Noise-benefit forbidden-interval theorems for threshold signal detectors based on cross correlations

Sanya Mitaim[1] and Bart Kosko[2]

[1]*Department of Electrical and Computer Engineering, Faculty of Engineering Thammasat University, Pathumthani 12120, Thailand*

[2]*Department of Electrical Engineering, Signal and Image Processing Institute, University of Southern California, Los Angeles, California 90089-2564, USA*

(Received 25 March 2014; published 14 November 2014)

We show that the main forbidden interval theorems of stochastic resonance hold for a correlation performance measure. Earlier theorems held only for performance measures based on mutual information or the probability of error detection. Forbidden interval theorems ensure that a threshold signal detector benefits from deliberately added noise if the average noise does not lie in an interval that depends on the threshold value. We first show that this result holds for correlation for all finite-variance noise and for all forms of infinite-variance stable noise. A second forbidden-interval theorem gives necessary and sufficient conditions for a local noise benefit in a bipolar signal system when the noise comes from a location-scale family. A third theorem gives a general condition for a local noise benefit for arbitrary signals with finite second moments and for location-scale noise. This result also extends forbidden intervals to forbidden bands of parameters. A fourth theorem gives necessary and sufficient conditions for a local noise benefit when both the independent signal and noise are normal. A final theorem derives necessary and sufficient conditions for forbidden bands when using arrays of threshold detectors for arbitrary signals and location-scale noise.

DOI: 10.1103/PhysRevE.90.052124 PACS number(s): 05.40.Ca, 87.19.lc, 87.19.ll, 84.35.+i

I. FORBIDDEN INTERVAL THEOREMS FOR CROSS CORRELATION

We show that "forbidden interval" theorems (FITs) hold for a cross-correlation performance measure. These correlation noise benefits also extend to arrays of threshold neurons. Earlier results found similar noise benefits only for mutual information [1–6] or probability of error [7]. We have found experimental evidence of both correlation-based and bit-error-based noise benefits for carbon nanotube detectors [8,9] when testing a FIT prediction of a mutual-entropy-based noise benefit in a threshold detector. Other stochastic resonance (SR) results have used a correlation performance measure to demonstrate a noise benefit [10–12] but not as a FIT. Moskowitz has recently extended FITs to algebraic information theory [13].

A forbidden interval theorem states a sufficient or necessary condition for a nonlinear signal system to benefit from added noise so long as the average noise does not fall in an interval of parameter values. A FIT acts as a type of screening device for a nonlinear system because it tells the user whether the system can have a noise benefit at all. This screening effect also holds for the related necessary and sufficient inequalities that ensure a noise benefit based on maximum likelihood or Neyman-Pearson signal detection [14]. Adaptive algorithms or other schemes can then find the optimal noise level in systems that possess noise benefits [3,11,21]. The first FITs applied only to nonlinear signal systems that used mutual information as the performance measure [1,3]. Corollary 1 in Ref. [7] was a necessary-condition FIT based not on mutual information but on the probability of detection error. All FITs give rigorous conditions for a noise benefit or SR [1–3,7,11,14–23]. This paper extends correlation-based SR to threshold systems and threshold arrays that obey quantitative FITs. These threshold systems model threshold-like behavior in a wide range of physical and biological systems [18,24–28].

Theorem 1 below gives a direct correlation FIT dual of our earlier mutual-information FIT for threshold signals. It holds for all possible additive noise that has a finite variance. It further holds for all infinite-variance noise from the general stable family of probability density functions that includes Cauchy and Gaussian noise as special cases. Non-Gaussian stable noise does not have a mean but it does have a location parameter that acts like the mean and that equals the median if the stable noise is symmetric. The FIT holds in the stable case for noise whose location parameter does not lie in the forbidden interval. This first correlation FIT produces *total* SR in the sense that added noise achieves the correlation maximum for the threshold system.

Figure 1 shows a simulation instance of Theorem 1. We converted the binary yin-yang image to a bipolar image with amplitude A and used it as input to the threshold system (1). The threshold θ is 1 while the bipolar signal amplitude A is 0.7. This gives the forbidden interval $(\theta - A, \theta + A) = (0.3, 1.7)$. The uniform noise mean μ_N is 2. So there is a noise benefit because $\mu_N \notin (0.3, 1.7)$. The second panel uses uniform noise with mean $\mu_N = 1$. So there is no noise benefit because the noise mean falls in the forbidden interval (0.3, 1.7).

Theorem 2 gives a new type of correlation FIT for *partial* SR or a local noise benefit as in [7,14,23]. The FIT gives necessary and sufficient conditions for a positive correlation derivative with respect to the standard deviation σ_N of the added noise N: $\frac{\partial C}{\partial \sigma_N} > 0$. The signal system is a more complex bipolar signal system but the FIT applies only to noise that comes from a location-scale family. Location-scale family noise includes many common types of noise such as uniform, Gaussian, and α-stable noise. It does not include Poisson noise.

The next section describes the stochastic threshold signal system and the basic properties of the cross-correlation performance measure. Section II states and proves the correlation FIT of Theorem 1 for finite-variance noise and for infinite-variance stable noise. Section III sets up and proves the local correlation FIT of Theorem 2 for location-scale noise and bipolar signals.

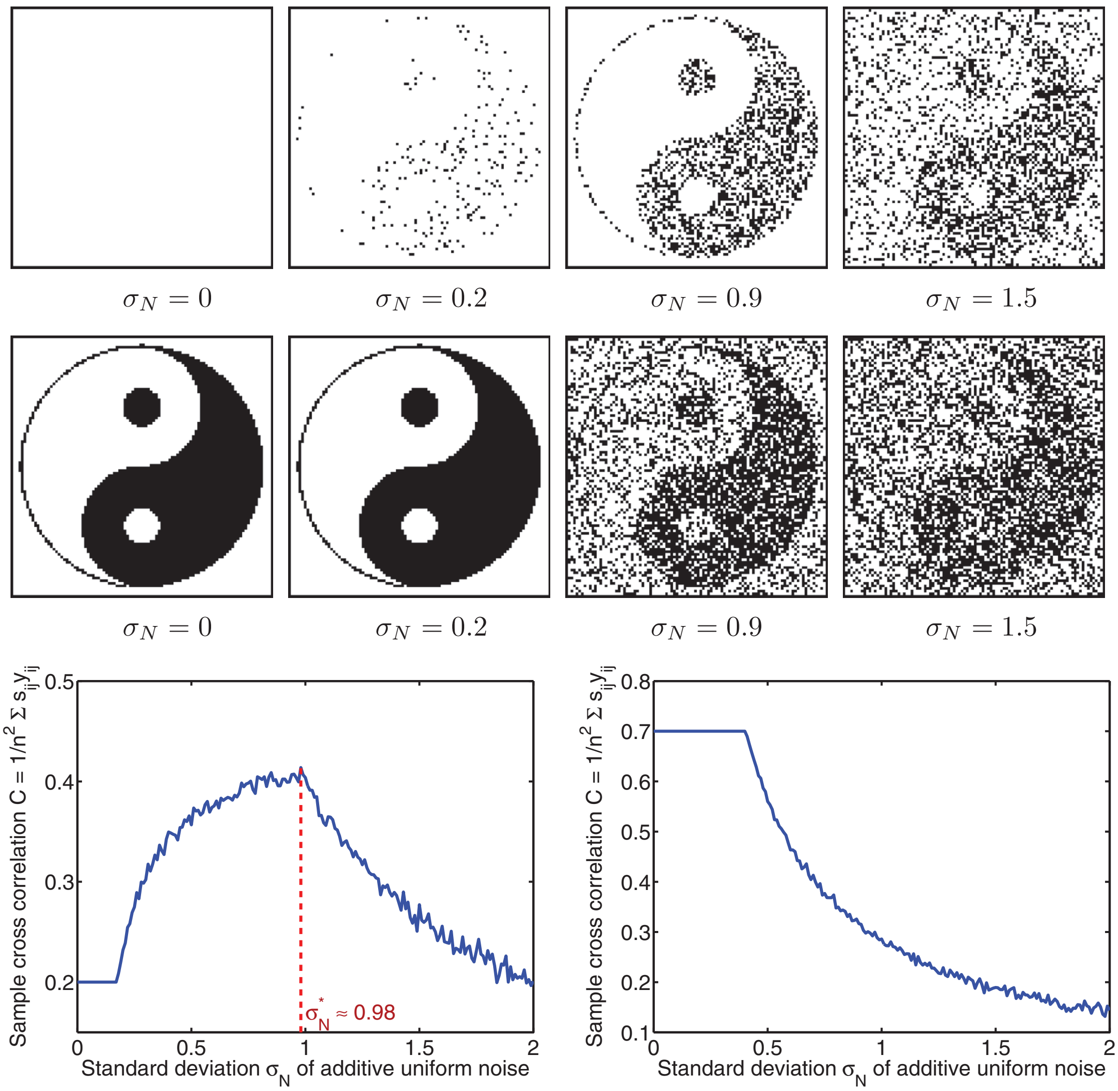


FIG. 1. (Color online) Forbidden-interval noise benefit in a subthreshold signal system (1) that uses the bipolar yin-yang image as its input. The threshold θ is $\theta = 1$. The bipolar signal has amplitude $A = 0.7$. Top panels: The noise is uniform with mean $\mu_N = 2$. So the mean does not lie in the forbidden interval $(\theta - A, \theta + A) = (0.3, 1.7)$. Thus the correlation system benefits from added noise and the correlation curve shows the signature nonmonotonic inverted U of a global SR noise benefit. The optimal noise is approximately $\sigma_N^* \approx 0.98$. Bottom panels: The noise is uniform with mean $\mu_N = 1$. So the mean lies in the forbidden interval $(\theta - A, \theta + A) = (0.3, 1.7)$. Thus there is no noise benefit.

The final sections extend the correlation FITs to different noise and signal types and to arrays of detectors.

II. CROSS CORRELATION IN A STOCHASTIC THRESHOLD DETECTOR WITH SUBTHRESHOLD BIPOLAR INPUT SIGNALS

This section describes the stochastic threshold signal system and introduces the performance measure of cross correlation. Consider the standard discrete-time threshold signal detector [2,3,11,15,17,18,21,29,30],

$$Y = \mathrm{sgn}(S + N - \theta) = \begin{cases} 1 & \text{if} \quad S + N \geqslant \theta \\ -1 & \text{if} \quad S + N < \theta \end{cases}. \qquad (1)$$

Here $\theta > 0$ is the system's threshold. We assume that the additive noise N is white with the probability density function (pdf) $f_N(n)$. This white-noise assumption does not limit the FIT analysis. The same system analysis applies to other types of noise with correlation functions $R_N(\tau)$ because these threshold systems are feedforward systems and thus have no dynamics. We consider bipolar input Bernoulli signal S with arbitrary success probability p such that $0 < p < 1$ with amplitude $A > 0$. Thus the signal's pdf has the form

$$f_S(s) = p\delta(s + A) + (1 - p)\delta(s - A), \qquad (2)$$

where δ denotes the Dirac δ function.

We assume *subthreshold* input signals for this threshold system: $A < \theta$. The output Y of the threshold system exactly matches the input signal S when $A > \theta$ and so there is no noise benefit. Let the symbol "0" denote the input signal $S = -A$ and output signal $Y = -1$. The symbol "1" denotes the input signal $S = A$ and output signal $Y = 1$. Then the conditional probabilities $P_{Y|S}(y|s)$ are

$$P_{Y|S}(0|0) = \Pr\{S + N < \theta\}|_{S=-A}, \tag{3}$$

$$= \Pr\{N < \theta + A\}, \tag{4}$$

$$= \int_{-\infty}^{\theta+A} f_N(n)dn = F_N(\theta + A), \tag{5}$$

$$P_{Y|S}(1|0) = 1 - P_{Y|S}(0|0) = 1 - F_N(\theta + A), \tag{6}$$

$$P_{Y|S}(0|1) = \Pr\{S + N < \theta\}|_{S=A}, \tag{7}$$

$$= \int_{-\infty}^{\theta-A} f_N(n)dn = F_N(\theta - A), \tag{8}$$

$$P_{Y|S}(1|1) = 1 - P_{Y|S}(0|1) = 1 - F_N(\theta - A), \tag{9}$$

where $F_N(z)$ is the cumulative distribution function (cdf) of N. The marginal density P_Y is

$$P_Y(y) = \sum_s P_{Y|S}(y|s)P_S(s) \tag{10}$$

by the theorem on total probability.

Other researchers have derived the conditional probabilities $P_{Y|S}(y|s)$ of the threshold system with Gaussian noise with bipolar inputs [15] and Gaussian inputs [31]. We neither restrict the noise density to be Gaussian nor require that the density have finite variance even if the density has a bell-curve shape.

We use cross correlation to measure the noise benefit or the SR effect. The cross correlation C of the input random variable S and output random variable Y is the expectation of the product of the two variables S and Y:

$$C = E[SY] = \sum_{s\in\mathcal{S}}\sum_{y\in\mathcal{Y}} sy P_{SY}(s,y), \tag{11}$$

$$= \sum_{s\in\mathcal{S}}\sum_{y\in\mathcal{Y}} sy P_{Y|S}(y|s)P_S(s). \tag{12}$$

We can also use the respective cross covariance:

$$K = E[(S - \mu_S)(Y - \mu_Y)] = C - \mu_S\mu_Y, \tag{13}$$

where $\mu_S = E[S] = \sum_{s\in\mathcal{S}} sP_S(s)$ and $\mu_Y = E[Y] = \sum_{y\in\mathcal{Y}} yP_Y(y)$. The cross correlation C and cross covariance K can be any real number. But $C \leqslant A$ for (1) with bipolar signals A and $-A$ since $Y = 1$ or $Y = -1$.

We next prove a lemma that allows a direct proof of Theorem 1. The lemma states that the signal S and output Y always have a lower-bounded correlation: $C \geqslant \mu_S\mu_Y$. S and Y are uncorrelated if and only if $C = \mu_S\mu_Y$. And independence implies uncorrelatedness. Then Theorem 1 states that the noise mean $\mu_N = E[N]$ does not lie in the "forbidden" subthreshold interval $(\theta - A,\theta + A)$ if and only if S and Y are asymptotically independent (and thus $C \to \mu_S\mu_Y$) as the noise standard deviation $\sigma \to 0$ for finite-variance noise (or as the noise dispersion $\gamma \to 0$ for α-stable noise with infinite variance).

The lemma further shows that $C > \mu_S\mu_Y$ when the forbidden interval $(\theta - A,\theta + A)$ has positive noise probability. The case $C = \mu_S\mu_Y$ holds just when the forbidden interval has zero noise probability. So the idea behind the proof of Theorem 1 is that *what goes down must go up* [1,2]: Increasing the noise variance or dispersion must necessarily increase the correlation C at some point. That increase is precisely an SR noise benefit. So the entire proof of Theorem 1 rests on showing that the forbidden interval has asymptotically zero noise probability as the noise variance or dispersion shrinks to zero.

Lemma. The threshold system (1) has an input-output cross correlation C that is at least $\mu_S\mu_Y$: $C \geqslant \mu_S\mu_Y$. Further: $C > \mu_S\mu_Y$ if $\lambda > 0$ and $C = \mu_S\mu_Y$ if $\lambda = 0$, where λ is the noise probability of the forbidden interval $(\theta - A,\theta + A)$:

$$\lambda = \int_{\theta-A}^{\theta+A} f_N(n)dn \geqslant 0. \tag{14}$$

Proof. The cross correlation has the form

$$C = \sum_{s\in\mathcal{S}}\sum_{y\in\mathcal{Y}} sy P_{Y|S}(y|s)P_S(s). \tag{15}$$

Note that $C = \mu_S\mu_Y$ if S and Y are uncorrelated where

$$\mu_S\mu_Y = \sum_{s\in\mathcal{S}}\sum_{y\in\mathcal{Y}} sy P_S(s)P_Y(y). \tag{16}$$

Thus (5)–(9) imply that

$$P_{Y|S}(0|0) - P_{Y|S}(0|1) = \int_{\theta-A}^{\theta+A} f_N(n)dn = \lambda. \tag{17}$$

$$P_{Y|S}(1|1) - P_{Y|S}(1|0) = \int_{\theta-A}^{\theta+A} f_N(n)dn = \lambda. \tag{18}$$

The two-symbol alphabet set $\mathcal{S}$ and the theorem on total probability give

$$P_Y(y) = \sum_s P_{Y|S}(y|s)P_S(s), \tag{19}$$

$$= P_{Y|S}(y|0)P_S(0) + P_{Y|S}(y|1)P_S(1), \tag{20}$$

$$= P_{Y|S}(y|0)P_S(0) + P_{Y|S}(y|1)(1 - P_S(0)), \tag{21}$$

$$= (P_{Y|S}(y|0) - P_{Y|S}(y|1))P_S(0) + P_{Y|S}(y|1), \tag{22}$$

$$= P_{Y|S}(y|0)(1 - P_S(1)) + P_{Y|S}(y|1)P_S(1), \tag{23}$$

$$= P_{Y|S}(y|0) - (P_{Y|S}(y|0) - P_{Y|S}(y|1))P_S(1). \tag{24}$$

It follows from (17) and (18) that

$$P_Y(0) = P_{Y|S}(0|1) + \lambda P_S(0) \geqslant P_{Y|S}(0|1), \tag{25}$$

$$P_Y(0) = P_{Y|S}(0|0) - \lambda P_S(1) \leqslant P_{Y|S}(0|0). \tag{26}$$

We have similarly that

$$P_Y(1) = P_{Y|S}(1|0) + \lambda P_S(1) \geqslant P_{Y|S}(1|0), \tag{27}$$

$$P_Y(1) = P_{Y|S}(1|1) - \lambda P_S(0) \leqslant P_{Y|S}(1|1). \tag{28}$$

Thus

$$P_{Y|S}(y|s) \geqslant P_Y(y) \quad \text{if} \quad sy = A, \tag{29}$$

$$P_{Y|S}(y|s) \leqslant P_Y(y) \quad \text{if} \quad sy = -A, \tag{30}$$

since either $y = 1$ or $y = -1$. Then (29) and (30) imply that

$$syP_S(s)P_{Y|S}(y|s) \geqslant syP_S(s)P_Y(s), \tag{31}$$

for all $s \in \{-A,A\}$ and $y \in \{-1,1\}$. So summing gives $C \geqslant \mu_S\mu_Y$.

Suppose last that the noise pdf $f_N(n)$ has nonzero measure in $(\theta - A,\theta + A)$. Then $\lambda > 0$ and the inequalities (25)–(28)

become strict inequalities as follows:

$$P_Y(0) > P_{Y|S}(0|1), \tag{32}$$

$$P_Y(0) < P_{Y|S}(0|0), \tag{33}$$

$$P_Y(1) > P_{Y|S}(1|0), \tag{34}$$

$$P_Y(1) < P_{Y|S}(1|1). \tag{35}$$

Then

$$C > \mu_S \mu_Y \quad \text{and} \quad K > 0. \tag{36}$$

Note that $\lambda = 0$ implies that $f_N(n)$ has zero mass in the interval $(\theta - A, \theta + A)$. Then

$$C = \mu_S \mu_Y \quad \text{and} \quad K = 0. \tag{37}$$

■

III. THE FIRST CORRELATION FORBIDDEN INTERVAL THEOREM: THRESHOLD SYSTEMS WITH BIPOLAR SUBTHRESHOLD SIGNALS

This section states and proves the first necessary and sufficient condition for SR based on correlation in a threshold system. The theorem holds for all noise with finite second moments and all α-stable impulsive noise. Stable noise has infinite variance if $\alpha < 2$ but such noise still has finite lower-order moments up to order α if $\alpha < 2$. Gaussian noise is stable with $\alpha = 2$. Cauchy noise is stable with $\alpha = 1$. A general α-stable pdf f has characteristic function or Fourier transform φ [32–36]:

$$\varphi(\omega) = \exp\{ia\omega - \gamma|\omega|^{\alpha}(1 + i\beta \operatorname{sgn}(\omega)\Gamma)\}, \tag{38}$$

where

$$\Gamma = \begin{cases} \tan\frac{\alpha\pi}{2} & \text{for} \quad \alpha \neq 1 \\ -\frac{2}{\pi}\ln|\omega| & \text{for} \quad \alpha = 1 \end{cases}, \tag{39}$$

and $i = \sqrt{-1}$, $0 < \alpha \leqslant 2$, $-1 \leqslant \beta \leqslant 1$, and $\gamma > 0$. The parameter α is the characteristic exponent. The variance of an α-stable density does not exist if $\alpha < 2$. The location parameter a acts like the "mean" of the density when $\alpha > 1$. β is a skewness parameter. The density is symmetric about a when $\beta = 0$. Then α controls the tail thickness. The bell curve has thicker tails as α falls. The theorem below still holds even when $\beta \neq 0$. The dispersion parameter $\gamma = |\kappa|^{\alpha}$ acts like a variance because it controls the width of a symmetric α-stable bell curve where κ is the scale parameter. There are few known closed forms of the α-stable densities for symmetric bell curves (when $\beta = 0$) [34]. Numerical integration of φ gives the probability densities $f(n)$. Figure 2 gives examples of α-stable pdfs and their white-noise realizations. Figure 2(c) shows that non-Gaussian bell curves can have infinite variance and yet have a finite dispersion.

Below we state the first noise-benefit theorem for the threshold system (1) for any finite-variance noise and α-stable noise. Note that when the noise standard deviation shrinks to zero $\sigma_N \to 0$ (or dispersion $\gamma_N \to 0$) the pdf $f_N(n) \to \delta(n - \mu_N)$, where μ_N is the noise mean [or $f_N(n) \to \delta(n - a)$ for α-stable noise with location a] for delta pulse δ. The lemma allows the proof of this noise-benefit theorem to turn on showing that S and Y are asymptotically independent (and thus asymptotically uncorrelated) as the noise probability of the forbidden interval shrinks to zero.

Figure 1 illustrates the sufficient and necessary conditions in Theorem 1 below for uniform noise added to a binary yin-yang image. The first panel shows sufficiency: There is a correlation noise benefit because the noise mean of 2 does not fall in the forbidden interval (0.3,1.7). The second panel shows necessity: There is no correlation noise benefit because the lower-intensity noise has mean 1 and thus the mean falls in the forbidden interval (0.3,1.7). The theorem's necessity result shows that the correlation maximum occurs when there is no noise. It does not rule out possible local noise fluctuations where a local increase in the noise intensity can produce a local increase in correlation.

Theorem 1. Suppose that the threshold signal system (1) has noise pdf $f_N(n)$ and that the input signal S is subthreshold ($A < \theta$). Suppose that the noise mean $\mu_N = E[N]$ does not lie in the signal-threshold interval $(\theta - A, \theta + A)$ if N has finite variance. Suppose that $a \notin (\theta - A, \theta + A)$ for the location parameter a of an α-stable noise density with characteristic function (38). Then the threshold system (1) exhibits the nonmonotone SR effect in the sense that $C \to \mu_S\mu_Y$ as $\sigma \to 0$ or $\gamma \to 0$. Conversely, there is no noise benefit in C if $\mu_N \in (\theta - A, \theta + A)$ or $a \in (\theta - A, \theta + A)$.

Proof. Sufficiency. Assume $0 < P_S(s) < 1$ to avoid triviality when $P_S(s) = 0$ or 1. The lemma implies that we need show only that S and Y are asymptotically independent and so $C \to \mu_S\mu_Y$ as $\sigma \to 0$ (or as $\gamma \to 0$). So we need to show only that $P_{SY}(s,y) = P_S(s)P_Y(y)$ or $P_{Y|S}(y|s) = P_Y(y)$ as $\sigma \to 0$ (or as $\gamma \to 0$) for all signal symbols $s \in \mathcal{S}$ and $y \in \mathcal{Y}$. Thus the result follows (similarly to the proofs in Refs. [1,2] for mutual information) if we can show that

$$\lambda = \int_{\theta - A}^{\theta + A} f_N(n)dn \to 0 \quad \text{as} \quad \sigma \to 0 \quad \text{or} \quad \gamma \to 0. \tag{40}$$

Case 1. Finite-variance noise.

Let the mean of the noise be $\mu_N = E[N]$ and the variance be $\sigma^2 = E[(N - \mu_N)^2]$. Then $\mu_N \notin (\theta - A, \theta + A)$ by hypothesis.

Now suppose that $\mu_N < \theta - A$. Pick $\varepsilon = \frac{1}{2}(\theta - A - \mu_N) > 0$. So $\theta - A - \varepsilon = \theta - A - \varepsilon + \mu_N - \mu_N = \mu_N + (\theta - A - \mu_N) - \varepsilon = \mu_N + 2\varepsilon - \varepsilon = \mu_N + \varepsilon$. Then

$$\lambda = \int_{\theta - A}^{\theta + A} f_N(n)dn, \tag{41}$$

$$\leqslant \int_{\theta - A}^{\infty} f_N(n)dn, \tag{42}$$

$$\leqslant \int_{\theta - A - \varepsilon}^{\infty} f_N(n)dn, \tag{43}$$

$$= \int_{\mu_N + \varepsilon}^{\infty} f_N(n)dn, \tag{44}$$

$$= \Pr\{N \geqslant \mu_N + \varepsilon\} = \Pr\{N - \mu_N \geqslant \varepsilon\}, \tag{45}$$

$$\leqslant \Pr\{|N - \mu_N| \geqslant \varepsilon\}, \tag{46}$$

$$\leqslant \frac{\sigma^2}{\varepsilon^2} \quad \text{by Chebyshev's inequality}, \tag{47}$$

$$\to 0 \quad \text{as } \sigma \to 0. \tag{48}$$

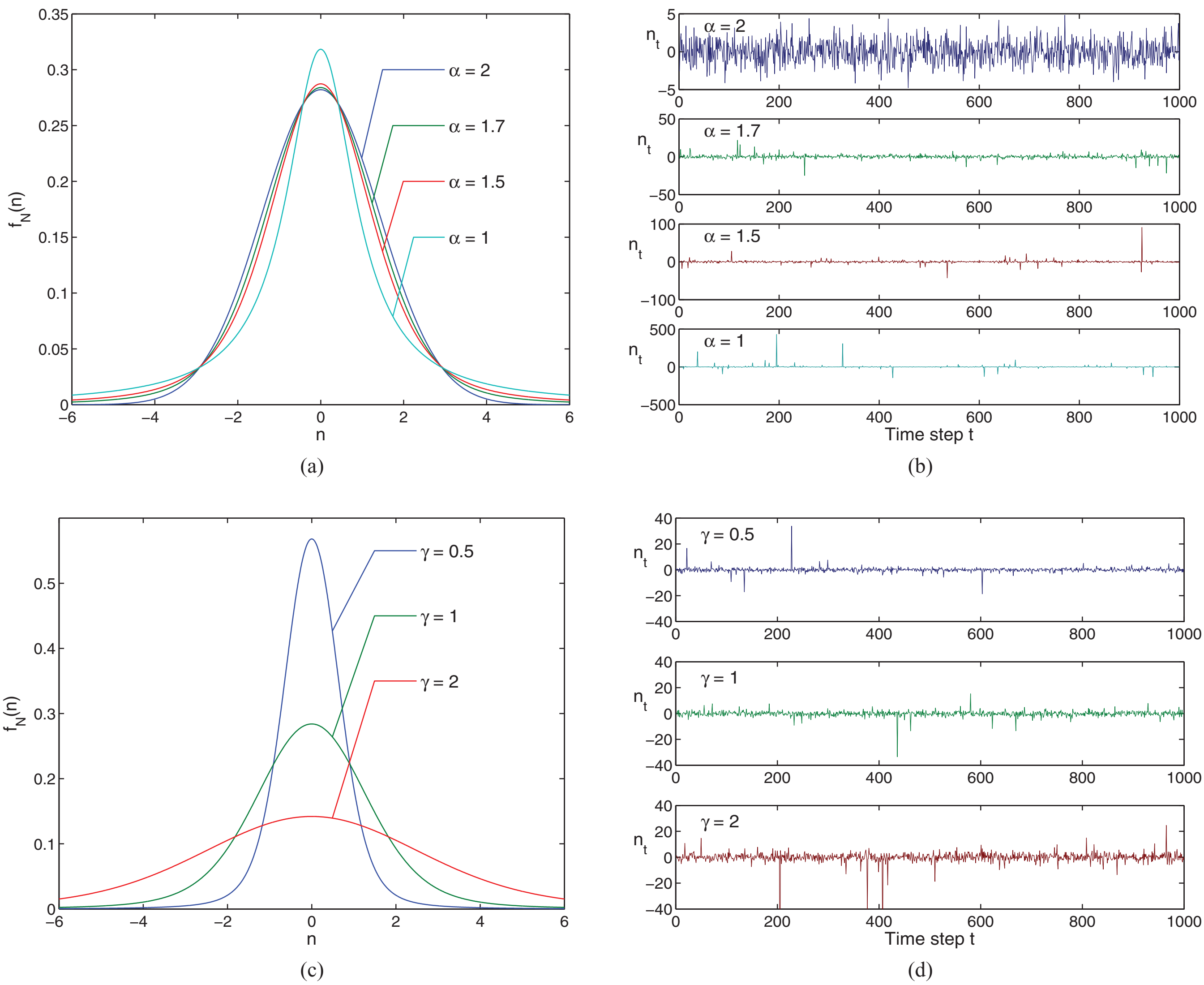


FIG. 2. (Color online) Samples of symmetric α-stable probability densities and their white-noise realizations. (a) Standard symmetric α-stable density functions with zero location ($a = 0$) and unit dispersion ($\gamma = 1$) for $\alpha = 2$, 1.7, 1.5, and 1. The densities are bell curves that have thicker tails as α decreases and thus that model increasingly impulsive noise as α decreases. The case $\alpha = 2$ gives a Gaussian density with variance two (or unit dispersion). The case $\alpha = 1$ gives the Cauchy density with infinite variance. (b) White-noise samples of α-stable random variables n_t with zero location and unit dispersion. The plots show realizations when $\alpha = 2$, 1.7, 1.5, and 1. Note the scale differences on the y axes. The α-stable variable N becomes more impulsive as the parameter α falls and thus as the bell curves get thicker. The algorithm in Refs. [37,38] generated these realizations. (c) α-stable density functions for $\alpha = 1.7$ with dispersions $\gamma = 0.5$, 1, and 2. (d) Samples of finite and infinite α-stable noise n_t for $\alpha = 1.7$ with finite dispersions $\gamma = 0.5$, 1, and 2.

Suppose next that $\mu_N > \theta + A$. Then pick $\varepsilon = \frac{1}{2}(\mu_N - \theta - A) > 0$ and so $\theta + A + \varepsilon = \theta + A + \varepsilon + \mu_N - \mu_N = \mu_N - (\mu_N - \theta - A) + \varepsilon = \mu_N - 2\varepsilon + \varepsilon = \mu_N - \varepsilon$. Then

$$\lambda = \int_{\theta-A}^{\theta+A} f_N(n)dn, \tag{49}$$

$$\leqslant \int_{-\infty}^{\theta+A+\varepsilon} f_N(n)dn, \tag{50}$$

$$= \int_{-\infty}^{\mu_N-\varepsilon} f_N(n)dn, \tag{51}$$

$$= \Pr\{N \leqslant \mu_N - \varepsilon\} = \Pr\{N - \mu_N \leqslant -\varepsilon\}, \tag{52}$$

$$\leqslant \Pr\{|N - \mu_N| \geqslant \varepsilon\}, \tag{53}$$

$$\leqslant \frac{\sigma^2}{\varepsilon^2} \quad \text{by Chebyshev's inequality,} \tag{54}$$

$$\to 0 \quad \text{as } \sigma \to 0. \tag{55}$$

Case 2. α-Stable noise.

The characteristic function $\varphi(\omega)$ of α-stable density $f_N(n)$ has the exponential form (38). This reduces to a simple complex exponential in the zero-dispersion limit:

$$\lim_{\gamma\to 0} \varphi(\omega) = \exp\{ia\omega\} \tag{56}$$

for each α, each skewness β, and each location a. So Fourier transformation gives the corresponding density function in the limiting case ($\gamma \to 0$) as a translated δ function:

$$\lim_{\gamma\to 0} f_N(n) = \delta(n-a). \quad (57)$$

Then

$$\lambda = \int_{\theta-A}^{\theta+A} f_N(n)dn, \quad (58)$$
$$= \int_{\theta-A}^{\theta+A} \delta(n-a)dn, \quad (59)$$
$$= 0 \quad \text{because } a \notin (\theta-A,\theta+A). \quad (60)$$

Then $P_Y(y) = P_{Y|S}(y|s)$ as $\gamma \to 0$.

Thus Cases 1 and 2 both imply that S and Y are asymptotically independent and so they are asymptotically uncorrelated: $C = \mu_S\mu_Y$ as $\sigma \to 0$ for finite-variance noise or as $\gamma \to 0$ for α-stable noise.

Necessity. We show that the cross correlation C is maximum ($C \to A$) as $\sigma_N \to 0$ (or $\gamma_N \to 0$) if $\mu_N \in (\theta-A,\theta+A)$ [or if $a \in (\theta-A,\theta+A)$]. Assume $0 < P_S(s) < 1$ to avoid triviality when $P_S(s) = 0$ or 1.

Case 1. Finite-variance noise.

We now show that $P_{Y|S}(y|s)$ is either 1 or 0 as $\sigma_N \to 0$. Let the noise mean be $\mu_N = E[N]$ and the variance be $\sigma_N^2 = E[(N-\mu_N)^2]$. Then again $\mu_N \in (\theta-A,\theta+A)$ by hypothesis.

Consider $P_{Y|S}(0|0)$. Pick $\varepsilon = \frac{1}{2}(\theta+A-\mu_N) > 0$. So $\theta + A - \varepsilon = \theta + A - \varepsilon + \mu_N - \mu_N = \mu_N + (\theta+A-\mu_N) - \varepsilon = \mu_N + 2\varepsilon - \varepsilon = \mu_N + \varepsilon$. Then

$$P_{Y|S}(0|0) = \int_{-\infty}^{\theta+A} f_N(n)dn, \quad (61)$$
$$\geqslant \int_{-\infty}^{\theta+A-\varepsilon} f_N(n)dn, \quad (62)$$
$$= \int_{-\infty}^{\mu_N+\varepsilon} f_N(n)dn, \quad (63)$$
$$= 1 - \int_{\mu_N+\varepsilon}^{\infty} f_N(n)dn, \quad (64)$$
$$= 1 - \Pr\{N \geqslant \mu_N + \varepsilon\}, \quad (65)$$
$$= 1 - \Pr\{N - \mu_N \geqslant \varepsilon\}, \quad (66)$$
$$\geqslant 1 - \Pr\{|N-\mu_N| \geqslant \varepsilon\}, \quad (67)$$
$$\geqslant 1 - \frac{\sigma_N^2}{\varepsilon^2} \quad \text{by Chebyshev's inequality}, \quad (68)$$
$$\to 1 \quad \text{as } \sigma_N \to 0. \quad (69)$$

So $P_{Y|S}(0|0) = 1$ and $P_{Y|S}(1|0) = 0$.

Similarly for $P_{Y|S}(1|1)$: Pick $\varepsilon = \frac{1}{2}(\mu_N - \theta + A) > 0$. So $\theta - A + \varepsilon = \theta - A + \varepsilon + \mu_N - \mu_N = \mu_N + (\theta - A - \mu_N) + \varepsilon = \mu_N - 2\varepsilon + \varepsilon = \mu_N - \varepsilon$. Then

$$P_{Y|S}(1|1) = \int_{\theta+A}^{\infty} f_N(n)dn, \quad (70)$$
$$\geqslant \int_{\theta-A+\varepsilon}^{\infty} f_N(n)dn, \quad (71)$$
$$= \int_{\mu_N-\varepsilon}^{\infty} f_N(n)dn, \quad (72)$$
$$= 1 - \int_{-\infty}^{\mu_N-\varepsilon} f_N(n)dn, \quad (73)$$
$$= 1 - \Pr\{N \leqslant \mu_N - \varepsilon\}, \quad (74)$$
$$= 1 - \Pr\{N - \mu_N \leqslant -\varepsilon\}, \quad (75)$$
$$\geqslant 1 - \Pr\{|N-\mu_N| \geqslant \varepsilon\}, \quad (76)$$
$$\geqslant 1 - \frac{\sigma^2}{\varepsilon^2} \quad \text{by Chebyshev's inequality}, \quad (77)$$
$$\to 1 \quad \text{as } \sigma_N \to 0. \quad (78)$$

So $P_{Y|S}(1|1) = 1$ and $P_{Y|S}(0|1) = 0$.

Case 2. α-Stable noise.

Again we have

$$\lim_{\gamma\to 0} f_N(n) = \delta(n-a). \quad (79)$$

Then

$$P_{Y|S}(0|0) = \int_{-\infty}^{\theta+A} f_N(n)dn, \quad (80)$$
$$\to \int_{-\infty}^{\theta+A} \delta(n-a)dn = 1 \quad \text{as } \gamma \to 0. \quad (81)$$

Similarly,

$$P_{Y|S}(1|1) = \int_{\theta-A}^{\infty} f_N(n)dn, \quad (82)$$
$$\to \int_{\theta-A}^{\infty} \delta(n-a)dn = 1 \quad \text{as } \gamma \to 0. \quad (83)$$

The four conditional probabilities for both finite-variance and infinite-variance cases imply that the cross correlation $C \to A$ as $\sigma \to 0$ (or $\gamma \to 0$) since then (12) gives $C = A(1)P(A) + (-A)(-1)P(-A) = A$. ■

IV. A FORBIDDEN INTERVAL THEOREM FOR LOCATION-SCALE NOISE AND BIPOLAR SIGNALS

This section derives necessary and sufficient conditions for the *local* noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$. Theorem 2 shows that this takes the form of a correlation FIT and that the conditions depend on the system parameters.

The signal system now is a threshold system with bipolar signals and additive white noise N that has pdf $f_N(n)$ that belongs to a location-scale family:

$$f_N(n) = \frac{1}{\sigma_N} f_{\tilde{N}}\left(\frac{n-\mu_N}{\sigma_N}\right) \quad (84)$$

with mean μ_N and variance σ_N^2. Thus the cdf is

$$F_N(n) = F_{\tilde{N}}\left(\frac{n-\mu_N}{\sigma_N}\right), \quad (85)$$

where $F_{\tilde{N}}$ is the cdf of the standardized random variable $\tilde{N} = \frac{N-\mu_N}{\sigma_N}$. Note that we can replace the mean μ_N with location parameter a and the standard deviation σ_N with the scale parameter $\kappa = \gamma^{1/\alpha}$ for α-stable noise. Thus the α-stable family belongs to the location-scale family for fixed α and β. We can rewrite the pdf of any random variable Y in terms of the pdf of the standard random variable

$X = (Y - a)/\kappa$:

$$f(y;a,\kappa,\alpha,\beta) = \frac{1}{\kappa} f\left(\frac{y-a}{\kappa};0,1,\alpha,\beta\right), \quad (86)$$

for fixed α and β. So the pdf $f(y;a,\kappa,\alpha,\beta)$ of an α-stable random variable Y with location a and scale κ has the form

$$f(y;a,\kappa,\alpha,\beta) = \int_{-\infty}^{\infty} \exp\{-ity + iat - |\kappa t|^{\alpha}(1 + i\beta \text{sgn}(t)\Gamma)\}dt. \quad (87)$$

Let $x = (y - a)/\kappa$ and $\tau = \kappa t$. Then

$$f(y;a,\kappa,\alpha,\beta) = \int_{-\infty}^{\infty} \exp\{-it(x\kappa + a) + ita - |\kappa t|^{\alpha}(1 - \beta \text{sgn}(t)\Gamma)\}dt, \quad (88)$$

$$= \frac{1}{\kappa}\int_{-\infty}^{\infty} \exp\{-i\tau x - |\tau|^{\alpha}(1 - \beta \text{sgn}(\tau)\Gamma)\}d\tau, \quad (89)$$

$$= \frac{1}{\kappa} f(x;0,1,\alpha,\beta) = \frac{1}{\kappa} f\left(\frac{y-a}{\kappa};0,1,\alpha,\beta\right). \quad (90)$$

The location-scale structure of the pdf and (5)–(9) give the conditional pdf $f_{Y|S}(y|s)$ as

$$f_{Y|S}(y|s) = \begin{cases} F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = -1 \\ 1 - F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = 1 \end{cases}, \quad (91)$$

$$= F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\delta(y+1) + \left(1 - F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\right)\delta(y-1). \quad (92)$$

Then the input-output cross correlation measure $C = E[SY]$ of the threshold system (1) with bipolar signal S with pdf $f_S(s) = p\delta(s + A) + (1 - p)\delta(s - A)$ and location-scale noise N has the form

$$C = \mu_S + 2pAF_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right) - 2(1-p)AF_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right). \quad (93)$$

Suppose the cross correlation (93) is differentiable with respect to σ_N. Theorem 2 below states the necessary and sufficient condition for the partial noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ for bipolar signal S and location-scale noise N.

Theorem 2. Suppose the signal S is bipolar with pdf $f_S(s) = p\delta(s + A) + (1 - p)\delta(s - A)$. Suppose the location-scale noise N has mean μ_N and variance σ_N^2 with pdf $f_N(n) = \frac{1}{\sigma_N} f_{\tilde{N}}(\frac{n-\mu_N}{\sigma_N})$.

Necessity. The threshold system (1) does not have a local noise benefit $\frac{\partial C}{\partial \sigma_N} < 0$ if $\mu_N \in (\theta - A, \theta + A)$.

Sufficiency. The threshold system (1) has a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if $\mu_N \notin (\theta - A, \theta + A)$ and if the system parameters satisfy inequalities (i) or (ii) as follows:

(i) $\mu_N > \theta + A$ and

$$\frac{p}{1-p} > \frac{(\theta - A - \mu_N)f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N)f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}. \quad (94)$$

or

(ii) $\mu_N < \theta - A$ and

$$\frac{p}{1-p} < \frac{(\theta - A - \mu_N)f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N)f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}. \quad (95)$$

Proof.

$$C = E[SY] \quad (96)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} syf_{S,Y}(s,y)\,ds\,dy \quad (97)$$

$$= \int_{-\infty}^{\infty} sf_S(s)\int_{-\infty}^{\infty} yf_{Y|S}(y|s)\,dy\,ds \quad (98)$$

$$= \int_{-\infty}^{\infty} sf_S(s)\left\{(-1)F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) + (1)\left[1 - F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\right]\right\}ds \quad (99)$$

$$= \int_{-\infty}^{\infty} sf_S(s)\left[1 - 2F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\right]ds \quad (100)$$

$$= \int_{-\infty}^{\infty} sf_S(s)ds - 2\int_{-\infty}^{\infty} sf_S(s)F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)ds \quad (101)$$

$$= \mu_S - 2\int_{-\infty}^{\infty} sf_S(s)F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)ds. \quad (102)$$

The input-output cross correlation measure for the threshold system with bipolar signal S with pdf $f_S(s) = p\delta(s + A) + (1 - p)\delta(s - A)$ and location-scale noise N follows from (102) as

$$C = \mu_S - 2\int_{-\infty}^{\infty} s[p\delta(s + A) + (1 - p)\delta(s - A)] \times F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)ds \quad (103)$$

$$= \mu_S + 2pAF_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right) - 2(1-p)AF_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right). \quad (104)$$

We first show that the local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ holds if and only if the following inequality holds:

$$(1-p)(\theta - A - \mu_N)f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right) > p(\theta + A - \mu_N)f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right). \quad (105)$$

Suppose the cdf $F_{\tilde{N}}$ is absolutely continuous and thus differentiable [39]: $\frac{dF_{\tilde{N}}(z)}{dz} = f_{\tilde{N}}(z)$. Then the partial derivative

$\frac{\partial C}{\partial \sigma_N}$ has the form

$$\frac{\partial C}{\partial \sigma_N} = -2pA\left(\frac{\theta + A - \mu_N}{\sigma_N^2}\right) f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right) + 2(1-p)A\left(\frac{\theta - A - \mu_N}{\sigma_N^2}\right) f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right). \tag{106}$$

Then set $\frac{\partial C}{\partial \sigma_N} > 0$ and (106) becomes (105).

Inequality (105) requires checking the four cases that correspond to whether $\theta - A - \mu_N$ is positive or negative and whether $\theta + A - \mu_N$ is positive or negative.

Case 1. $\theta - A - \mu_N < 0$ and $\theta + A - \mu_N < 0$.

The conditions imply that $\mu_N > \theta + A$. Thus we have a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if and only if $\mu_N > \theta + A$ and the signal-noise-system parameters in Eq. (105) that satisfy

$$\frac{p}{1-p} > \frac{(\theta - A - \mu_N) f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N) f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}. \tag{107}$$

Case 2. $\theta - A - \mu_N > 0$ and $\theta + A - \mu_N > 0$.

The conditions imply that $\mu_N < \theta - A$. Thus we have a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if and only if $\mu_N < \theta - A$ and the signal-noise-system parameters in Eq. (105) that satisfy

$$\frac{p}{1-p} < \frac{(\theta - A - \mu_N) f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N) f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}. \tag{108}$$

Thus Cases 1 and 2 give sufficient conditions for a local noise benefit (i) and (ii).

Case 3. $\theta - A - \mu_N < 0$ and $\theta + A - \mu_N > 0$.

The conditions imply that $\theta - A < \mu_N < \theta + A$. The left-hand side of (105) is negative while the right-hand side of (105) is positive and thus (105) does not hold. So there is no noise benefit.

The condition $\theta - A < \mu_N < \theta + A$ results in a negative partial derivative $\frac{\partial C}{\partial \sigma_N}$:

$$\frac{\partial C}{\partial \sigma_N} = -2pA\left(\frac{\theta + A - \mu_N}{\sigma_N^2}\right) f_{\tilde{N}}\left(\frac{\theta + A - \mu_N}{\sigma_N}\right) + 2(1-p)A\left(\frac{\theta - A - \mu_N}{\sigma_N^2}\right) f_{\tilde{N}}\left(\frac{\theta - A - \mu_N}{\sigma_N}\right) \tag{109}$$

$$< 0. \tag{110}$$

Thus $\theta - A < \mu_N < \theta + A$ is a necessary condition for a local noise benefit.

Case 4. $\theta - A - \mu_N > 0$ and $\theta + A - \mu_N < 0$.

The conditions imply that $\mu_N < \theta - A$ and $\mu_N > \theta + A$. This case is logically impossible since $A > 0$. ■

The proof also shows that the noise benefit is a local maximum when equality replaces the inequalities in Eq. (94) or in Eq. (108).

V. LOCAL NOISE BENEFITS FOR ARBITRARY SIGNAL AND LOCATION-SCALE NOISE

We next consider the threshold system (1) when the signal S has arbitrary pdf $f_S(s)$ and the noise N comes from the location-scale family $f_N(n) = \frac{1}{\sigma_N} f_{\tilde{N}}(\frac{n-\mu_N}{\sigma_N})$. Thus the conditional pdf of Y given a signal value s is

$$f_{Y|S}(y|s) = \begin{cases} F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = -1 \\ 1 - F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = 1 \end{cases}, \tag{111}$$

$$= F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\delta(y+1) + \left[1 - F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\right]\delta(y-1), \tag{112}$$

where $F_{\tilde{N}}$ is the cdf of the standardized random variable $\tilde{N} = \frac{N-\mu_N}{\sigma_N}$.

The input-output cross correlation measure C for the threshold system with arbitrary signal pdf $f_S(s)$ and location-scale noise with pdf $f_N(n) = \frac{1}{\sigma_N} f_{\tilde{N}}(\frac{n-\mu_N}{\sigma_N})$ has the form (102)

$$C = \mu_S - 2\int_{-\infty}^{\infty} s f_S(s) F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds. \tag{113}$$

Theorem 3 below states a necessary and sufficient condition for a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ for an arbitrary signal S and location-scale noise N.

Theorem 3. Suppose the input signal S has pdf $f_S(s)$. Suppose the location-scale noise N has mean μ_N and variance σ_N^2 with pdf $f_N(n) = \frac{1}{\sigma_N} f_{\tilde{N}}(\frac{n-\mu_N}{\sigma_N})$. Then the threshold system (1) has the local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if and only if

$$r_{SN}^2 < (\theta - \mu_N)\mu_{SN}, \tag{114}$$

where

$$\mu_{SN} = \int_{-\infty}^{\infty} s \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{115}$$

$$r_{SN}^2 = \int_{-\infty}^{\infty} s^2 \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{116}$$

and the normalizer $k_{SN} > 0$ when the the product $f_S(s) f_{\tilde{N}}(\frac{\theta - s - \mu_N}{\sigma_N})$ has nonzero support:

$$k_{SN} = \int_{-\infty}^{\infty} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds. \tag{117}$$

Proof. Suppose C is differentiable with respect to the noise standard deviation σ_N. Then

$$\frac{\partial C}{\partial \sigma_N} = -2\int_{-\infty}^{\infty} s f_S(s) \frac{\partial}{\partial \sigma_N} F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds. \tag{118}$$

The partial derivative $\frac{\partial F_{\tilde{N}}}{\partial \sigma_N}$ has the form

$$\frac{\partial}{\partial \sigma_N} F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) = \frac{\partial}{\partial \sigma_N} \int_{\infty}^{\frac{\theta - s - \mu_N}{\sigma_N}} f_{\tilde{N}}(z) dz = -\frac{\theta - s - \mu_N}{\sigma_N^2} f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right). \tag{119}$$

Then

$$\frac{\partial C}{\partial \sigma_N} = 2\int_{-\infty}^{\infty} \frac{s(\theta - s - \mu_N)}{\sigma_N^2} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{120}$$

$$= 2\int_{-\infty}^{\infty} \frac{s(\theta - \mu_N)}{\sigma_N^2} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds - 2\int_{-\infty}^{\infty} \frac{s^2}{\sigma_N^2} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds. \tag{121}$$

Consider the product $f_S(s) f_{\tilde{N}}(\frac{\theta - s - \mu_N}{\sigma_N})$. Suppose the supports of both pdfs overlap. We can define $f_{SN}(\theta;s) = \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}(\frac{\theta - s - \mu_N}{\sigma_N})$ as a pdf, where $k_{SN}(\theta) > 0$ is the normalizer such that $\int_{-\infty}^{\infty} \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}(\frac{\theta - s - \mu_N}{\sigma_N}) ds = 1$. So the partial derivative $\frac{\partial C}{\partial \sigma_N}$ has the form

$$\frac{\partial C}{\partial \sigma_N} = k_{SN}\int_{-\infty}^{\infty} \frac{s(\theta - \mu_N)}{\sigma_N^2} \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds - 2k_{SN}\int_{-\infty}^{\infty} \frac{s^2}{\sigma_N^2} \frac{1}{k_{SN}} f_S(s) f_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{122}$$

$$= \frac{2k_{SN}(\theta - \mu_N)}{\sigma_N^2}\int_{-\infty}^{\infty} s f_{SN}(\theta;s) ds - \frac{2k_{SN}}{\sigma_N^2}\int_{-\infty}^{\infty} s^2 f_{SN}(\theta;s) ds, \tag{123}$$

$$= \frac{2(\theta - \mu_N) k_{SN}\, \mu_{SN}}{\sigma_N^2} - \frac{2\, k_{SN}\, r_{SN}^2}{\sigma_N^2} \tag{124}$$

$$= \frac{2\, k_{SN}\left((\theta - \mu_N)\mu_{SN} - r_{SN}^2\right)}{\sigma_N^2}. \tag{125}$$

A local noise benefit occurs when $\frac{\partial C}{\partial \sigma_N} > 0$. This holds in (125) if and only if

$$r_{SN}^2 < (\theta - \mu_N)\mu_{SN} \tag{126}$$

since $k_{SN} > 0$ and $\sigma_N^2 > 0$. ■

The necessary and sufficient condition (114) characterizes noise benefits in threshold systems with arbitrary input signals and location-scale noise. But the first moment μ_{SN} and the second moment r_{SN}^2 depend on $\theta - \mu_N$. So specific forms of (114) require knowledge of μ_{SN} and r_{SN}^2.

VI. FORBIDDEN INTERVAL THEOREM FOR GAUSSIAN SIGNAL AND NOISE

Suppose the threshold system (1) has Gaussian input signal S with pdf $f_S(s) = N(\mu_S,\sigma_S^2)$ and Gaussian noise N with pdf $f_N(n) = N(\mu_N,\sigma_N^2)$. Then the input-output cross correlation measure C (113) becomes

$$C = \mu_S - 2\int_{-\infty}^{\infty} s f_S(s)\Phi\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{127}$$

where $\Phi(z)$ is the standard normal cdf.

Theorem 4 states a forbidden interval theorem for a local noise benefit in a threshold system with Gaussian signal and Gaussian noise.

Theorem 4. Suppose the signal S is Gaussian with mean μ_S and variance σ_S^2: $S \sim N(\mu_S,\sigma_S^2)$ and the noise N is Gaussian with mean μ_N and variance σ_N^2: $N \sim N(\mu_N,\sigma_N^2)$. Then the threshold system (1) has a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if and only if $\mu_N \notin (\theta - a_2, \theta - a_1)$, where

$$a_1 = -\frac{1}{2}\mu_S\left(\frac{\sigma_N^2}{\sigma_S^2} - 1\right) - \frac{1}{2}\sqrt{\mu_S^2(\sigma_N^2)^2 + 4\left(\sigma_S^2 + \sigma_N^2 + \mu_S\frac{\sigma_N^2}{\sigma_S^2}\right)}, \tag{128}$$

$$a_2 = -\frac{1}{2}\mu_S\left(\frac{\sigma_N^2}{\sigma_S^2} - 1\right) + \frac{1}{2}\sqrt{\mu_S^2(\sigma_N^2)^2 + 4\left(\sigma_S^2 + \sigma_N^2 + \mu_S\frac{\sigma_N^2}{\sigma_S^2}\right)}. \tag{129}$$

Proof. There is a local noise benefit if and only if $\frac{\partial C}{\partial \sigma_N} > 0$. The partial derivative (125) is

$$\frac{\partial C}{\partial \sigma_N} = \frac{2\, k_{SN}\left((\theta - \mu_N)\mu_{SN} - r_{SN}^2\right)}{\sigma_N^2}, \tag{130}$$

where the first moment μ_{SN} and second moment r_{SN}^2 for Gaussian signal and noise are (similar to the parameters of the posterior density in Bayes theorem) [40]

$$\mu_{SN} = \frac{\mu_S\sigma_N^2 + (\theta - \mu_N)\sigma_S^2}{\sigma_S^2 + \sigma_N^2}, \tag{131}$$

$$\sigma_{SN}^2 = \frac{\sigma_S^2\sigma_N^2}{\sigma_S^2 + \sigma_N^2}, \tag{132}$$

and

$$r_{SN}^2 = \sigma_{SN}^2 + \mu_{SN}^2. \tag{133}$$

The normalizer is

$$k_{SN} = \frac{\sigma_{SN}}{\sqrt{2\pi}\sigma_S} e^{-\frac{1}{2}\left(\frac{\mu_S^2\sigma_N^2 + (\theta - \mu_N)^2\sigma_S^2}{\sigma_S^2\sigma_N^2} - \frac{\mu_S\sigma_N^2 + (\theta - \mu_N)\sigma_S^2}{\sigma_S^2\sigma_N^2(\sigma_S^2 + \sigma_N^2)}\right)}. \tag{134}$$

Substitute (131)–(133) into (114) to obtain

$$\frac{\sigma_S^2\sigma_N^2}{\sigma_S^2 + \sigma_N^2} + \frac{\left(\mu_S\sigma_N^2 + (\theta - \mu_N)\sigma_S^2\right)^2}{\left(\sigma_S^2 + \sigma_N^2\right)^2} < (\theta - \mu_N)\frac{\mu_S\sigma_N^2 + (\theta - \mu_N)\sigma_S^2}{\sigma_S^2 + \sigma_N^2}. \tag{135}$$

This holds if and only if

$$\sigma_S^2\sigma_N^2\left(\sigma_S^2 + \sigma_N^2\right) + \mu_S^2\sigma_N^4 + 2(\theta - \mu_N)\mu_S\sigma_S^2\sigma_N^2 + (\theta - \mu_N)^2\sigma_S^4 < (\theta - \mu_N)\mu_S\left(\sigma_S^2\sigma_N^2 + \sigma_N^4\right) + (\theta - \mu_N)^2\sigma_S^4 + (\theta - \mu_N)^2\mu_S\sigma_N^2. \tag{136}$$

This holds in turn if and only if

$$(\theta - \mu_N)^2 + \mu_S\left(\frac{\sigma_N^2}{\sigma_S^2} - 1\right)(\theta - \mu_N) - \left(\sigma_S^2 + \sigma_N^2 + \frac{\mu_S^2\sigma_S^4}{\sigma_S^4}\right) > 0. \tag{137}$$

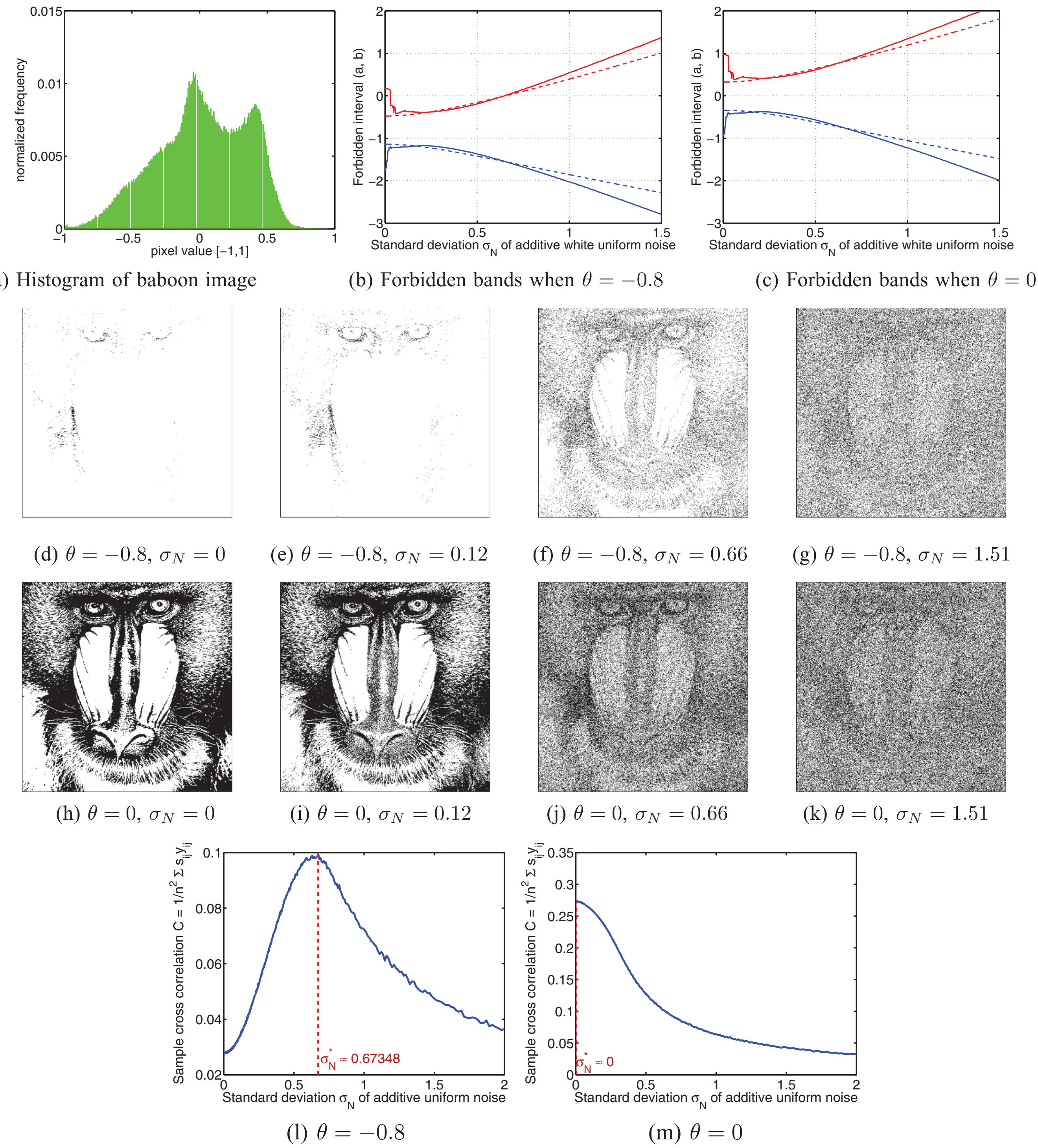


FIG. 3. (Color online) Forbidden-interval noise benefit in an array of threshold detectors (1) that uses the baboon image as its input. (a) Histogram of the gray-scale baboon image. We scale the pixel values to lie in the interval [−1,1] and use these values as input signals to the array of threshold detectors. [(b) and (c)] Forbidden intervals (bands) using the Gaussian approximation of Theorem 5 for $\theta = -0.8$ and $\theta = 0$. Each interval depends on the signal and noise statistics. Solid lines are intervals from the statistics of the baboon image with uniform noise. Dashed lines are estimated intervals using Gaussian approximation of both signal and noise. [(d)–(g)] Quantized (thresholded) baboon images when the detectors have threshold $\theta = -0.8$. The noise N is uniform with zero mean. Thus $\mu_N \notin (\theta - a_2, \theta - a_1)$ and this gives an SR noise benefit. [(h)–(k)] The detectors have threshold $\theta = 0$ and the noise N is uniform with zero mean. Thus $\mu_N \in (\theta - a_2, \theta - a_1)$ and there is no noise benefit. (l) Cross correlation C when $\theta = -0.8$: This gives an SR noise benefit. (m) Cross correlation C when $\theta = 0$: There is no noise benefit.

But (137) is quadratic in $(\theta - \mu_N)$. So we can replace the inequality in Eq. (137) with an equality and find its roots from the quadratic formula:

$$\theta - \mu_N = -\frac{1}{2}\mu_S\left(\frac{\sigma_N^2}{\sigma_S^2} - 1\right) \pm \frac{1}{2}\sqrt{\mu_S^2\left(\frac{\sigma_N^2}{\sigma_S^2} - 1\right)^2 + 4\left(\sigma_S^2 + \sigma_N^2 + \frac{\mu_S^2\sigma_S^4}{\sigma_S^4}\right)} \quad (138)$$

$$= a_1(\mu_S,\sigma_S^2,\sigma_N^2), a_2(\mu_S,\sigma_S^2,\sigma_N^2), \quad (139)$$

where the roots a_1 and a_2 depend on μ_S, σ_S^2, and σ_N^2 and $a_1(\mu_S,\sigma_S^2,\sigma_N^2) < 0 < a_2(\mu_S,\sigma_S^2,\sigma_N^2)$. Then a local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ holds if and only if

$$(\theta - \mu_N - a_1(\mu_S,\sigma_S^2,\sigma_N^2))(\theta - \mu_N - a_2(\mu_S,\sigma_S^2,\sigma_N^2)) > 0. \quad (140)$$

This is equivalent to

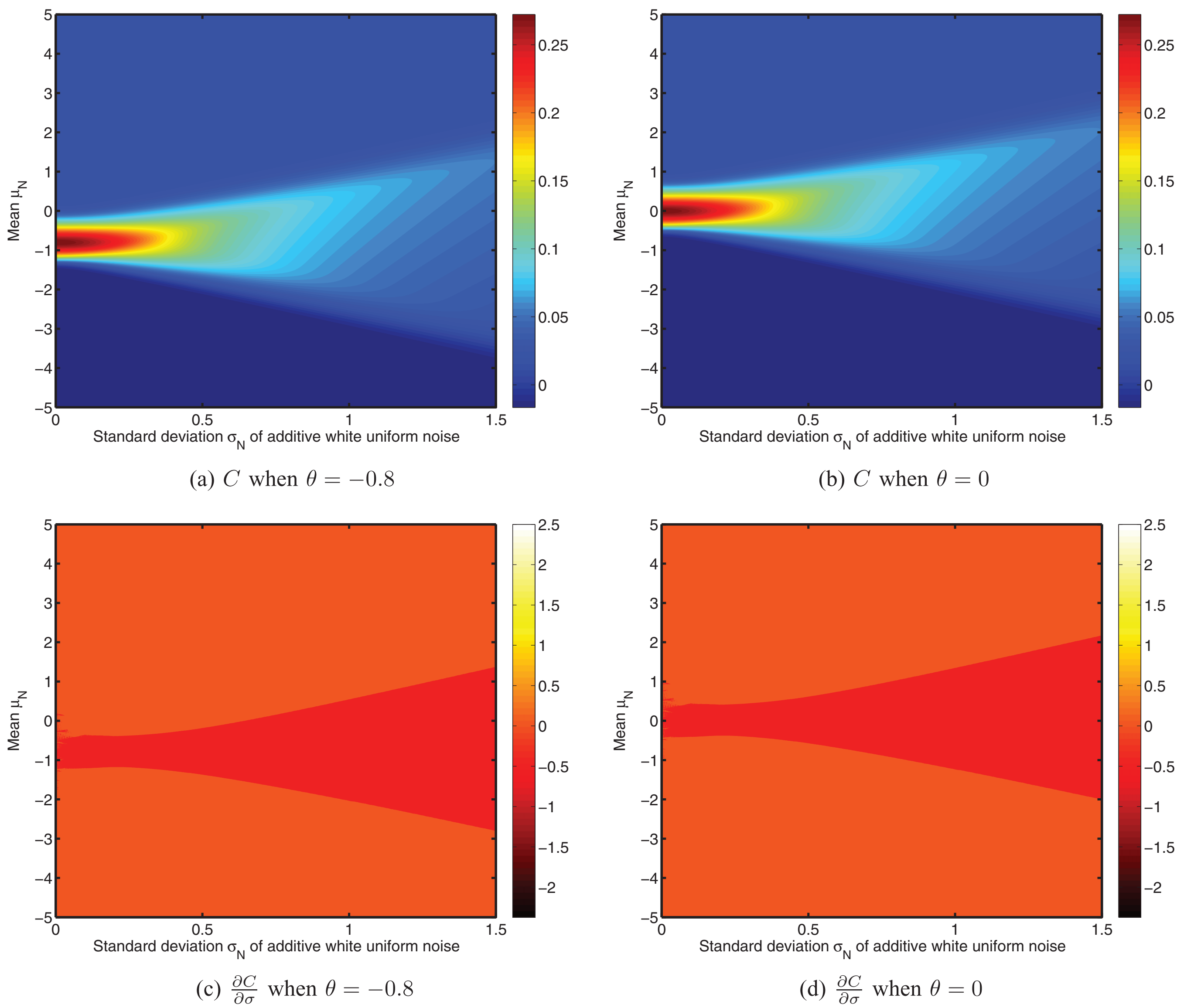
$$\mu_N < \theta - a_2(\mu_S,\sigma_S^2,\sigma_N^2) \quad (141)$$

or

$$\mu_N > \theta - a_1(\mu_S,\sigma_S^2,\sigma_N^2). \quad (142)$$

■

Note that the correlation-based forbidden interval changes as the noise variance σ_N^2 changes for the given signal's mean μ_S and variance σ_S^2. This also applies to forbidden intervals for all signal and noise pdfs.



(a) C when $\theta = -0.8$

(b) C when $\theta = 0$

(c) $\frac{\partial C}{\partial \sigma}$ when $\theta = -0.8$

(d) $\frac{\partial C}{\partial \sigma}$ when $\theta = 0$

FIG. 4. (Color online) Cross correlations C in (a) and (b) and derivatives $\frac{\partial C}{\partial \sigma}$ in (c) and (d) of an array of threshold detectors for the baboon image with uniform noise. The forbidden intervals (bands) for noise benefits in an array of threshold detectors (1) in Figure 2 derive from the contours of $\frac{\partial C}{\partial \sigma} = 0$.

VII. NOISE BENEFITS FOR ARRAYS OF THRESHOLD DETECTORS

This section derives correlation-based noise benefits for an array of Q threshold detectors. The independent and identically distributed input sequence S_i has arbitrary pdf with finite mean and variance and the independent and identically distributed noise N_i has location-scale pdf with finite mean and variance. The input signal S_i and noise N_i are independent. Each threshold detector has the input-output relationship

$$Y_i = \mathrm{sgn}(S_i + N_i - \theta_i) \tag{143}$$

for $i = 1,\dots,Q$. This resembles the case when we collect Q samples of output signals of a threshold detector (1) from the input signal sequence S_i. But the structure of the array of threshold detectors here differs from the parallel summing array of noisy threshold elements in Ref. [31] and the array of signal quantizers used for maximum-likelihood detection and Neyman-Pearson detection in Ref. [14]. Those systems use an array of threshold detectors to process a signal sample S with independent additive noise N_i and then sum the results as an output Y.

Figure 3 shows examples of noise benefits for the array of thresholds on the 512×512 baboon image with uniform noise. The histogram of the baboon image approximates the pdf $f_S(s)$ of the input signal. We obtain the first and second moments μ_{SN} and r^2_{SN} from the histogram and the noise pdf. Then we solve (148) to obtain the *forbidden bands* in Figs. 3(b) and 3(c). Figures 3(b) and 3(c) show how well the Gaussian approximation applies to the baboon images with uniform noise. Figure 4 shows the regions where the conditions in Theorem 3 hold.

We use the sample cross-correlation measure $\hat{C}$ of two Q-dimensional random vectors $S = [S_1 \cdots S_Q]^t$ and $Y = [Y_1 \cdots Y_Q]^t$ as a performance measure of the array:

$$\hat{C} = \frac{1}{Q}\sum_{i=1}^{Q} S_i Y_i. \tag{144}$$

Thus $\hat{C}$ is random with mean

$$\mu_{\hat{C}} = E[\hat{C}] = C, \tag{145}$$

$$= \mu_S - 2\int_{-\infty}^{\infty} s f_S(s) F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds, \tag{146}$$

and variance

$$\sigma^2_{\hat{C}} = \frac{1}{Q}\Bigg[\sigma_S^2 + 4\mu_S \int_{-\infty}^{\infty} s f_S(s) F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds - 4\left(\int_{-\infty}^{\infty} s f_S(s) F_{\tilde{N}}\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds\right)^2\Bigg]. \tag{147}$$

So we have on average a local noise benefit if $\frac{\partial \mu_{\hat{C}}}{\partial \sigma_N} = \frac{\partial C}{\partial \sigma_N} > 0$. Thus the condition for an *average* local noise benefit for the sample cross correlation $\hat{C}$ has the same form as the condition in Theorem 3. We state this condition as Theorem 5.

Theorem 5. Suppose the signal S has arbitrary pdf with finite mean μ_S and finite variance σ_S^2. Suppose the noise N also has an arbitrary pdf with finite mean μ_N and finite variance σ_N^2. Then the Q array of threshold systems (1) has an *average* local noise benefit $\frac{\partial \mu_{\hat{C}}}{\partial \sigma_N} > 0$ if and only if

$$r^2_{SN} < (\theta - \mu_N)\mu_{SN}, \tag{148}$$

where μ_{SN} and r^2_{SN} have the same form as in Eqs. (115) and (116).

VIII. CONCLUSION

We have found necessary and sufficient conditions for five correlation forbidden interval theorems. All theorems find a correlation-based noise benefit for one of three types of forbidden interval: $(\theta - A, \theta + A)$ with bipolar input signals, $r^2_{SN} < (\theta - \mu_N)\mu_{SN}$ with arbitrary signal and location-scale noise, or $(\theta - a_2(\mu_S,\sigma_S^2,\sigma_N^2), \theta - a_1(\mu_S,\sigma_S^2,\sigma_N^2))$ with a Gaussian input signal and Gaussian noise for the stochastic threshold signal detector in Eq. (1). Earlier FITs applied only to two-valued signals and not to continuous signals. Correlation noise benefits may hold for other combinations of random signals and noise. Adaptive algorithms should help find the optimal noise level in all such cases.

[1] B. Kosko and S. Mitaim, Neural Networks **16**, 755 (2003).
[2] B. Kosko and S. Mitaim, Phys. Rev. E **70**, 031911 (2004).
[3] S. Mitaim and B. Kosko, Transact. Neural Networks **15**, 1526 (2004).
[4] A. Patel and B. Kosko, Neural Networks **18**, 467 (2005).
[5] A. Patel and B. Kosko, IEEE Transact. Neural Networks **19**, 1993 (2008).
[6] M. M. Wilde and B. Kosko, J. Phys. A: Math. Theor. **42**, 465309 (2009).
[7] A. Patel and B. Kosko, Neural Networks **22**, 697 (2009).
[8] B. Kosko, I. Lee, S. Mitaim, A. Patel, and M. M. Wilde, in *Applications of Nonlinear Dynamics Model and Design of Complex Systems*, edited by V. In, P. Longhini, and A. Palacios (Springer, Berlin, 2009), pp. 71–89.
[9] I. Lee, X. Liu, C. Zhou, and B. Kosko, IEEE Transact. Nanotechnol. **5**, 613 (2006).
[10] J. J. Collins, C. C. Chow, and T. T. Imhoff, Nature **376**, 236 (1995).
[11] B. Kosko and S. Mitaim, Phys. Rev. E **64**, 051110 (2001).
[12] M. D. McDonnell and D. Abbott, Fluct. Noise Lett. **2**, L205 (2002).
[13] I. S. Moskowitz, P. N. Safier, and P. Cotae, in *Proceedings-(799) of Modelling, Identification and Control/801: Advances in Computer Science-2013* (Acta Press, Canada, 2013).
[14] A. Patel and B. Kosko, IEEE Transact. Signal Process. **59**, 488 (2011).
[15] A. R. Bulsara and A. Zador, Phys. Rev. E **54**, R2185 (1996).
[16] B. Franzke and B. Kosko, Phys. Rev. E **84**, 041112 (2011).
[17] L. Gammaitoni, Phys. Lett. A **208**, 315 (1995).
[18] P. Jung, Phys. Lett. A **207**, 93 (1995).
[19] B. Kosko, *Noise* (Viking/Penguin, New York, 2006).
[20] M. McDonnell, N. Stocks, C. Pearce, and D. Abbott, *Stochastic Resonance: From Suprathreshold Stochastic Resonance to*

Stochastic Signal Quantization (Cambridge University Press, Cambridge, 2008).

[21] S. Mitaim and B. Kosko, Proc. IEEE **86**, 2152 (1998).

[22] O. Osoba, S. Mitaim, and B. Kosko, Fluct. Noise Lett. **12**, 1350012 (2013).

[23] A. Patel and B. Kosko, IEEE Signal Process. Lett. **17**, 1005 (2010).

[24] C. Y. Cai, Li-Ping Yang, and C. P. Sun, Phys. Rev. A **89**, 012128 (2014).

[25] L. Gammaitoni, Phys. Rev. E **52**, 4691 (1995).

[26] Z. Gingl, L. B. Kiss, and F. Moss, Europhys. Lett. **29**, 191 (1995).

[27] P. Jung, Phys. Rev. E **50**, 2513 (1994).

[28] K. Szczepaniec and B. Dybiec, Eur. Phys. J. B, **86**, 468 (2013).

[29] J. J. Hopfield, Proc. Natl Acad. Sci. USA **79**, 2554 (1982).

[30] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence* (Prentice Hall, Inc., Upper Saddle River, NJ, 1991).

[31] N. G. Stocks, Phys. Rev. E, **63**, 041114 (2001).

[32] V. Akgiray and C. G. Lamoureux, J. Bus. Econ. Stat. **7**, 85 (1989).

[33] H. Bergstrom, Arkiv Matematik **2**, 375 (1952).

[34] K. Górska and K. A. Penson, Phys. Rev. E **83**, 061125 (2011).

[35] M. Grigoriu, *Applied Non-Gaussian Processes* (Prentice Hall, New York, 1995).

[36] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications* (John Wiley & Sons, New York, 1995).

[37] J. M. Chambers, C. L. Mallows, and B. W. Stuck, J. Am. Stat. Assoc. **71**, 340 (1976).

[38] P. Tsakalides and C. L. Nikias, IEEE Transact. Signal Process. **44**, 1623 (1996).

[39] P. Billingsley, *Probability and Measure* (John Wiley & Sons, New York, 2012).

[40] R. V. Hogg, J. W. McKean, and A. T. Craig, *Introduction to Mathematical Statistics*, 7th ed. (Prentice Hall, New York, 2013).