NOISE-BENEFIT FORBIDDEN INTERVAL THEOREMS FOR THRESHOLD SIGNAL DETECTORS BASED ON CROSS CORRELATION

Sanya Mitaim and Bart Kosko

†Department of Electrical and Computer Engineering, Faculty of Engineering
Thammasat University, Pathumthani 12120, Thailand
*Department of Electrical Engineering, Signal and Image Processing Institute
University of Southern California, Los Angeles, California 90089-2564, USA

ABSTRACT

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 also holds 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.

Keywords: cross correlation, forbidden interval theorem, noise benefit, stable noise, stochastic resonance, threshold systems.

I. FORBIDDEN INTERVAL THEOREMS FOR CROSS CORRELATION

We show for the first time 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], [2], [3], [4], [5], [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 SR results have used a correlation performance measure to demonstrate a noise benefit [10], [11], [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. The first FITs applied only to nonlinear signal systems that used mutual information as the performance measure [1], [3]. Corollary 1 in [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 so-called stochastic resonance or SR [1], [2], [3], [7], [11], [14], [15], [16], [17], [18], [19], [20], [21], [22], [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], [25], [26], [27], [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 $\theta$ is 1 while the
bipolar signal amplitude $A$ is 0.7. This gives the FIT
of $(\theta - A, \theta + A) = (0.3, 1.7)$. The uniform noise mean
$\mu_N$ 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 FIT $(0.3, 1.7)$.

Theorem 2 gives a new type of correlation FIT for
partial SR or a local noise benefit as in [7], [23],
[14]. The FIT gives necessary and sufficient conditions
for a positive correlation derivative with respect to the
partial SR or a local noise benefit as in [7], [23],
[14]. The FIT of Theorem 2 for location-scale noise and bipolar
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 alpha-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. 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]

$$ Y = \text{sgn}(S + N - \theta) = \begin{cases} 1 & \text{if } S + N \geq \theta \\ -1 & \text{if } S + N < \theta \end{cases} $$

(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

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

(2)

where $\delta$ denotes the Dirac delta 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

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

(3)

$$ = Pr\{N < \theta + A\} $$

(4)

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

(5)

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

(6)

$$ P_{Y|S}(0|1) = P_r\{S + N < \theta\} \bigg|_{S=A} $$

(7)

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

(8)

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

(9)

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

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

(10)

by the theorem on total probability.

Other researchers have derived the conditional proba-
bilities $P_{Y|S}(y|s)$ 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 of the input
random variable $S$ and output random variable $Y$ is
the expectation of the product of the two variables $S$
Fig. 1. Forbidden-interval noise benefit in a subthreshold signal system (1) that uses the bipolar yin-yang image as its input. The threshold \( \theta = 1 \). 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.

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: \( C \geq \mu_S \mu_Y \). \( S \) 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 alpha-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
Proof. The threshold system has the form

\[ C = \sum_{s \in S} \sum_{y \in Y} sy P_{Y|S}(y|s)P_{S}(s). \]  

(15)

Note that \( C = \mu_{SY} \) if \( S \) and \( Y \) are uncorrelated where

\[ \mu_{SY} = \sum_{s \in S} \sum_{y \in Y} sy P_{S}(s)P_{Y}(y). \]  

(16)

Thus (5)–(9) imply that

\[ P_{Y|S}(0|0) - P_{Y|S}(0|1) = \int_{\theta-A}^{\theta+A} f_{n}(n)dn = \lambda. \]  

(17)

\[ P_{Y|S}(1|1) - P_{Y|S}(1|0) = \int_{\theta-A}^{\theta+A} f_{n}(n)dn = \lambda. \]  

(18)

The two-symbol alphabet set \( S \) and the theorem on total probability give

\[ P_{Y}(y) = \sum_{s} P_{Y|S}(y|s)P_{S}(s) \]  

(19)

\[ = P_{Y|S}(y|0)P_{S}(0) + P_{Y|S}(y|1)P_{S}(1) \]  

(20)

\[ = P_{Y|S}(y|0)P_{S}(0) + P_{Y|S}(y|1)(1 - P_{S}(0)) \]  

(21)

\[ = (P_{Y|S}(y|0) - P_{Y|S}(y|1))P_{S}(0) + P_{Y|S}(y|1) \]  

(22)

\[ = P_{Y|S}(y|0)(1 - P_{S}(1)) + P_{Y|S}(y|1)P_{S}(1) \]  

(23)

\[ = P_{Y|S}(y|0) - (P_{Y|S}(y|0) - P_{Y|S}(y|1))P_{S}(1). \]  

(24)

It follows from (17)–(18) that

\[ P_{Y}(0) = P_{Y|S}(0|0) + \lambda P_{S}(0) \geq P_{Y|S}(0|0). \]  

(25)

\[ P_{Y}(0) = P_{Y|S}(0|0) - \lambda P_{S}(1) \leq P_{Y|S}(0|0). \]  

(26)

We have similarly that

\[ P_{Y}(1) = P_{Y|S}(1|0) + \lambda P_{S}(1) \geq P_{Y|S}(1|0). \]  

(27)

\[ P_{Y}(1) = P_{Y|S}(1|1) - \lambda P_{S}(0) \leq P_{Y|S}(1|1). \]  

(28)

Thus

\[ P_{Y|S}(y|s) \geq P_{Y}(y) \quad \text{if } sy = A \]  

(29)

\[ P_{Y|S}(y|s) \leq P_{Y}(y) \quad \text{if } sy = -A \]  

(30)

since either \( y = 1 \) or \( y = -1 \). Then (29)–(30) imply that

\[ sy P_{S}(s)P_{Y|S}(y|s) \geq sy P_{S}(s)P_{Y}(y) \]  

(31)

for all \( s \in \{-A, A\} \) and \( y \in \{-1, 1\} \). So summing gives

\[ C \geq \mu_{SY}. \]

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:

\[ P_{Y}(0) > P_{Y|S}(0|1) \]  

(32)

\[ P_{Y}(0) < P_{Y|S}(0|0) \]  

(33)

\[ P_{Y}(1) > P_{Y|S}(1|0) \]  

(34)

\[ P_{Y}(1) < P_{Y|S}(1|1). \]  

(35)

Then

\[ C > \mu_{SY} \quad \text{and} \quad K > 0. \]  

(36)

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

\[ C = \mu_{SY} \quad \text{and} \quad K = 0. \]  

(37)

Q.E.D.

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 alpha-stable impulsive noise. Stable noise has infinite variance if \( \alpha < 2 \) but such noise still has finite lower-order moments up to order \( \alpha \) if \( \alpha < 2 \). Gaussian noise is stable with \( \alpha = 2 \). Cauchy noise is stable with \( \alpha = 1 \). A general alpha-stable pdf \( f \) has characteristic function or Fourier transform \( \varphi \) [32], [33], [34], [35], [36]:

\[ \varphi(\omega) = \exp \{i\omega - \gamma|\omega|^{\alpha} (1 + i\beta \text{sgn}(\omega)\Gamma)\} \]  

(38)

where

\[ \Gamma = \begin{cases} \tan \frac{\alpha \pi}{2} & \text{for } \alpha \neq 1 \\ -\frac{2}{\pi} \ln |\omega| & \text{for } \alpha = 1 \end{cases} \]  

(39)

and \( i = \sqrt{-1}, \ 0 < \alpha < 2, -1 \leq \beta \leq 1, \) and \( \gamma > 0 \). The parameter \( \alpha \) is the characteristic exponent. Again the variance of an alpha-stable density does not exist if
\( \alpha < 2 \). The location parameter \( \alpha \) acts like the “mean” of the density when \( \alpha > 1 \). \( \beta \) is a skewness parameter. The density is symmetric about \( \alpha \) when \( \beta = 0 \). Then \( \alpha \) controls the tail thickness. The bell curve has thicker tails as \( \alpha \) falls. The theorem below still holds even when \( \beta \neq 0 \). The dispersion parameter \( \gamma = |\kappa|^a \) acts like a variance because it controls the width of a symmetric alpha-stable bell curve where \( \kappa \) is the scale parameter. There are few known closed forms of the \( \alpha \)-stable densities for symmetric bell curves (when \( \beta = 0 \)) [34]. Numerical integration of \( \varphi \) gives the probability densities \( f(n) \). Figure 2 gives examples of alpha-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 alpha-stable noise. Note that when the noise standard deviation shrinks to zero \( \sigma_N \to 0 \) (or dispersion \( \gamma_N \to 0 \)) then the pdf \( f_N(n) \to \delta(n - \mu_N) \) where \( \mu_N \) is the noise mean (or \( f_N(n) \to \delta(n - a) \) for alpha-stable noise with location \( a \)) for delta pulse \( \delta \). 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 alpha-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 \( P_{YS}(y|s) = P_Y(y) \) 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 (similar to the proofs in [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

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

\[
\leq \int_{\theta-A}^{\theta-A+\varepsilon} f_N(n)dn \tag{42}
\]

\[
\leq \int_{\theta-A-\varepsilon}^{\theta-A} f_N(n)dn \tag{43}
\]

\[
= \int_{\mu_N+\varepsilon}^{\mu_N} f_N(n)dn \tag{44}
\]

\[
= \Pr\{N \geq \mu_N + \varepsilon\} = \Pr\{N - \mu_N \geq \varepsilon\} \tag{45}
\]

\[
\leq \Pr\{|N - \mu_N| \geq \varepsilon\} \tag{46}
\]

\[
\leq \frac{\sigma^2}{\varepsilon^2} \quad \text{by Chebyshev’s inequality} \tag{47}
\]

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

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

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

\[
\leq \int_{\theta-A+\varepsilon}^{\theta+A+\varepsilon} f_N(n)dn \tag{50}
\]

\[
= \int_{\mu_N-\varepsilon}^{\mu_N} f_N(n)dn \tag{51}
\]

\[
= \Pr\{N \leq \mu_N - \varepsilon\} = \Pr\{N - \mu_N \leq -\varepsilon\} \tag{52}
\]

\[
\leq \Pr\{|N - \mu_N| \geq \varepsilon\} \tag{53}
\]

\[
\leq \frac{\sigma^2}{\varepsilon^2} \quad \text{by Chebyshev’s inequality} \tag{54}
\]

\[
\to 0 \quad \text{as} \quad \sigma \to 0. \tag{55}
\]
Case 2: Alpha-stable noise

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

\[
\lim_{\gamma \to 0} \varphi(\omega) = \exp \{ i\alpha \omega \} \tag{56}
\]

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

\[
\lim_{\gamma \to 0} f_N(n) = \delta(n - a). \tag{57}
\]

Then

\[
\lambda = \int_{\theta - A}^{\theta + A} f_N(n)dn = \int_{\theta - A}^{\theta + A} \delta(n - a)dn \tag{58}
\]

\[
= \int_{\theta - A}^{\theta + A} \delta(n - a)dn = 0 \quad \text{because} \quad a \notin (\theta - A, \theta + A). \tag{59}
\]

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 alpha-stable noise.

Necessity. We show that the cross correlation \( C \) is max-
imum 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 $\epsilon = \frac{1}{2}(\theta - A - \mu_N) > 0$.

So $\theta + A - \epsilon = \theta + A - \epsilon + \mu_N - \mu_N = \mu_N + (\theta - A - \mu_N) + \epsilon = \mu_N - 2\epsilon - \epsilon = \mu_N + \epsilon$. Then

$$P_{Y|S}(0|0) = \int_{-\infty}^{\theta + A} f_N(n) dn$$

Lower bound (70)

$$\geq \int_{-\infty}^{\theta + A - \epsilon} f_N(n) dn$$

Upper bound (71)

$$= \int_{-\infty}^{\mu_N + \epsilon} f_N(n) dn$$

$$= 1 - \int_{\mu_N + \epsilon}^{\infty} f_N(n) dn$$

$$= 1 - \Pr\{N \geq \mu_N + \epsilon\}$$

$$= 1 - \Pr\{N - \mu_N \geq \epsilon\}$$

$$\geq 1 - \frac{\sigma_N^2}{\epsilon^2} \text{ by Chebyshev's inequality}$$

$$\to 1 \text{ as } \sigma_N \to 0.$$ 

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

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

So $\theta - A + \epsilon = \theta - A + \epsilon + \mu_N - \mu_N = \mu_N + (\theta - A - \mu_N) + \epsilon = \mu_N - 2\epsilon + \epsilon = \mu_N - \epsilon$. Then

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

$$\geq \int_{\theta - A + \epsilon}^{\theta + A} f_N(n) dn$$

$$= \int_{\mu_N - \epsilon}^{\infty} f_N(n) dn$$

$$= 1 - \int_{-\infty}^{\mu_N - \epsilon} f_N(n) dn$$

$$= 1 - \Pr\{N \leq \mu_N - \epsilon\}$$

$$= 1 - \Pr\{|N - \mu_N| \leq \epsilon\}$$

$$\geq 1 - \frac{\sigma_N^2}{\epsilon^2} \text{ by Chebyshev's inequality}$$

$$\to 1 \text{ as } \sigma_N \to 0.$$ 

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

**Case 2: Alpha-stable noise**

Again we have

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

Then

$$P_{Y|S}(0|0) = \int_{-\infty}^{\theta + A} f_N(n) dn$$

$$\rightarrow \int_{-\infty}^{\theta + A} \delta(n - a) dn = 1 \text{ as } \gamma \to 0.$$

Similarly:

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

$$\rightarrow \int_{\theta - A}^{\theta + A} \delta(n - a) dn = 1 \text{ as } \gamma \to 0.$$

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 $\gamma \to 0$). Q.E.D.

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

This section derives necessary and sufficient conditions for the local noise benefit $\frac{\partial^4}{\partial \sigma^4}$ > 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:

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

with mean $\mu_N$ and variance $\sigma_N^2$. Thus the cdf is

$$F_N(n) = F_{\tilde{N}}(\frac{n - \mu_N}{\sigma_N})$$

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

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

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

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

(87)
Let $x = (y - a)/\kappa$ and $\tau = \kappa t$. Then

$$f(y; a, \kappa, \alpha, \beta) = \int_{-\infty}^{\infty} \exp \left\{ -i\kappa x + a + i\alpha \right\} d\tau$$

$$-i\kappa|\alpha(1 - \beta \text{sgn}(\tau)) dt$$

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

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

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

$$f_{Y|S}(y|s) = \begin{cases} F_N(\frac{\theta - s - \mu_N}{\sigma_N}) & y = -1 \\ 1 - F_N(\frac{\theta - s - \mu_N}{\sigma_N}) & y = 1 \\ (1 - F_N(\frac{\theta - s + \mu_N}{\sigma_N}))N \delta(y + 1) + (1 - F_N(\frac{\theta - s - \mu_N}{\sigma_N}))N \delta(y - 1). & \end{cases}$$

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

$$C = \mu_S + 2pA F_N(\frac{\theta + A - \mu_N}{\sigma_N}) - 2(1 - p)A F_N(\frac{\theta - A - \mu_N}{\sigma_N}).$$

Suppose the cross correlation (93) is differentiable with respect to $\sigma_N$. 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 $\mu_N$ and variance $\sigma_N^2$ with pdf $f_N(n) = \frac{1}{\sigma_N} f_N(\frac{n - \mu_N}{\sigma_N})$.

**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):

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

$$\frac{p}{1 - p} > (\theta - A - \mu_N) f_N(\frac{\theta - A - \mu_N}{\sigma_N}) - (\theta + A + \mu_N) f_N(\frac{\theta + A + \mu_N}{\sigma_N}).$$

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

$$\frac{p}{1 - p} < (\theta - A + \mu_N) f_N(\frac{\theta - A + \mu_N}{\sigma_N}) - (\theta + A - \mu_N) f_N(\frac{\theta + A - \mu_N}{\sigma_N}).$$

**Proof:**

$$C = E[SY]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} syf_{SY}(s, y) ds dy$$

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

$$= \int_{-\infty}^{\infty} sf_S(s) \left[ \int_{-\infty}^{\infty} sf_S(s)F_N(\frac{\theta - s - \mu_N}{\sigma_N}) - 1 \right] ds$$

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

$$= \mu_S - 2\int_{-\infty}^{\infty} sf_S(s)F_N(\frac{\theta - s + \mu_N}{\sigma_N}) ds.$$

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

$$C = \mu_S - 2\int_{-\infty}^{\infty} s(\theta s(a + A) + (1 - \delta(s - A) + (1 - p)\delta(s - A)) F_N(\frac{\theta - s - \mu_N}{\sigma_N}) ds$$

$$= \mu_S + 2pA F_N(\frac{\theta + A - \mu_N}{\sigma_N})$$

$$- 2(1 - p)A F_N(\frac{\theta - A - \mu_N}{\sigma_N}).$$

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

$$(1 - p)(\theta - A - \mu_N)f_N(\frac{\theta - A - \mu_N}{\sigma_N})$$

$$> p(\theta + A - \mu_N)f_N(\frac{\theta + A - \mu_N}{\sigma_N}).$$

Suppose the cdf $F_N$ is absolutely continuous and thus differentiable [39]: $\frac{dF_N(z)}{dz} = f_N(z)$. Then the partial
derivative $\frac{\partial C}{\partial \sigma_N}$ has the form

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

(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 $\mu_N > \theta + A$ and the signal-noise-system parameters in (105) that satisfy

$$
\frac{p}{1-p} > \frac{(\theta - A - \mu_N)f_N\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N)f_N\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}.
$$

(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 $\mu_N < \theta - A$ and the signal-noise-system parameters in (105) that satisfy

$$
\frac{p}{1-p} < \frac{(\theta - A - \mu_N)f_N\left(\frac{\theta - A - \mu_N}{\sigma_N}\right)}{(\theta + A - \mu_N)f_N\left(\frac{\theta + A - \mu_N}{\sigma_N}\right)}.
$$

(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 $\frac{\partial C}{\partial \sigma_N}$:

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

(109)

$$
< 0.
$$

(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 $A > 0$.

Q.E.D.

The proof also shows that the noise benefit is a local maximum when equality replaces the inequalities in (94) or in (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 $f_N(n) = \frac{1}{\sigma_n}f_N\left(\frac{n-\mu_N}{\sigma_N}\right)$. Thus the conditional pdf of $Y$ given a signal value $s$ is

$$
f_{Y|S}(y|s) = \begin{cases} 
F_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = -1 \\
1 - F_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) & y = 1
\end{cases}
$$

(111)

$$
= F_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right)\delta(y+1) + (1 - F_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right))\delta(y-1)
$$

(112)

where $F_N$ 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 $f_N(n) = \frac{1}{\sigma_n}f_N\left(\frac{n - \mu_N}{\sigma_N}\right)$ has the form (102):

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

(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 $\mu_N$ and variance $\sigma_N^2$. Suppose the location-scale noise $N$ has mean $\mu_N$ and variance $\sigma_N^2$. Then the threshold system (1) has the local noise benefit $\frac{\partial C}{\partial \sigma_N} > 0$ if and only if

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

(114)

where

$$
\mu_{SN} = \int_{-\infty}^{\infty} s k_{SN} f_S(s) f_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds
$$

(115)

$$
r_{SN}^2 = \int_{-\infty}^{\infty} s^2 k_{SN} f_S(s) f_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds
$$

(116)

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

$$
k_{SN} = \int_{-\infty}^{\infty} f_S(s) f_N\left(\frac{\theta - s - \mu_N}{\sigma_N}\right) ds.
$$

(117)
Proof: Suppose $C$ is differentiable with respect to the noise standard deviation $\sigma_N$. Then

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

The partial derivative $\frac{\partial F_N}{\partial \sigma_N}$ has the form

$$\frac{\partial}{\partial \sigma_N} F_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right) = \frac{\partial}{\partial \sigma_N} \int_{-\infty}^{s + \mu_N} f_N(z) dz = -\frac{\theta - s - \mu_N}{\sigma_N^2} f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right). \quad (119)$$

Then

$$\frac{\partial C}{\partial \sigma_N} =
2 \int_{-\infty}^{\infty} s \left( \frac{\theta - \mu_N}{\sigma_N} \right) f_S(s) f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right) ds
= 2 \int_{-\infty}^{\infty} s \left( \frac{\theta - \mu_N}{\sigma_N^2} \right) f_S(s) f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right) ds
- 2 \int_{-\infty}^{\infty} \frac{s^2}{\sigma_N^2} f_S(s) f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right) ds. \quad (120)$$

Consider the product $f_S(s) f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right)$. Suppose the supports of both pdfs overlap. We can define $f_{SN}(\theta; s) = \frac{1}{k_{SN}} f_S(s) f_N \left( \frac{\theta - s - \mu_N}{\sigma_N} \right)$ as a pdf where $k_{SN}(\theta) > 0$ is the normalizer such that

$$\int_{-\infty}^{\infty} f_{SN}(\theta; s) ds = 1.$$ 

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

$$\frac{\partial C}{\partial \sigma_N} =
2 \int_{-\infty}^{\infty} \frac{s (\theta - \mu_N)}{\sigma_N^2} f_{SN}(\theta; s) ds
- 2 \int_{-\infty}^{\infty} \frac{s^2}{\sigma_N^2} f_{SN}(\theta; s) ds \quad (121)$$

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

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

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

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 $\mu_S$ and variance $\sigma_S^2$: $S \sim N(\mu_S, \sigma_S^2)$ and the noise $N$ is Gaussian with mean $\mu_N$ and variance $\sigma_N^2$: $N \sim N(\mu_N, \sigma_N^2)$. 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

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

and $\mu_N \notin N(\mu_N, \sigma_N^2)$. 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

$$a_2 = -\frac{1}{2} \mu_S \left( \frac{\sigma_S^2}{\sigma_S^2} - 1 \right)$$

$$+ \frac{1}{2} \sqrt{\mu_S^2 (\sigma_S^2)^2 + 4(\sigma_S^2 + \sigma_S^2 + \mu_S \sigma_S^2) \left( \frac{\sigma_S^2}{\sigma_S^2} \right)} \quad (129)$$

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

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

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

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

$$\tau_{SN}^2 = \frac{\sigma_S^2 \mu_S^2}{\sigma_S^2 + \sigma_N^2} \quad (132)$$
forbidden intervals for all signal and noise pdfs.

The normalizer is

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

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

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

This holds in turn if and only if

$$\sigma_S^2 \sigma_N^2 (\sigma_S^2 + \sigma_N^2) + \mu_S^2 \sigma_N^2 + 2(\theta - \mu_N) \mu_S \sigma_N \sigma_S^2 + (\theta - \mu_N)^2 \sigma_S^4 + (\theta - \mu_N)^2 \mu_S \sigma_N^2.$$  

This holds if and only if

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

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

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

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

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

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

This is equivalent to

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

or  

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

Q.E.D.

Note that the correlation-based forbidden interval changes as the noise variance $\sigma_N^2$ changes for the given signal’s mean $\mu_S$ and variance $\sigma_S^2$. This also applies to forbidden intervals for all signal and noise pdfs.

### VII. Noise Benefits for Arrays of Threshold Detectors

This section derives correlation-based noise benefits for an array of $Q$ threshold detectors. The i.i.d. input sequence $S_i$ has arbitrary pdf with finite mean and variance and the i.i.d. 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 input-output relationship

$$Y_i = \text{sgn}(S_i + N_i - \theta_i)$$  

for $i = 1, \ldots, Q$. 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 [31] and the array of signal quantizers used for maximum-likelihood detection and Neyman-Pearson detection in [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 $\times$ 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 $\mu_{SN}$ and $\sigma_{SN}^2$ from the histogram and the noise pdf. Then we solve (148) to obtain the forbidden bands in Figures 3b and 3c. Figures 3b and 3c 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-d random vectors $S = [S_1 \cdots S_Q]^t$ and $Y = [Y_1 \cdots Y_Q]^t$ as a performance measure of the array:

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

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

$$\mu_{\hat{C}} = E[\hat{C}] = C$$

$$= \mu_S - 2 \int_{-\infty}^{\infty} f_S(s)F_N(\frac{\theta - s - \mu_N}{\sigma_N})ds$$

and variance

$$\sigma_{\hat{C}}^2 = \frac{1}{Q} \sigma_S^2 + 4 \mu_S \int_{-\infty}^{\infty} f_S(s)F_N(\frac{\theta - s - \mu_N}{\sigma_N})ds$$

$$- 4 \left( \int_{-\infty}^{\infty} f_S(s)F_N(\frac{\theta - s - \mu_N}{\sigma_N})ds \right)^2.$$
Fig. 3. 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 be in the interval [-1,1] and use these values as input signals to the array of threshold detectors. (b)-(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. Dash 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.
So on average we have a local noise benefit if \( \frac{\partial \mu_C}{\partial \sigma_N} = \frac{\partial C}{\partial \sigma} > 0 \). Thus the condition for an average local noise benefit for the sample cross correlation \( \bar{C} \) 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 \( \mu_S \) and finite variance \( \sigma_S^2 \). Suppose the noise \( N \) also has an arbitrary pdf with finite mean \( \mu_N \) and finite variance \( \sigma_N^2 \). Then the Q-array of threshold systems (1) has an average local noise benefit \( \frac{\partial \mu_C}{\partial \sigma_N} > 0 \) if and only if

\[
\frac{r_{SN}^2}{\mu_{SN}^2} < (\theta - \mu_N)\mu_{SN}
\]  

(148)

where \( \mu_{SN} \) and \( r_{SN}^2 \) have the same form as in (115)-(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, \( r_{SN}^2 < (\theta - \mu_N)\mu_{SN} \) 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 (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.

**REFERENCES**
