Applications of Forbidden Interval Theorems in Stochastic Resonance

Bart Kosko, Ian Lee, Sanya Mitaim, Ashok Patel and Mark M. Wilde

Abstract Forbidden interval theorems state whether a stochastic-resonance noise benefit occurs based on whether the average noise value falls outside or inside an interval of parameter values. Such theorems act as a type of screening device for mutual-information noise benefits in the detection of subthreshold signals. Their proof structure reduces the search for a noise benefit to the often simple task of showing that a zero limit exists. This chapter presents the basic forbidden interval theorem for threshold neurons and four applications of increasing complexity. The first application shows that small amounts of electrical noise can help a carbon nanotube detect faint electrical signals. The second application extends the basic forbidden interval theorem to quantum communication through the judicious use of squeezed light. The third application extends the theorems to noise benefits in standard models of spiking retinas. The fourth application extends the noise benefits in retinal and other neuron models to Levy noise that generalizes Brownian motion and allows for jump and impulsive noise processes.

1 Forbidden Interval Theorems for Stochastic Resonance

Stochastic resonance (SR) occurs in a nonlinear system when noise benefits the system [3, 17, 33]. The noise benefit can take the form of an increase in mutual information or a signal-to-noise ratio or correlation or a decrease in an error measure. But when will such a noise benefit occur?

Forbidden interval theorems answer that SR question for several nonlinear systems. The theorems act as a type of SR *screening device* because they can give sufficient or necessary conditions for an SR noise benefit. We here restrict noise benefits

B. Kosko (✉)
Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA, e-mail: kosko@usc.edu

to improvements in the system's input-output Shannon mutual information given an increase in the variance or dispersion of the exogenous noise. But forbidden interval theorems do not show how to find such a noise benefit. Nor do they indicate the magnitude of such a noise benefit if it occurs. They merely indicate whether such a noise benefit exists in theory for a given combination of parameters.

Stochastic learning algorithms can often find the actual noise benefit that a forbidden interval theorem predicts. These adaptive algorithms can find the local noise benefits if we take enough samples from the process and if we apply a sufficiently robust SR learning algorithm of the form $\sigma_{k+1} = \sigma_k + \mu_k \phi\left(\frac{\partial I}{\partial \sigma}\right)$ where I is mutual information or some other performance measure, σ is the noise intensity, and ϕ is a statistical "robustifier" such as signum or other impulse suppressor [18, 23, 24]. Both the learning algorithms and the forbidden interval theorems apply to many other system performance measures other than mutual information.

The simplest forbidden interval theorem is the strongest because it gives both necessary and sufficient conditions for an SR noise benefit while it requires only a threshold nonlinearity. It applies to a threshold neuron or other threshold system with threshold θ for subthreshold bipolar signal amplitudes A and $-A$: $-A < A < \theta$. The additive noise n has either a finite mean $E[n]$ or comparable location parameter a in the case of infinite-variance stable noise. Then the simplest forbidden interval theorem fully characterizes the presence or absence of a noise benefit because it gives both necessary and sufficient conditions for this SR effect: SR occurs if and only if $E[n] \notin (\theta - A, \theta + A)$. So a noise benefit occurs just in case the mean
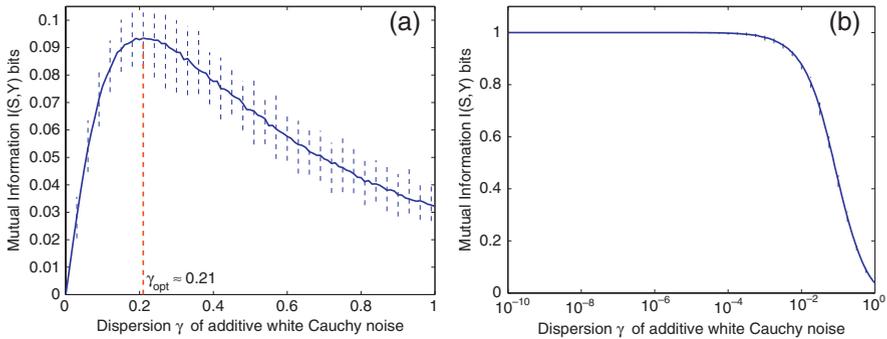

Fig. 1 Forbidden interval effects. The graphs show the smoothed input-output mutual information of a threshold system (1) as a function of the dispersion of additive white alpha-stable noise n_t with $\alpha = 1$ (infinite-variance Cauchy noise). The system has threshold $\theta = 0.5$. The bipolar input Bernoulli signal s_t has amplitude $A = 0.4$ with success probability $p_A = \frac{1}{2}$. Each trial produced 10,000 input-output samples $\{s_t, y_t\}$ that estimated the probability densities to obtain the mutual information. The vertical dashed lines show the absolute deviation between the smallest and largest outliers in each sample average of 100 outcomes. (**a**): Stochastic resonance (SR) when the Cauchy noise n_t has location $a = 0$ and thus a lies outside the forbidden interval: $a = 0 \notin (\theta - A, \theta + A) = (0.1, 0.9)$. The system has a nonzero noise optimum at $\gamma_{opt} \approx 0.21$ and thus shows the SR effect. (**b**): No SR when the Cauchy noise n_t has location $a = 0.2$ that lies in the forbidden interval: $a = 0.2 \in (\theta - A, \theta + A) = (0.1, 0.9)$. The system is optimal when $\gamma \to 0$ and thus *does not* show the SR effect: The mutual information $I(S, Y)$ is maximum because it equals the input entropy $H(S) = 1$

or location-scale noise does not fall in an interval that depends on the threshold and signal amplitudes (see Fig. 1). A reviewer of [20] referred to this interval as a "forbidden interval." We liked the colorful term and have used it ever since. The theorems below show just how complicated these interval conditions can become due to the complexity of the system's nonlinear dynamics and the nature of the noise or diffusion processes.

2 Proof Strategy: *What Goes Down Must Go Up*

This section presents sufficient and necessary forbidden interval theorems for SR in threshold systems. We use the discrete-time threshold system [4, 10, 16]

$$y_t = \text{sgn}(s_t + n_t - \theta) \tag{1}$$

where $\theta > 0$ is the system's threshold, y_t is the system's output, s_t is the bipolar input Bernoulli signal with amplitude $A > 0$ and with success probability $p_A = \frac{1}{2}$, n_t is the additive white noise with probability density $p(n)$, and the signum function is

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases} . \tag{2}$$

Note that the theorems below hold for any two-symbol alphabet set $\mathscr{Y}$ for the output y_t. The Shannon mutual information $I(S,Y)$ has the form [8]

$$I(S,Y) = H(Y) - H(Y|S) = \sum_{s \in \mathscr{S}} \sum_{y \in \mathscr{Y}} P_{SY}(s,y) \log \frac{P_{SY}(s,y)}{P_S(s)P_Y(y)} . \tag{3}$$

The idea behind forbidden interval theorems is that we can cast the proof in terms of showing that a limit must go to zero as the noise intensity goes to zero.

Theorem 1. *Suppose that the threshold system* (1) *has noise probability density function p(n) and that the input signal is subthreshold $A < \theta$. Suppose that there is some statistical dependence between input random variable S and output random variable Y (so that $I(S,Y) > 0$). Suppose that the noise mean $E[n]$ does not lie in the signal-threshold interval $(\theta - A, \theta + A)$ if p(n) has finite variance. Then the threshold system* (1) *exhibits the nonmonotone SR effect in the sense that $I(S,Y) \to 0$ as $\sigma \to 0$.*

Proof. Assume $0 < P_S(s) < 1$ to avoid triviality when $P_S(s) = 0$ or 1. We show that S and Y are asymptotically independent: $I(\sigma) \to 0$ as $\sigma \to 0$. Recall that $I(S,Y) = 0$ if and only if S and Y are statistically independent [8]. 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 $\sigma \to 0$ for all signal symbols $s \in \mathscr{S}$ and $y \in \mathscr{Y}$. The two-symbol alphabet set $\mathscr{S} = \{0, 1\}$ gives

$$P_Y(y) = \sum_{s \in \mathscr{S}} P_{Y|S}(y|s)P_S(s) = (P_{Y|S}(y|0) - P_{Y|S}(y|1))P_S(0) + P_{Y|S}(y|1) . \tag{4}$$

So we need to show only that $P_{Y|S}(y|0) - P_{Y|S}(y|1) = 0$ as $\sigma \to 0$. This condition implies that $P_Y(y) = P_{Y|S}(y|1)$ and $P_Y(y) = P_{Y|S}(y|0)$. We assume for simplicity that

the noise density $p(n)$ is integrable. The argument below still holds if $p(n)$ is discrete and if we replace integrals with appropriate sums.

Consider $y =$ "0." Then

$$P_{Y|S}(0|0) - P_{Y|S}(0|1) = \int_{-\infty}^{\theta+A} p(n)dn - \int_{-\infty}^{\theta-A} p(n)dn = \int_{\theta-A}^{\theta+A} p(n)dn \quad . \quad (5)$$

Similarly for $y =$ "1":

$$P_{Y|S}(1|0) - P_{Y|S}(1|1) = \int_{\theta+A}^{\infty} p(n)dn - \int_{\theta-A}^{\infty} p(n)dn = -\int_{\theta-A}^{\theta+A} p(n)dn \quad . \quad (6)$$

The result follows if we can show that

$$\int_{\theta-A}^{\theta+A} p(n)dn \to 0 \qquad \text{as } \sigma \to 0 \quad . \tag{7}$$

Denote the mean of the noise by $\mu = E[n]$ and the variance by $\sigma^2 = E[(n-\mu)^2]$. Then $\mu \notin (\theta - A, \theta + A)$ by hypothesis.

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

$$P_{Y|S}(0|0) - P_{Y|S}(0|1) = \int_{\theta-A}^{\theta+A} p(n)dn \leq \int_{\theta-A}^{\infty} p(n)dn \leq \int_{\theta-A-\varepsilon}^{\infty} p(n)dn \tag{8}$$

$$= \int_{\mu+\varepsilon}^{\infty} p(n)dn = P(n \geq \mu + \varepsilon) = P(n - \mu \geq \varepsilon) \leq P(|n-\mu| \geq \varepsilon) \tag{9}$$

$$\leq \frac{\sigma^2}{\varepsilon^2} \qquad \text{by Chebyshev's inequality} \tag{10}$$

$$\to 0 \qquad \text{as } \sigma \to 0 \quad . \tag{11}$$

A symmetric argument shows that for $\mu > \theta + A$

$$P_{Y|S}(0|0) - P_{Y|S}(0|1) \leq \frac{\sigma^2}{\varepsilon^2} \to 0 \quad \text{as } \sigma \to 0 \quad . \quad \text{QED} \tag{12}$$

The next forbidden interval shows that a structurally similar interval condition holds for *all* types of infinite-variance stable noise [19]. Stable models apply to a quantum alpha-stable noise source. Stable models apply to diverse physical phenomena that include impulsive interrupts in phone lines, underwater acoustics, low-frequency atmospheric signals, and gravitational fluctuations [25]. Symmetric alpha-stable noise [25, 30] results from an impulsive noise source and describes a family of thick-tailed bell-curve densities per (13) below. The parameter α (which differs from the coherent state α in Sect. 4) lies in $(0,2]$ and governs the thickness of the distribution's tail: $\alpha = 1$ corresponds to the thick-tailed Cauchy random variable and $\alpha = 2$ corresponds to the familiar thin-tailed Gaussian random variable. The bell curve's tail thickness increases as α decreases. The generalized central limit theorem states that all and only normalized stable random variables converge in distribution to a stable random variable [2].

Theorem 2. *Suppose $I(S,Y) > 0$ and the threshold system* (1) *uses alpha-stable noise with location parameter $a \notin (\theta - A, \theta + A)$. Then the system exhibits the non-monotone SR effect if the input signal is subthreshold.*

Proof. Again the result follows if $\int_{\theta-A}^{\theta+A} p(n)dn \to 0$ as $\gamma \to 0$. The characteristic function $\varphi(\omega)$ of alpha-stable density $p(n)$ has the exponential form

$$\varphi(\omega) = \begin{cases} \exp\left\{ia\omega - \gamma|\omega|^\alpha \left(1 + i\beta \operatorname{sign}(\omega)\tan\frac{\alpha\pi}{2}\right)\right\} & \text{for } \alpha \neq 1 \\ \exp\left\{ia\omega - \gamma|\omega|(1 - i\frac{2}{\pi}\beta \ln|\omega|\operatorname{sign}(\omega))\right\} & \text{for } \alpha = 1 \end{cases}. \quad (13)$$

This reduces to a simple complex exponential in the zero-dispersion limit: $\lim_{\gamma\to 0} \varphi(\omega)$ $= \exp\{ia\omega\}$ for *all* α's, skewness β's, and location a's. So Fourier transformation gives the corresponding density function in the limiting case ($\gamma \to 0$) as a translated delta function $\lim_{\gamma\to 0} p(n) = \delta(n-a)$. Then

$$P_{Y|S}(0|0) - P_{Y|S}(0|1) = \int_{\theta-A}^{\theta+A} p(n)dn = \int_{\theta-A}^{\theta+A} \delta(n-a)dn = 0 \quad (14)$$

because $a \notin (\theta - A, \theta + A)$. QED

Similar proofs give converses in Theorems 3 and 4 [20].

Theorem 3. *Suppose that the threshold system* (1) *has noise probability density function p(n) and that the input signal S is subthreshold. Suppose that the noise mean E[n] lies in the signal-threshold interval* $(\theta - A, \theta + A)$ *if p(n) has finite variance. Then the threshold system* (1) *does not exhibit the nonmonotone SR effect in the sense that I(S,Y) achieves its maximum when* $\sigma \to 0$: $I(S,Y) = H(Y) = H(S)$ *when* $\sigma \to 0$.

Theorem 4. *Suppose that the threshold system* (1) *has subthreshold input signal and use alpha-stable noise with location parameter* $a \in (\theta - A, \theta + A)$. *Then the threshold system* (1) *does not exhibit the nonmonotone SR effect:* $I(S,Y)$ *achieves its maximum when* $\gamma \to 0$: $I(S,Y) = H(Y) = H(S)$ *when* $\gamma \to 0$.

3 SR in a Carbon Nanotube Signal Detector

A carbon nanotube signal detector benefits from small amounts of added electrical noise in accord with the forbidden interval Theorems 1 and 2 above [21]. Our experiments used a carbon nanotube field-effect transistor to detect noisy subthreshold electrical signals. Two new SR hypothesis tests in [21] also confirmed the SR effect in the nanotube transistor. Three measures of detector performance showed the SR effect: Shannon's mutual information, the normalized correlation measure, and an inverted bit error rate compared the input and output discrete-time random sequences. The nanotube detector had a threshold-like input-output characteristic in its gate effect (see Fig. 2). It produced little current for subthreshold digital input voltages that fed the transistor's gate. Three types of synchronized white noise corrupted the subthreshold Bernoulli sequences that fed the detector. Gaussian, uniform, and impulsive Cauchy noise combined with the random input voltage sequences to help the detector produce random output current sequences.
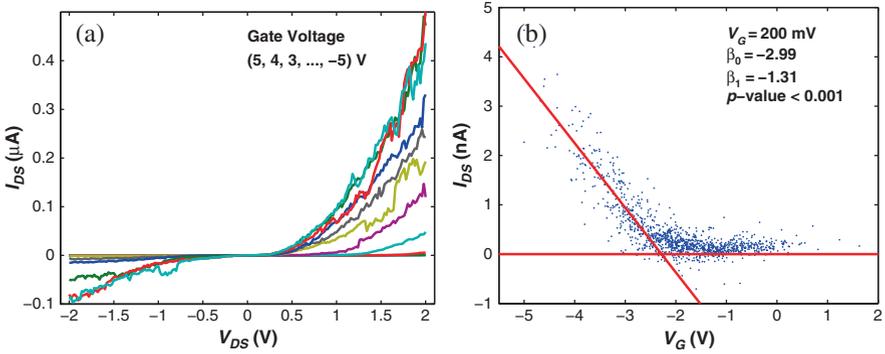
Fig. 2 Less than ideal carbon nanotube transistor gate effect. (**a**) plots the drain-source current versus the drain-source voltage for different gate voltages. (**b**) plots the experimental input-output pairs and shows the transistor's current-voltage I-V_G characteristics in response to noisy input signals. Linear regression fit the data to the transistor equation, extrapolated the nonlinearity, and estimated the threshold voltage $V_T \approx -2.3$ V

The experiments observed the SR effect by measuring how well an output sequence matched its input sequence. Shannon's mutual information used histograms to estimate the probability densities and computed the entropies. The correlation measure was a scalar inner product of the input and output sequences. The inverted bit error rate computed how often the bits matched between the input and output sequences. The observed nanotube SR effect was robust: it persisted even when infinite-variance Cauchy noise corrupted the signal stream.

Simulations and lab experiments both showed that the above forbidden interval theorems apply to threshold detectors. The simulations modeled transistors in general. They used a threshold-like ramp function that often models an ideal transistor's current-to-gate-voltage $(I$-$V_G)$ characteristics: $Y = G(S - V_T)$ where Y is the output current, S is the input voltage, V_T is the threshold voltage, and G is a nonzero gain for suprathreshold inputs and zero otherwise. The negative threshold voltage V_T gives the forbidden voltage intervals of the form $(V_T + x, V_T - x)$ for threshold voltage $V_T = -2.3$ for $x = -2, -1.8, -1.6$, and -1.4 volts. So $E[n] = 0$ fell outside all intervals.

The experiments tested single-walled carbon nanotube transistors. The prototype transistors had non-ideal characteristics such as the I-V curves in Fig. 2. The laboratory data generated SR-curves Fig. 3(b), (c), (d) that qualitatively agreed with those generated in simulation Fig. 3(a).

Experiments confirmed the SR prediction: noise helped a pristine (undoped) single-walled carbon nanotube transistor [32] detect subthreshold signals. The experiments applied different Bernoulli input sequences that used different combinations of subthreshold gate voltages as their ON/OFF symbols. Synchronized Gaussian, uniform, and infinite-variance Cauchy noise added to the input sequences and helped the nanotube transistor detect the subthreshold input. The performance
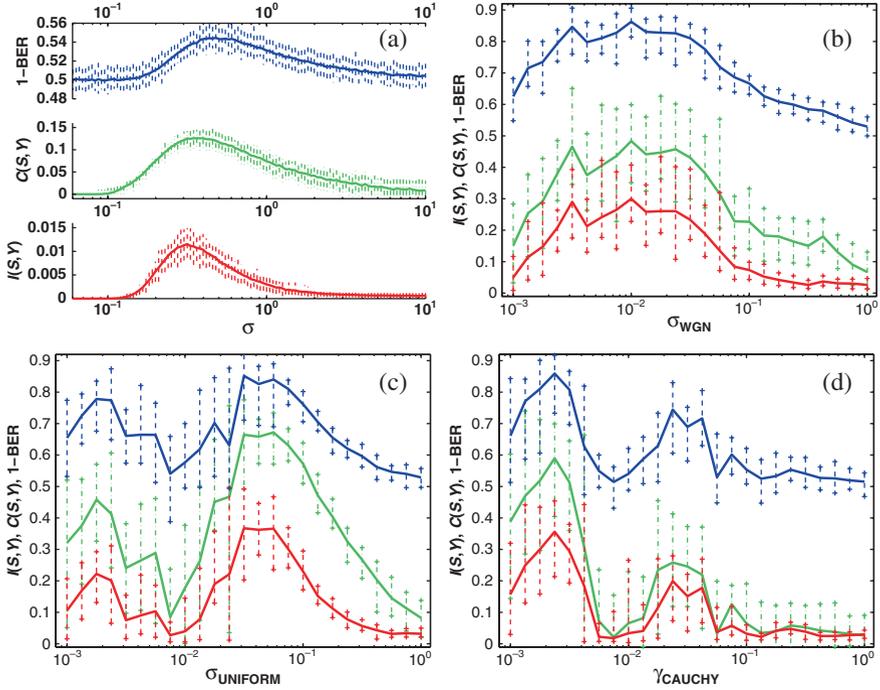
Fig. 3 Noise-enhanced detector performance. (**a**) A simulated threshold detector exhibits the signature SR modes with additive white Gaussian noise. The experiments found the SR effect for three different measures, for three different types of noise, and for multiple combinations of voltage-symbols. (**b**) shows the SR effect for Gaussian noise and binary symbols $(-1.6, -1.4)$ V, (**c**) for uniform noise and binary symbols $(-2.0, -1.8)$ V, and (**d**) for impulsive Cauchy noise and binary symbols $(-2.0, -1.8)$ V. The subthreshold symbols were more positive than the threshold voltage ($V_T = -2.3$ V) consistent with p-type semiconductors. Each pair of voltage symbols had a 0.2V separation because sensitivity analysis showed that the separation gave complete SR modes within the range of noise levels. A linear regression of the transistor's gate effect estimated the threshold voltage and aided the selection of the subthreshold ON/OFF symbols. (**b**) The SR mode of the mutual-information curve is six times the value at minimal noise. The SR mode of the correlation-measure curve is three times the value at minimal noise. The SR mode of the inverted bit-error-rate curve shows a 40% improvement over the value at minimal noise. (**d**) Cauchy noise. We passed impulsive or infinite-variance white Cauchy noise through the nanotube detector to test whether it was robust to occasional large noise spikes. We chose the highly impulsive Cauchy noise for this task. Not all Cauchy experiments produced a measurable SR effect

measures were Shannon's mutual information $I(S, Y)$ in (3), an input-output correlation measure $C(S, Y)$, and an inverted bit error rate $1 - BER$.

Both the experimental and simulated nanotube detectors had correlated SR curves. The correlation coefficient R measured the strength of the correlation. The simulated SR curves in Fig. 3(d) had correlation coefficients of $R = 0.9367$ for $I(S, Y)$ and $C(S, Y)$, $R = 0.8265$ for $I(S, Y)$ and $1 - BER$, and $R = 0.9541$ for $C(S, Y)$ and $1 - BER$. The experimental SR curves in Fig. 3(a) had correlation coefficients

of $R = 0.9830$ for $I(S,Y)$ and $C(S,Y)$, $R = 0.9774$ for $I(S,Y)$ and $1 - BER$, and $R = 0.9877$ for $C(S,Y)$ and $1 - BER$. The correlations were statistically significant for p-value < 0.001.

We observed the nanotube SR effect in Fig. 3(b) as one of four such successful combinations of input binary values with the parameter choices ON = -1.6 V and OFF = -1.4 V. Figures 3(c) and (d) each represents a selection from successful combinations. [21] contains additional details of the experimental setup and parameters. This SR effect occurred despite the nanotube instabilities that caused fluctuations in the stochastic I-V_G curve in Fig. 2(b). [21] discusses the generation of the stochastic current-voltage curve in the presence of hysteresis.

The nanotube experiments produced the SR effect for Shannon's mutual information [31], an input-output correlation measure [6, 7], and an inverted bit error rate that measured how well the output sequences matched the input Bernoulli sequences. The input signal S was a sequence of random binary voltages that produced a random output sequence in the form of a transistor current. Histograms of the sequences estimated the probability density functions that computed the entropies. The correlation measure normalized the zero-lag value ($l = 0$) of the cross-correlation sequence

$$r_{SY}(l) = \sum_{k=1}^{N} s(k)y(k-l) \tag{15}$$

of the two sequences with subtracted means. These two measures did not assume that the nanotube detector had a special structure and did not impose a threshold scheme on the experiment. But the inverted bit error rate $1 - BER$ decided whether each output y_i was a "0" or "1" by applying a threshold scheme: a Baye's discriminant function in a two-class minimum-distance classifier [9] that used complete knowledge of the input. The rate $1 - BER$ measured how often the input and output bits agreed: $1 - BER = 1 - N_{error}/N = N_{correct}/N$ where N_{error} counted the number of bits that differed between the length-N input and output sequences and $N_{correct}$ counted the number of bits that agreed.

The experimental observation of nanotube SR used the nonlinear field effect of carbon nanotubes. A semiconductor single-walled carbon nanotube (SWNT) can change its conductivity in response to an external electric field in a gate effect [32]. The SR experiments used a chemical-vapor-deposition (CVD) grown SWNT [15]. The reference contains the details of the fabrication.

The semiconductor SWNT forms a Schottky diode at the interface with metal so that a metal-nanotube-metal contact forms a field-effect transistor (FET) with an adjacent gate electrode [35]. The typical current-voltage I-V_G characteristics

$$I = \begin{cases} G(V_G - V_T) & \text{for } V_G < V_T \\ 0 & \text{else} \end{cases} \tag{16}$$

indicate that the pristine semiconductor nanotubes act as hole-doped semiconductors at room temperatures and that the nanotube devices are p-type FETs [32]. The transconductance G is negative and the gate voltage $V_G < V_T$ is suprathreshold for p-type FETs. Equation (16) modeled the ideal p-type FET in the simulated experiments.

We provide a terse summary of the laboratory setup— [21] contains the details. Each of the nanotube experiments applied 25 sampled noise levels that ranged from 0.001 to 1 standard deviation σ (dispersion γ for infinite-variance Cauchy) linearly in logarithmic scale. The noisy input S was a *synchronized* Bernoulli sequence $s_i = b_i + n_i$ of the sum of random subthreshold binary values b_i and additive white noise n_i of three types. So there was no *timing* noise in the pulse train as in the FHN neuron model [28, 29]. Synchronization allows the nanotube systems to implement a variety of algorithms from signal processing and communications.

The experiments updated the noisy input symbols s_i about once every 10 ms. A 200 mV drain-source voltage biased the nanotube at room temperature in vacuum. The experiments measured and averaged ten samples of the detector output at 100 kilosamples/s near the end of each symbol interval to estimate the output symbols y_i. This allows the transistor output to settle. A gate voltage is subthreshold if it is more positive than a p-type FET's threshold voltage and produces picoamp current in an OFF state. The experiment tested whether noise could enhance subthreshold signals to produce measurable currents.

The detector consisted of a single-walled semiconductor carbon nanotube bridging two electrodes [15]. The CVD technique combined with e-beam lithography to grow a single-walled nanotube that was 3–5 μm long and less than 2 nm in diameter between two electrodes. The gap between the electrodes was approximately 3 μm wide but the single-walled nanotube was not straight as it spanned the gap. Atomic force microscopy examined the detector and showed that the nanotube had a diameter $d < 2$ nm that was consistent with a single-walled nanotube.

A PC-based National Instruments PCI-MI0-16XE-10 data acquisition (DAQ) board converted the noise-corrupted signal S from digital to analog (DA) and converted the conditioned noisy output Y from analog to digital (AD). The AD-DA conversion has a 16-bit resolution and a 10 μs rise time. A DL 1211 current-voltage preamplifier conditioned the detector output current before data acquisition sampled it as a voltage. The amplifier converts a small current (10 nA) into a large voltage (1 V) with the 10^{-8} A/V gain setting. The analog voltage has a maximal time delay of 0.1 ms with the 0.1 ms rise-time setting.

Two new hypothesis tests verified that the SR curves were nonmonotonic and confirmed the nanotube SR effect. A chi-square test and a Kolmogorov-Smirnov test both rejected the similarity between a monotonically decreasing β-probability density function and each of the three SR curves with p-value < 0.001. The statistical tests were goodness-of-fit tests that treated the SR curves as candidate pdfs and compared them against the benchmark pdf. The parameters of the benchmark β-pdf were $\alpha = 0.5$ and $\theta = 5$ among many others that we tested [21].

4 Forbidden Interval Theorem for Quantum Communication

Quantum optics and "squeezed light" [22] extend the basic or classical forbidden interval theorems above for detecting subthreshold signals [34]. The new quantum forbidden interval theorem shows that a noise benefit occurs if and only if the

position quadrature of the noise mean or location does not lie in a given interval. The result holds for all types of finite-variance noise and for all infinite-variance stable noise.

The quantum forbidden-interval theorem applies to the quantum-optical communication system of Fig. 4. This system assumes weak or subthreshold signals and further assumes that noise corrupts these signals. Alice sends a squeezed displaced vacuum as a binary signal to Bob. Then Bob decodes the binary message by position-quadrature homodyning and thresholding. Figure 5 shows the non-monotonic signature of the predicted SR noise benefit in the quantum-optical communication system.

We develop the quantum-optical model in the Heisenberg picture. Suppose Alice possesses a vacuum mode. Let $\hat{x}$ denote the position-quadrature operator of this vacuum state. This operator collapses to a zero-mean $1/2$-variance Gaussian random variable X if Alice measures her state. Suppose that Alice does not measure it. Suppose instead that she sends her mode through a squeezer. Suppose further that she can control the strength of squeezing with a squeezing parameter r. Her position-quadrature operator $\hat{x}$ evolves under the squeezer to become $\hat{x}e^{-r}$. She encodes a random message bit $S \in \{0,1\}$ by displacing her state by $\alpha \in \mathbb{C}$ if $S = 1$ or by $-\alpha$ if $S = 0$. Her operator evolves under the displacement to become $\hat{x}e^{-r} + (-1)^{S+1}\alpha_x$ where $\alpha_x = Re\{\alpha\}$. She sends her state to Bob over an additive noisy bosonic channel [12]. A noisy bosonic channel affects any annihilation operator $\hat{a}_{in}$ by $\hat{a}_{out} = \hat{a}_{in} + v$. Annihilation operator $\hat{a}_{out}$ represents the output mode. Complex random variable v represents the noisy effects of the bosonic channel. Random variable v need not be Gaussian for the SR effect to occur—it can have finite variance or possess an alpha-stable distribution [25].

Bob receives the state $\hat{x}e^{-r} + (-1)^{S+1}\alpha_x + v_x$ from the noisy channel where $v_x = Re\{v\}$. Bob performs a position-quadrature homodyne detection so that the state collapses to the random variable $(-1)^{S+1}\alpha_x + N$ where $N = Xe^{-r} + v_x$ sums both noisy random variables. Bob thresholds the result of this homodyne detection with threshold θ to retrieve a bit Y. This bit Y should be the message bit S that Alice first sent.

Random variables Xe^{-r} and v_x are independent because random variable Xe^{-r} originates from vacuum fluctuations and because v_x represents Bob's loss of knowledge due to the state's propagation through a noisy quantum channel. The density
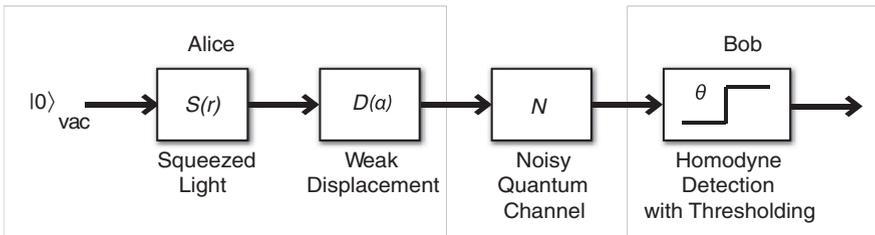

Fig. 4 Noisy quantum-optical communication channel for stochastic resonance
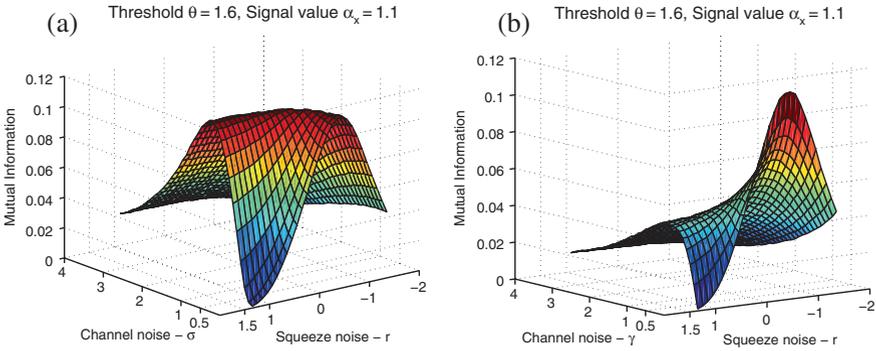
Fig. 5 Quantum SR noise benefits for (**a**) Gaussian noise and (**b**) Cauchy noise

$p_N(n)$ of random variable N is $p_N(n) = (p_{Xe^{-r}} * p_{v_x})(n)$ where $p_{Xe^{-r}}(n)$ is the density of a zero-mean $(e^{-2r}/2)$-variance Gaussian random variable, $p_{v_x}(n)$ is the density of v_x, and $*$ denotes convolution.

The quantum forbidden interval theorem below gives necessary and sufficient conditions for the nonmonotone SR effect to occur under both finite-variance and infinite-variance alpha-stable noise. We state the parameters for the finite-variance case without parentheses and the parameters for the infinite-variance case with parentheses.

Theorem 5. *Suppose the channel noise's position quadrature has finite variance* σ_x^2 *(dispersion* γ_x*) and mean* μ_x *(location* a_x*). Suppose the input signal's position quadrature* α_x *is subthreshold:* $\alpha_x < \theta$*. Suppose there is some statistical dependence between input signal S and output signal Y so that the mutual information obeys* $I(S,Y) > 0$*. Then the quantum communication system exhibits the nonmonotone SR effect if and only if the position quadrature of the noise mean does not lie in the forbidden interval:* $\mu_x \notin (\theta - \alpha_x, \theta + \alpha_x)$ *(or* $a_x \notin (\theta - \alpha_x, \theta + \alpha_x)$ *in the stable case). The nonmonotone SR effect is that* $I(S,Y) \to 0$ *as* $\sigma_x \to 0$ *(or* $\gamma_x \to 0$*) and as squeezing parameter* $r \to \infty$*.*

Proof. The finite-variance proof for sufficiency and necessity follows the proof method in [19] and [20] respectively if we use $p_N(n)$ as the noise density. The infinite-variance proof for sufficiency and necessity follows the stable proof method in [19] and [20] respectively if we use $p_N(n)$ as the noise density and if v_x is an alpha-stable random variable. QED

Figure 5 shows simulation instances of Theorem 5. Figure 5 displays the full "inverted-U" curve for realistic squeezing values [36]. The theorem guarantees only that the nonmonotone SR effect occurs. It does not give the optimal combination of channel noise and squeezing or guarantee a large increase in mutual information.

One criticism is that the theorem is not realistic because it requires infinite squeezing and thus requires infinite energy to produce the SR effect. But the theorem guarantees that the SR effect occurs for some finite squeezing. The simulations in

Fig. 5 display the full joint σ^2 and r nonmonotone SR signature for experimentally plausible squeezing values and for realistic channel noise levels.

5 Forbidden Interval Theorem for Spiking Retinal Neurons

A different forbidden interval theorem guarantees an SR noise benefit for spiking retinal neurons [26]. Figure 6 below shows an SR noise benefit in a spiking retinal neuron. The neuron should emit more spikes when the brightness contrast level is low rather than high. The right amount of Gaussian noise helps the neuron discriminate between two levels of brightness contrast. The retinal neuron emits too few spikes if no noise corrupts the Bernoulli sequence of contrast levels. The neuron also emits too many spikes and emits many of them at the wrong time if too much noise corrupts the sequence.

The retina model of Fig. 6 is a noisy version of a common Wiener-type cascade model [5, 14]:

$$r(t) \ = \ r_0 h\left[\int_{-\infty}^{\infty} f(z)\{S(t-z)+n_1(t)\}dz+n_2(t)\right] \tag{17}$$

where S is the input stimulus defined below, r is the instantaneous Poisson spike rate that gives the exponential interspike-interval density function as $p(t) = r(t)$ $\exp[-\int_0^t r(\tau)d\tau]$, f is a band-pass linear filter function, and h is a memoryless monotone–nondecreasing function. Here n_1 denotes the combined stimulus and photoreceptor noise and n_2 denotes the combined ion-channel noise and the synaptic noise.

The input stimulus S is Michelson's visual contrast signal: $S = (L_c - L_s)/(L_c + L_s)$. L_c is the amount of light that falls on the center of the ganglion cell's receptive field. L_s is the light that falls on its surround region. The sigmoid-shaped memoryless function h approximates the spike threshold and saturation level. We define h as a piecewise-linear approximation of a sigmoidal nonlinearity [37]:

$$h(x) \ = \ \begin{cases} \theta_2 - \theta_1 & \text{if } x > \theta_2 \\ x - \theta_1 & \text{if } \theta_1 \leq x \leq \theta_2 \\ 0 & \text{if } x < \theta_1 \end{cases} \tag{18}$$

and so

$$r(w(t)) = \begin{cases} r_0(\theta_2 - \theta_1) & \text{if } w(t) > \theta_2 \\ r_0(w(t) - \theta_1) & \text{if } \theta_1 \leq w(t) \leq \theta_2 \\ 0 & \text{if } w(t) < \theta_1 \end{cases} \quad . \tag{19}$$

The subthreshold contrast signal $S(t) \in \{A, B\}$ is a random Bernoulli sequence with $P(S(t) = A) = p$ and $P(S(t) = B) = 1 - p$. The time duration of each signal value A and B in $S(t)$ is much larger than the time constant of the linear filter $f(t)$. We define $v(t)$ as the filtered output of the contrast signal $S(t)$ without noise $n_1(t)$ and such that

$$v(t)|_{S(t)=A} \ = \ v_1 \quad \text{and} \quad v(t)|_{S(t)=B} \ = \ v_2 \tag{20}$$
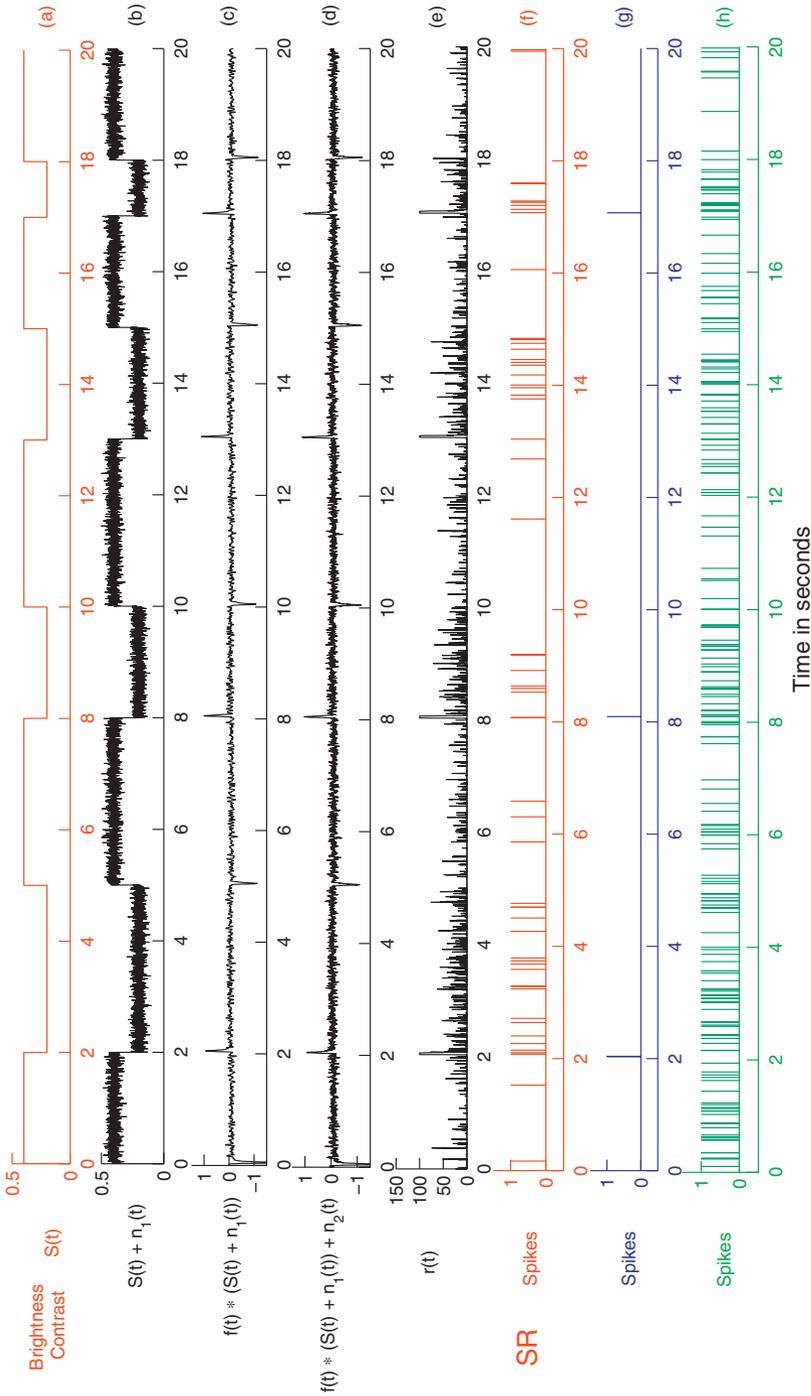
Fig. 6 Stochastic resonance in a spiking retinal neuron. The neuron should emit more spikes when the brightness contrast level is low rather than high. Noise improves the discrimination of subthreshold contrast stimuli in the retina model (17), (18), (19). (**a**) Bernoulli contrast signal S as a function of time t. (**b**) Contrast signal S plus Gaussian white noise n_1 with variance $\sigma_1 = 0.03$. (**c**) Signal in plot (**b**) filtered with f in (1). (**d**) Filtered noisy signal in (**c**) plus noise n_2 (synaptic and ion-channel noise) with variance $\sigma_2 = 0.06$. (**e**) Noisy spike rate $r(t)$. (**f**) SR effect: Output Poisson spikes that result from the noisy spike rate $r(t)$. (**g**) Output spikes in the absence of noise. (**h**) Output spikes in the presence of too much noise

in steady-state where $v_1 > v_2$ and $\max(v_1, v_2) < \theta_1 < \theta_2$. So the input signal $S(t)$ is subthreshold. We measure the average spike rate for each symbol only when the corresponding value of $v(t)$ is in steady-state. Theorem 6 below gives necessary and sufficient conditions for an SR noise effect in the retina neuron model (17), (18), (19) for either noise source n_1 or n_2. It states that standard spiking retinal models benefit from additive white noise if and only if a joint noise mean or location parameter does not fall in a forbidden interval of threshold-based values. Theorem 6 holds for all finite-variance noise and for all impulsive or infinite-variance stable noise [26].

Theorem 6. *Suppose that the noise sources n_1 and n_2 in the retina model* (17), (18), (19) *are white and have finite-variance (or finite-dispersion in the stable case) probability density functions $p_1(n)$ and $p_2(n)$ with corresponding variances (dispersions) σ_1^2 and σ_2^2 (γ_1 and γ_2). Suppose that the input signal S is subthreshold ($v_2 < v_1 < \theta_1 < \theta_2$) and that there is some statistical dependence between the input contrast random variable S and the output random variable R so that $I(S, R) > 0$. Then the retina model* (17), (18), (19) *exhibits the nonmonotone SR effect in the sense that $I(S, R) \to 0$ as σ_1^2 and σ_2^2 (or γ_1 and γ_2) decrease to zero if and only if the mean sum $E[n_1] \int f(\tau) d\tau + E[n_2]$ (or the location parameter sum in the stable case) does not lie in the interval $(\theta_1 - v_1, \theta_2 - v_2)$. The only-if part holds in the sense that the system performs better without noise than with it when the interval condition fails.*

6 Forbidden Interval Theorems for Levy Noise Diffusions

The most complex forbidden interval theorems apply to nonlinear stochastic differential equations with additive Levy diffusions. Levy processes generalize Brownian motions to allow for jumps and other impulsive behavior [1]. Levy processes include not only Brownian processes but also compound Poisson processes, infinite-variance α-stable processes, generalized inverse Gausssian processes, and generalized hyperbolic processes. A random process L_t is a Levy process if it has *independent increments* $L_t - L_s$, if it is *stationary* ($L_t - L_s$ has the same distribution as L_{t-s}), and if it is *continuous in probability* ($L_s \to L_t$ in probability if $s \to t$). A Levy process L_t has a drift component, a Brownian (Gaussian) component, and a jump component. Figure 7 shows sample paths from four different types of Levy processes.

The forbidden interval results of Theorems 7 and 8 below show that a broad class of additive white Levy noise [27] with finite second moments can benefit a wide range of noisy feedback continuous and sensory spiking neuron models of the general form

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \tag{21}$$
$$Y_t = g(X_t) \quad . \tag{22}$$

Here $c(X_{t-})$ is a bounded Levy diffusion term, dL_t is a white Levy noise with noise scale κ, and $b(X_{t-})$ is a Lipschitz continuous drift term that has the additive net
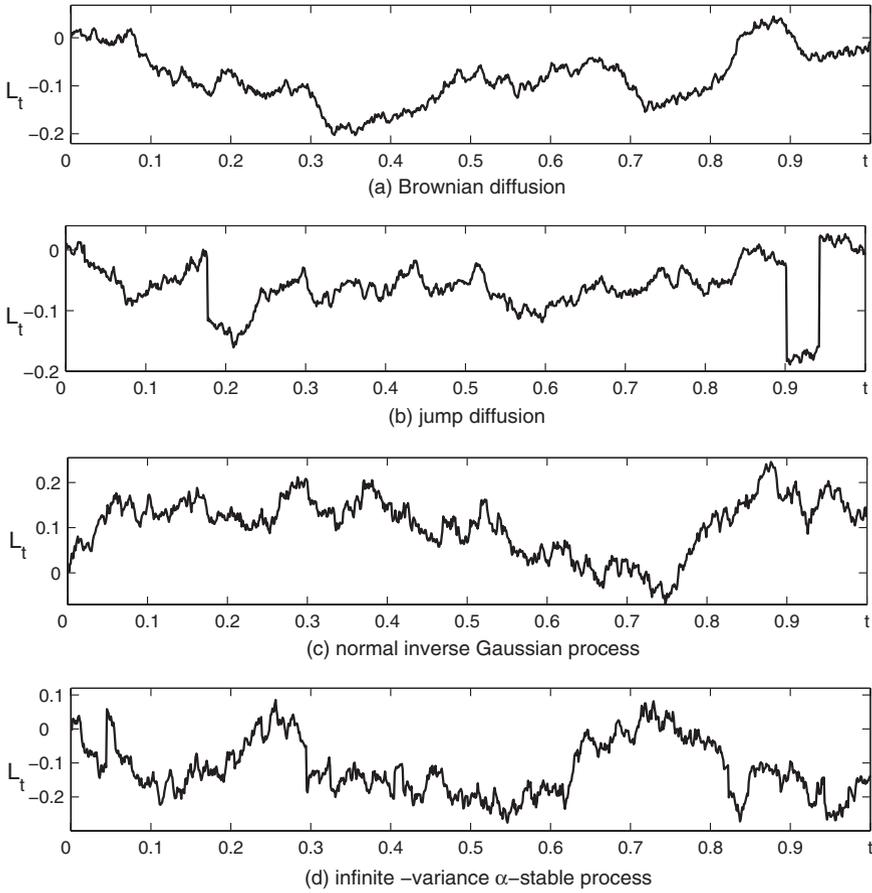
Fig. 7 Sample paths from one-dimensional Levy processes: (**a**) Brownian motion with drift $\mu = 0.1$ and variance $\sigma = 0.15$, (**b**) jump diffusion with $\mu = 0.1$, $\sigma = 0.225$, Poisson jump rate $\lambda = 3$, and uniformly distributed jump magnitudes in the interval $[-0.2, 0.2]$ (and so with Levy measure $\nu(dy) = (3/0.4)dy$ for $y \in [-0.2, 0.2]$ and zero else in [27]]), (**c**) normal inverse Gaussian (NIG) process with parameters $\alpha = 20$, $\beta = 0$, $\delta = 0.1$, and $\mu = 0$, (**d**) infinite-variance α-stable process with $\alpha = 1.9$ and dispersion $\kappa = 0.0272$ ($\mu = 0$, $\sigma = 0$, and $\nu(dy)$ is of the form $\frac{k}{|y|^{1+\alpha}} dy$ [27])

excitatory or inhibitory input forcing signal S_t—either s_1 or s_2. The neuron feeds its activation or membrane potential signal X_t back to itself through the drift term $b(X_{t-})$ and emits the (observable) thresholded or spike signal Y_t as output. Here g is a static transformation. We use the threshold $g(X_t) = 1$ if $X_t > 0$ and zero else for continuous neuron models. We use a related threshold g in spiking neuron models where g determines the spike occurrence. The drift term $b(X_{t-})$ is $-X_{t-} + f(X_{t-}) + S_t$ for continuous neuron models where the neuronal signal function $f(x)$ includes common signal functions such as the logistic, bistable, linear-threshold, and Gaussian or "radial basis" [16]. The drift term $b(X_{t-})$ of sensory spiking neurons includes popular spiking neuron models such as the FitzHugh-Nagumo (FHN) model,

the leaky integrate-and-fire model, and the reduced Type I neuron model. There exist θ_1 and θ_2 for continuous neuron models such that the input S_t is subthreshold when $\theta_1 \leq s_1 < s_2 \leq \theta_2$ while there exists B for spiking neuron models such that S_t is subthreshold when $S_t < B$. The values of θ_1, θ_2, and B depend on the model parameters.

Equation (21) is shorthand for the system of stochastic differential equations

$$dX_t^i = b^i(X_{t-})dt + \sum_{j=1}^{m} c_j^i(X_{t-})dL_t^j \qquad \text{for} \quad i = 1, .., d \tag{23}$$

with initial condition X_0^i. Here $X_t = (X_t^1, ..., X_t^d)'$, $b(X_t) = (b^1(X_t), ..., b^d(X_t))'$, and c is a $d \times m$ matrix with rows $c^i(X_t) = (c_1^i(X_t), ..., c_m^i(X_t))$. The functions $b^i \colon \mathbb{R}^d \to \mathbb{R}$ are locally or globally Lipschitz measurable functions. The functions $c_j^i \colon \mathbb{R}^d \to \mathbb{R}$ are bounded globally Lipschitz measurable functions such that $|c_j^i|^2 \leq H_j^i \in \mathbb{R}^+$. The L_t^j terms are independent Levy processes for $j = 1, ..., m$.

Levy noise has advantages over standard Gaussian noise in neuron models despite its increased mathematical complexity. A Levy noise model more accurately describes how the neuron's membrane potential evolves than does a simpler diffusion model because the more general Levy model includes not only pure-diffusion and pure-jump models but jump-diffusion models as well. Neuron models with additive Gaussian noise are pure-diffusion models. These neuron models rely on the classical central limit theorem for their Gaussian structure and thus they rely on special limiting-case assumptions of incoming Poisson spikes from other neurons. These assumptions require at least that the number of impinging synapses is large and that the synapses have small membrane effects due to the small coupling coefficient or the synaptic weights [11]. The Gaussian noise assumption may be more appropriate for signal inputs from dendritic trees because of the sheer number of dendrites. But often fewer inputs come from synapses near the post-synaptic neuron's trigger zone and these inputs produce impulses in noise amplitudes because of the higher concentration of voltage-sensitive sodium channels in the trigger zone [13]. Engineering applications also favor the more general Levy model because physical devices may be limited in their number of model-neuron connections and because real signals and noise can often be impulsive. Adding Levy noise to enhance faint signals could apply to a variety of signal and image processing problems that include low-light imaging in satellites and other sensor devices, night vision, artificial vision and olfaction, neural prosthetics, infrared imaging, impulsive signal detection, and some types of pattern recognition.

Theorems 7 and 8 below generalize the forbidden interval Theorem 6 for continuous and sensory spiking neuron models [26] to a broad class of Levy noise that may depend on the neuron's membrane potential. The next two forbidden interval theorems require the following Ito-theoretic lemma and corresponding limitation that the Levy process have finite second moments [27].

Lemma. *Let $b^i \colon \mathbb{R}^d \to \mathbb{R}$ and $c_j^i \colon \mathbb{R}^d \to \mathbb{R}$ in (23) be measurable functions that satisfy the respective local and global Lipschitz conditions*

$$\|b^i(z) - b^i(y)\| \leq C_n\|z - y\| \quad \text{when } \|z\| \leq n \text{ and } \|y\| \leq n \quad (24)$$

$$\|c^i_j(z) - c^i_j(y)\| \leq K_1\|z - y\| \text{ for all } z \text{ and } y \in \mathbb{R}^d \quad (25)$$

$$\text{and} \quad |c^i_j|^2 \leq H^i_j \quad \text{for} \quad i = 1, ..., d \text{ and } j = 1, ..., m. \quad (26)$$

Suppose $dX_t = b(X_t)dt + c(X_{t-})dL_t$ and $d\hat{X}_t = b(\hat{X}_t)dt$ where dL_t is a Levy noise with $\mu = \mathbf{0}$ and finite second moments. Then for every $T \in \mathbb{R}^+$ and for every $\varepsilon > 0$:

$$E[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \varepsilon] \rightarrow 0 \quad \text{as} \quad \sigma^j \rightarrow 0 \text{ and } v^j \rightarrow 0 \text{ for all} \quad j = 1, ..., m, \quad (27)$$

and hence

$$P(\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \varepsilon) \rightarrow 0 \quad \text{as} \quad \sigma^j \rightarrow 0 \text{ and } v^j \rightarrow 0 \text{ for all } j = 1, ..., m \quad (28)$$

since mean-square convergence implies convergence in probability.

Theorem 7. *Suppose that the continuous neuron models of the form (21) and (22) have a bounded globally Lipschitz Levy diffusion term $c(X_{t-}) \leq H$ and that the additive Levy noise has drift velocity μ. Suppose also that the input signal $S(t) \in \{s_1, s_2\}$ is subthreshold: $\theta_1 \leq s_1 < s_2 \leq \theta_2$ and that there is some statistical dependence between the input random variable S and the output spike-rate random variable R so that $I(S, R) > 0$. Then such continuous neuron models exhibit the nonmonotone SR effect in the sense that $I(S, R) \rightarrow 0$ as the Levy noise parameters $\sigma \rightarrow 0$ and $v \rightarrow 0$ if $\theta_1 - s_1 \leq H\mu \leq \theta_2 - s_2$.*

Theorem 8. *Suppose that the spiking neuron models of the form (21) and (22) have a locally Lipschitz drift term $b(X_{t-})$ and a bounded globally Lipschitz Levy diffusion term $c(X_{t-}) \leq H$. Suppose also that the additive Levy noise has drift velocity μ and that the input signal $S(t) \in \{s_1, s_2\}$ is subthreshold: $S(t) < B$. Suppose there is some statistical dependence between the input random variable S and the output spike-rate random variable R so that $I(S, R) > 0$. Then such spiking neuron models exhibit the SR effect in the sense that $I(S, R) \rightarrow 0$ as the Levy noise parameters $\sigma \rightarrow 0$ and $v \rightarrow 0$ if $H\mu < B - s_2$.*

These forbidden interval theorems still require that the Levy process have a finite second moment. Simulations show that often we can drop this condition in the case of infinite-variance stable processes and still produce an SR noise benefit [27]]. But it is an open question whether some form of forbidden interval theorem holds for more general Levy diffusions.

References

1. D. Applebaum, "Levy Processes–From Probability to Finance and Quantum Groups," *Notices of American Mathematical Society*, vol. 51, no. 11, pp. 1336–1347, December 2004.
2. L. Breiman, *Probability*, Addison-Wesley, 1968.
3. A. R. Bulsara and L. Gammaitoni, "Tuning in to Noise," *Physics Today*, vol. 49, pp. 39–45, March 1996.

4. A. R. Bulsara and A. Zador, "Threshold Detection of Wideband Signals: A Noise-Induced Maximum in the Mutual Information," *Physical Review E*, vol. 54, no. 3, pp. R2185–R2188, September 1996.

5. D. Chander and E. J. Chichilnisky, "Adaptation to Temporal Contrast in Primate and Salamander Retina," *Journal of Neuroscience*, vol. 21, no. 24, pp. 9904–9916, December 2001.

6. J. J. Collins, C. C. Chow, A. C. Capela, and T. T. Imhoff, "Aperiodic Stochastic Resonance," *Physical Review E*, vol. 54, no. 5, pp. 5575–5584, November 1996.

7. J. J. Collins, C. C. Chow, and T. T. Imhoff, "Aperiodic Stochastic Resonance in Excitable Systems," *Physical Review E*, vol. 52, no. 4, pp. R3321–R3324, October 1995.

8. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.

9. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

10. L. Gammaitoni, "Stochastic Resonance and the Dithering Effect in Threshold Physical Systems," *Physical Review E*, vol. 52, no. 5, pp. 4691–4698, November 1995.

11. W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.

12. A. S. Holevo and R. F. Werner, "Evaluating Capacities of Bosonic Gaussian Channels," *Physical Review A*, vol. 63, pp. 032312, 2001.

13. E. R. Kandel, J. H. Schwartz, and Thomas M. Jessell, *Principles of Neuroscience*, McGraw-Hill, 4th revised edition, 2000.

14. K. J. Kim and F. Rieke, "Temporal Contrast Adaptation in the Input and Output Signals Signals of Salamander Retinal Ganglion Cell," *Journal of Neuroscience*, vol. 21, no. 23, pp. 287–299, December 2001.

15. J. Kong, H. Soh, A. Cassell, C. F. Quate, and H. Dai, "Synthesis of Individual Single-Walled Carbon Nanotubes on Patterned Silicon Wafers," *Nature*, vol. 395, pp. 878–881, 1998.

16. B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall, 1991.

17. B. Kosko, *Noise*, Viking/Penguin, 2006.

18. B. Kosko and S. Mitaim, "Robust Stochastic Resonance: Signal Detection and Adaptation in Impulsive Noise," *Physical Review E*, vol. 64, no. 051110, pp. 1–11, October 2001.

19. B. Kosko and S. Mitaim, "Stochastic Resoannce in Noisy Threshold Neurons," *Neural Networks*, vol. 16, no. 5–6, pp. 755–761, June-July 2003.

20. B. Kosko and S. Mitaim, "Robust Stochastic Resonance for Simple Threshold Neurons," *Physical Review E*, vol. 70, no. 031991, pp. 1–10, September 2004.

21. I. Lee, X. Liu, C. Zhou, and B. Kosko, "Noise-Enhanced Detection of Subthreshold Signals with Carbon Nanotubes," *IEEE Transactions on Nanotechnology*, vol. 5, no. 6, pp. 613–627, November 2006.

22. R. Loudon and P. L. Knight, "Squeezed Light," *Modern Optics*, vol. 34, pp. 709–759, 1987.

23. S. Mitaim and B. Kosko, "Adaptive Stochastic Resonance," *Proceedings of the IEEE: Special Issue on Intelligent Signal Processing*, vol. 86, no. 11, pp. 2152–2183, November 1998.

24. S. Mitaim and B. Kosko, "Adaptive Stochastic Resonance in Noisy Neurons Based on Mutual Information," *IEEE Transactions on Neural Networks*, vol. 15, no. 6, pp. 1562–1540, November 2004.

25. C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, John Wiley and Sons, 1995.

26. A. Patel and B. Kosko, "Stochastic Resonance in Noisy Spiking Retinal and Sensory Neuron Models," *Neural Networks*, vol. 18, pp. 467–478, August 2005.

27. A. Patel and B. Kosko, "Stochastic Resonance in Continuous and Spiking Neuron Models with Levy Noise," *IEEE Transactions on Neural Networks*, 2008, to appear.

28. X. Pei, K. Bachmann, and F. Moss, "The Detection Threshold, Noise and Stochastic Resonance in the Fitzhugh-Nagumo Neuron Model," *Physics Letters A*, vol. 206, pp. 61–65, October 1995.

29. X. Pei, L. Wilkens, and F. Moss, "Noise-Mediated Spike Timing Precision from Aperiodic Stimuli in an Array of Hodgkin-Huxley-Type Neurons," *Physical Review Letters*, vol. 77, no. 22, pp. 4679–4682, November 1996.

30. S. I. Resnick, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, 2007.
31. N. G. Stocks, "Information Transmission in Parallel Threshold Arrays," *Physical Review E*, vol. 63, no. 041114, 2001.
32. S. J. Tans, R. M. Verschueren, and C. Dekker, "Room Temperature Transistor Based on a Single Carbon Nanotube," *Nature*, vol. 393, pp. 49–52, 1998.
33. K. Wiesenfeld and F. Moss, "Stochastic Resonance and the Benefits of Noise: From Ice Ages to Crayfish and SQUIDs," *Nature*, vol. 373, pp. 33–36, January 1995.
34. M. Wilde and B. Kosko, "Quantum Forbidden-Interval Theorems for Stochastic Resonance with Squeezed Light," in *Proceedings of the 8th International Conference on Quantum Communication, Measurement, and Computing*, November 2006, pp. 553–556.
35. J. W. G. Wildoer, L. C. Venema, A. G. Rinzler, R. E. Smalley, and C. Dekker, "Electronic Structure of Atomically Resolved Carbon Nanotubes," *Nature*, vol. 391, pp. 59–62, 1998.
36. L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, "Generation of Squeezed States by Parametric Down Conversion," *Physical Review Letters*, vol. 57, pp. 2520, 1986.
37. Y. Yu, B. Potetz, and T. S. Lee, "The Role of Spiking Nonlinearity in Contrast Gain Control and Information Transmission," *Vision Research*, vol. 45, no. 5, pp. 583–592, March 2005.