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 $\alpha$-stable ($\alpha$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 $\times$ 64 binary watermark message embedded in the midfrequency DCT-2 coefficients of the 512 $\times$ 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...
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 (\(\lim_{\sigma \to 0} dP_d/d\sigma\)) 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 \((Q \to \infty)\) 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 \(\mathcal{S}_{\alpha}\) 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 \(\alpha\)-STABLE NOISE

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

Fig. 1. Noise-enhanced digital watermark extraction using a noise-based algorithm: SR noise benefits in quantizer-array ML watermark decoding: (a) Binary 64 x 64 watermark yin-yang image. (b) Watermarked 512 x 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 \((Q \to \infty)\) correlation detector. This dashed U-shaped line gives the lower bound on the pixel-detection error for any finite \((Q < \infty)\) quantizer-array detector. (d) Retrieved yin-yang image using the ML linear correlation detector. (e) 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 \(\sigma\) 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).
detect a known deterministic signal \( s_k \) with amplitude \( A \) in additive white symmetric \( \alpha \)-stable (S\&S) channel noise \( V_k \) given \( K \) random samples \( X_1, \ldots, X_K \):

\[
\begin{align*}
H_0 & : \quad X_k = V_k \\
H_1 & : \quad X_k = As_k + V_k
\end{align*}
\]

(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 \( \tau \). The \( V_k \) are independent and identically distributed (i.i.d.) zero-location S\&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 \( \varphi \) of the S\&S noise random variable \( V_k \) has the exponential form [39], [40]

\[
\varphi(\omega) = \exp(i\delta \omega - \gamma |\omega|^\alpha)
\]

(2)

where real \( \delta \) 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 \( \sigma \) is the scale parameter. The bell curve’s tails get thicker as \( \alpha \) falls from 2 to near zero. So energetic impulses become more frequent for smaller values of \( \alpha \).

S\&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\&S pdfs are the thick-tailed Cauchy with \( \alpha = 1 \) and the thin-tailed Gaussian with \( \alpha = 2 \). The Gaussian pdf alone among S\&S pdfs has a finite variance and finite higher-order moments. The \( m \)th lower-order moments of an \( \alpha \)-stable pdf with \( \alpha < 2 \) exist if and only if \( m < \alpha \). The location parameter \( \delta \) serves as a proxy for the mean if \( 1 < \alpha < 2 \) and as a proxy for the median if \( 0 < \alpha < 1 \). The uniformly most powerful detector for the hypotheses in (1) is a NP log-likelihood ratio test [45], [46]

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

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

\[
\Lambda_{Lin}(X) = \sum_{k=1}^{K} s_k X_k \overset{H_1}{\underset{H_0}{\gtrless}} \lambda_\tau
\]

(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

\[
f_{V_k}(v_k) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + v_k^2}
\]

(5)
for real \( v_k \) and positive dispersion \( \gamma \). The NP detector is non-linear for such Cauchy channel noise and has the form

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

(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\&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\&S noise is mildly impulsive (when \( \alpha \geq 1.6 \)).

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

\[
\Lambda_{LO}(X) = \sum_{k=1}^{K} s_k f_{\alpha}(X_k) \overset{H_1}{\underset{H_0}{\gtrless}} \lambda_\tau
\]

(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 \( f_\alpha \) does not have a closed form because \( g_{LO} \) requires both \( f_\alpha \) and \( f_{\alpha'} \). 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

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

(8)
and hole-puncher functions

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

(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 \( g_{NQ} \). 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:

\[ \Lambda_{\text{NQ}}(X) = \sum_{k=1}^{K} s_k g_{\text{NQ}}(X_k) H_i \overset{H_0}{\approx} \lambda \]  
\[ \text{(10)} \]

where  
\[ g_{\text{NQ}}(X_k) = \frac{1}{Q} \sum_{q=1}^{Q} \text{sign}(X_k + N_q - \theta), \]  
\[ \text{(11)} \]

Here  \( \lambda \) is the detection threshold,  \( \theta \) is the quantization threshold, and  \( \text{sign}(X_k + N_q - \theta) = \pm 1 \) 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  \( f_N(n; \sigma_N) = 1/\sigma_N f_{N/\sigma_N}(n/\sigma_N) \). Here  \( \sigma_N \) is the noise standard deviation and  \( f_{N/\sigma_N} \) is the standard pdf for the whole family [53]. Then the noise cumulative distribution function (CDF) is  \( F_N(n; \sigma_N) = F_{N/\sigma_N}(n/\sigma_N) \) where  \( F_{N/\sigma_N} \) 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  \( \Lambda_{\text{NQ}} \) under the hypothesis  \( H_i \) ( \( i = 0 \) or  \( i = 1 \) when  \( \sigma_N \) is the quantizer noise intensity:  \( \mu_i(\sigma_N) = E(\Lambda_{\text{NQ}}|H_i) \) and  \( \sigma_i^2(\sigma_N) = \text{Var}(\Lambda_{\text{NQ}}|H_i) \). Then  \( \mu_0(\sigma_N) = -\mu_1(\sigma_N) \) and  \( \sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N) \) for all  \( \sigma_N \) because both the additive channel noise  \( V \) and the quantizer noise  \( N \) are symmetric. The mean  \( \mu_i \) and variance  \( \sigma_i^2 \) of the test statistic  \( \Lambda_{\text{NQ}} \) depend on both  \( V \) and  \( N \). So  \( \mu_i \) and  \( \sigma_i^2 \) depend on the noise intensities  \( \sigma_V \) and  \( \sigma_N \). We write these two terms as  \( \mu_i(\sigma_N) \) and  \( \sigma_i^2(\sigma_N) \) because we control only the quantizer noise intensity  \( \sigma_N \) and not the channel noise intensity  \( \sigma_V \). 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  \( \Lambda_{\text{NQ}} \) in (10) gives rise to a key simplification. The pdf of  \( \Lambda_{\text{NQ}} \) 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  \( d/d\sigma_N(h_1\mu_1(\sigma_N)/\sigma_1(\sigma_N)) > 0 \) since  \( \mu_1 > 0 \) in (72).

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

\[ \sigma_1(\sigma_N) \mu_0'(\sigma_N) > \mu_1(\sigma_N) \sigma_1'(\sigma_N) \]  
\[ \text{(12)} \]

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  \( \Lambda_{\text{NQ}} \).

Proof: We first derive an approximate linear form for the detection threshold  \( \lambda_r: \lambda_r \approx \lambda_0 \sigma_0(\sigma_N) + \mu_0(\sigma_N) \). The NP detection rule based on  \( \Lambda_{\text{NQ}} \) rejects  \( H_0 \) if  \( \Lambda_{\text{NQ}} > \lambda_r \) because we choose the detection threshold  \( \lambda_r \) such that  \( P(\Lambda_{\text{NQ}} > \lambda_r | H_0) = \tau \). We also need to define the constant  \( z_r \) so that  \( 1 - \Phi(z_r) = \tau \) where  \( \Phi(z) = \int_{-\infty}^{\infty} (1/\sqrt{2\pi}) e^{-z^2/2}dz \). Then standardizing  \( \Lambda_{\text{NQ}} \) under the assumption that the null hypothesis  \( H_0 \) is true gives

\[ P(\Lambda_{\text{NQ}} > \lambda_r | H_0) = P \left( \frac{\Lambda_{\text{NQ}} - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} > \frac{\lambda_r - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} \right) \]  
\[ \approx P \left( Z > \frac{\lambda_r - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} \right) \]  
\[ \text{(13)} \]

for  \( Z \sim N(0,1) \) by the central limit theorem

\[ 1 - \Phi \left( \frac{\lambda_r - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} \right) \]  
\[ \text{(15)} \]

so  \( 1 - \Phi(z_r) \approx 1 - \Phi \left( \frac{\lambda_r - \mu_0(\sigma_N)}{\sigma_0(\sigma_N)} \right) \) and thus  \( z_r \approx \lambda_r - \mu_0(\sigma_N)/\sigma_0(\sigma_N) \). So the detection threshold  \( \lambda_r \) has the approximate linear form  \( \lambda_r \approx z_r \sigma_0(\sigma_N) + \mu_0(\sigma_N) \). Standardizing  \( \Lambda_{\text{NQ}} \) under the assumption that the alternative hypothesis  \( H_1 \) is true likewise gives the detection probability  \( P_D \) as

\[ P_D(\sigma_N) = P(\Lambda_{\text{NQ}} > \lambda_r | H_1) \]  
\[ \text{(16)} \]

\[ = P \left( \frac{\Lambda_{\text{NQ}} - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} > \frac{\lambda_r - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ \approx P \left( Z > \frac{\lambda_r - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ \text{(17)} \]

for  \( Z \sim N(0,1) \) by the central limit theorem

\[ 1 - P \left( Z \leq \frac{\lambda_r - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ = 1 - \Phi \left( \frac{\lambda_r - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ \text{(19)} \]

because  \( \lambda_r \approx z_r \sigma_0(\sigma_N) + \mu_0(\sigma_N) \)

\[ = 1 - \Phi \left( \frac{z_r \sigma_0(\sigma_N) + \mu_0(\sigma_N) - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ \text{(21)} \]

\[ = 1 - \Phi \left( \frac{z_r \sigma_0(\sigma_N) + \mu_0(\sigma_N) - \mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \]  
\[ \text{(22)} \]

because  \( \mu_0(\sigma_N) = -\mu_1(\sigma_N) \) and  \( \sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N) \)

\[ = 1 - \Phi \left( \frac{z_r}{\sigma_1(\sigma_N)} \right) \]  
\[ \text{(23)} \]
Then the normal pdf $\phi(\tilde{z}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{z}^2}{2}}$ and the chain rule of differential calculus give

$$\frac{dP_D}{d\sigma_N} \approx 2\phi \left( z_r - \frac{2\mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \sigma_1(\sigma_N) \frac{\mu_1'(\sigma_N) - \mu_1(\sigma_N) \sigma_1'(\sigma_N)}{\sigma_1^2(\sigma_N)}$$

(24)

because $z_r$ 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 $\phi$ is a pdf and thus $\phi > 0$.

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 $\alpha = 1.85$, $\gamma = 1.71.85 = 2.67$, 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 $\Delta$ 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 $\sigma_N$ of the additive uniform quantizer noise. We used $10^6$ simulation trials to estimate $\mu_1(\sigma_N)$ and $\sigma_1(\sigma_N)$ and then used the difference quotients $\frac{\mu_1(\sigma_{N_j}) - \mu_1(\sigma_{N_{j-1}})}{\sigma_{N_j} - \sigma_{N_{j-1}}}$ and $\frac{\sigma_1(\sigma_{N_j}) - \sigma_1(\sigma_{N_{j-1}})}{\sigma_{N_j} - \sigma_{N_{j-1}}}$ 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 $\Lambda_{\text{NOQ}}$’s pdf. Circle marks show the detection probabilities computed from the $10^6$ 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 $\sigma_N \to 0$. 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 $\Lambda_{\text{NOQ}}$. (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).
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. 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.

So the strong law of large numbers [53] implies that the sample mean \( g_{NS}(X_k) \) in (11) converges with probability one to its population mean in (26):
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 $\theta = A/2$:

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

(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 $f_{\alpha S}$ for $\alpha = 1.85$. Fig. 4 shows that the limiting $(Q \to \infty)$ 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 $g_{LCO}(X_k) \approx (\alpha + 1)/X_k$. 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]

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

(32)

or such as [46]

$$g(X_k) = \frac{aA_{sk}X_k}{1 + bA_{sk}X_k^2}.$$ 

(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 $\alpha$-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 $H_1$:

$$H_0 : X_k = As_k + V_k$$
$$H_1 : X_k = As_k + V_k.$$ 

(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]:

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

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

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

$$\Lambda_{GGM}(X) = \sum_{k=1}^{K} \frac{1}{2} \left[ |X_k + A_{s_k}|^r - |X_k - A_{s_k}|^r \right] H_1 \geq H_0$$  (36)

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

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

$$f(\sigma, v) = \frac{1}{\sigma} f_{bsg} \left( \frac{v}{\sigma} \right) = \sqrt{\frac{\pi}{2}} \frac{1}{\Gamma \left( \frac{3}{2} \right)} \sigma^{-1/2} e^{-B|\log(\sigma)|^r},$$  (38)

where $f_{bsg}$ is the standard pdf of the family, $\sigma$ is the standard deviation, and $I$ is the gamma function. This family of pdfs includes all normal ($r = 2$) and Laplace ($r = 1$) pdfs. It includes in the limit ($r \to \infty$) 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 $\sigma$ but we still need joint estimates of $\sigma$ 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 Briassous and Tsakalis [26] have proposed using instead the Cauchy pdf to model the DCT coefficients.

The ML Cauchy detector has the form [40]

$$A_{CM}(X) = \sum_{k=1}^{K} \log \left( \frac{\gamma^2 + (X_k + A_{s_k})^2}{\gamma^2 + (X_k - A_{s_k})^2} \right) H_1 \geq H_0$$  (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 $\Lambda_{NQ}$ and its limit $H_{NQ\infty}$. 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 $\Lambda_{NQ}$ and $\Lambda_{NQ\infty}$.

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 $\Lambda_{NQ}(X)$ and its limit $\Lambda_{NQ\infty}(X)$ with $\theta = 0$ for the ML detection of the signal $s_k$ in (34):

$$A_{NQ}(X) = \sum_{k=1}^{K} s_k \Lambda_{NQ}(X_k) H_1 \geq H_0$$  (40)

where $\Lambda_{NQ}(X_k) = 1$ if $k \in Q$ and $\Lambda_{NQ}(X_k) = 0$ if $k \not\in Q$.

$$A_{NQ\infty}(X) = \sum_{k=1}^{K} s_k \Lambda_{NQ\infty}(X_k) H_1 \geq H_0$$  (42)

where $\Lambda_{NQ\infty}(X_k) = 2F_N \left( \frac{X_k}{\sigma_N} \right) - 1.$  (43)

We use $\theta = 0$ because the pdfs of the random samples $X_k$ are symmetric about $-A_{s_k}$ and $A_{s_k}$ 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 $\sigma_N$. But they do not require that we know the signal amplitude $A$ or the channel noise pdf parameters $\gamma$ or $\beta$. 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 $\Lambda_{NQ}$ and $\Lambda_{NQ\infty}$ in (40) and (42) obey $\mu(\sigma_N) = -\mu(\sigma_N)$ and $\sigma^2(\sigma_N) = \sigma^2(\sigma_N)$ for all $\sigma_N$. These equalities hold because (40)–(43) imply that $\Lambda_{NQ}(H_0) = -\Lambda_{NQ}(H_1)$ and $\Lambda_{NQ\infty}(H_0) = -\Lambda_{NQ\infty}(H_1)$. The pdf of $\Lambda_{NQ}$ 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 $s_k \Lambda_{NQ}(X_k)$ and $s_k \Lambda_{NQ\infty}(X_k)$ are independent even though they are not identically distributed. This holds for uniformly bounded random variables that satisfy the Lindeberg condition [70]: $\lim_{K \to \infty} \text{Var}(\Lambda_{NQ}) = \infty$ and $\lim_{K \to \infty} \text{Var}(\Lambda_{NQ\infty}) = \infty$. The variables $s_k \Lambda_{NQ}(X_k)$ and $s_k \Lambda_{NQ\infty}(X_k)$ 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 $\mathcal{S}_\mathcal{Q}$. 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 ($Q \to \infty$) 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 $\Lambda_{NQ}$ in (40) has sufficiently large sample size $K$ so that it is approximately conditionally normal: $\Lambda_{NQ}(H_0) \approx N(\mu(\sigma_N), \sigma^2(\sigma_N))$ and $\Lambda_{NQ}(H_1) \approx N(\mu(\sigma_N), \sigma^2(\sigma_N))$ where $\mu(\sigma_N) = -\mu(\sigma_N)$ and $\sigma^2(\sigma_N) = \sigma^2(\sigma_N)$. Then the inequality

$$\sigma(\sigma_N) \mu(\sigma_N) > \mu(\sigma_N) \sigma(\sigma_N)$$  (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 \( \Lambda_{NQ} \) 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

\[
P_e(\sigma_N) = p_0P(\Lambda_{NQ} > 0| H_0) + p_0P(\Lambda_{NQ} < 0| H_1) \tag{45}
\]

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

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

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

\[
= \frac{1}{2} \Phi \left( \frac{-\mu_0(\sigma_N)}{\sigma_0(\sigma_N)} \right) + \frac{1}{2} \Phi \left( \frac{-\mu_1(\sigma_N)}{\sigma_1(\sigma_N)} \right) \tag{49}
\]

\[
\text{since } \mu_0(\sigma_N) = -\mu_1(\sigma_N) \text{ and } \sigma_0^2(\sigma_N) = \sigma_1^2(\sigma_N) = \sigma_N^2(\sigma_N) \quad \text{(50)}
\]

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

\[
dP_e(\sigma_N)/d\sigma_N \approx -\phi \left( \frac{-\mu_0(\sigma_N)}{\sigma_1(\sigma_N)} \right) \times \frac{\sigma_1(\sigma_N)}{\sigma_1(\sigma_N)} \times \mu_0'(\sigma_N) - \mu_1(\sigma_N) \sigma_1'(\sigma_N) \sigma_0^2(\sigma_N) \tag{51}
\]

So \( \sigma_1(\sigma_N)\mu_0'(\sigma_N) > \mu_1(\sigma_N)\sigma_1'(\sigma_N) \) is necessary and sufficient for the SR effect \( (dP_e/d\sigma_N < 0) \) because \( \phi \) is a pdf and thus \( \phi(0) \geq 0 \).

The SR condition (44) in Corollary 1 allows us to find a near-optimal quantizer-noise standard deviation \( \sigma_{NQ}^{\text{opt}} \) for the ML detector if we have enough samples of the received signal vector \( X \) under both hypotheses \( H_0 \) and \( H_1 \) in (34). Then the pdf of \( \Lambda_{NQ} \) is a mixture of two equally likely Gaussian pdfs: \( 1/2N(-\mu_0(\sigma_N), \sigma_0^2(\sigma_N)) + 1/2N(\mu_1(\sigma_N), \sigma_1^2(\sigma_N)) \). So \( \Lambda_{NQ}/\sigma_1^2(\sigma_N) \) is a non-central chi-square random variable [53] with noncentrality parameter \( \mu_0^2(\sigma_N) \) and 1 degree of freedom. Then

\[
\text{Var}(\Lambda_{NQ}) = \sigma_1^2(\sigma_N) + \mu_0^2(\sigma_N) \tag{52}
\]

\[
\text{Var}(\Lambda_{NQ}^2) = 2\sigma_1^2(\sigma_N) + 4\sigma_1^2(\mu_0^2(\sigma_N)) \tag{53}
\]

Putting \( \mu_0^2(\sigma_N) \) from (52) in (53) gives a quadratic equation in \( \sigma_1^2(\sigma_N) \) with solution

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

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 \( \Lambda_{NQ} \) and \( \Lambda_{NQ}^2 \) 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 \( \hat{\mu}_0^2(\sigma_N) \) of \( \mu_0^2(\sigma_N) \). The same received vector \( X \) allows us to compute \( \hat{\sigma}_1^2(\sigma_N) \) and \( \hat{\mu}_0^2(\sigma_N) \) for all values of \( \sigma_N \) by (40)–(43). Then a zero-crossing of \( \text{li}_1(\mu_0(\sigma_N)/\sigma_1(\sigma_N)) - \text{li}_1(\mu_1(\sigma_N)/\sigma_1(\sigma_N)) \) estimates the optimal quantizer-noise standard deviation \( \sigma_{NQ}^{\text{opt}} \) for a small step-size of \( \Delta \sigma_N = \sigma_N - \sigma_{N-1} \) because the SR condition (44) is equivalent to \( d/d\sigma_N(\text{li}_1(\mu_0(\sigma_N)/\sigma_1(\sigma_N))) > 0 \) 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 \( P_B \) 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 \( s_0 \) in generalized-Gaussian channel noise and in \( s_{0,\delta} \) 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 \( s_{0,\delta} \) 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 \( (Q \rightarrow \infty) \) ML correlation detector (42) with limiting-array Gaussian-quantizer-noise nonlinearity \( g_{NQ}(x_k) = 2B(x_k/\sigma_N) - 1 \) where we have replaced \( F_N \) in (43) with the standard normal CDF \( \Phi \). The thick horizontal dash-dot line shows the error probability of the optimal generalized-Gaussian ML detector (36). The limiting-array \( (Q \rightarrow \infty) \) 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...
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 \( r = 1, 2, \sigma = 2 \). (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 \( N \to \infty \) 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.

Fig. 6. Initial SR effects in the quantizer-array ML detection of a known bipolar sequence of unknown amplitude \( A \) in infinite-variance \( S \alpha S \) channel noise. (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 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 $W[k]$ of an $M_1 \times M_2$ binary image $B(n)$ in the DCT-2 coefficients $V[k]$ of an $I_1 \times I_2$ host image $H(n)$. Here $k = (k_1,k_2)$ and $n = (n_1,n_2)$ are the respective 2-D indices in the transform domain and in the spatial domain. We apply an $8 \times 8$ 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 $b = (b_1,\ldots,b_M)$ with $M(= M_1 \times M_2)$ bipolar ($\pm 1$) 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 $I_i$ as the set of 2-D indices of the DCT-coefficient set $D_i$; $I_i = \{k : V[k] \in D_i\}$. Then $|D_i| = |I_i| = K$. The secret key also gives the initial seed to a pseudorandom sequence generator that produces a bipolar spreading sequence $\hat{s}_i[k]$ of length $K$ ($k \in I_i$) for each message bit $b_i$ [33], [59]. Each $\hat{s}_i$ is a pseudorandom sequence of $K$ i.i.d. random variables in $\{1,-1\}$. The cross-correlation among these $M$ spreading sequences $\{\hat{s}_i\}_{i=1}^M$ equals zero while their autocorrelation equals the Kronecker delta function.

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

Retrieving the hidden message $b$ requires that we know the pseudorandom assignment of DCT-coefficients $\{V_i[k] : k \in D_i\}$ to each message bit $b_i$ in $b$ and that we also know the pseudorandom sequence $\hat{s}_i$ 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:

$$H_0 (b_i = -1) : X_i[k] = -a[k] \hat{s}_i[k] + V_i[k]$$
$$H_1 (b_i = +1) : X_i[k] = a[k] \hat{s}_i[k] + V_i[k]$$

for all $k \in I_i$ and for $i = 1, \ldots, M$. Here $s_i = \hat{s}_i[k]^2$ is the known signal sequence and $A$ is the known or unknown scaling factor of the perceptual mask. Define $X_i = \{X_i[k] : k \in I_i\}$. The optimal decision rule to decode the message bit $b_i$ is the ML rule

$$L_W(X_i) = \sum_{k \in I_i} \log \left( \frac{f(X_i[k]|H_1)}{f(X_i[k]|H_0)} \right) \overset{H_1}{\gtrless} 0$$

because we assume that the DCT-2 coefficients $V_i[k]$ 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]

$$L_{Lin}(X_i) = \sum_{k \in I_i} s_i[k] X_i[k] \overset{H_1}{\gtrless} 0$$

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

$$L_{GG,ML}(X_i) = \sum_{k \in I_i} \left[ |X_i[k]| + a[k] \right] - |X_i[k] - a[k]|^r$$

$$\overset{H_1}{\gtrless} 0$$

or the Cauchy detector

$$L_{Cauchy,ML}(X_i) = \sum_{k \in I_i} \log \left( \frac{\gamma^2 + (X_i[k] + a[k])^2}{\gamma^2 + (X_i[k] - a[k])^2} \right)$$

$$\overset{H_1}{\gtrless} 0$$

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 $\gamma$. Both the suboptimal quantizer-array detector (40)–(41)

$$L_{NQ}(X_i) = \sum_{k \in I_i} s_i[k] g_{NQ}(X_i[k]) \overset{H_1}{\gtrless} 0$$

where $g_{NQ}(X_i[k]) = \frac{1}{Q} \sum_{q=1}^{Q} sign(X_i[k] + N_q)$

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

$$L_{N\infty}(X_i) = \sum_{k \in I_i} s_i[k] g_{N\infty}(X_i[k]) \overset{H_1}{\gtrless} 0$$

where $g_{N\infty}(X_i[k]) = 2F_X \left( \frac{|X_i[k]|}{\sigma_N} \right) - 1$
require that we know the quantizer noise $N_q$ and its intensity $\sigma_J$. They do not require that we know the scaling factor $A$ of the perceptual mask or the channel noise pdf parameters $r$ or $\gamma$. 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 $8 \times 8$ block-wise DCT-2 transform $V[k]$ of the $L_1 \times L_2$ host image $H[n]$.
2. Let $D$ be the set of mid-frequency DCT-2 coefficients of all $8 \times 8$ DCT blocks of $V[k]$.
3. Convert the $M_1 \times M_2$ binary (black-and-white) watermark image $B(n)$ into an $M_1 \times M_2 = M$-bit bipolar watermark message $b = (b_1, \ldots, b_M)$ 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 $T_i$ be the set of two-dimensional indexes of the DCT-coefficient set $D_i$.
6. Use the secret key to generate $M$ pseudorandom bipolar spreading sequences $s_i[k]$ of length $K$ where $s_i[k] = \pm 1$ for all $k \in T_i$ and $i = 1, \ldots, M$.
7. For each message bit $b_i$: compute the watermark signals $W_i[k] = b_i s_i[k] s_i[k]$ for all $k \in T_i$ where $a[k]$ is the perceptual mask in [26] with scale factor $A$.
8. For each message bit $b_i$: compute the DCT-2 coefficients $X_i[k] = V[k] + b_i a[k] s_i[k]$ for all $k \in T_i$.
9. Compute the inverse block-wise DCT-2 transform using the watermarked DCT-coefficients $X[k]$ to get the watermarked image $H_W[n]$.

**Watermark Decoding**

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

**Without Noise Injection**

- Use the Cauchy detector (60) (estimate the dispersion $\gamma$) or
- Use the GG detector (60) (estimate the shape parameter $\gamma$)

Else

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

**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 quantizer-noise intensity $\sigma$.

Each DCT-coefficient set $D_i$ 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 $s_i$. The watermark had a constant amplitude $(a[k] = A)$ 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 ($Q \leftrightarrow \infty$) correlation detector (62)–(63) where $\mathcal{M}_\infty$ is the soft-limiter nonlinearity $\mathcal{G}_SL$ 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 $\sigma$ 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

<table>
<thead>
<tr>
<th>Images</th>
<th>Minimal Average Pixel Errors in Watermark Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>limiting array Detector</td>
</tr>
<tr>
<td>Elaine</td>
<td>59 (131)</td>
</tr>
<tr>
<td>Goldhill</td>
<td>112 (350)</td>
</tr>
<tr>
<td>Pirate</td>
<td>105 (357)</td>
</tr>
<tr>
<td>Bird</td>
<td>178 (317)</td>
</tr>
<tr>
<td>Peppers</td>
<td>26 (120)</td>
</tr>
<tr>
<td>Lena</td>
<td>7 (79)</td>
</tr>
<tr>
<td>Tiffany</td>
<td>0.3 (27)</td>
</tr>
</tbody>
</table>

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 → ∞) detector for the symmetric uniform quantizer noise, and the generalized-Gaussian detector for various values of their respective pdf parameters γ, σ_N, and τ 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 psychovisual 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.

Psycovisual watermarking greatly reduced decoding errors compared with constant-amplitude watermark. Consider the Lena image: Table I and Fig. 1(c) and (d) shows 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 communications 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 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 \( 2\pi(\sigma_N) \) and use the chain rule to get

\[
\mu_1(\sigma_N)\sigma_\xi^2(\sigma_N) < 2\sigma_\xi^2(\sigma_N)\mu_1'(\sigma_N)
\]

(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 \( \sigma_N \) in some positive interval if \( Q = 1 \). We first derive the equations for \( \mu_1(\sigma_N), \mu_1'(\sigma_N), \sigma_\xi^2(\sigma_N), \) and \( \sigma_\xi^2(\sigma_N) \). 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

\[
\mu_1(\sigma_N) = E(\Lambda_{NQ}|H_1) = \sum_{k=1}^{K} E(Y_k|H_1),
\]

(66)

The \( Y_k|H_1 \) are i.i.d. random variables with population mean

\[
E(Y_k|H_1) = \int_{-\infty}^{\infty} E(Y_k|X_k = x; H_1)f_{X_k}(x)dx
\]

(67)

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

(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 \( E(Y_k|V_k = v; H_1) = E(Y_k|V_k = v; H_1) \) for brevity.

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

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

(69)

\[
= 1 - 2F_N\left(\frac{A}{2} - v\right)
\]

(70)

where \( F_N \) is the standard CDF for the scale family of the quantizer noise. So (66), (68), and (70) imply that

\[
\mu_1(\sigma_N) = \int_{-\infty}^{\infty} K\left[1 - 2F_N\left(\frac{A}{2} - v\right)\right] f_V(v - A)dv.
\]

(71)

\[
> 0
\]

(72)

because \( 1 - 2F_N(A/2 - v/\sigma_N) \) 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

\[
\mu_1'(\sigma_N) = \frac{\partial}{\partial \sigma_N} \int_{-\infty}^{\infty} K\left[1 - 2F_N\left(\frac{A}{2} - v\right)\right] f_V(v - A)dv
\]

(73)

\[
= -\int_{-\infty}^{\infty} nK f_V(-\frac{A}{2} - n\sigma_N) f_N(n)dn
\]

(74)

because of the distributional derivatives [83] of \( \mu_1(\sigma_N) \) and \([1 - 2F_N(A/2 - v/\sigma_N)]\) with respect to \( \sigma_N \). 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

\[
\mu_1'(\sigma_N) = \int_{-\infty}^{\infty} nK f_V\left(-\frac{A}{2} - n\sigma_N\right) f_N(n)dn
\]

(75)

\[
= -\int_{-\infty}^{\infty} nK f_V\left(-\frac{A}{2} + n\sigma_N\right) f_N(n)dn.
\]

(76)
because $f_N$ is a symmetric pdf. The mean-value theorem [84] implies that for any \( \epsilon > 0 \)
\[
\left| f_{V Y} \left( -\frac{A}{2} + n\sigma_N \right) - f_{V Y} \left( -\frac{A}{2} - n\sigma_N \right) \right| \leq 2n \sup_{\left| (A/2) - u \right| \leq \epsilon} f'_{V Y}(u) \tag{77}
\]
for all \( |n\sigma_N| \leq \epsilon \). The supremum in (77) is finite for some small \( \epsilon \) because the pdf derivative \( f'_{V Y} \) 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:
\[
\lim_{\sigma_N \to 0} \frac{\mu'_1(\sigma_N)}{\sigma_N} = \int_0^\infty 2nK \lim_{\sigma_N \to 0} \frac{f_{V Y} \left( -\frac{A}{2} + n\sigma_N \right) - f_{V Y} \left( -\frac{A}{2} - n\sigma_N \right)}{\sigma_N} f_N(n)dn \tag{78}
\]
by dominated convergence
\[
= -K \int_0^\infty 2n^2 f'_{V Y} \left( -\frac{A}{2} \right) f_N(n)dn \tag{79}
\]
by L’Hospital’s rule [85]
\[
= -2K f'_{V Y} \left( -\frac{A}{2} \right) \int_0^\infty 2n^2 f_N(n)dn \tag{80}
\]
\[
= -2K f'_{V Y} \left( -\frac{A}{2} \right) \tag{81}
\]
because \( f_N \) is the symmetric pdf of the zero-mean and unit-variance quantizer noise \( \bar{N} \). The unimodality of the symmetric channel noise \( V \) implies that (82) is negative. Then \( \mu'_1(\sigma_N) \) is negative for all noise intensities \( \sigma_N \) in some interval \((0, h)\). Then (82) also implies that
\[
\lim_{\sigma_N \to 0} \frac{\mu'_1(\sigma_N)}{\sigma_N} = 0 \tag{82}
\]
because \( \sigma_N \to 0 \) in \( \mu'_1(\sigma_N)/\sigma_N \).

We now derive expressions for the population variance \( \sigma^2_0(\sigma_N) \) of \( \Lambda_{\text{NQ}}|H_1 \) and its distributional derivative \( \sigma^2_0'(\sigma_N) \). The \( Y_k|H_1 \) are i.i.d. random variables. So
\[
\sigma^2_0(\sigma_N) = \text{Var}(\Lambda_{\text{NQ}}|H_1) = KV_{\text{ar}}(Y_k|H_1) = K \left[ E(Y^2_k|H_1) - E^2(Y_k|H_1) \right] \tag{83}
\]
where
\[
E(Y_k|H_1) = \frac{\mu_1(\sigma_N)}{K} \tag{84}
\]
and
\[
E(Y^2_k|H_1) = \int_{-\infty}^{\infty} E(Y^2_{k q}|v, H_1) f_Y(v - A)dv. \tag{85}
\]
Expand the integrand term \( E(Y^2_{k q}|v, H_1) \) as follows:
\[
E(Y^2_{k q}|v, H_1) = \frac{1}{Q} E(Y_{k q}^2|v, H_1) + \frac{Q - 1}{Q} E^2(Y_{k q}|v, H_1) \tag{86}
\]
due to the definition of \( Y_{k q} \)
\[
= \frac{1}{Q} E + \frac{Q - 1}{Q} E^2(\text{I}_{k q}|v, H_1) \tag{87}
\]
because \( E(Y^2_{k q}|v, H_1) = 1 \) by the definition of \( Y_{k q} \)
\[
= \frac{1}{Q} + \frac{Q - 1}{Q} E^2(Y_{k q}|v, H_1) \tag{88}
\]
because of (70). Put (90) in (87) and then put (86) and (87) in (85). This gives
\[
\sigma^2_0(\sigma_N) = \int_{-\infty}^{\infty} f_Y(v - A)dv - \frac{\mu_2(\sigma_N)}{K} \tag{89}
\]
\[
+ \frac{1}{Q} \int_{-\infty}^{\infty} \left[ 1 - 2F_N \left( \frac{A - v}{\sigma_N} \right) \right]^2 f_Y(v - A)dv. \tag{90}
\]
Then the distributional derivative of \( \sigma^2_0(\sigma_N) \) with respect to \( \sigma_N \) is
\[
\sigma^2_0'(\sigma_N) = \int_{-\infty}^{\infty} K(Q - 1) \frac{1}{Q} \left[ 1 - 2F_N \left( \frac{A - v}{\sigma_N} \right) \right] f_Y(v - A)dv - \frac{2\mu_3(\sigma_N)\mu'_1(\sigma_N)}{K}. \tag{91}
\]
Equation (91) implies that \( \sigma^2_0(\sigma_N) \) 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 \( \Lambda_{\text{NQ}} \).

**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
\[
\lim_{\sigma_N \to 0} \frac{dP_0}{d\sigma_N} = 2\phi \left( z_T - \frac{2\mu_1(0)}{\sigma_1(0)} \right) \times \frac{\sigma_1(0)\mu'_1(0)}{\sigma_0'^2(0)} - \frac{\mu_1(0)\mu'_1(0)}{\sigma_0'^2(0)} \tag{92}
\]
because of (82). We know that
\[
\lim_{\sigma_N \to 0} \sigma_0'^2(\sigma_N) = 2\sigma_1(0) \lim_{\sigma_N \to 0} \frac{\sigma_1'(\sigma_N)}{\sigma_0'(0)} \tag{93}
\]
because \( \sigma_0'^2(\sigma_N) = 2\sigma_1(\sigma_N)\sigma_1'(\sigma_N) \). Equations (94) and (95) imply that the rate of the initial SR effect increases if \( \lim_{\sigma_N \to 0} \sigma_0'^2(\sigma_N) \) decreases. Further
\[
\lim_{\sigma_N \to 0} \left[ \sigma_0'^2(\sigma_N) - \int_{-\infty}^{\infty} K(Q - 1) \frac{1}{Q} \left[ 1 - 2F_N(\sigma_N|n) \right]^2 f_Y(v - A)dv \right] = \frac{2K(Q - 1)\mu_2(0)}{Q} \tag{94}
\]
if we substitute \( A/2 - v/\sigma_N = n \) in (92). Then
\[
\lim_{\sigma_N \to 0} \sigma_0'^2(\sigma_N) = \frac{2K(Q - 1)\mu_2(0)}{Q} \times \int_{-\infty}^{\infty} \left[ 1 - 2F_N(\sigma_N|n) \right] f_Y(n)dn. \tag{95}
\]
The integral of (97) is negative because \([1 - 2F_N(\sigma_N|n)]n \) is nonpositive. Then \( \lim_{\sigma_N \to 0} \sigma_0'^2(\sigma_N) \) 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 \( \sigma_N \) is maximal if \( \lim_{n \to \infty} \sigma_N \to 0 \). Equation (107) states that the symmetric standard quantizer noise \( \hat{N} \) that minimizes the expectation \( \int_{-\infty}^{\infty} [1 - 2F_N(n)] n f_N(n) \) maximizes the rate of increase for all \( \ell \in \mathbb{N} \). Equation (108) states that the symmetric standard quantizer noise \( \hat{N} \) that minimizes the expectation \( \int_{-\infty}^{\infty} [1 - 2F_N(n)] n f_N(n) \) maximizes the rate of increase for all \( \ell \in \mathbb{N} \).

\[
\sum_{\ell=0}^{\infty} [F_{\hat{N}}(n_\ell) - F_{\hat{N}}(-n_\ell)] n_\ell P_{\hat{N}}(n_\ell) = 2 \left( \sum_{\ell=1}^{\infty} [F_{\hat{N}}(n_\ell) - F_{\hat{N}}(-n_\ell)]^2 P_{\hat{N}}(n_\ell) \right)^{1/2} \frac{1}{\sqrt{2}}
\]

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

\[
\sqrt{2} \left( \sum_{\ell=1}^{\infty} p_\ell + \sum_{\ell=0}^{\infty} 2p_\ell \right)^{1/2} \text{ where } p_\ell = P_{\hat{N}}(n_\ell) P_{\hat{N}}(-n_\ell) \text{ and so }
\]

\[
\int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn \leq \frac{1}{\sqrt{2}} \left[ \left( \sum_{\ell=0}^{\infty} p_\ell \right)^2 + \left( \sum_{\ell=0}^{\infty} p_\ell \right)^2 \right]^{1/2}
\]

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

\[
\int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn = \frac{1}{\sqrt{3}} \text{ because } \sum_{\ell=0}^{\infty} p_\ell = 1
\]

\[
\int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn = \int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn
\]

Again the Cauchy-Schwarz inequality [84] gives

\[
\int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn \leq \left[ \int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn \right]^{1/2} \left[ \int_{-\infty}^{\infty} n^2 f_{\hat{N}}(n) \, dn \right]^{1/2}
\]

because \( \hat{N} \) has zero mean and unit variance.

\[
\int_{-\infty}^{\infty} [1 - 2F_{\hat{N}}(n)] n f_{\hat{N}}(n) \, dn = \frac{1}{\sqrt{3}}
\]

Equations (113)–(114) follow because \( U = F_{\hat{N}}(\hat{N}) \) is a uniform random variable in \([0, 1]\) for any continuous quantizer noise \( \hat{N} \). Inequality (110) becomes an equality if and only
if $F_N$ obeys $[2F_N(n) - 1]^2 = \sigma^2$ for some constant $\sigma$ on the support of $f_N^\pm$ [84]. Then $F_N(n) = 1/2 + \sqrt{\sigma}/2$ for all $n \in [-1/\sqrt{\sigma}, 1/\sqrt{\sigma}]$ for $\sigma = 1/3$ because $F_N$ is the CDF of standard quantizer noise. The same CDF implies that $N$ is uniformly distributed in $[-\sqrt{3}, \sqrt{3}]$. So symmetric uniform quantizer noise achieves the upper bound $1/3$ 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


