2 > Q_1$.

(c) Zero-mean uniform noise is the optimal finite-variance symmetric scale-family quantizer noise in that it gives the maximal rate of the initial SR effect among all possible finite-variance quantizer noise in the ML quantizer-array detector (40)–(41).

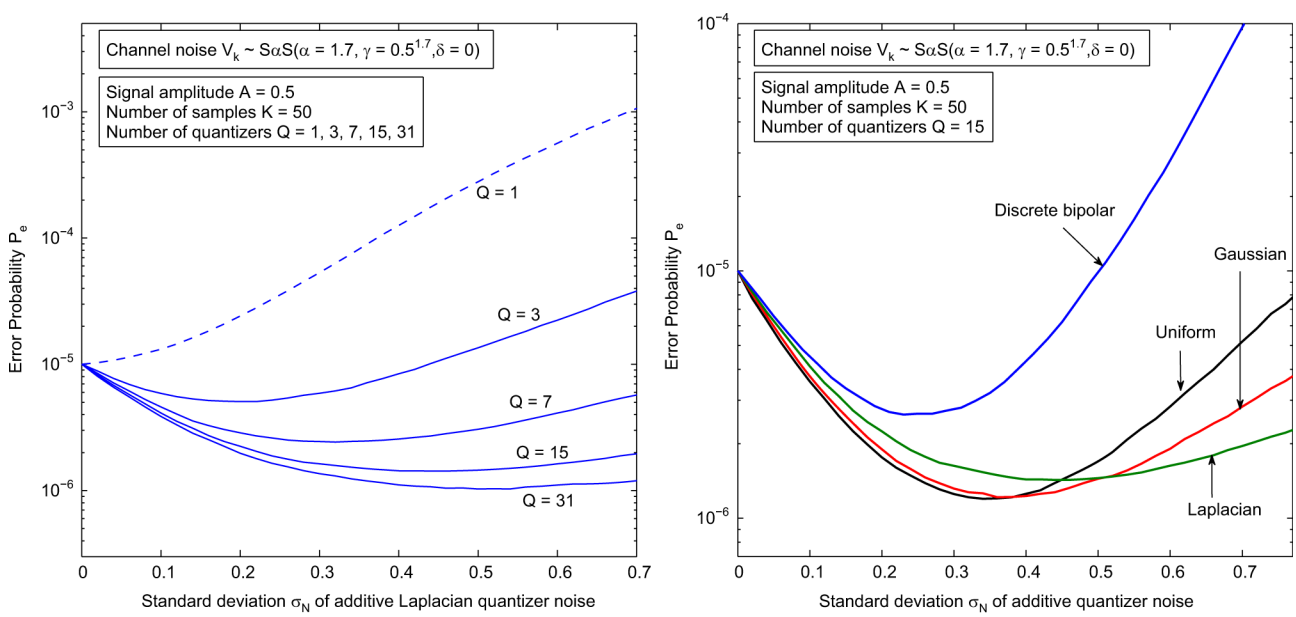
The simulation results in Figs. 5 and 6 show the respective predicted SR effects in the ML detection (34) of signal s_k in generalized-Gaussian channel noise and in $S\alpha S$ channel noise for quantizer-array detectors. The signal was a bipolar sequence with amplitude $A = 0.5$. The respective sample sizes were $K = 75$ and $K = 50$ for these detectors. The channel noise was generalized-Gaussian with parameters $r = 1.2$ and $\sigma = 2$ in Fig. 5. It was $S\alpha S$ with $\alpha = 1.7$ and $\gamma = 0.5^{1.7}$ in Fig. 5. The thin dashed lines in Figs. 5(a) and 6(a) show that the SR effect does not occur if $Q = 1$ as Theorem 2(a) predicts. Figs. 5(a) and 6(a) also show that the rate of the initial SR effect increases as the number Q of quantizers increases as Corollary 2(b) predicts. The thick dashed line in Fig. 5(a) shows the error probability of the limiting-array $(Q \to \infty)$ ML correlation detector (42) with limiting-array Gaussian-quantizer-noise nonlinearity $\bar{g}_{N\infty}(X_k) = 2\Phi(X_k/\sigma_N) - 1$ where we have replaced $F_{\tilde{N}}$ in (43) with the standard normal CDF Φ. The thick horizontal dash-dot line shows the error probability of the optimal generalized-Gaussian ML detector (36). The limiting-array $(Q \to \infty)$ detector does not require that we know the signal amplitude A. It still gave almost the same detection performance as the optimal generalized-Gaussian detector gave.

The simulation results in Figs. 5(b) and 6(b) show the initial SR-rate optimality of symmetric uniform quantizer noise. The symmetric uniform quantizer noise gave the maximal rate of the initial SR effect as Corollary 2(c) predicts. Gaussian noise gave the best peak SR effect in Fig. 5(b) in the sense that it had the



(a) (b)

Fig. 5. Initial SR effects in ML quantizer-array detection. The signal was a bipolar sequence with amplitude $A = 0.5$. The channel noise was generalized-Gaussian with parameters $r = 1.2$ and $\sigma = 2$. (a) Initial SR effects for zero-mean Gaussian quantizer noise. The initial SR effect does not occur if $Q = 1$ as Corollary 2(a) predicts. The rate of the initial SR effect increased as the number Q of quantizers increased as Corollary 2(b) predicts. The thick dashed line shows the error probability P_e of the respective limiting-array $(Q \to \infty)$ detector. This detection performance was nearly optimal compared to the optimal generalized-Gaussian detector (36) (thick horizontal dashed-dot line). (b) Comparison of initial SR noise benefits in the ML quantizer-array detector (40)–(41) for four different types of quantizer noise. Symmetric uniform noise gave the maximal rate of the initial SR effect as Corollary 2(c) predicts. But Gaussian noise gave the best peak SR effect because it had the largest decrease in error probability. Laplacian quantizer noise gave the most robust SR effect and had almost the same peak SR effect as Gaussian noise had. Symmetric discrete bipolar noise gave the smallest SR effect and was least robust.



(a) (b)

Fig. 6. Initial SR effects in the quantizer-array ML detection of a known bipolar sequence s_k of unknown amplitude A in infinite-variance $S\alpha S$ channel noise with $\alpha = 1.7$ and $\gamma = 0.5^{1.7}$. (a) Initial SR effects for Laplacian quantizer noise. The initial SR effect did not occur if $Q = 1$ as Corollary 2(a) predicts. The rate of initial SR effect increased as the number Q of quantizers increased as Corollary 2(b) predicts. (b) Comparison of initial SR effects in the quantizer-array ML detection of a deterministic bipolar sequence s_k of unknown amplitude A in infinite-variance $S\alpha S$ channel noise with different types of quantizer noise. The symmetric uniform noise gave the maximal rate of the initial SR effect as Corollary 2(c) predicts. Symmetric discrete bipolar noise gave the smallest SR effect and was least robust. Symmetric uniform noise had the peak SR effect in the sense that it had the largest decrease in error probability. Gaussian noise gave almost the same peak SR effect as uniform noise did. The SR effect was most robust against Laplacian quantizer noise.

maximal decrease in error probability. Symmetric uniform noise gave the highest peak SR effect in Fig. 6(b) when compared with symmetric Laplacian, Gaussian, and discrete bipolar noise even though Corollary 2(c) does not guarantee such optimality for the peak SR effect. Symmetric discrete bipolar noise gave the smallest SR effect and was least robust.

VI. Watermark Decoding Using the Noisy Nonlinear-Correlation Detector

The above noisy nonlinear detectors and their limiting-array $(Q \to \infty)$ detectors can benefit digital watermark extraction. We will demonstrate this for blind watermark decoding based on direct-sequence spread spectrum in the DCT domain [33], [71], [72]. Digital watermarking helps protect copyrighted multimedia data because it hides or embeds a mark or signal in the data without changing the data too much [73], [74]. Transform-domain watermarking techniques tend to be more robust and tamper-proof than direct spatial watermarking [74], [75]. So we use the popular direct-sequence spread-spectrum approach to watermarking in the DCT domain because such spreading gives a robust but invisible watermark and because it allows various types of detectors for blind watermark extraction [33], [71], [72].

DCT watermarking adds or embeds the watermark signals $W[\mathbf{k}]$ of an $M_1 \times M_2$ binary watermark image $B(\mathbf{n})$ in the DCT-2 coefficients $V[\mathbf{k}]$ of an $L_1 \times L_2$ host image $H(\mathbf{n})$. Here $\mathbf{k} = (k_1, k_2)$ and $\mathbf{n} = (n_1, n_2)$ are the respective 2-D indices in the transform domain and in the spatial domain. We apply an 8×8 block-wise DCT-2 transform [72]. Using a block-wise DCT involves less computation than using a full-image DCT.

The watermark image B is an $M_1 \times M_2$ black-and-white image such that about half its pixels are black and half are white. This watermark image B gives a watermark message $\mathbf{b} = (b_1, \ldots, b_M)$ with $M(= M_1 \times M_2)$ bipolar (± 1) bits such that $b_i = 1$ if $B(n_1, n_2) = 255$ (white) and such that $b_i = -1$ if $B(n_1, n_2) = 0$ (black) where $i = n_1 + (n_2 - 1) \times M_1$.

A secret key picks a pseudorandom assignment of disjoint subsets D_i of size or cardinality K from the set D of mid-frequency DCT-coefficients for each message bit b_i. Denote $\mathcal{I}_i$ as the set of 2-D indices of the DCT-coefficient set $\mathcal{D}_i$: $\mathcal{I}_i = \{\mathbf{k} : V_i[\mathbf{k}] \in \mathcal{D}_i\}$. Then $|\mathcal{D}_i| = |\mathcal{I}_i| = K$. The secret key also gives the initial seed to a pseudorandom sequence generator that produces a bipolar spreading sequence $\tilde{s}_i[\mathbf{k}]$ of length K $(\mathbf{k} \in \mathcal{I}_i)$ for each message bit b_i [33], [59]. Each $\tilde{s}_i$ is a pseudorandom sequence of K i.i.d. random variables in $\{1, -1\}$. The cross-correlation among these M spreading sequences $\{\tilde{s}_i\}_{i=1}^{M}$ equals zero while their autocorrelation equals the Kronecker delta function.

We embed the information of bit b_i in each DCT-coefficient $V_i[\mathbf{k}]$ of the set $\mathcal{D}_i$ using the corresponding spreading sequence $\tilde{s}_i[\mathbf{k}]$ of length K. Each bipolar message bit b_i multiplies its biploar pseudorandom spreading sequence $\tilde{s}_i$ to "spread" the spectrum of the original message signal over many frequencies. We use the psychovisual DCT-domain perceptual mask $a[\mathbf{k}] = A\tilde{T}[\mathbf{k}]$ of [26] to obtain the watermark embedding strength that reduces the visibility of the watermark in the watermarked image [76]. Here $\tilde{T}[\mathbf{k}]$ is the known shape of the perceptual mask and $0 < A < 1$ is the known or unknown scaling factor. This perceptual mask $a[\mathbf{k}]$ also multiplies the pseudorandom sequence $\tilde{s}_i[\mathbf{k}]$ to give the watermark signal $W_i[\mathbf{k}] = b_i a[\mathbf{k}] \tilde{s}_i[\mathbf{k}]$ for each $\mathbf{k}$-pixel in the DCT domain. We then add $W_i[\mathbf{k}]$ to the host-image DCT-2-coefficient $V_i[\mathbf{k}] \in \mathcal{D}_i$. This gives the watermarked DCT-2 coefficient $X_i[\mathbf{k}] = V_i[\mathbf{k}] + b_i a[\mathbf{k}] \tilde{s}_i[\mathbf{k}]$. Then the inverse block-wise DCT-2 transform gives a watermarked image $H_W[\mathbf{n}]$.

Retrieving the hidden message $\mathbf{b}$ requires that we know the pseudorandom assignment of DCT-coefficients $\{V_i[\mathbf{k}] | \mathbf{k} \in \mathcal{D}_i\}$ to each message bit b_i in $\mathbf{b}$ and that we also know the pseudorandom sequence $\tilde{s}_i$ for each b_i. Then an attacker cannot extract the watermark without the secret key. So suppose that we do know the secret key. Then watermark decoding just tests M binary hypotheses:

$$\begin{aligned} H_0\ (b_i = -1) &: X_i[\mathbf{k}] = -As_i[\mathbf{k}] + V_i[\mathbf{k}] \\ H_1\ (b_i = +1) &: X_i[\mathbf{k}] = As_i[\mathbf{k}] + V_i[\mathbf{k}] \end{aligned} \tag{55}$$

for all $\mathbf{k} \in \mathcal{I}_i$ and for $i = 1, \ldots, M$. Here $s_i = \tilde{s}_i[\mathbf{k}]\tilde{T}[\mathbf{k}]$ is the known signal sequence and A is the known or unknown scaling factor of the perceptual mask. Define $\mathbf{X}_i = \{X_i[\mathbf{k}] | \mathbf{k} \in \mathcal{I}_i\}$. The optimal decision rule to decode the message bit b_i is the ML rule

$$\Lambda_W(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} \log\left(\frac{f(X_i[\mathbf{k}]|H_1)}{f(X_i[\mathbf{k}]|H_0)}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{56}$$

because we assume that the DCT-2 coefficients $V_i[\mathbf{k}]$ are i.i.d. random variables and that the message bits b_i are equally likely to be -1 or 1. So the ML detection rule (56) becomes a simple linear correlator [59]

$$\Lambda_{Lin}(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} s_i[\mathbf{k}] X_i[\mathbf{k}] \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{57}$$

if the DCT-2 coefficients are Gaussian random variables. The ML detection rule (56) becomes the generalized-Gaussian (GG) detector or decoder

$$\Lambda_{GG_{ML}}(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} \left[|X_i[\mathbf{k}] + As_i[\mathbf{k}]|^r - |X_i[\mathbf{k}] - As_i[\mathbf{k}]|^r\right] \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{58}$$

or the Cauchy detector

$$\Lambda_{C_{ML}}(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} \log\left(\frac{\gamma^2 + (X_i[\mathbf{k}] + As_i[\mathbf{k}])^2}{\gamma^2 + (X_i[\mathbf{k}] - As_i[\mathbf{k}])^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{59}$$

if the respective DCT-2 coefficients have a generalized Gaussian or Cauchy pdf. But these optimal ML detectors require that we know the scaling factor A of the perceptual mask and that we know the pdf parameters r and γ. Both the suboptimal quantizer-array detector (40)–(41)

$$\Lambda_{NQ}(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} s_i[\mathbf{k}] \bar{g}_{NQ}(X_i[\mathbf{k}]) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{60}$$

$$\text{where} \quad \bar{g}_{NQ}(X_i[\mathbf{k}]) = \frac{1}{Q} \sum_{q=1}^{Q} sign(X_i[\mathbf{k}] + N_q) \tag{61}$$

and its limiting-array nonlinear correlation detector (42)–(43)

$$\Lambda_{N\infty}(\mathbf{X}_i) = \sum_{\mathbf{k} \in \mathcal{I}_i} s_i[\mathbf{k}] \bar{g}_{N\infty}(X_i[\mathbf{k}]) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{62}$$

$$\text{where} \quad \bar{g}_{N\infty}(X_i[\mathbf{k}]) = 2F_{\tilde{N}}\left(\frac{X_i[\mathbf{k}]}{\sigma_N}\right) - 1 \tag{63}$$

require that we know the quantizer noise N_q and its intensity σ_N. They do not require that we know the scaling factor A of the perceptual mask or the channel noise pdf parameters r or γ. We use the zero-crossing method of Section V to find a near-optimal quantizer-noise standard deviation for these array-based detectors. The noise-based algorithm below summarizes the processes of watermark embedding and watermark decoding.

Watermark Embedding

1. Compute the 8×8 block-wise DCT-2 transform $V[\mathbf{k}]$ of the $L_1 \times L_2$ host image $H[\mathbf{n}]$.
2. Let D be the set of mid-frequency DCT-2 coefficients of all 8×8 DCT blocks of $V[\mathbf{k}]$.
3. Convert the $M_1 \times M_2$ binary (black-and-white) watermark image $B(\mathbf{n})$ into an $M_1 \times M_2 = M$-bit bipolar watermark message $\mathbf{b} = (b_1, \ldots, b_M)$ such that $b_i = 1$ if $B(n_1, n_2) = 255$ (white) or $b_i = -1$ if $B(n_1, n_2) = 0$ (black) where $i = n_1 + (n_2 - 1) \times M_1$.
4. Use a secret key to generate M pseudorandom disjoint subsets D_i of size K from the set D.
5. Let $\mathcal{I}_i$ be the set of two-dimensional indexes of the DCT-coefficient set $\mathcal{D}_i$.
6. Use the secret key to generate M pseudorandom bipolar spreading sequences $\tilde{s}_i[\mathbf{k}]$ of length K where $\tilde{s}_i[\mathbf{k}] = \pm 1$ for all $\mathbf{k} \in \mathcal{I}_i$ and $i = 1, \ldots, M$.
7. For each message bit b_i: compute the watermark signals $W_i[\mathbf{k}] = b_i a[\mathbf{k}] \tilde{s}_i[\mathbf{k}]$ for all $\mathbf{k} \in \mathcal{I}_i$ where $a[\mathbf{k}]$ is the perceptual mask in [26] with scale factor A.
8. For each message bit b_i: compute the watermarked DCT-2 coefficients $X_i[\mathbf{k}] = V_i[\mathbf{k}] + b_i a[\mathbf{k}] \tilde{s}_i[\mathbf{k}]$ for all $\mathbf{k} \in \mathcal{I}_i$.
9. Compute the inverse block-wise DCT-2 transform using the watermarked DCT-coefficients $X[\mathbf{k}]$ to get the watermarked image $H_W[\mathbf{n}]$.

Watermark Decoding

1. Compute the 8×8 block-wise DCT-2 transform coefficients $X[\mathbf{k}]$ of the $L_1 \times L_2$ watermarked host image $H_W[\mathbf{n}]$.
2. Use the secret key to obtain the index sets $\mathcal{I}_i$ for $i = 1, \ldots, M$.
3. Obtain M sets of watermarked DCT-coefficients $\mathbf{X}_i = \{X_i[\mathbf{k}] | \mathbf{k} \in \mathcal{I}_i\}$ for $i = 1, \ldots, M$.
4. Use the secret key to reproduce the pseudo spreading sequences $\{\tilde{s}_i\}_{i=1}^M$.
5. Find the decoded message bits $\{\hat{b}_i\}_{i=1}^M$ using one of the following ML decoders for M binary hypothesis tests in (55):

Without Noise Injection

<u>IF</u> the user knows the scaling factor A of the perceptual mask

- Use the Cauchy detector (60) (estimate the dispersion γ) <u>or</u>
- Use the GG detector (60) (estimate the shape parameter r)

<u>Else</u>

- Use the Cauchy detector (60) (estimate both A and γ) <u>or</u>
- Use the limiting-array $(Q \to \infty)$ correlation detector (62)–(63) for uniform quantizer noise and use the zero-crossing method of Section V to find a near-optimal qunatizer-noise intensity <u>or</u>

With Noise Injection

- Use the noisy quantizer-array detector (61)–(62) with quantizer number $Q = 100$ and use the zero-crossing method of Section V to find a near-optimal qunatizer-noise intensity.

<u>End</u>

Fig. 1 shows a pronounced SR noise benefit in the watermark decoding of the quantizer-array detector (60)–(61) and its limiting-array nonlinear correlation detector (62)–(63) when the quantizer noise is symmetric uniform noise. The simulation software was Matlab. Fig. 1(a) shows the yin-yang image. We used this binary (black and white) 64×64 image as a hidden message to watermark the 512×512 gray-scale Lena image. Fig. 1(b) shows the Lena image watermarked with the yin-yang image such that its peak signal-to-noise ratio (PSNR) was 46.6413 dB. We define the PSNR as the ratio of the maximum power of the original host image's pixels $H[\mathbf{n}]$ to the average power of the difference between the original image H and the watermarked image H_W:

$$PSNR = 10 \log_{10} \left[\frac{\max\limits_{\mathbf{n} \in \{1,\ldots,L1\} \times \{1,\ldots,L_2\}} H^2[\mathbf{n}]}{\frac{|H[\mathbf{n}] - H_W[\mathbf{n}]|^2}{(L_1 \times L_2)}} \right]. \quad (64)$$

Each DCT-coefficient set $\mathcal{D}_i$ of the Lena image in Fig. 1(a) hides one message bit of the yin-yang image using a Matlab-generated pseudorandom bipolar spreading sequence $\tilde{s}_i$. The watermark had a constant amplitude $(a[\mathbf{k}] = A)$ and so did not involve psychovisual properties. The solid U-shaped line in Fig. 1(c) shows the average pixel-detection errors of the ML noisy quantizer-array detector (61)–(62) for 200 randomly generated secret keys. The dashed vertical lines show the total min-max deviation of the pixel-detection errors in these simulation trials. The dashed U-shaped line shows the average pixel-detection errors of the limiting-array $(Q \to \infty)$ correlation detector (62)–(63) where $\bar{g}_{N\infty}$ is the soft-limiter nonlinearity $\bar{g}_{SL}$ of (8) because the quantizer noise is symmetric uniform noise. So the thick dashed line gives the lower bound on the pixel-detection error that any quantizer-array detector with symmetric uniform quantizer noise can achieve by increasing the number Q of quantizers in its array.

Figs. 1(d)–1(g) show the extracted yin-yang image using the ML linear correlation detector (57) and the ML noisy quantizer-array detector (60)–(61) according to the noise-based algorithm. The noisy quantizer-array nonlinear detector outperforms the linear correlation detector. Figs. 1(e) and 1(f) show that adding uniform quantizer noise improved the watermark decoding. The pixel-detection errors decreased by more than 33% as the uniform quantizer noise standard deviation σ increased from $\sigma = 0$ to $\sigma = 1$. Fig. 1(g) shows that too much quantizer noise degraded the watermark detection. But the SR effect was robust against the quantizer noise intensity because the pixel-detection error in (g) was still less than the pixel-detection errors in (d) and (e).

We also watermarked Lena and six other known images (Elaine, Goldhill, Pirate, Peppers, Bird, and Tiffany) with the yin-yang image using a perceptual mask based on the psychovisual properties in [26]. Fig. 7 shows these six watermarked images. We used the scaling factor A of the psychovisual perceptual mask such that the PSNRs of all these watermarked images remained between 43 dB and 47 dB. We used over 200 simulation trials for each of these seven images.

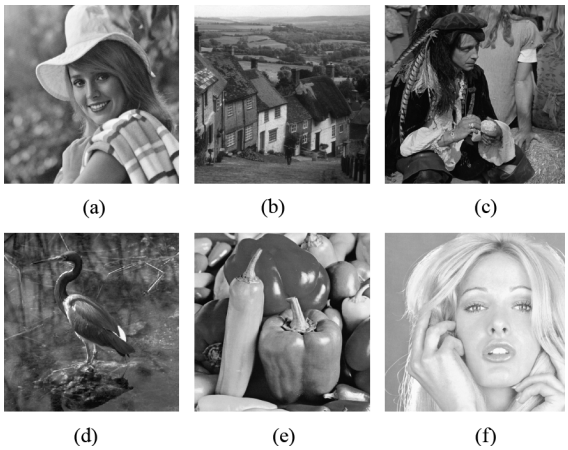


Fig. 7. Six different 512×512 host images watermarked with the yin-yang image in Fig. 1(a) using a perceptual mask [26] based on the psychovisual properties. (a) Elaine. (b) Goldhill. (c) Pirate. (d) Bird. (e) Peppers. (f) Tiffany. The PSNR of each watermarked image was between 43 to 47 dB.

TABLE I
MINIMAL VALUES OF AVERAGE WATERMARK-DECODING PERFORMANCE FOR THE CAUCHY DETECTOR, THE LIMITING-ARRAY ($Q \rightarrow \infty$) DETECTOR, AND THE GENERALIZED-GAUSSIAN (GG) DETECTOR. THE WATERMARKING USED A PERCEPTUAL MASK [26] BASED ON PSYCHOVISUAL PROPERTIES

Images	Minimal Average Pixel Errors in Watermark Extraction			
	limiting array Detector	Cauchy Detector	GG Detector	Linear Detector
Elaine	59	131 (61)	144 (63)	156
Goldhill	112	350 (109)	390 (120)	560
Pirate	105	357 (90)	392 (98)	645
Bird	178	317 (173)	360 (190)	561
Peppers	26	120 (24)	153 (29)	376
Lena	7	79 (5)	100 (8)	242
Tiffany	0.3	27 (0.2)	39 (0.22)	151

A set of 200 randomly generated secret keys allocated the DCT-2-coefficients and produced the spreading sequences for the watermarking process of each host image. We then applied the ML Cauchy detector, the limiting-array ($Q \rightarrow \infty$) detector for the symmetric uniform quantizer noise, and the generalized-Gaussian detector for various values of their respective pdf parameters γ, σ_N, and r to each of the seven watermarked images to decode the watermark using each secret key.

Table I shows the minimal values of the average pixel-detection errors for each host image and for each detector. The numbers in parentheses show the pixel decoding errors for the Cauchy detector and the generalized-Gaussian detector when the detectors used the scale factor A of the pyschovisual perceptual mask. The linear detector and the limiting array detector did not need the watermark scale factor A. The limiting array detector for the symmetric uniform quantizer noise performed substantially better than did the other two nonlinear detectors when the detectors did not use the scale factor A. The Cauchy detector gave the best performance otherwise. The computational complexity for all the detectors was only of order K for all arithmetic and logical operations where the length of the signal was K. The limiting array detector also had the least complexity among the nonlinear detectors while the generalized-Gaussian detector had the most.

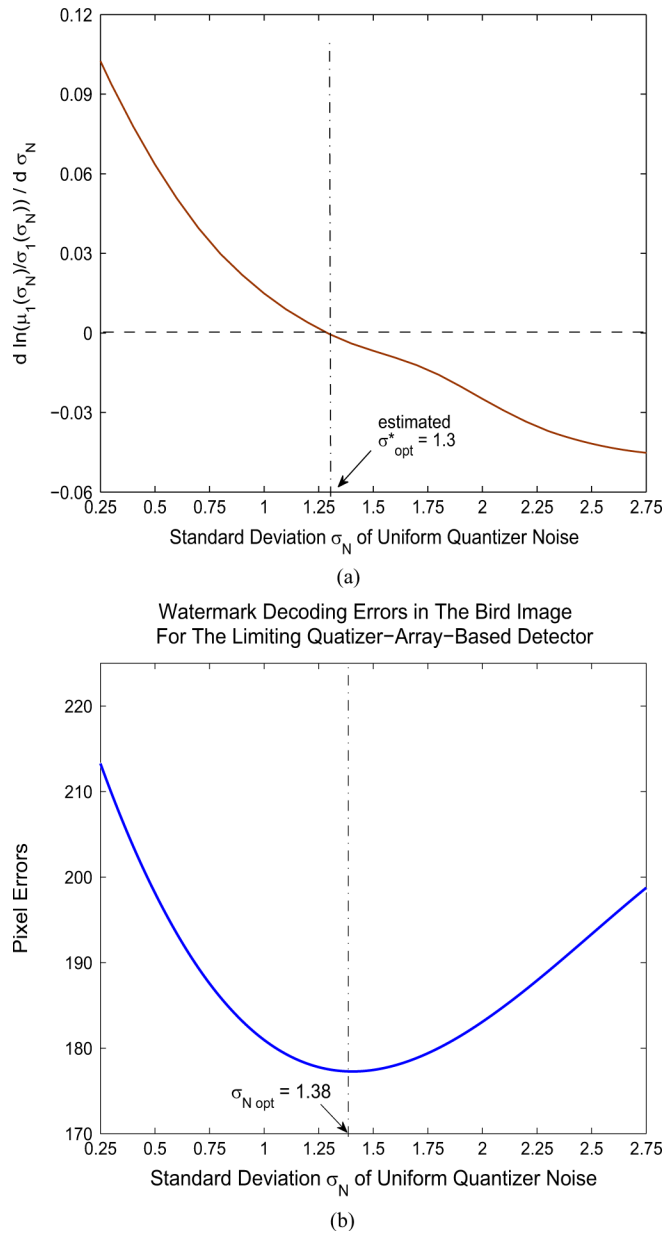


Fig. 8. Finding a near-optimal quantizer-noise standard deviation for the limiting-array detector (62)–(63) in the watermark decoding of the 'Bird' image.

Psychovisual watermarking greatly reduced decoding errors compared with constant-amplitude watermark. Consider the Lena image: Table I and Fig. 1(c) and (d) show that the limiting-array detector had 7 versus 43 pixel errors and the linear detector had 242 versus 672 pixel errors.

Fig. 8 shows the estimation accuracy of the zero-crossing method in Section V for the limiting-array detector (62)–(63) in the watermark decoding of the Bird image. The method gives the near-optimal quantizer-noise standard deviation of 1.3 in Fig. 8(a). Fig. 8(b) shows the true optimal quantizer-noise standard deviation of 1.38 for the pixel detection error.

Noise can also benefit DCT-domain watermark decoding for some other types of nonlinear detectors. Researchers [77]–[79] have found noise benefits in DCT-domain watermark decoding based on parameter-induced stochastic resonance [80]. Their approaches differed from ours in two main ways. They used a

pulse-amplitude-modulated antipodal watermark signal but did not use pseudorandom bipolar spreading sequences to embed this watermark signal in the DCT coefficients. They further used nonlinear but *dynamical* detectors to process the watermarked DCT coefficients. Sun *et al.* [77] used a monostable system with selected parameters for a given watermark. Wu and Qiu [81] and Sun and Lei [78] used a single bistable detector while Duan *et al.* [79] used an array of bistable saturating detectors. These bistable detectors require tuning two system parameters besides finding the optimal additive noise type and its variance. These detectors are also subthreshold systems. Their dynamical nature made the error probability in the decoding of one watermark bit depend on the value of the previous watermark bit. Our noisy quantizer-array detectors produced *suprathreshold* SR noise benefits [5]. The detection error probabilities for the watermark bits were also independent.

VII. Conclusion

Noise-enhanced quantizer-array correlation detection leads naturally to noise-enhanced watermark decoding because we can view digital watermarking systems as digital communication systems [82] with statistical hypothesis testing at the pixel level. Such noise benefits will occur if the symmetric unimodal channel noise is not Gaussian and only if the array contains more than one quantizer. This noise-enhancement technique should apply to other problems of signal processing and communications that involve some type of nonlinear statistical hypothesis testing. The paper also showed that uniform noise gives the optimal initial rate of the SR noise benefit among all finite-variance symmetric scale-family quantizer noise for both NP and ML detection. Finding practical algorithms for the overall optimal quantizer noise remains an open research question.

Appendix
Proof of Theorem 2

Proof: **Part (a)** The definition of the initial SR effect and Theorem 1 imply that the initial SR effect occurs if and only if there exists some $b > 0$ such that inequality (12) holds for all quantizer-noise intensities $\sigma_N \in (0, b)$. Then multiply both sides of (12) by $2\sigma_1(\sigma_N)$ and use the chain rule to get

$$\mu_1(\sigma_N)\sigma_1^{2\prime}(\sigma_N) < 2\sigma_1^2(\sigma_N)\mu_1'(\sigma_N) \tag{65}$$

for all $\sigma_N \in (0, b)$ as a necessary condition for the initial SR effect. We will prove that inequality (65) does not hold for all quantizer-noise intensities σ_N in some positive interval if $Q = 1$. We first derive the equations for $\mu_1(\sigma_N)$, $\mu_1'(\sigma_N)$, $\sigma_1^2(\sigma_N)$, and $\sigma_1^{2\prime}(\sigma_N)$. We then show that the RHS of (65) is negative in some noise-intensity interval $(0, h)$ while the left-hand side (LHS) of (65) is positive in the same interval.

Define first the random variables $Y_k = g_{NQ}(X_k)$ for $k = 1, \ldots, K$. Then $\Lambda_{NQ} = \sum_{k=1}^{K} Y_k$ and the population mean of $\Lambda_{NQ}|H_1$ is

$$\mu_1(\sigma_N) = E(\Lambda_{NQ}|H_1) = \sum_{k=1}^{K} E(Y_k|H_1). \tag{66}$$

The $Y_k|H_1$ are i.i.d. random variables with population mean

$$E(Y_k|H_1) = \int_{-\infty}^{\infty} E(Y_k|X_k = x; H_1) f_{X_k|H_1}(x)dx \tag{67}$$

$$= \int_{-\infty}^{\infty} E(Y_k|V_k = v; H_1) f_V(v - A)dv. \tag{68}$$

Here A is the signal amplitude and $E(Y_k|X_k = x; H_1)$ is the conditional mean with received signal $X_k = x$ when the alternative hypothesis H_1 is true. Then (68) follows because $X_k|H_1 = V_k + A$ and because the V_k are i.i.d. channel-noise random variables with common pdf f_V. Write $E(Y_k|V_k = v; H_1) = E(Y_k|v; H_1)$ for brevity.

Define next $Y_{k,q} = sign(X_k + N_q - \theta)$ where $\theta = A/2$ and the N_q are finite-variance i.i.d. scale-family quantizer-noise random variables with variance σ_N^2 and CDF F_N. Then

$$E(Y_k|v; H_1) = E(Y_{k,q}|v, H_1) = 1 - 2F_N\left(\frac{A}{2} - v\right) \tag{69}$$

$$= 1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2} - v}{\sigma_N}\right) \tag{70}$$

where $F_{\tilde{N}}$ is the standard CDF for the scale family of the quantizer noise. So (66), (68), and (70) imply that

$$\mu_1(\sigma_N) = \int_{-\infty}^{\infty} K\left[1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2} - v}{\sigma_N}\right)\right] f_V(v - A)dv. \tag{71}$$

$$> 0 \tag{72}$$

because $1 - 2F_{\tilde{N}}(A/2 - v/\sigma_N)$ is a nondecreasing odd function around $A/2$ and because $f_V(v - A)$ is a symmetric unimodal density with mode A.

We next derive an expression for $\mu_1'(\sigma_N)$ and then find the limit $\lim_{\sigma_N \to 0} \mu_1'(\sigma_N)/\sigma_N$. Note first that

$$\mu_1'(\sigma_N) = \frac{\partial}{\partial \sigma_N} \int_{-\infty}^{\infty} K\left[1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2} - v}{\sigma_N}\right)\right] \times f_V(v - A)dv \tag{73}$$

$$= \int_{-\infty}^{\infty} \left(\frac{\frac{A}{2} - v}{\sigma_N}\right) \frac{K}{\sigma_N} f_{\tilde{N}}\left(\frac{\frac{A}{2} - v}{\sigma_N}\right) \times f_V(v - A)dv \tag{74}$$

because of the distributional derivatives [83] of $\mu_1(\sigma_N)$ and $[1 - 2F_{\tilde{N}}(A/2 - v/\sigma_N)]$ with respect to σ_N. This allows us to interchange the order of integration and differentiation [83] in (73)–(74). Next substitute $A/2 - v/\sigma_N = n$ in (74) to get

$$\frac{\mu_1'(\sigma_N)}{\sigma_N} = \int_{-\infty}^{\infty} nK \frac{f_V\left(-\frac{A}{2} - n\sigma_N\right)}{\sigma_N} f_{\tilde{N}}(n)dn \tag{75}$$

$$= -\int_{0}^{\infty} nK \frac{f_V\left(-\frac{A}{2} + n\sigma_N\right) - f_V\left(-\frac{A}{2} - n\sigma_N\right)}{\sigma_N} f_{\tilde{N}}(n)dn \tag{76}$$

because $f_{\tilde{N}}$ is a symmetric pdf. The mean-value theorem [84] implies that for any $\epsilon > 0$

$$\left|\frac{f_V\left(-\frac{A}{2}+n\sigma_N\right)-f_V\left(-\frac{A}{2}-n\sigma_N\right)}{\sigma_N}\right| \leq 2n \sup_{|-(A/2)-u|\leq\epsilon} f_V'(u) \quad (77)$$

for all $|n\sigma_N| \leq \epsilon$. The supremum in (77) is finite for some small ϵ because the pdf derivative f_V' is continuous at $-A/2$. Then Lebesgue's dominated convergence theorem [84] allows us to commute the limit and integral in (76) because the RHS of (77) bounds or dominates the LHS:

$$\lim_{\sigma_N \to 0} \frac{\mu_1'(\sigma_N)}{\sigma_N} = \int_0^\infty -nK \lim_{\sigma_N \to 0} \frac{f_V\left(-\frac{A}{2}+n\sigma_N\right)-f_V\left(-\frac{A}{2}-n\sigma_N\right)}{\sigma_N} f_{\tilde{N}}(n)dn \quad \text{by dominated convergence} \quad (78)$$

$$= -K\int_0^\infty 2n^2 f_V'\left(-\frac{A}{2}\right) f_{\tilde{N}}(n)dn \quad (79)$$

by L'Hospital's rule [85]

$$= -Kf_V'\left(-\frac{A}{2}\right)\int_0^\infty 2n^2 f_{\tilde{N}}(n)dn \quad (80)$$

$$= -2Kf_V'\left(-\frac{A}{2}\right) \quad (81)$$

because $f_{\tilde{N}}$ is the symmetric pdf of the zero-mean and unit-variance quantizer noise $\tilde{N}$. The unimodality of the symmetric channel noise V implies that (82) is negative. Then $\mu_1'(\sigma_N)$ is negative for all noise intensities σ_N in some interval $(0, h)$. Then (82) also implies that

$$\lim_{\sigma_N \to 0} \mu_1'(\sigma_N) = 0 \quad (82)$$

because $\sigma_N \to 0$ in $\mu_1'(\sigma_N)/\sigma_N$.

We now derive expressions for the population variance $\sigma_1^2(\sigma_N)$ of $\Lambda_{NQ}|H_1$ and its distributional derivative $\sigma_1^{2\prime}(\sigma_N)$. The $Y_k|H_1$ are i.i.d. random variables. So

$$\sigma_1^2(\sigma_N) = Var\left(\Lambda_{NQ}|H_1\right) \quad (83)$$
$$= KVar(Y_k|H_1) \quad (84)$$
$$= K\left[E(Y_k^2|H_1) - E^2(Y_k|H_1)\right] \quad (85)$$

where

$$E(Y_k|H_1) = \frac{\mu_1(\sigma_N)}{K} \quad (86)$$

and

$$E(Y_k^2|H_1) = \int_{-\infty}^\infty E(Y_k^2|v, H_1) f_V(v-A)dv. \quad (87)$$

Expand the integrand term $E(Y_k^2|v, H_1)$ as follows:

$$E(Y_k^2|v; H_1) = \frac{1}{Q}E(Y_{k,q}^2|v, H_1) + \frac{Q-1}{Q}E^2(Y_{k,q}|v, H_1) \quad (88)$$

$$= \frac{1}{Q} + \frac{Q-1}{Q}E^2(Y_{k,q}|v, H_1) \quad (89)$$

because $E(Y_{k,q}^2|v, H_1) = 1$ by the definition of $Y_{k,q}$

$$= \frac{1}{Q} + \frac{Q-1}{Q}\left[1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2}-v}{\sigma_N}\right)\right]^2 \quad (90)$$

because of (70). Put (90) in (87) and then put (86) and (87) in (85). This gives

$$\sigma_1^2(\sigma_N) = \int_{-\infty}^\infty \frac{K}{Q} f_V(v-A)dv - \frac{\mu_1^2(\sigma_N)}{K} + \int_{-\infty}^\infty \frac{K(Q-1)}{Q}\left[1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2}-v}{\sigma_N}\right)\right]^2 f_V(v-A)dv. \quad (91)$$

Then the distributional derivative of $\sigma_1^2(\sigma_N)$ with respect to σ_N is

$$\sigma_1^{2\prime}(\sigma_N) = \left[\int_{-\infty}^\infty \frac{K(Q-1)}{Q} 2\left[1 - 2F_{\tilde{N}}\left(\frac{\frac{A}{2}-v}{\sigma_N}\right)\right]\left(\frac{\frac{A}{2}-v}{\sigma_N}\right) \frac{1}{\sigma_N} f_{\tilde{N}}\left(\frac{\frac{A}{2}-v}{\sigma_N}\right) f_V(v-A)dv\right] - \frac{2\mu_1(\sigma_N)\mu_1'(\sigma_N)}{K}. \quad (92)$$

Equation (92) implies that $\sigma_1^{2\prime}(\sigma_N)$ is positive for all $\sigma_N \in (0, h)$ if $Q = 1$ because $\mu_1'(\sigma_N)$ is negative in the quantizer-noise-intensity interval $(0, h)$ and because of (72). So the LHS of (65) is positive and the RHS of (65) is negative for all $\sigma_N \in (0, h)$ if $Q = 1$. Hence $Q > 1$ is necessary for the initial SR effect in the NP detection of a dc signal A in symmetric unimodal channel noise V using the nonlinear test statistic Λ_{NQ}.

Part (b) Take the limit $\sigma_N \to 0$ on both sides of (24) to find the rate of the initial SR effect near a zero quantizer-noise intensity

$$\lim_{\sigma_N \to 0} \frac{dP_D}{d\sigma_N} = 2\phi\left(z_\tau - \frac{2\mu_1(0)}{\sigma_1(0)}\right) \times \frac{\sigma_1(0)\mu_1'(0) - \mu_1(0)\lim_{\sigma_N \to 0}\sigma_1{}'(\sigma_N)}{\sigma_1^2(0)} \quad (93)$$

$$= -2\phi\left(z_\tau - \frac{2\mu_1(0)}{\sigma_1(0)}\right)\frac{\mu_1(0)\lim_{\sigma_N \to 0}\sigma_1{}'(\sigma_N)}{\sigma_1^2(0)} \quad (94)$$

because of (82). We know that

$$\lim_{\sigma_N \to 0} \sigma_1^{2\prime}(\sigma_N) = 2\sigma_1(0) \lim_{\sigma_N \to 0} \sigma_1{}'(\sigma_N) \quad (95)$$

because $\sigma_1^{2\prime}(\sigma_N) = 2\sigma_1(\sigma_N)\sigma_1{}'(\sigma_N)$. Equations (94) and (96) imply that the rate of the initial SR effect increases if $\lim_{\sigma_N \to 0}\sigma_1^{2\prime}(\sigma_N)$ decreases. Further

$$\lim_{\sigma_N \to 0} \sigma_1^{2\prime}(\sigma_N) = \lim_{\sigma_N \to 0}\left[\int_{-\infty}^\infty \frac{K(Q-1)}{Q} 2[1 - 2F_{\tilde{N}}(n)] \times n f_{\tilde{N}}(n) f_V\left(-\frac{A}{2} - \sigma_N n\right)dn\right] \quad (96)$$

if we substitute $A/2 - v/\sigma_N = n$ in (92). Then

$$\lim_{\sigma_N \to 0} \sigma_1^{2\prime}(\sigma_N) = \frac{2K(Q-1)f_V\left(-\frac{A}{2}\right)}{Q} \times \int_{-\infty}^\infty [1 - 2F_{\tilde{N}}(n)] n f_{\tilde{N}}(n)dn. \quad (97)$$

The integral of (97) is negative because $[1 - 2F_{\tilde{N}}(n)]n$ is non-positive. Then $\lim_{\sigma_N \to 0}\sigma_1^{2\prime}(\sigma_N)$ decreases as the number Q

of quantizers increases because $Q_2 - 1/Q_2 > Q_1 - 1/Q_1$ if $Q_2 > Q_1$. So the initial SR effect with Q_2 quantizers is larger than that of the detector with Q_1 quantizers if $Q_2 > Q_1$.

Part (c) Fix the channel noise V and the number Q of quantizers and choose the symmetric scale-family quantizer noise N. Equation (94) and (96) imply that the rate of increase in the detection probability with respect to σ_N is maximal if $\lim_{\sigma_N \to 0} \sigma_1^{2\prime}(\sigma_N)$ achieves its minimal value. We want to find the symmetric standard (zero-mean and unit-variance) pdf $f_{\tilde{N}}$ of the zero-mean unit-variance scale-family quantizer noise $\tilde{N}$ that minimizes the expectation $\int_{-\infty}^{\infty}[1 - 2F_{\tilde{N}}(n)]nf_{\tilde{N}}(n)dn$. But this is equivalent to maximizing

$$\int_{-\infty}^{\infty}[2F_{\tilde{N}}(n) - 1]nf_{\tilde{N}}(n)dn \tag{98}$$

because $[1 - 2F_{\tilde{N}}(n)]n$ is nonpositive and $f_{\tilde{N}}$ is a probability density function.

Suppose first that the quantizer noise $\tilde{N}$ is a symmetric discrete random variable on the sample space $\Omega_{\tilde{N}} = \{n_\ell \in \mathbf{R} : \ell \in \mathcal{L} \subseteq \mathbb{Z}\}$ where $n_0 = 0$. Suppose also that $n_{-\ell} = -n_\ell$ and $n_\ell > 0$ for all $\ell \in \mathbb{N} = \{1, 2, 3, \ldots\}$. Let $P_{\tilde{N}}(n_\ell) = P(\tilde{N} = n_\ell) = p_\ell/2$ denote the probability density of this symmetric standard discrete quantizer noise $\tilde{N}$ where $P_{\tilde{N}}(n_{-\ell}) = p_\ell/2 = P_{\tilde{N}}(n_\ell)$ for all $\ell \in \mathbb{N}$ and $P_{\tilde{N}}(n_0) = p_0$. Then the finite variance of $\tilde{N}$ lets us replace (98) with the appropriate convergent series $\sum_{\ell=-\infty}^{\infty}[2F_{\tilde{N}}(n_\ell) - 1]n_\ell P_{\tilde{N}}(n_\ell)$. Write

$$\begin{aligned}
&\sum_{\ell=-\infty}^{\infty}[2F_{\tilde{N}}(n_\ell) - 1]n_\ell P_{\tilde{N}}(n) \\
&\quad = 2\sum_{\ell=-\infty}^{\infty} F_{\tilde{N}}(n_\ell)n_\ell P_{\tilde{N}}(n_\ell) - \sum_{\ell=-\infty}^{\infty} n_\ell P_{\tilde{N}}(n_\ell) && (99)\\
&\quad = 2\sum_{\ell=-\infty}^{\infty} F_{\tilde{N}}(n_\ell)n_\ell P_{\tilde{N}}(n_\ell) && (100)\\
&\quad = 2\sum_{\ell=1}^{\infty}[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]n_\ell P_{\tilde{N}}(n_\ell). && (101)
\end{aligned}$$

Equality (100) follows because $\tilde{N}$ is zero-mean while (101) follows because the density $P_{\tilde{N}}(n)$ is symmetric ($n_{-\ell} = -n_\ell$ and $P_{\tilde{N}}(-n_\ell) = P_{\tilde{N}}(n_\ell)$) and $n_0 = 0$. Then we need to find the density $P_{\tilde{N}}$ that maximizes (101). So write

$$\begin{aligned}
&2\sum_{\ell=1}^{\infty}[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]n_\ell P_{\tilde{N}}(n_\ell) \\
&= 2\sum_{\ell=1}^{\infty}\left|[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]\sqrt{P_{\tilde{N}}(n_\ell)}n_\ell\sqrt{P_{\tilde{N}}(n_\ell)}\right|. && (102)
\end{aligned}$$

Then the Cauchy-Schwarz inequality [84] gives

$$\begin{aligned}
&2\sum_{\ell=1}^{\infty}[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]n_\ell P_{\tilde{N}}(n_\ell) \\
&\le 2\left[\sum_{\ell=1}^{\infty}[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]^2 P_{\tilde{N}}(n_\ell)\right]^{1/2}\left[\sum_{\ell=1}^{\infty} n_\ell^2 P_{\tilde{N}}(n_\ell)\right]^{1/2} && (103)\\
&= 2\left[\sum_{\ell=1}^{\infty}[F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]^2 P_{\tilde{N}}(n_\ell)\right]^{1/2}\frac{1}{\sqrt{2}} && (104)
\end{aligned}$$

because $\tilde{N}$ is symmetric zero mean and unit variance noise and so $\sum_{\ell=1}^{\infty} n_\ell^2 P_{\tilde{N}}(n_\ell) = \frac{1}{2}$

$$= \sqrt{2}\left[\sum_{\ell=1}^{\infty}\frac{\left(p_\ell + \sum_{i=0}^{\ell-1} 2p_i\right)^2}{4}\frac{p_\ell}{2}\right]^{1/2} \tag{105}$$

because $P_{\tilde{N}}(n_\ell) = P_{\tilde{N}}(-n_\ell) = \frac{p_\ell}{2}$ and so

$$F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell) = \frac{\left(p_\ell + \sum_{i=0}^{\ell-1} 2p_i\right)}{2} \text{ for all } \ell \in \mathbb{N}$$

$$\le \sqrt{2}\left[\frac{(p_0 + p_1 + p_2 + \cdots)^3 + \frac{(p_0+p_1+p_2+\cdots)^3}{3}}{8}\right]^{1/2} \tag{106}$$

by expanding the quadratics in (105) and using the multinomial theorem

$$\begin{aligned}
&= \sqrt{2}\left[\frac{1 + \frac{1}{3}}{8}\right]^{1/2} \quad \text{because } \sum_{\ell=0}^{\infty} p_\ell = 1 && (107)\\
&= \frac{1}{\sqrt{3}}. && (108)
\end{aligned}$$

So (108) gives the upper bound on (98) for any discrete symmetric standard quantizer noise $\tilde{N}$.

Suppose now that the quantizer noise pdf f_N is continuous. Write

$$\begin{aligned}
&\int_{-\infty}^{\infty}[2F_{\tilde{N}}(n) - 1]nf_{\tilde{N}}(n)dn \\
&\quad = \int_{-\infty}^{\infty}\left|[2F_{\tilde{N}}(n) - 1]\sqrt{f_{\tilde{N}}(n)}n\sqrt{f_{\tilde{N}}(n)}\right|dn. && (109)
\end{aligned}$$

Again the Cauchy-Schwarz inequality [84] gives

$$\begin{aligned}
&\int_{-\infty}^{\infty}[2F_{\tilde{N}}(n) - 1]nf_{\tilde{N}}(n)dn \\
&\le \left[\int_{-\infty}^{\infty}[2F_{\tilde{N}}(n) - 1]^2 f_{\tilde{N}}(n)dn\right]^{1/2}\left[\int_{-\infty}^{\infty} n^2 f_{\tilde{N}}(n)dn\right]^{1/2} && (110)\\
&= \left[\int_{-\infty}^{\infty}[2F_{\tilde{N}}(n) - 1]^2 f_{\tilde{N}}(n)dn\right]^{1/2} && (111)
\end{aligned}$$

because $\tilde{N}$ has zero mean and unit variance

$$\begin{aligned}
&= \left[E\left[\left(2F_{\tilde{N}}(\tilde{N}) - 1\right)^2\right]\right]^{1/2} && (112)\\
&= \left[E\left[(2U - 1)^2\right]\right]^{1/2} && (113)\\
&= \frac{1}{\sqrt{3}}. && (114)
\end{aligned}$$

Equalities (113)–(114) follow because $U = F_{\tilde{N}}(\tilde{N})$ is a uniform random variable in $[0, 1]$ for any continuous quantizer noise $\tilde{N}$ [53]. Inequality (110) becomes an equality if and only

if $F_{\tilde{N}}$ obeys $[2F_{\tilde{N}}(n) - 1]^2 = \kappa n^2$ for some constant κ on the support of $f_{\tilde{N}}$ [84]. Then $F_{\tilde{N}}(n) = 1/2 + \sqrt{\kappa}n/2$ for all $n \in [-1/\sqrt{\kappa}, 1/\sqrt{\kappa}]$ for $\kappa = 1/3$ because $F_{\tilde{N}}$ is the CDF of standard quantizer noise. The same CDF implies that $\tilde{N}$ is uniformly distributed in $[-\sqrt{3}, \sqrt{3}]$. So symmetric uniform quantizer noise achieves the upper bound $1/\sqrt{3}$ of (114) and (108). So zero-mean uniform noise gives the maximal rate of the initial SR effect among all finite-variance continuous scale-family quantizer noise. □

REFERENCES

[1] 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, pp. 3172–3184, Jul. 2007.

[2] M. Guerriero, P. Willett, S. Marano, and V. Matta, "Speedier sequential tests via stochastic resonance," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. ICASSP* , 4, 2008, pp. 3901–3904.

[3] S. Kay, "Can detectability be improved by adding noise?," *IEEE Signal Process. Lett.*, vol. 7, pp. 8–10, Jan. 2000.

[4] B. Kosko, *Noise*. New York: Viking/Penguin, 2006.

[5] M. D. McDonnell and D. Abbott, "What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology," *PLoS Computat. Biol.*, vol. 5, no. 5, p. E1000348, May 2009.

[6] S. Mitaim and B. Kosko, "Adaptive stochastic resonance," *Proc. IEEE*, vol. 86, pp. 2152–2183, 1998.

[7] A. Patel and B. Kosko, "Stochastic resonance in noisy spiking retinal and sensory neuron models," *Neural Netw.*, vol. 18, no. 5–6, pp. 467–478, 2005.

[8] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained statistical signal detection," *IEEE Trans. Signal Process.*, vol. 57, pp. 1655–1669, 2009.

[9] D. Rousseau and F. Chapeau-Blondeau, "Constructive role of noise in signal detection from parallel array of quantizers," *Signal Process.*, vol. 85, no. 3, pp. 571–580, Mar. 2005.

[10] N. G. Stocks, "Information transmission in parallel threshold arrays: Suprathreshold stochastic resonance," *Phys. Rev. E*, vol. 63, no. 4, p. 041114, Mar. 2001.

[11] A. R. Bulsara and A. Zador, "Threshold detection of wideband signals: A noise-induced maximum in the mutual information," *Phys. Rev. E*, vol. 54, no. 3, pp. R2185–R2188, Sep. 1996.

[12] A. Patel and B. Kosko, "Error-probability noise benefits in threshold neural signal detection," *Neural Netw.*, vol. 22, no. 5–6, pp. 697–706, Jul./Aug. 2009.

[13] 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.

[14] A. A. Saha and G. V. Anand, "Design of detectors based on stochastic resonance," *Signal Process.*, vol. 83, no. 6, pp. 1193–1212, 2003.

[15] P.-O. Amblard, S. Zozor, M. D. McDonnell, and N. G. Stocks, "Pooling networks for a discrimination task: Noise-enhanced detection," *Proc. SPIE*, vol. 6602, no. 1, pp. 66020S-1–12, 2007.

[16] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," *Phys. Rev. E*, vol. 52, no. 5, pp. 4691–4698, Nov. 1995.

[17] P. Hänggi, "Stochastic resonance in biology," *Chem. Phys. Chem.*, vol. 3, pp. 285–290, 2002.

[18] 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.

[19] 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.

[20] D. Applebaum, "Extending stochastic resonance for neuron models to general Levy noise," *IEEE Trans. Neural Netw.*, vol. 20, pp. 1993–1995, Dec. 2009.

[21] W. C. Stacey and D. M. Durand, "Stochastic resonance improves signal detection in hippocampal CA1 neurons," *J. Neurophysiol.*, vol. 83, pp. 1394–1402, 2000.

[22] N. Stocks, D. Appligham, and R. Morse, "The application of suprathreshold stochastic resonance to cochlear implant coding," *Fluctuation and Noise Lett.*, vol. 2, pp. L169–L181, 2002.

[23] K. Wiesenfeld and F. Moss, "Stochastic resonance and the benefits of noise: From ice ages to crayfish and squid," *Nature (London)*, vol. 373, p. 33, 1995.

[24] A. Benerjee, P. Burlina, and R. Chellappa, "Adaptive target detection in foliage-penetrating SAR images using alpha-stable models," *IEEE Trans. Image Process.*, vol. 8, pp. 1823–1831, Dec. 1999.

[25] M. Bouvet and S. C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-37, no. 4, p. 621, 1989.

[26] A. Briassouli, P. Tsakalides, and A. Stouraitis, "Hidden messages in heavy-tails: DCT-domain watermark detection using alpha-stable models," *IEEE Trans. Multimedia*, vol. 7, pp. 700–715, 2005.

[27] S. Kay, J. Michels, H. Chen, and P. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Signal Process. Lett.* , vol. 13, no. 11, pp. 695–698, Nov. 2006.

[28] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Noise enhanced nonparametric detection," *IEEE Trans. Inf. Theory*, vol. 55, pp. 499–506, 2009.

[29] M. D. McDonnell, N. G. Stocks, and D. Abbott, "Optimal stimulus and noise distributions for information transmission via suprathreshold stochastic resonance," *Phys. Rev. E*, vol. 75, no. 6, pp. 061 105–+, 2007.

[30] S. Zozor and P.-O. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177–3181, Dec. 2003.

[31] F. Chapeau-Blondeau and D. Rousseau, "Noise-enhanced performance for an optimal Bayesian estimator," *IEEE Trans. Signal Process.*, vol. 52, pp. 1327–1334, May 2004.

[32] D. C. Munson, "A note on Lena," *IEEE Trans. Image Process.*, vol. 5, p. 3, 1996.

[33] J. R. Hernandez, A. Member, F. Perez-Gonzalez, M. Amado, and O. Pérez-gonzález, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Trans. Image Process.*, vol. 9, pp. 55–68, 2000.

[34] M. Thompson, D. Halverson, and G. Wise, "Robust detection in nominally Laplace noise," *IEEE Trans. Commun.*, vol. 42, pp. 1651–1660, 1994.

[35] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class a and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1129–1149, May 1999.

[36] S. M. Saberali, H. Amindavar, and J. Ritcey, "Blind detection in symmetric non-Gaussian noise with unknown pdf using maximum entropy method with moment generating function constraints," *Signal Process.*, vol. 90, no. 3, pp. 891–899, 2010.

[37] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Mar. 2000.

[38] 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 University Press, 2008.

[39] M. Grigoriu, *Applied Non-Gaussian Processes*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[40] C. L. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*. New York: Wiley-Intersci., 1995.

[41] , R. J. Adler, R. E. Feldman, and M. S. Taqqu, Eds., *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Boston, MA: Birkhauser, 1998.

[42] A. Achim, S. Member, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 772–783, 2001.

[43] P. G. Georgiou, S. Member, P. Tsakalides, and C. Kyriakakis, "Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise," *IEEE Trans. Multimedia*, vol. 1, pp. 291–301, 1999.

[44] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson, "Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level," *IEEE/ACM Trans. Netw.*, vol. 5, pp. 71–86, 1997.

[45] G. Tsihrintzis and C. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process," *IEEE Trans. Commun.*, vol. 43, pp. 904–914, Mar. 1995.

[46] S. Zozor, J.-M. Brossier, and P.-O. Amblard, "A parametric approach to suboptimal signal detection in α-stable noise," *IEEE Trans. Signal Process.*, vol. 54, pp. 4497–4509, Dec. 2006.

[47] S. Ambike, J. Ilow, and D. Hatzinakos, "Detection for binary transmission in a mixture of Gaussian noise and impulsive noise modeled as an alpha-stable process," *IEEE Signal Process. Lett.*, vol. 1, no. 3, pp. 55–57, Mar. 1994.
[48] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer, 2008.
[49] C. Brown and A. Zoubir, "A nonparametric approach to signal detection in impulsive interference," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2665–2669, Sep. 2000.
[50] C. Brown, "Score functions for locally suboptimum and locally suboptimum rank detection in alpha-stable interference," in *Proc. 11th IEEE Signal Process. Workshop on Statist. Signal Process. (SSP)*, 2001, pp. 58–61.
[51] E. Kuruoglu, W. Fitzgerald, and P. Rayner, "Near optimal detection of signals in impulsive noise modeled with a symmetric α-stable distribution," *IEEE Commun. Lett.*, vol. 2, no. 10, pp. 282–284, Oct. 1998.
[52] X. Li, L. Jin, and S. Wang, "A simplified non-Gaussian mixture model for signal lo detection in α-stable interference," in *Proc. 2008 Congress on Image and Signal Process.*, 2008, vol. 5, pp. 403–407.
[53] G. Casella and R. Berger, *Statistical Inference*. Pacific Grove, CA: Duxbury Resource Center, 2001.
[54] B. Hayes, "Randomness as a resource," *Amer. Scientist*, vol. 89, no. 4, pp. 300–304, 2001.
[55] M. Epstein, L. Hars, R. Krasinski, M. Rosner, and H. Zheng, "Design and implementation of a true random number generator based on digital circuit artifacts," *Lecture Notes in Comput. Sci.*, vol. 2779, pp. 152–165, 2003.
[56] N. Kimura and S. Latifi, "A survey on data compression in wireless sensor networks," in *Proc. Int. Conf. Inf. Technol.: Coding and Comput. (ITCC'05)—Vol. II*, 2005, pp. 8–13.
[57] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, "Low-complexity digital encoding strategies for wireless sensor networks," *Proc. ICASSP98*, pp. 3273–3276, 1998.
[58] J. Nolan, "Numerical calculation of stable densities and distribution functions," *Commun. Statist.*, vol. 13, pp. 759–774, 1997.
[59] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York: McGraw-Hill, 2008.
[60] S. Nadarajah, "A generalized normal distribution," *J. Appl. Statist.*, vol. 32, no. 7, pp. 685–694, Sep. 2005.
[61] Y. Bazi, L. Bruzzone, and F. Melgani, "Image thresholding based on the EM algorithm and the generalized Gaussian distribution," *Pattern Recogn.*, vol. 40, no. 2, pp. 619–634, 2007.
[62] M. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.
[63] S. Gazor and W. Zhang, "Speech probability distribution," *IEEE Signal Process. Lett.*, vol. 10, no. 7, pp. 204–207, Jul. 2003.
[64] R. Krupinski and J. Purczynski, "Approximated fast estimator for the shape parameter of generalized Gaussian distribution," *Signal Process.*, vol. 86, no. 2, pp. 205–211, 2006.
[65] Q. Cheng and T. Huang, "An additive approach to transform-domain information hiding and optimum detection structure," *IEEE Trans. Multimedia*, vol. 3, no. 3, pp. 273–284, Sep. 2001.
[66] A. Nikolaidis and I. Pitas, "Asymptotically optimal detection for additive watermarking in the DCT and DWT domains," *IEEE Trans. Image Process.*, vol. 12, no. 5, pp. 563–571, May 2003.
[67] J. Wang, G. Liu, Y. Dai, J. Sun, Z. Wang, and S. Lian, "Locally optimum detection for Barni's multiplicative watermarking in DWT domain," *Signal Process.*, vol. 88, no. 1, pp. 117–130, 2008.
[68] A. C. Bovik, *The Essential Guide to Image Processing*. New York: Academic , 2009.
[69] G. V. D. Wouwer, P. Scheunders, and D. V. Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, pp. 592–598, 1999.
[70] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley-Intersci., 1995.
[71] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 357–372, 1998.
[72] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2nd ed. Boston, MA: Morgan Kaufmann, 2007.
[73] M. Arnold, S. D. Wolthusen, and M. Schmucker, *Techniques and Applications of Digital Watermarking and Content Protection*. Dedham, MA: Artech House, 2003.
[74] B. Furht and D. Kirovski, *Multimedia Security Handbook*. Boca Raton, FL: CRC, 2005.
[75] G. Langelaar, I. Setyawan, and R. Lagendijk, "Watermarking digital image and video data. A state-of-the-art overview," *IEEE Signal Process. Mag.*, vol. 17, no. 5, pp. 20–46, Sep. 2000.
[76] A. J. Ahumada and H. A. Peterson, B. E. Rogowitz, Ed., "Luminance-model-based DCT quantization for color image compression," in *Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) Conf. Ser.*, Aug. 1992, vol. 1666, pp. 365–374.
[77] S. Sun and P. Qui, "Algorithm for digital watermarking based on parameter-induced stochastic resonance," *J. Commun.*, vol. 88, no. 8, pp. 48–55, 2005.
[78] S. Sun and B. Lei, "On an aperiodic stochastic resonance signal processor and its application in digital watermarking," *Signal Process.*, vol. 88, no. 8, pp. 2085–2094, 2008.
[79] F. Duan, D. Abbott, and F. M. C. Chapeau-Blondeau, "The application of saturating detectors to a DCT-domain watermarking scheme," *Fluctuation and Noise Lett.*, vol. 8, no. 1, pp. L65–L79, 2008.
[80] F. Duan and B. Xu, "Parameter-induced stochastic resonance and baseband binary PAM signals transmission over an awgn channel," *I. J. Bifurc. Chaos*, vol. 13, no. 2, pp. 411–425, 2003.
[81] G. Wu and Z. Qiu, "A novel watermarking scheme based on stochastic resonance," in *Proc. 8th Int. Conf. Signal Process.*, 2006, vol. 2.
[82] S. Baudry, J. Delaigle, B. Sankur, B. Macq, and H. Maitre, "Analysis of error correction strategies for typical communication channels in watermarking," *Signal Process.*, vol. 81, pp. 1239–1250, 2001.
[83] D. Jones, *The Theory of Generalized Functions*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1982.
[84] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. New York: Wiley-Interscience, 1999.
[85] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.

Ashok Patel (M'10) received the B.E. and M.E. degrees in electrical engineering from Gujarat University, Ahmedabad, India. He received the M.A. degree in applied mathematics, the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering, all from the University of Southern California (USC).

He is a Postdoctoral Researcher and a lecturer with the Department of Electrical Engineering and the Department of Industrial and Systems Engineering, USC.

Bart Kosko (M'85–SM'07–F'10) received degrees in philosophy, economics, mathematics, electrical engineering, and law.

He is a Professor of electrical engineering and law with 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*, the novel *Nanotime*, edited the volume *Neural Networks for Signal Processing*, and co-edited the volume *Intelligent Signal Processing*. His most recent book is *Noise*.