

Stochastic Resonance in Continuous and Spiking Neuron Models With Levy Noise

Ashok Patel and Bart Kosko

Abstract—Levy noise can help neurons detect faint or subthreshold signals. Levy noise extends standard Brownian noise to many types of impulsive jump-noise processes found in real and model neurons as well as in models of finance and other random phenomena. Two new theorems and the Itô calculus show that white Levy noise will benefit subthreshold neuronal signal detection if the noise process's scaled drift velocity falls inside an interval that depends on the threshold values. These results generalize earlier “forbidden interval” theorems of neuronal “stochastic resonance” (SR) or noise-injection benefits. Global and local Lipschitz conditions imply that additive white Levy noise can increase the mutual information or bit count of several feedback neuron models that obey a general stochastic differential equation (SDE). Simulation results show that the same noise benefits still occur for some infinite-variance stable Levy noise processes even though the theorems themselves apply only to finite-variance Levy noise. The Appendix proves the two Itô-theoretic lemmas that underlie the new Levy noise-benefit theorems.

Index Terms—Levy noise, jump diffusion, mutual information, neuron models, signal detection, stochastic resonance (SR).

I. STOCHASTIC RESONANCE IN NEURAL SIGNAL DETECTION

STOCHASTIC RESONANCE (SR) occurs when noise benefits a system rather than harms it. Small amounts of noise can often enhance some forms of nonlinear signal processing while too much noise degrades it [12], [13], [22], [27], [45], [49], [58], [60], [61], [69], [71], [72], [84]. SR has many useful applications in physics, biology, and medicine [5]–[7], [11], [14], [17], [18], [21], [23], [32], [40], [41], [43], [52], [53], [55], [56], [62], [70], [75], [83], [85], [89], [91]. SR in neural networks is itself part of the important and growing area of stochastic neural networks [9], [10], [38], [86]–[88], [90]. We show that a wide range of general feedback continuous neurons and spiking neurons benefit from a broad class of additive white Levy noise. This appears to be the first demonstration of the SR effect for neuron models subject to Levy noise perturbations.

Fig. 1 shows how impulsive Levy noise can enhance the Kanizsa square visual illusion in which four dark-corner figures give rise to an illusory bright interior square. Each pixel is the thresholded output of a noisy bistable neuron whose input signals are subthreshold and quantized pixel values of the original noise-free Kanizsa image. The outputs of the bistable neurons do not depend on the input signals if there is no additive noise

because the input signals are subthreshold. Fig. 1(a) shows that adding infinite-variance Levy noise induces a slight correlation between the pixel input and output signals. More intense Levy noise increases this correlation in Fig. 1(b) and (c). Still more intense Levy noise degrades the image and undermines the visual illusion in Fig. 1(d) and (e). Fig. 2 shows typical sample paths from different types of Levy noise. Fig. 3 shows the characteristic inverted-U or nonmonotonic signature of SR for white Levy noise that perturbs a continuous bistable neuron.

We generalize the recent “forbidden interval” theorems [50], [51], [61], [65], [66] for continuous and spiking neuron models to a broad class of finite-second-moment Levy noise that may depend on the neuron's membrane potential. The original forbidden interval theorem [50], [51] states that simple threshold neurons will have an SR noise benefit in the sense that noise increases the neuron's mutual information or bit count if and only if the noise mean or location parameter μ does not fall in a threshold-related interval: SR occurs if and only if $\mu \notin (T - A, T + A)$ for threshold T where $-A < A < T$ for bipolar subthreshold signal $\pm A$. The theorems below show that such an SR noise benefit will occur if the additive white Levy noise process has a bounded scaled drift velocity that does not fall within a threshold-based interval. This holds for general feedback continuous neuron models that include common signal functions such as logistic sigmoids or Gaussians. It also holds for spiking neurons such as the FitzHugh–Nagumo, leaky integrate-and-fire, and reduced type I neuron models. We used the Itô stochastic calculus to prove our results under the assumption that the Levy noise has a finite second moment. But Fig. 1 and Figs. 3(c), 4(c), 5(c), 6(c), 7(c), and 8(c) all show that the SR noise benefit still occurs in the more general infinite-variance case of some types of α -stable Levy noise. So the SR effect is not limited to finite-second-moment Levy noise. We were not able to prove that these stable infinite-variance SR effects must occur as we did prove with simpler neuron models [50], [51], [65].

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 [35], [74]. 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

Manuscript received May 31, 2007; revised December 09, 2007; accepted March 23, 2008. First published November 17, 2008; current version published November 28, 2008.

The authors are with the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail:kosko@usc.edu).

Digital Object Identifier 10.1109/TNN.2008.2005610

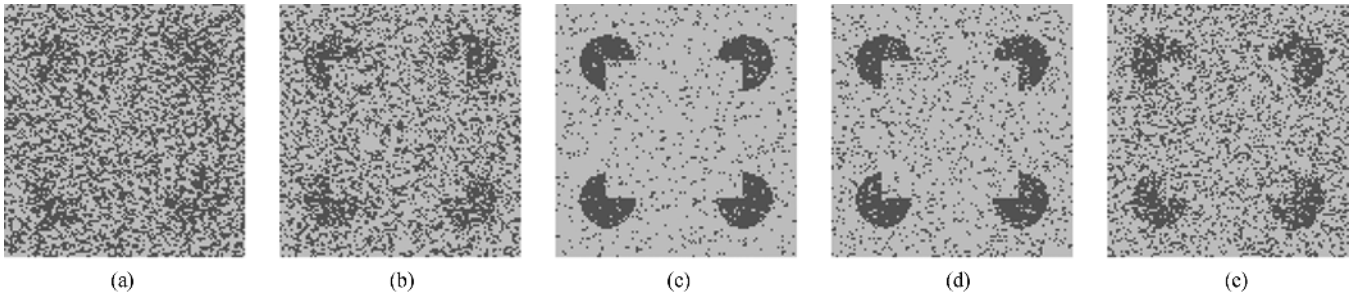


Fig. 1. SR in the Kanisza square illusion with symmetric α -stable noise ($\alpha = 1.9$) in a thick-tailed bell curve with infinite-variance but finite intensity or dispersion γ [30]. The Kanisza square illusion improves as the noise dispersion γ increases from 0.047 to 0.3789 and then it degrades as the dispersion increases further. Each pixel represents the output of the noisy bistable potential neuron model (1)–(2) and (5) that uses the pixel values of the original Kanisza square image as subthreshold input signals. The additive α -stable noise dispersions are as follows: (a) $\gamma = 0.047$, (b) $\gamma = 0.1015$, (c) $\gamma = 0.3789$, (d) $\gamma = 1$, and (e) $\gamma = 3.7321$.

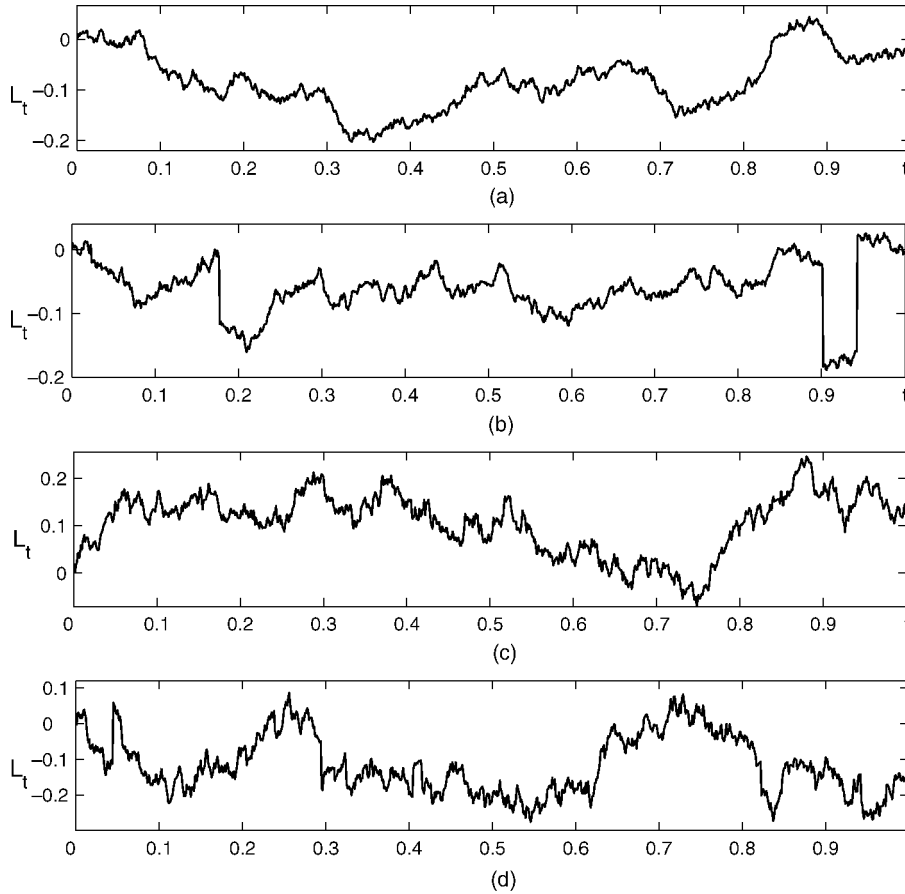


Fig. 2. Sample paths from 1-D Levy processes: (a) Brownian motion with drift $\mu = 0.1$ and variance $\sigma = 0.15$; (b) jump diffusion with $\mu = 0.1$, $\sigma = 0.225$, Poisson jump rate $\lambda = 3$, and uniformly distributed jump magnitudes in the interval $[-0.2, 0.2]$ (and so with Levy measure $\nu(dy) = (3/0.4)dy$ for $y \in [-0.2, 0.2]$ and zero else); (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 $(k/|y|^{1+\alpha})dy$].

be large and that the synapses have small membrane effects due to the small coupling coefficient or the synaptic weights [28], [47]. 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 postsynaptic 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 [29], [46], [64]. Engineering applications also favor the more general Levy model because physical devices may be limited in their number of model-neuron connections

[59] and because real signals and noise can often be impulsive [30], [63], [76].

II. NOISY FEEDBACK NEURON MODELS

We study Levy SR noise benefits in the noisy feedback neuron models of the general form

$$\dot{x} = -x(t) + f(x(t)) + s(t) + n(t) \tag{1}$$

$$y(t) = g(x(t)) \tag{2}$$

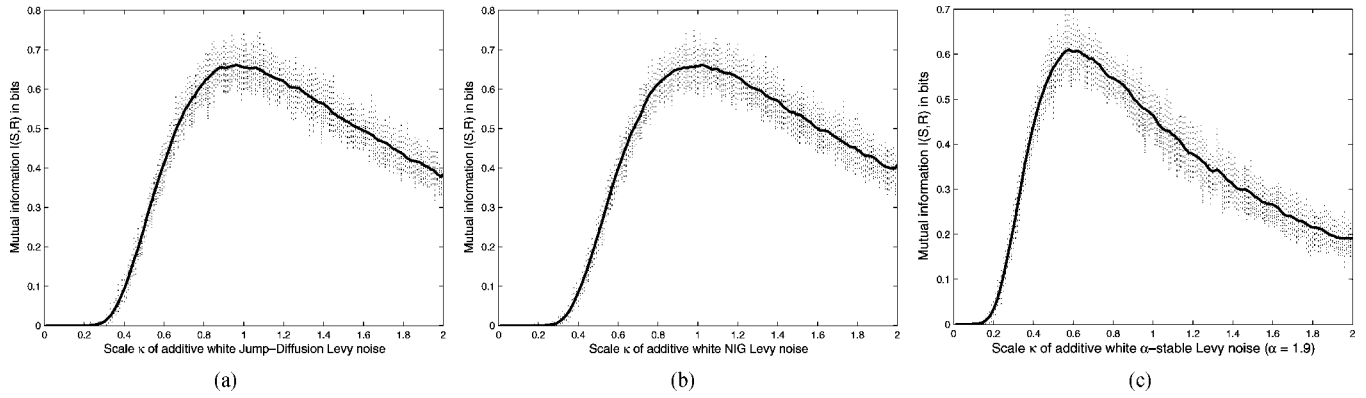


Fig. 3. Mutual information Levy noise benefits in the continuous bistable neuron (1)–(2) and (5). Additive white Levy noise dL_t increases the mutual information of the bistable potential neuron for the subthreshold input signals $s_1 = -0.3$ and $s_2 = 0.4$. The types of Levy noise dL_t are as follows: (a) Gaussian with uniformly distributed jumps, (b) pure-jump normal inverse Gaussian (NIG), and (c) symmetric α -stable noise with $\alpha = 1.9$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

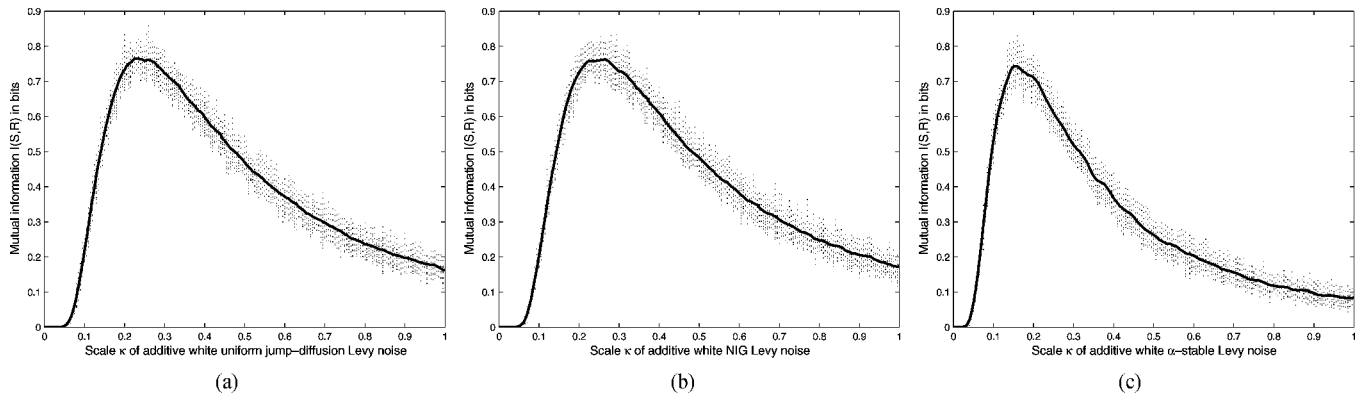


Fig. 4. Mutual information SR Levy noise benefits in the logistic continuous neuron (1)–(2) and (4). Additive white Levy noise dL_t increases the mutual information of the logistic neuron for the subthreshold input signal $s_1 = -0.6$, $s_2 = -0.4$, and $c = 8$. The types of Levy noise dL_t are as follows: (a) Gaussian with uniformly distributed jumps, (b) pure-jump normal inverse Gaussian (NIG), and (c) symmetric α -stable noise with $\alpha = 1.95$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

with initial condition $x(t_0) = x_0$. Here $s(t)$ is the additive net excitatory or inhibitory input forcing signal—either s_1 or s_2 . The additive noise term $n(t)$ is Levy noise with mean or location μ and intensity scale κ (or dispersion γ for symmetric α -stable noise where $\gamma = \kappa^\alpha$ with characteristic function $\phi(u) = e^{i\mu u - \gamma|u|^\alpha}$). The neuron feeds its activation or membrane potential signal $x(t)$ back to itself through $-x(t) + f(x(t))$ and emits the (observable) thresholded or spike signal $y(t)$ as output. Here g is a static transformation function. We use the threshold

$$g(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{else} \end{cases} \quad (3)$$

for continuous neuron models. We use a related threshold g in spiking neuron models where g determines the spike occurrence. The neuronal signal function $f(x)$ of (1) can be of quite general form for continuous neuron models [66].

- *Logistic*. The logistic signal function [48] is sigmoidal and strictly increasing

$$f(x) = \frac{1}{1 + e^{-cx}} \quad (4)$$

for scaling constant $c > 0$. We use $c = 8$. This signal function gives a bistable additive neuron model.

- *Hyperbolic Tangent*. This signal function is also sigmoidal and gives a bistable additive neuron model [2], [15], [37], [48]

$$f(x) = 2 \tanh x. \quad (5)$$

- *Linear Threshold*. This linear-threshold signal has the form [48]

$$f(x) = \begin{cases} cx & |cx| \leq 1 \\ 1 & cx > 1 \\ -1 & cx < -1 \end{cases} \quad (6)$$

for constant $c > 0$. We use $c = 2$.

- *Exponential*. This signal function is asymmetric and has the form [48]

$$f(x) = \begin{cases} 1 - \exp\{-cx\} & \text{if } x > 0 \\ 0 & \text{else} \end{cases} \quad (7)$$

for constant $c > 0$. We use $c = 8$.

- *Gaussian*. The Gaussian or “radial basis” signal function [48] differs in form from the signal functions above because it is nonmonotonic

$$f(x) = \exp\{-cx^2\} \quad (8)$$

for constant $c > 0$. We use $c = 8$.

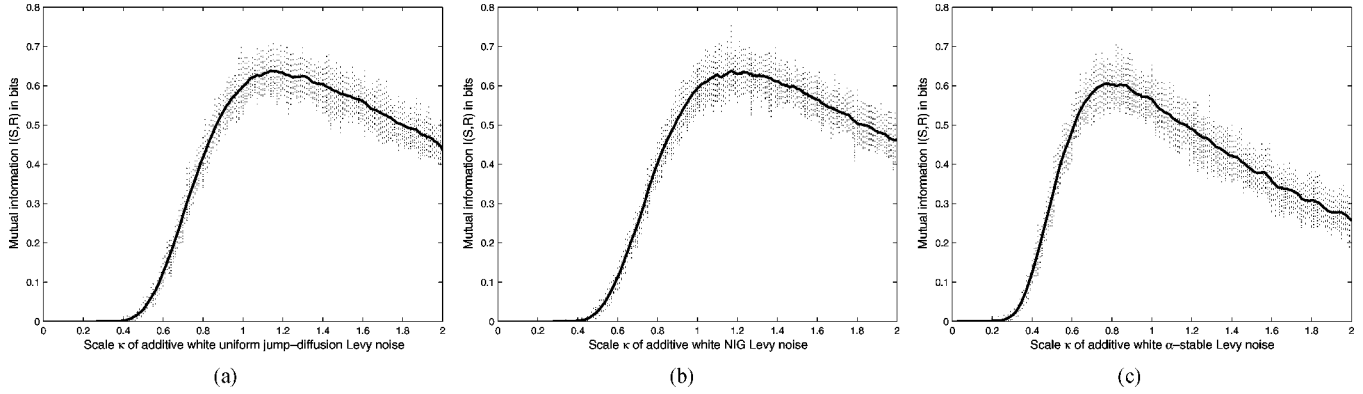


Fig. 5. Mutual information Levy noise benefits in the linear-threshold continuous neuron (1)–(2) and (6). Additive white Levy noise dL_t increases the mutual information of the linear-threshold neuron for the subthreshold input signal $s_1 = -0.4$, $s_2 = 0.4$, and $c = 2$. The types of Levy noise dL_t are as follows: (a) Gaussian with uniformly distributed jumps, (b) pure-jump normal inverse Gaussian (NIG), and (c) symmetric α -stable noise with $\alpha = 1.95$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

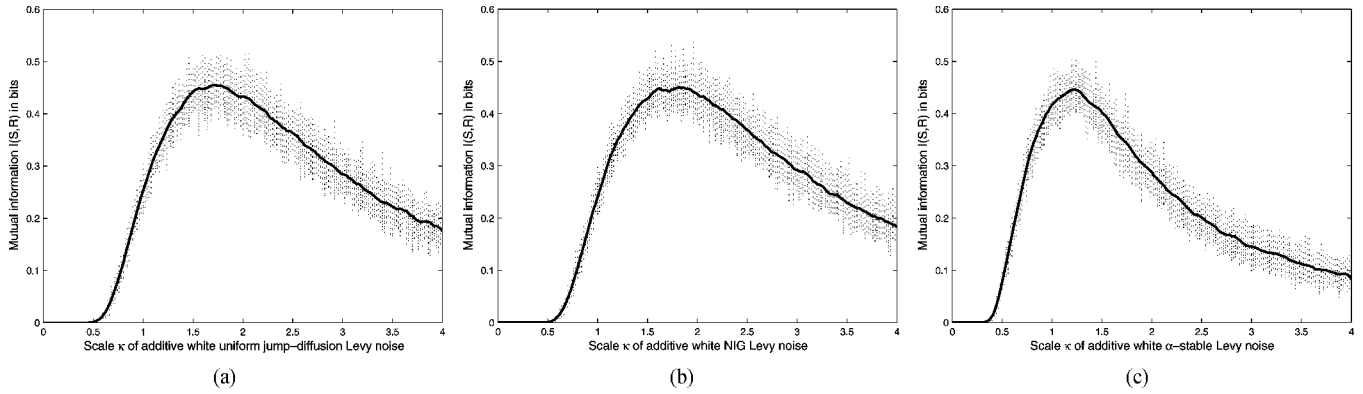


Fig. 6. Mutual information Levy noise benefits in the Gaussian or “radial basis” continuous neuron (1)–(2) and (8). Additive white Levy noise dL_t increases the mutual information of the Gaussian neuron for the subthreshold input signal $s_1 = -0.4$, $s_2 = 0.4$, and $c = 8$. The types of Levy noise dL_t are as follows: (a) Gaussian with uniformly distributed jumps, (b) pure-jump normal inverse Gaussian (NIG), and (c) symmetric α -stable noise with $\alpha = 1.95$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

The above neuron models can have up to three fixed points depending on the input signal and the model parameters. The input signal is subthreshold in the sense that switching it from s_1 to s_2 or vice versa does not change the output Y_t of (22). There exist θ_1 and θ_2 such that the input S is subthreshold when $\theta_1 \leq s_1 < s_2 \leq \theta_2$. The values of θ_1 and θ_2 depend on the model parameters. Consider the linear-threshold neuron model (1)–(2) and (6) with $c = 2$. A simple calculation shows that if the input signal $S_t \in \{s_1, s_2\}$ satisfies $-0.5 < s_1 < s_2 < 0.5$ then the linear-threshold neuron has two stable fixed points (one positive and the other negative) and has one unstable fixed point between them. The Gaussian neuron model (1)–(2) and (8) has only one fixed point if $0 < s_1 < s_2$. So the input is subthreshold because switching it from s_1 to s_2 or vice versa does not change the output Y_t . Fig. 3 shows the mutual information noise benefits in the bistable neuron model (1)–(2) and (5) for three different additive white Levy noise cases when the input signals are subthreshold. Note the signature nonmonotonic shape of all three SR noise-benefit curves in Fig. 3.

The membrane potential dynamics (1) is 1-D for all our neuron models except for the 2-D FitzHugh–Nagumo (FHN) spiking neuron model below. So next we briefly describe

multidimensional Levy processes and set up a general multidimensional Levy stochastic differential equation (SDE) framework for our feedback continuous and spiking neuron models.

III. LEVY PROCESSES AND STOCHASTIC DIFFERENTIAL EQUATIONS

Levy processes [68], [77] form a wide class of random processes that include Brownian motion, α -stable processes, compound Poisson processes, generalized inverse Gaussian processes, and generalized hyperbolic processes. Fig. 2 shows some typical scalar Levy sample paths. Levy processes can account for the impulsiveness or discreteness of both signals and noise. Researchers have used Levy processes to model diverse phenomena in economics [4], [78], physics [81], electrical engineering [1], [4], [63], [67], biology [80], and seismology [82]. A Levy process $L_t = (L_t^1, \dots, L_t^m)'$ for $t \geq 0$ in a given probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t < \infty}, P)$ is a stochastic process taking values in \mathbf{R}^m with stationary and independent increments (we assume that $L_0 = 0$ with probability 1). The Levy process L_t obeys the following three properties:

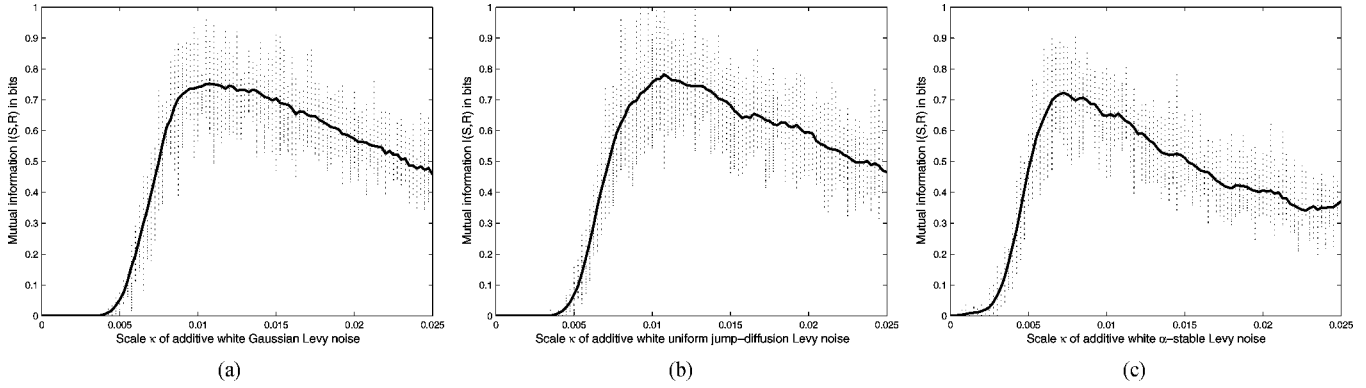


Fig. 7. Mutual information Levy noise benefits in the leaky integrate-and-fire (IF) spiking neuron model (40). Additive white Levy noise dL_t increases the mutual information of the IF neuron with parameters $a = 0.5$ and $\delta = 0.02$ for the subthreshold input signal $s_1 = 0.005$ and $s_2 = 0.012$. The types of Levy noise dL_t are as follows: (a) Gaussian, (b) Gaussian with uniformly distributed jumps, and (c) symmetric α -stable noise with $\alpha = 1.95$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

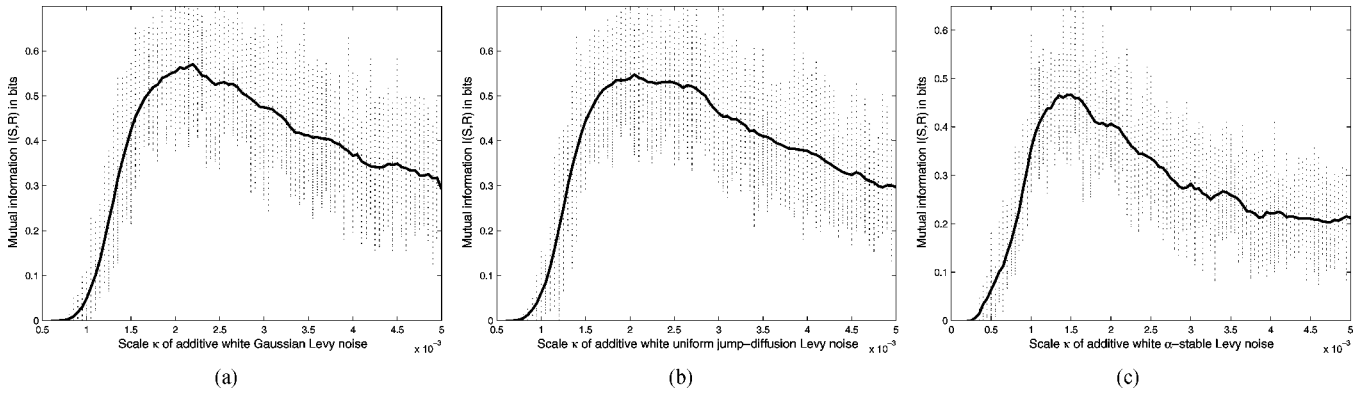


Fig. 8. Mutual information Levy noise benefits in the FHN spiking neuron (42)–(43). Additive white Levy noise dL_t increases the mutual information of the FHN neuron for the subthreshold input signal $s_1 = -0.0045$ and $s_2 = 0.0045$. The types of Levy noise dL_t are as follows: (a) Gaussian, (b) Gaussian with uniformly distributed jumps, and (c) symmetric α -stable noise with $\alpha = 1.9$ (thick-tailed bell curve with infinite variance [63]). The dashed vertical lines show the total min–max deviations of the mutual information in 100 simulation trials.

- 1) $L_t - L_s$ is independent of sigma-algebra \mathcal{F}_s for $0 \leq s < t \leq \infty$;
- 2) $L_t - L_s$ has the same distribution as L_{t-s} ;
- 3) $L_s \rightarrow L_t$ in probability if $s \rightarrow t$.

The Levy–Khinchine formula gives the characteristic function ϕ of L_t as [3]

$$\phi(u) = E(e^{i\langle u, L_t \rangle}) = e^{t\eta(u)}, \quad \text{for } t \geq 0 \text{ and } u \in \mathbf{R}^m \quad (9)$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product (so $|u| = \langle u, u \rangle^{1/2}$). The characteristic exponent or the so-called Levy exponent is

$$\eta(u) = i\langle \mu, u \rangle - \frac{1}{2}\langle u, \mathbf{K}u \rangle + \int_{\mathbf{R}^m \setminus \{0\}} [e^{i\langle u, y \rangle} - 1 - i\langle u, y \rangle \chi_{|y| < 1}(y)] \nu(dy) \quad (10)$$

for some $\mu \in \mathbf{R}^m$, a positive–definite symmetric $m \times m$ matrix \mathbf{K} , and measure ν on Borel subsets of $\mathbf{R}_0^m = \mathbf{R}^m \setminus \{0\}$ (or $\nu(\{0\}) = 0$). Then ν is a Levy measure such that

$$\int_{\mathbf{R}_0^m} \min\{1, |y|^2\} \nu(dy) < \infty. \quad (11)$$

A Levy process L_t combines a drift component, a Brownian motion (Gaussian) component, and a jump component. The Levy–Khinchine triplet (μ, \mathbf{K}, ν) completely determines these components. The Levy measure ν determines both the average number of jumps per unit time and the distribution of jump magnitudes in the jump component of L_t . Jumps of any size in a Borel set B form a compound Poisson process with rate $\int_B \nu(dy)$ and jump density $\nu(dy) / \int_B \nu(dy)$ if the closure \bar{B} does not contain $\mathbf{0}$. μ gives the velocity of the drift component. \mathbf{K} is the covariance matrix of the Gaussian component. If $\mathbf{K} = \mathbf{0}$ and $\nu = 0$ then (9) becomes $E(e^{i\langle u, L_t \rangle}) = e^{it\langle \mu, u \rangle}$. Then $L_t = \mu t$ is a simple m -dimensional deterministic motion (drift) with velocity vector μ . If $\mathbf{K} \neq \mathbf{0}$ and $\nu = 0$ then L_t is an m -dimensional Brownian motion with drift because (9) takes the form $E(e^{i\langle u, L_t \rangle}) = e^{t[i\langle \mu, u \rangle - (1/2)\langle u, \mathbf{K}u \rangle]}$ and because this exponential is the characteristic function of a Gaussian random vector with mean vector μt and covariance matrix $t\mathbf{K}$. If $\mathbf{K} \neq \mathbf{0}$ and $\nu(\mathbf{R}^m) < \infty$ then L_t is a jump-diffusion process while $\mathbf{K} = \mathbf{0}$ and $\nu(\mathbf{R}^m) < \infty$ give a compound Poisson process. If $\mathbf{K} = \mathbf{0}$ and $\nu(\mathbf{R}^m) = \infty$ then L_t is a purely discontinuous jump process and has an infinite number of small jumps in any time interval of positive length.

We consider only the Levy processes whose components L_t^k have finite second moments: $E[(L_t^k)]^2 < \infty$. This excludes

the important family of infinite-variance α -stable processes (including the $\alpha = 0.5$ Levy stable case) where $\alpha \in (0, 2]$ measures the tail thickness and where symmetric α -stable distributions have characteristic functions $\phi(u) = e^{iu\mu - \gamma|u|^\alpha}$ [30], [50], [63], [76]. But a finite-moment assumption does not itself imply that the Levy measure is finite: $(\nu(\mathbf{R}) < \infty)$. Normal inverse Gaussian $\text{NIG}(\alpha, \beta, \delta, \mu)$ distributions are examples of semithick-tailed pure-jump Levy processes that have infinite Levy measure and yet have finite moments of all order [33], [73]. They can model the risks of options hedging and of credit default in portfolios of risky debt obligations [42], [79]. They have characteristic functions of the form $\phi(u) = e^{iu\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2})}$, where $0 \leq |\beta| < \alpha$ and $\delta > 0$.

Let $L_t(\mu, \mathbf{K}, \nu) = (L_t^1, \dots, L_t^m)'$ be a Levy process that takes values in \mathbf{R}^m , where $L_t^j(\mu^j, \sigma^j, \nu^j)$ are real-valued independent Levy processes for $j = 1, \dots, m$. We denote the Levy-Itô decomposition [3] of L_t^j for each $j = 1, \dots, m$ and $t \geq 0$ as

$$L_t^j = \mu^j t + \sigma^j B_t^j + \int_{|y^j| < 1} y^j \tilde{N}^j(t, dy^j) + \int_{|y^j| \geq 1} y^j N^j(t, dy^j) \quad (12)$$

$$= \mu^j t + \sigma^j B_t^j + \int_0^t \int_{|y^j| < 1} y^j \tilde{N}^j(ds, dy^j) + \int_0^t \int_{|y^j| \geq 1} y^j N^j(ds, dy^j). \quad (13)$$

Here μ^j determines the velocity of the deterministic drift process $\mu^j t$ while the B_t^j are real-valued independent standard Brownian motions. Then $\mu = (\mu^1, \dots, \mu^m)'$ and $\mathbf{K} = \text{diag}[(\sigma^1)^2, \dots, (\sigma^m)^2]$. The N^j are independent Poisson random measures on $\mathbf{R}^+ \times \mathbf{R}_0$ with compensated (mean-subtracted) Poisson processes \tilde{N}^j and intensity/Levy measures ν^j . Define the Poisson random measure as

$$N^j(t, B) = \# \{ \Delta L_s \in B \text{ for } 0 \leq s \leq t \} \quad (14)$$

for each Borel set B in \mathbf{R}_0 . The Poisson random measure gives the random number of jumps of L_t in the time interval $[0, t]$ with jump size ΔL_t in the set B . $N^j(t, B)$ is a Poisson random variable with intensity $\nu^j(B)$ if $\nu^j(B) < \infty$ and if we fix t and B . But $N^j(t, \cdot)(\omega)$ is a measure if we fix $\omega \in \Omega$ and $t \geq 0$. This measure is not a martingale but the compensated Poisson random measure

$$\tilde{N}^j(t, B) = N^j(t, B) - t\nu^j(B) \quad (15)$$

is a martingale and gives the compensated Poisson integral [12] [the second term on the right-hand side of (12)] as

$$\int_{|y^j| < 1} y^j \tilde{N}^j(t, dy^j) = \int_{|y^j| < 1} y^j N^j(t, dy^j) - t \int_{|y^j| < 1} y^j \nu^j(dy^j), \quad \text{for } j = 1, \dots, m. \quad (16)$$

We assume again that each L_t^j has a finite second moment ($E|L_t^j|^2 < \infty$). But if L_t^j is a Levy process with triplet

(μ^j, σ^j, ν^j) then L_t^j has a finite p^{th} moment for $p \in \mathbf{R}^+$ if and only if $\int_{|y^j| > 1} |y^j|^p \nu^j(dy^j) < \infty$ [77]. The drift velocity μ^j relates to the expected value of a Levy process L_t^j by $E(L_t^j) = \mu^j t + \int_{|y^j| > 1} y^j \nu^j(dy^j)$ and $E(L_t^j) = tE(L_1^j)$. So if L_t^j is a standard Brownian motion then $\nu^j = 0$, $E(L_t^j) = 0$, and $\text{Var}(L_t^j) = t(\sigma^j)^2$.

The variance of the Levy process in (12) is

$$\text{Var}(L_t^j) = \text{Var}(\sigma^j B_t^j) + \text{Var} \left(\int_{|y^j| < 1} y^j \tilde{N}^j(t, dy^j) \right) + \text{Var} \left(\int_{|y^j| \geq 1} y^j N^j(t, dy^j) \right) \quad (17)$$

because the underlying processes are independent. The variance terms on the right-hand side of (17) have the following form [3]:

$$\text{Var}(L_t^j) = t(\sigma^j)^2 \quad (18)$$

$$\text{Var} \left(\int_{|y^j| \geq 1} y^j N^j(t, dy^j) \right) = t \int_{|y^j| \geq 1} |y^j|^2 \nu^j(dy^j) \quad (19)$$

$$\text{Var} \left(\int_{|y^j| < 1} y^j \tilde{N}^j(t, dy^j) \right) \leq E \left(\int_{|y^j| < 1} y^j \tilde{N}^j(t, dy^j) \right)^2 = t \int_{|y^j| < 1} |y^j|^2 \nu^j(dy^j). \quad (20)$$

The last equality follows from the Itô isometry identity [19, Prop. 8.8]. Then (17) and (18)–(20) imply that the $\text{Var}(L_t^j) \rightarrow 0$ if and only if $\sigma^j \rightarrow 0$ and $\nu^j \rightarrow 0$.

We can rewrite (1)–(2) as a more general Itô SDE [3]

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (21)$$

$$Y_t = g(X_t) \quad (22)$$

with initial condition X_0 . Here $b(X_{t-}) = -X_{t-} + f(X_{t-}) + S_t$ is a Lipschitz continuous drift term, $c(X_{t-})$ is a bounded Levy diffusion term, and dL_t is a white Levy noise with noise scale κ .

Our continuous neuron models are again 1-D but the spiking FHN neuron model is 2-D. So consider the general d -dimensional SDE in the matrix form with m -dimensional Levy noise $L_t = (L_t^1, \dots, L_t^m)'$

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (23)$$

which is shorthand for the system of SDEs

$$dX_t^i = b^i(X_{t-})dt + \sum_{j=1}^m c_j^i(X_{t-})dL_t^j, \quad \text{for } i = 1, \dots, d \quad (24)$$

with initial conditions X_0^i . Here $X_t = (X_t^1, \dots, X_t^d)'$, $b(X_t) = (b^1(X_t), \dots, b^d(X_t))'$, and c is a $d \times m$ matrix with rows $c^i(X_t) = (c_1^i(X_t), \dots, c_m^i(X_t))$. The functions $b^i: \mathbf{R}^d \rightarrow \mathbf{R}$ are locally or globally Lipschitz measurable functions. The functions $c_j^i: \mathbf{R}^d \rightarrow \mathbf{R}$ are bounded globally Lipschitz measurable functions such that $|c_j^i|^2 \leq H_j^i \in \mathbf{R}^+$.

The L_t^j terms are independent Levy processes as in (13) with $\mu^j = 0$ for $j = 1, \dots, m$. Then

$$\begin{aligned}
 dX_t^i &= b^i(X_{t-})dt + \sum_{j=1}^m [c_j^i(X_{t-})\mu^j]dt + \sum_{j=1}^m [c_j^i(X_{t-})\sigma^j]dB_t^j \\
 &\quad + \sum_{j=1}^m \int_{|y^j| < 1} [c_j^i(X_{t-})y^j]\tilde{N}^j(dt, dy^j) \\
 &\quad + \sum_{j=1}^m \int_{|y^j| \geq 1} [c_j^i(X_{t-})y^j]N^j(dt, dy^j) \\
 &= b^i(X_{t-})dt + \sum_{j=1}^m [\mu_j^i(X_{t-})]dt + \sum_{j=1}^m \sigma_j^i(X_{t-})dB_t^j \\
 &\quad + \sum_{j=1}^m \int_{|y^j| < 1} F_j^i(X_{t-}, y^j)\tilde{N}^j(dt, dy^j) \\
 &\quad + \sum_{j=1}^m \int_{|y^j| \geq 1} G_j^i(X_{t-}, y^j)N^j(dt, dy^j) \quad (25)
 \end{aligned}$$

where $\mu_j^i(X_{t-}) = c_j^i(X_{t-})\mu^j = 0$, $\sigma_j^i(X_{t-}) = c_j^i(X_{t-})\sigma^j$, $F_j^i(X_{t-}, y^j) = c_j^i(X_{t-})y^j$, and $G_j^i(X_{t-}, y^j) = c_j^i(X_{t-})y^j$ are all globally Lipschitz functions. This equation has the integral form with initial condition X_0^i

$$\begin{aligned}
 X_t^i &= X_0^i + \int_0^t b^i(X_{s-})ds + \sum_{j=1}^m \int_0^t \sigma_j^i(X_{s-})dB_s^j \\
 &\quad + \sum_{j=1}^m \int_0^t \int_{|y^j| < 1} F_j^i(X_{s-}, y^j)\tilde{N}^j(ds, dy^j) \\
 &\quad + \sum_{j=1}^m \int_0^t \int_{|y^j| \geq 1} G_j^i(X_{s-}, y^j)N^j(ds, dy^j). \quad (26)
 \end{aligned}$$

IV. LEVY NOISE BENEFITS IN CONTINUOUS NEURON MODELS

We now prove that Levy noise can benefit the noisy continuous neurons (21)–(22) with signal functions (4)–(8) and sub-threshold input signals. We assume that the neuron receives a constant subthreshold input signal $S_t \in \{s_1, s_2\}$ for time T . Let S denote the input signal and let Y denote the output signal Y_t for a sufficiently large randomly chosen time $t \leq T$.

Noise researchers have used various system performance measures to detect SR noise benefits [8], [17], [45], [52], [58], [60], [61], [65], [72]. These include the output signal-to-noise ratio, cross correlation, error probability, and Shannon mutual information between input and output signals. We use Shannon mutual information to measure the Levy noise benefits. Mutual information measures the information that the neuron's output conveys about the input signal. It is a common detection performance measure when the input signal is random [8], [39], [61], [84].

Define the Shannon mutual information $I(S, Y)$ of the discrete input random variable S and the output random variable

Y as the difference between the output's unconditional and conditional entropy [20]

$$I(S, Y) = H(Y) - H(Y|S) \quad (27)$$

$$\begin{aligned}
 &= - \sum_y P_Y(y) \log P_Y(y) \\
 &\quad + \sum_s \sum_y P_{SY}(s, y) \log P_{Y|S}(y|s) \quad (28)
 \end{aligned}$$

$$\begin{aligned}
 &= - \sum_y P_Y(y) \log P_Y(y) \\
 &\quad + \sum_s P_S(s) \sum_y P_{Y|S}(y|s) \log P_{Y|S}(y|s) \quad (29)
 \end{aligned}$$

$$= \sum_{s, y} P_{SY}(s, y) \log \frac{P_{SY}(s, y)}{P_S(s)P_Y(y)}. \quad (30)$$

So the mutual information is the expectation of the random variable $\log P_{SY}(s, y)/P_S(s)P_Y(y)$

$$I(S, Y) = E \left[\log \frac{P_{SY}(s, y)}{P_S(s)P_Y(y)} \right]. \quad (31)$$

Here $P_S(s)$ is the probability density of the input S , $P_Y(y)$ is the probability density of the output Y , $P_{Y|S}(y|s)$ is the conditional density of the output Y given the input S , and $P_{SY}(s, y)$ is the joint density of the input S and the output Y . An SR *noise benefit* occurs in a system if and only if an increase in the input noise variance or dispersion increases the system's mutual information (31).

We need the following lemma to prove that noise improves the continuous neuron's mutual information or bit count. The Appendix gives the proof of Lemma 1.

Lemma 1: Let $b^i : \mathbf{R}^d \rightarrow \mathbf{R}$ and $c_j^i : \mathbf{R}^d \rightarrow \mathbf{R}$ in (23) and (24) be measurable functions that satisfy the global Lipschitz conditions

$$|b^i(x_1) - b^i(x_2)| \leq K_1 \|x_1 - x_2\| \quad (32)$$

$$|c_j^i(x_1) - c_j^i(x_2)| \leq K_2 \|x_1 - x_2\| \quad (33)$$

and

$$|c_j^i(x_1)|^2 \leq H_j^i \quad (34)$$

for all $x_1, x_2 \in \mathbf{R}^d$ and for $i = 1, \dots, d$ and $j = 1, \dots, m$. Suppose

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (35)$$

$$d\hat{X}_t = b(\hat{X}_t)dt \quad (36)$$

where dL_t is a Levy noise with $\mu = \mathbf{0}$ and finite second moments. Then, for every $T \in \mathbf{R}^+$ and for every $\varepsilon > 0$

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0, \quad (37)$$

for all $j = 1, \dots, m$

and hence

$$P \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \varepsilon \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0, \\ \text{for all } j = 1, \dots, m \quad (38)$$

because mean square convergence implies convergence in probability.

We prove the Levy SR theorem with the stochastic calculus and a special limiting argument. This avoids trying to solve for the process X_t in (21). The proof strategy follows that of the ‘‘forbidden interval’’ theorems [50], [51], [65]: *what goes down must go up*. Jensen’s inequality [20] implies that $I(S, Y) \geq 0$. Random variables S and Y are statistically independent if and only if $I(S, Y) = 0$. Hence $I(S, Y) > 0$ implies some degree of statistical dependence. So the system must exhibit the SR noise benefit if $I(S, Y) > 0$ and if $I(S, Y) \rightarrow 0$ when noise parameters $\sigma \rightarrow 0$ and $\nu \rightarrow 0$. Theorem 1 uses Lemma 1 to show that $I(S, Y) \rightarrow 0$ when noise parameters $\sigma \rightarrow 0$ and $\nu \rightarrow 0$. So some increase in the noise parameters must increase the mutual information.

Theorem 1: Suppose that the continuous neuron models (21)–(22) and (4)–(8) 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 the neuron models (21)–(22) with signal functions including (4)–(8) exhibit the nonmonotone SR effect in the sense that $I(S, Y) \rightarrow 0$ as the Levy noise parameters $\sigma \rightarrow 0$ and $\nu \rightarrow 0$ if $\theta_1 - s_1 \leq H\mu \leq \theta_2 - s_2$.

Proof: Let $\{\sigma_k, \nu_k\}_{k=1}^{\infty}$ be any decreasing sequence of Levy noise parameters such that $\sigma_k \rightarrow 0$ and $\nu_k \rightarrow 0$ as $k \rightarrow \infty$. Define $X(t)_k$ and $Y(t)_k$ as solution processes of the continuous neuron models with Levy noise parameters σ_k and ν_k instead of σ and ν .

Suppose that $\mu \neq 0$. We can absorb the drift $c(X_{t-})\mu$ into the input signal S because the Levy noise L_t is additive in the neuron models. Then the new input signal $S' = S + c(X_{t-})\mu$ and it does not affect the Lipschitz continuity of $b(X_{t-})$ in (21). Note that S' is subthreshold ($\theta_1 \leq S' \leq \theta_2$) if $\theta_1 - s_1 \leq H\mu \leq \theta_2 - s_2$. So we lose no generality if we consider the noise dL_t with $\mu = 0$ and let $S \in \{s_1, s_2\}$ be subthreshold in the continuous neuron models (21). This allows us to use Lemma 1.

Let the symbol ‘‘0’’ denote the input signal $S = s_1$ and the output signal $Y = 0$. Let the symbol ‘‘1’’ denote the input signal $S = s_2$ and the output signal $Y = 1$. Assume that $0 < P_S(s) < 1$ to avoid triviality when $P_S(s) = 0$ or 1. We show that S and Y are asymptotically independent by using the fact that $I(S, Y) = 0$ if and only if S and Y are statistically independent [20]. 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 $\sigma_k \rightarrow 0$ and $\nu_k \rightarrow 0$ as $k \rightarrow \infty$ for signal symbols $s \in S$ and $y \in Y$. The theorem of total probability and the two-symbol alphabet set S give

$$P_Y(y) = \sum_s P_{Y|S}(y|s)P_S(s) \\ = P_{Y|S}(y|0)P_S(0) + P_{Y|S}(y|1)P_S(1)$$

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

so we need to show only that $P_{Y_k|S}(y|0) - P_{Y_k|S}(y|1) = 0$ as $\sigma_k \rightarrow 0$ and $\nu_k \rightarrow 0$ for $y \in \{0, 1\}$. We prove the case for $y = 0$ only: $\lim_{k \rightarrow \infty} \{P_{Y_k|S}(0|0) - P_{Y_k|S}(0|1)\} = 0$ since the proof for $y = 1$ is similar. Then the desired limit goes to zero because

$$\lim_{k \rightarrow \infty} \{P_{Y_k|S}(0|0) - P_{Y_k|S}(0|1)\} \\ = \lim_{k \rightarrow \infty} P_{Y_k|S}(0|0) - \lim_{k \rightarrow \infty} P_{Y_k|S}(0|1) \\ = \lim_{k \rightarrow \infty} P[Y_k = 0|S = 0] - \lim_{k \rightarrow \infty} P[Y_k = 0|S = 1] \\ = \lim_{k \rightarrow \infty} P[X(t)_k < 0|S = 0] \\ - \lim_{k \rightarrow \infty} P[X(t)_k < 0|S = 1], \quad \text{for large } t \\ = \lim_{k \rightarrow \infty} \left\{ P[X(t)_k < 0, \hat{X}_t < 0|S = 0] \right. \\ \left. + P[X(t)_k < 0, \hat{X}_t > 0|S = 0] \right\} \\ - \lim_{k \rightarrow \infty} \left\{ P[X(t)_k > 0, \hat{X}_t < 0|S = 1] \right. \\ \left. + P[X(t)_k > 0, \hat{X}_t > 0|S = 1] \right\}, \quad \text{for large } t \\ = \lim_{k \rightarrow \infty} \left\{ P[X(t)_k < 0|\hat{X}_t < 0, S = 0]P[\hat{X}_t < 0|S = 0] \right. \\ \left. + P[X(t)_k < 0|\hat{X}_t > 0, S = 0]P[\hat{X}_t > 0|S = 0] \right\} \\ - \lim_{k \rightarrow \infty} \left\{ P[X(t)_k > 0|\hat{X}_t < 0, S = 1]P[\hat{X}_t < 0|S = 1] \right. \\ \left. + P[X(t)_k > 0|\hat{X}_t > 0, S = 1]P[\hat{X}_t > 0|S = 1] \right\} \\ \text{for large } t \\ = \left\{ 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \right\} - \left\{ 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \right\} \\ \text{by Lemma 1 and the assumption that} \\ P[\hat{X}_t < 0|S = s_i] = P[\hat{X}_t > 0|S = s_i] = \frac{1}{2} \text{ for } i = 1, 2 \\ = 0 \quad \text{Q.E.D.}$$

Figs. 4(a) and (b), 5(a) and (b), and 6(a) and (b) show simulation instances of Theorem 1 for finite-variance jump-diffusion and pure-jump additive white Levy noise in logistic, linear-threshold, and Gaussian neuron models. Small amounts of additive Levy noise in continuous neuron models produce the SR effect by increasing the Shannon mutual information $I(S, Y)$ between realizations of a random (Bernoulli) subthreshold input signal S and the neuron’s thresholded output random variable Y . The SR effect in Figs. 3(c), 4(c), 5(c), and 6(c) lies outside the scope of the theorem because it occurs for infinite-variance α -stable noise. Thus the SR effect in continuous neurons is not limited to finite-second-moment Levy noise.

V. LEVY NOISE BENEFITS IN SPIKING NEURON MODELS

We next demonstrate Levy SR noise benefits in three popular *spiking* neuron models: the leaky integrate-and-fire model [17], [28], the reduced type I neuron model [54], and the FHN model [26], [16]. This requires the use of Lemma 2 as we discuss below. These neuron models have a 1-D or 2-D form of (1). A spike occurs when the membrane potential $x(t)$ crosses a threshold value from below. We measure the mutual information $I(S, R)$ between the input signal $s(t)$ and the output spike-rate

response R of these spiking neuron models. We define the average output spike-rate response R in the time interval $(t_1, t_2]$ as

$$R = \frac{N(t_1, t_2]}{t_2 - t_1} \quad (39)$$

where $N(t_1, t_2]$ is the number of spikes in the time interval $(t_1, t_2]$.

A. The Leaky Integrate-and-Fire Neuron Model

The leaky integrate-and-fire neuron model has the form [17]

$$\dot{v} = -av + a - \delta + S + n \quad (40)$$

where v is the membrane voltage, a and δ are constants, δ/a is the barrier height of the potential, S is an input signal, and n is independent Gaussian white noise in the neural literature but here is Levy white noise. The input signal S is subthreshold when $S < \delta$. The neuron emits a spike when the membrane voltage v crosses the threshold value of 1 from below to above. The membrane voltage v resets to $1 - \delta/a$ just after the neuron emits a spike.

B. The Reduced Type I Neuron Model

The reduction procedure in [31], [36] gives a simple 1-D normal form [54] of the multidimensional dynamics of Type I neuron models

$$\dot{v} = \beta + v^2 + \sigma n \quad (41)$$

where v is the membrane potential, β is the value of input signal, and σ is the standard deviation of Gaussian white noise n in the neural literature but here is Levy white noise. This reduced model (41) operates in a subthreshold or excitable regime when the input $\beta < 0$.

C. The FHN Neuron Model

The FHN neuron model [16], [26], [28] is a 2-D simplification of the Hodgkin and Huxley neuron model [34]. It describes the response of a so-called type II excitable system [28], [55] that undergoes a Hopf bifurcation. The system first resides in the stable rest state for subthreshold inputs as do multistable systems. Then the system leaves the stable state in response to a strong input but returns to it after passing through firing and refractory states in a manner that differs from the behavior of multistable systems. The FHN neuron model is a limit-cycle oscillator of the form

$$\epsilon \dot{v} = -v \left(v^2 - \frac{1}{4} \right) - w + A + s(t) + n(t) \quad (42)$$

$$\dot{w} = v - w \quad (43)$$

where $v(t)$ is a fast (voltage) variable, $w(t)$ is slow (recovery) variable, A is a constant (tonic) activation signal, and $\epsilon = 0.005$. $n(t)$ is a white Levy noise and $s(t)$ is a subthreshold input signal—either s_1 or s_2 . We measure the neuron’s response to the input signal $s(t)$ in terms of the transition (firing) rate $r(t)$.

We can rewrite (42)–(43) as

$$\epsilon \dot{v} = -v \left(v^2 - \frac{1}{4} \right) - w + A_T - (B - s(t)) + n(t) \quad (44)$$

$$\dot{w} = v - w \quad (45)$$

where B is a positive constant parameter that corresponds to the distance that the input signal $s(t)$ must overcome to cross the threshold. Then $B - s(t)$ is the signal-to-threshold distance and so $s(t)$ is subthreshold when $B - s(t) > 0$. Our simulations used $B = 0.007$ and hence $A = -(5/(12\sqrt{3} + 0.007))$.

The deterministic FHN model [$n(t) \equiv 0$ in (44)] performs relaxation oscillations and has an action potential $v(t)$ that lies between -0.6 and 0.6 . The system emits a spike when $v(t)$ crosses the threshold value $\theta = 0$. We use a low-pass-filtered version of $v(t)$ to avoid false spike detections due to the additive noise. The low-pass filter is a 100-point moving-average smoother with a 0.001 second time step.

We rewrite (42)–(43) as

$$\dot{x}^1 = -\frac{x^1}{\epsilon} \left((x^1)^2 - \frac{1}{4} \right) - \frac{x^2}{\epsilon} + \frac{A}{\epsilon} + \frac{s(t)}{\epsilon} + \frac{n(t)}{\epsilon} \quad (46)$$

$$\dot{x}^2 = x^1 - x^2. \quad (47)$$

Here $x^1 = v$ and $x^2 = w$. The corresponding matrix Itô SDE is

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (48)$$

where $X_t = (X_t^1, X_t^2)^T$, $L_t = L_t^1$

$$b(X_{t-}) = \begin{bmatrix} b^1(X_{t-}^1, X_{t-}^2) \\ b^2(X_{t-}^1, X_{t-}^2) \end{bmatrix} = \begin{bmatrix} \frac{-X_{t-}^1}{\epsilon} \left((X_{t-}^1)^2 - \frac{1}{4} \right) - \frac{X_{t-}^2}{\epsilon} + \frac{A}{\epsilon} + \frac{s_t}{\epsilon} \\ X_{t-}^1 - X_{t-}^2 \end{bmatrix}$$

and

$$c(X_{t-}) = \begin{bmatrix} \sigma/\epsilon \\ 0 \end{bmatrix}$$

Thus all of the above spiking neuron models have the SDE form (23). Note that the drift term of the leaky integrate-and-fire neuron model is globally Lipschitz while the drift term of the reduced type I neuron model is locally Lipschitz. The Lipschitz condition is not easy to verify in the FHN model.

We now show that the drift term $b^1(X_t)$ in the preceding equation does not satisfy the global Lipschitz condition. Note that $b^1(X_t)$ is differentiable on \mathbf{R}^2 because the partial derivatives of $b^1(X_{t-}^1, X_{t-}^2)$ exist and are continuous on \mathbf{R}^2 . Suppose that $b^1(X_t)$ satisfies the following global Lipschitz condition: There exists a constant $K > 0$ such that

$$|b^1(Z_t) - b^1(Y_t)| \leq K \|Z_t - Y_t\|$$

for all Z_t and $Y_t \in \mathbf{R}^2$ and $t \in [0, T]$. Then the mean value theorem gives

$$\begin{aligned} b^1(Z_t) - b^1(Y_t) &= \begin{bmatrix} \frac{\partial b^1(\zeta)}{\partial X_t^1} & \frac{\partial b^1(\zeta)}{\partial X_t^2} \end{bmatrix} \cdot [Z_t - Y_t] \\ &= \frac{\partial b^1(\zeta)}{\partial X_t^1} (Z_t^1 - Y_t^1) + \frac{\partial b^1(\zeta)}{\partial X_t^2} (Z_t^2 - Y_t^2) \end{aligned} \quad (49)$$

for some ζ between Z_t and Y_t in \mathbf{R}^2 . Then

$$\begin{aligned} |b^1(Z_t) - b^1(Y_t)| &\geq \left| \frac{\partial b^1(\zeta)}{\partial X_t^1} \right| |Z_t^1 - Y_t^1| - \left| \frac{\partial b^1(\zeta)}{\partial X_t^2} \right| |Z_t^2 - Y_t^2| \\ &= \left| \frac{\partial b^1(\zeta)}{\partial X_t^1} \right| |Z_t^1 - Y_t^1| - \frac{1}{\epsilon} |Z_t^2 - Y_t^2| \\ &\quad \text{because } \frac{\partial b^1}{\partial X_t^2} = -\frac{1}{\epsilon} \\ &= \left| \frac{\partial b^1(\zeta)}{\partial X_t^1} \right| |Z_t^1 - Y_t^1| \quad \text{choosing} \\ &\quad Z_t \text{ and } Y_t \text{ such that } Z_t^2 = Y_t^2 \\ &= \left| \frac{\partial b^1(\zeta)}{\partial X_t^1} \right| \|Z_t - Y_t\| > K \|Z_t - Y_t\| \end{aligned}$$

for some $Z_t \in \mathbf{R}^2$ and $Y_t \in \mathbf{R}^2$ because $|\partial b^1/\partial X_t^1| = |-3(X_t^1)^2/\epsilon + 1/4\epsilon|$ is unbounded and continuous on \mathbf{R}^2 and so there is a domain $\mathcal{D} \subset \mathbf{R}^2$ such that $|\partial b^1(\zeta)/\partial X_t^1| > K$ for all $\zeta \in \mathcal{D}$. Thus $b^1(X_t)$ is not globally Lipschitz. So we cannot use Lemma 1 to prove the sufficient condition for the SR effect in the FHN neuron model (44)–(45).

But $b^1(X_t)$ is *locally* Lipschitz. The partial derivatives of $b^1(X_t^1, X_t^2)$ exist and are continuous on \mathbf{R}^2 . So $\partial b^1/\partial X_t^1$ and $\partial b^1/\partial X_t^2$ achieve their respective maxima on the compact set $\{\zeta \in \mathbf{R}^2 : \|\zeta\| \leq n\}$. Then (49) gives the required local Lipschitz condition

$$\begin{aligned} |b^1(Z_t) - b^1(Y_t)| \\ \leq \max \left\{ \sup_{\zeta} \left| \frac{\partial b^1(\zeta)}{\partial X_t^1} \right|, \sup_{\zeta} \left| \frac{\partial b^1(\zeta)}{\partial X_t^2} \right| \right\} \|Z_t - Y_t\| \\ = K'_n \|Z_t - Y_t\| \end{aligned}$$

for all Z_t and $Y_t \in \mathbf{R}^2$ such that $\|Z_t\| \leq n$, $\|Y_t\| \leq n$, and $\|\zeta\| \leq n$. Lemma 2 extends the conclusion of Lemma 1 to the locally Lipschitz drift terms $b^i(X_t)$.

Theorem 2 gives a “forbidden-interval” sufficient condition for a Levy SR noise benefits in spiking neuron models such as the leaky integrate-and-fire model [17], [28], the reduced type I neuron model [54], and the FHN model [26], [16]. It shows that these neuron models enjoy SR noise benefits if the noise mean μ falls to the left of a bound. Theorem 2 requires Lemma 2 to extend the conclusion of Lemma 1 to the locally Lipschitz drift terms $b^i(X_t)$. The Appendix gives the proof of Lemma 2.

Lemma 2: Let $b^i : \mathbf{R}^d \rightarrow \mathbf{R}$ and $c_j^i : \mathbf{R}^d \rightarrow \mathbf{R}$ in (23)–(24) [(40)–(43) for spiking neuron models] be measurable functions that satisfy the respective local and global Lipschitz conditions

$$|b^i(z) - b^i(y)| \leq C_n \|z - y\| \quad (50)$$

when $\|z\| \leq n$ and $\|y\| \leq n$

$$|c_j^i(z) - c_j^i(y)| \leq K_1 \|z - y\| \quad (51)$$

$$|c_j^i(z)|^2 \leq H_j^i \quad (52)$$

for all z and $y \in \mathbf{R}^d$, and for $i = 1, \dots, d$ and $j = 1, \dots, m$. Suppose

$$dX_t = b(X_t)dt + c(X_t)dL_t \quad (53)$$

$$d\hat{X}_t = b(\hat{X}_t)dt \quad (54)$$

where dL_t is a Levy noise with $\mu = \mathbf{0}$ and finite second moments. Then for every $T \in \mathbf{R}^+$ and for every $\epsilon > 0$

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \quad (55)$$

for all $j = 1, \dots, m$

and hence

$$P \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \epsilon \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \quad (56)$$

for all $j = 1, \dots, m$

because mean square convergence implies convergence in probability.

We can now state and prove Theorem 2.

Theorem 2: Suppose that the spiking neuron models (40)–(41) and (42)–(43) have the form of the Levy SDE (23) with a bounded globally Lipschitz Levy diffusion term $c(X_{t-}) \leq H$ and that the additive Levy noise has drift velocity μ . Suppose 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 the spiking neuron models (40)–(41) and (42)–(43) exhibit the SR effect in the sense that $I(S, R) \rightarrow 0$ as the Levy noise parameters $\sigma \rightarrow 0$ and $\nu \rightarrow 0$ if $H\mu < B - s_2$.

Proof: Let $\{\sigma_k, \nu_k\}_{k=1}^\infty$ be any decreasing sequence of Levy noise parameters such that $\sigma_k \rightarrow 0$ and $\nu_k \rightarrow 0$ as $k \rightarrow \infty$. Define $X(t)_k$ and R_k as the respective solution process and spiking rate process of the FHN spiking neuron model (48) with Levy noise parameters σ_k and ν_k instead of σ and ν .

Suppose that $\mu \neq 0$. We can absorb the drift $c(X_{t-})\mu$ into the input signal S because the Levy noise L_t is additive in all the neuron models. Then the new input signal $S' = S + c(X_{t-})\mu$ and this does not affect the Lipschitz continuity of $b(X_{t-})$ in (21). S' is subthreshold ($S' < B$) because $c(X_{t-})\mu < H\mu < B - s_2$ where $s_2 = \max\{s_1, s_2\}$. So we lose no generality if we consider the noise dL_t with $\mu = 0$ and let $S \in \{s_1, s_2\}$ be subthreshold in the continuous neuron model (21). This allows us to use Lemma 2.

Recall that $I(S, R) = 0$ if and only if S and R are statistically independent [20]. So we need to show only that $f_{SR}(s, r) = P_S(s)f_R(r)$ or $f_{R|S}(r|s) = f_R(r)$ as $\sigma \rightarrow 0$ and $\nu \rightarrow 0$ for signal symbols $s \in \{s_1, s_2\}$ and for all $r \geq 0$. Here f_{SR} is the joint probability density function and $f_{R|S}$ is the conditional density function. This is logically equivalent to $F_{R|S} = F_R$ as $\sigma_k \rightarrow 0$ and $\nu_k \rightarrow 0$ as $k \rightarrow 0$ where $F_{R|S}$ is the conditional distribution function [25]. Again the theorem of total probability and the two-symbol alphabet set $\{s_1, s_2\}$ give

$$\begin{aligned} F_R(r) &= \sum_s F_{R|S}(r|s)P_S(s) \\ &= F_{R|S}(r|s_1)P_S(s_1) + F_{R|S}(r|s_2)P_S(s_2) \\ &= F_{R|S}(r|s_1)P_S(s_1) + F_{R|S}(r|s_2)(1 - P_S(s_1)) \\ &= (F_{R|S}(r|s_1) - F_{R|S}(r|s_2))P_S(s_1) + F_{R|S}(r|s_2). \end{aligned} \quad (57)$$

So we need to show that $\lim_{k \rightarrow \infty} F_{R_k|S}(r|s_1) - F_{R_k|S}(r|s_2) = 0$ for all $r \geq 0$. This holds if and only if

$$\lim_{k \rightarrow \infty} P[R_k > r|S = s_1] - P[R_k > r|S = s_2] = 0. \quad (59)$$

We prove that $\lim_{k \rightarrow \infty} P[R_k > r|S = s_i] = 0$ for $i = 1$ and $i = 2$. Note that if $r > 0$ for (48) then $X^1(t)_k$ must cross the firing or spike threshold θ . Then

$$P[R_k > r|S = s_i] \leq P \left[\sup_{t_1 \leq t \leq t_2} X^1(t)_k > \theta | S = s_i \right].$$

Then Lemma 2 shows that the required limit goes to zero

$$\begin{aligned} & \lim_{k \rightarrow \infty} P[R_k > r|S = s_i] \\ & \leq \lim_{k \rightarrow \infty} P \left[\sup_{t_1 \leq t \leq t_2} X^1(t)_k > \theta | S = s_i \right] \\ & = \lim_{n \rightarrow \infty} P \left[\sup_{t_1 \leq t \leq t_2} X^1(t)_k > \theta, \hat{X}^1(t) < \theta | S = s_i \right] \\ & \quad \text{because } \hat{X}^1(t) \text{ converges to the FHN fixed point} \\ & \quad Z_{F_i} < \theta \quad \text{for large } t \\ & = 0 \quad \text{by Lemma 2.} \end{aligned}$$

Q.E.D.

Figs. 7(a) and (b) and 8(a) and (b) show simulation instances of Theorem 2 for finite-variance diffusion and jump-diffusion white Levy noise in the leaky integrate-and-fire and the FHN neuron models. Small amounts of additive Levy noise in these spiking neuron models produce the SR effect in terms of the noise-enhanced Shannon mutual information $I(S, Y)$ between realizations of a random (Bernoulli) subthreshold input signal S and the neuron's thresholded output random variable Y . The SR effects in Figs. 7(c) and 8(c) again lie outside the scope of Theorem 2 because they occur for infinite-variance α -stable noise and because Theorem 2 requires noise with finite second moments. Thus the SR effect in spiking neurons is not limited to finite-second-moment Levy noise.

VI. CONCLUSION

Levy noise processes can benefit several continuous and spiking neuron models because general forms of the SR "forbidden interval" theorem hold for several types of Levy noise. The generality of Levy noise extends simple Brownian models of noise to more complex and realistic Poisson jump models of noise that can affect biological and model neurons. But both Levy SR theorems require the finite-second-moment restrictions of the two lemmas. This rules out the important class of stable noise distributions in all but the Gaussian or pure-diffusion case.

Relaxing the second-moment assumption may produce SDEs that are not mathematically tractable. Yet the simulation evidence of Fig. 1 and Figs. 3(c), 4(c), 5(c), 6(c), 7(c), and 8(c) shows that the SR noise benefit continues to hold for several stable models where the noise has infinite variance and infinite higher order moments. It is an open research question whether a more general Levy SR result can include these and other observed noise benefits in continuous and spiking neuron models.

APPENDIX PROOFS OF LEMMAS

The proof of Lemma 2 relies on the proof technique of Lemma 1 in which we bound a mean squared term by four additive terms and then show that each of the four terms goes to zero in the limit.

Lemma 1: Let $b^i : \mathbf{R}^d \rightarrow \mathbf{R}$ and $c_j^i : \mathbf{R}^d \rightarrow \mathbf{R}$ in (23)–(24) be measurable functions that satisfy the global Lipschitz conditions

$$|b^i(x_1) - b^i(x_2)| \leq K_1 \|x_1 - x_2\| \quad (60)$$

$$|c_j^i(x_1) - c_j^i(x_2)| \leq K_2 \|x_1 - x_2\| \quad (61)$$

and

$$|c_j^i(x_1)|^2 \leq H_j^i \quad (62)$$

for all $x_1, x_2 \in \mathbf{R}^d$ and for $i = 1, \dots, d$ and $j = 1, \dots, m$. Suppose

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (63)$$

$$d\hat{X}_t = b(\hat{X}_t)dt \quad (64)$$

where dL_t is a Levy noise with $\mu = \mathbf{0}$ and finite second moments. Then for every $T \in \mathbf{R}^+$ and for every $\varepsilon > 0$

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \text{ for all } j = 1, \dots, m \quad (65)$$

and hence

$$P \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \varepsilon \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \text{ for all } j = 1, \dots, m \quad (66)$$

because mean square convergence implies convergence in probability.

Proof: The Lipschitz conditions (60) and (61) ensure that the process X_t exists [3] for $t \geq 0$ in (63). Then the proof commences with the inequality

$$\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \leq \sum_{i=1}^d \sup_{0 \leq t \leq T} (X_t^i - \hat{X}_t^i)^2 \quad (67)$$

which implies that

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \leq \sum_{i=1}^d E \left[\sup_{0 \leq t \leq T} (X_t^i - \hat{X}_t^i)^2 \right]. \quad (68)$$

Equations (26) and (63)–(64) imply

$$\begin{aligned} X_t^i - \hat{X}_t^i &= \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})]ds + \sum_{j=1}^m \int_0^t \sigma_j^i(X_{s-})dB_s^j \\ &+ \sum_{j=1}^m \int_0^t \int_{|y^j| < 1} F_j^i(X_{s-}, y^j) \tilde{N}^j(ds, dy^j) \\ &+ \sum_{j=1}^m \int_0^t \int_{|y^j| \geq 1} G_j^i(X_{s-}, y^j) N^j(ds, dy^j). \end{aligned} \quad (69)$$

This gives an upper bound on the squared difference as

$$\begin{aligned}
& (X_t^i - \hat{X}_t^i)^2 \\
& \leq (3m+1) \left(\left\{ \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})] ds \right\}^2 \right. \\
& \quad + \sum_{j=1}^m \left\{ \int_0^t \sigma_j^i(X_{s-}) dB_s^j \right\}^2 \\
& \quad + \sum_{j=1}^m \left\{ \int_0^t \int_{|y^j| < 1} F_j^i(X_{s-}, y^j) \tilde{N}^j(ds, dy^j) \right\}^2 \\
& \quad \left. + \sum_{j=1}^m \left\{ \int_0^t \int_{|y^j| \geq 1} G_j^i(X_{s-}, y^j) N^j(ds, dy^j) \right\}^2 \right) \quad (70)
\end{aligned}$$

because $(u_1 + \dots + u_n)^2 \leq n(u_1^2 + \dots + u_n^2)$. The Cauchy–Schwartz inequality gives

$$\begin{aligned}
& \left\{ \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})] ds \right\}^2 \\
& \leq \left(\int_0^t ds \right) \left(\int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})]^2 ds \right). \quad (71)
\end{aligned}$$

Now put (71) in the first term of (70) and then take expectations of the supremum on both sides to get four additive terms as an upper bound

$$\begin{aligned}
& E \left[\sup_{0 \leq t \leq T} (X_t^i - \hat{X}_t^i)^2 \right] \\
& \leq (3m+1) \left(E \left[\sup_{0 \leq t \leq T} t \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})]^2 ds \right] \right. \\
& \quad + \sum_{j=1}^m E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \sigma_j^i(X_{s-}) dB_s^j \right\}^2 \right] \\
& \quad + \sum_{j=1}^m E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \int_{|y^j| < 1} F_j^i(X_{s-}, y^j) \tilde{N}^j(ds, dy^j) \right\}^2 \right] \\
& \quad \left. + \sum_{j=1}^m E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \int_{|y^j| \geq 1} G_j^i(X_{s-}, y^j) N^j(ds, dy^j) \right\}^2 \right] \right). \quad (72)
\end{aligned}$$

We next show that each of the four terms goes to zero. Consider the first term on the right-hand side of (72)

$$\begin{aligned}
& E \left[\sup_{0 \leq t \leq T} t \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})]^2 ds \right] \\
& \leq TE \left[\sup_{0 \leq t \leq T} \int_0^t [b^i(X_{s-}) - b^i(\hat{X}_{s-})]^2 ds \right] \\
& \leq TK_1^2 E \left[\sup_{0 \leq t \leq T} \int_0^t \|X_{s-} - \hat{X}_{s-}\|^2 ds \right] \\
& \quad \text{by the Lipschitz condition (60)} \\
& \leq TK_1^2 \int_0^T E \left[\sup_{0 \leq u \leq s} \|X_{u-} - \hat{X}_{u-}\|^2 \right] ds. \quad (73)
\end{aligned}$$

The second term

$$\begin{aligned}
& E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \sigma_j^i(X_{s-}) dB_s^j \right\}^2 \right] \\
& \leq 4E \left[\left\{ \int_0^T \sigma_j^i(X_{s-}) dB_s^j \right\}^2 \right]
\end{aligned}$$

because $\int_0^t \sigma_j^i(X_{s-}) dB_s^j$ is a martingale and so we can apply Doob's L^p inequality [57]: $E[\sup_{a \leq t \leq b} |U_t|^p] \leq (p/(p-1))^p E|U_b|^p$ if $\{U_t\}_{t \geq 0}$ is a real-valued martingale, $[a, b]$ is a bounded interval of \mathbf{R}^+ , $U_t \in L^p(\Omega, \mathbf{R})$, and if $p > 1$ ($p = 2$ in our case). But

$$4E \left[\left\{ \int_0^T \sigma_j^i(X_{s-}) dB_s^j \right\}^2 \right] = 4 \int_0^T E[\{\sigma_j^i(X_{s-})\}^2] ds$$

by Itô isometry [3]

$$E \left(\left\{ \int_0^T f(s, w) dB_s \right\}^2 \right) = \int_0^T E(|f(s, w)|^2) ds$$

if $f \in \mathcal{H}_2([0, T])$ where $\mathcal{H}_2([0, T])$ is the space of all real-valued measurable $\{\mathcal{F}_t\}$ -adapted processes such that $E \left(\int_0^T |f(s, w)|^2 ds \right) < \infty$. Then

$$\begin{aligned}
& 4 \int_0^T E[\{\sigma_j^i(X_{s-})\}^2] ds \leq 4(\sigma^j)^2 \int_0^T E[\{c_j^i(X_{s-})\}^2] ds \\
& \quad \text{by definition of } \sigma_j^i(X_{s-}) \\
& \leq 4(\sigma^j)^2 TH_j^i \text{ because } |c_j^i|^2 \leq H_j^i. \quad (74)
\end{aligned}$$

Note that

$$\begin{aligned}
& E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \int_{|y^j| < 1} F_j^i(X_{s-}, y^j) \tilde{N}^j(ds, dy^j) \right\}^2 \right] \\
& \leq 4E \left[\left\{ \int_0^T \int_{|y^j| < 1} F_j^i(X_{s-}, y^j) \tilde{N}^j(ds, dy^j) \right\}^2 \right] \\
& \quad \text{by Doob's } L^p \text{ inequality} \\
& = 4E \left[\left\{ \int_0^T \int_{|y^j| < 1} c_j^i(X_{s-}) y^j \tilde{N}^j(ds, dy^j) \right\}^2 \right] \\
& \quad \text{by definition of } F_j^i(X_{s-}, y^j) \\
& \leq 4H_j^i E \left[\left\{ \int_0^T \int_{|y^j| < 1} y^j \tilde{N}^j(ds, dy^j) \right\}^2 \right] \\
& \quad \text{because } |c_j^i|^2 \leq H_j^i \\
& \leq 4H_j^i E \left[\left\{ \int_{|y^j| < 1} y^j \tilde{N}^j(T, dy^j) \right\}^2 \right] \\
& = 4H_j^i T \int_{|y^j| < 1} |y^j|^2 \nu^j(dy^j) \quad (75)
\end{aligned}$$

by Itô isometry and (20). Similar arguments and (19) give

$$E \left[\sup_{0 \leq t \leq T} \left\{ \int_0^t \int_{|y^j| \geq 1} G_j^i(X_{s-}, y^j) N^j(ds, dy^j) \right\}^2 \right] \leq 4H_j^i T \int_{|y^j| \geq 1} |y^j|^2 \nu^j(dy^j). \quad (76)$$

Substituting the estimates (73)–(76) in inequality (72) gives

$$E \left[\sup_{0 \leq t \leq T} (X_t^i - \hat{X}_t^i)^2 \right] \leq (3m + 1) \left(TK_1^2 \cdot \int_0^T E \left[\sup_{0 \leq u \leq s} \|X_{u-} - \hat{X}_{u-}\|^2 \right] ds + \sum_{j=1}^m 4TH_j^i \left[(\sigma^j)^2 + \int_{\mathbf{R}} |y^j|^2 \nu^j(dy^j) \right] \right). \quad (77)$$

Inequalities (68) and (77) imply that we can write

$$z(T) \leq A + Q \int_0^T z(s) ds$$

where

$$z(T) = E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \quad A = \sum_{i=1}^d \sum_{j=1}^m (3m+1) 4TH_j^i \left[(\sigma^j)^2 + \int_{\mathbf{R}} |y^j|^2 \nu^j(dy^j) \right] \quad (78)$$

and $Q = (3m + 1)dTK_1^2$. Then we get $z(T) \leq Ae^{QT}$ by Gronwall’s inequality [24]: $\phi(t) \leq \alpha e^{\beta t}$ for all $t \in [0, T]$ and for real continuous $\phi(t)$ in $[0, T]$ such that $\phi(t) \leq \alpha + \beta \int_0^t \phi(\tau) d\tau$ where $t \in [0, T]$ and $\beta > 0$. Note that $A \rightarrow 0$ as $\sigma^j \rightarrow 0$ and $\nu^j \rightarrow 0$. Hence

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \quad (79)$$

for each $j = 1, \dots, m$. This implies the claim (66). Q.E.D.

Lemma 2: Let $b^i : \mathbf{R}^d \rightarrow \mathbf{R}$ and $c_j^i : \mathbf{R}^d \rightarrow \mathbf{R}$ in (23)–(24) [(40)–(43) for spiking neuron models] be measurable functions that satisfy the respective local and global Lipschitz conditions

$$|b^i(z) - b^i(y)| \leq C_n \|z - y\| \quad (80)$$

when $\|z\| \leq n$ and $\|y\| \leq n$

$$|c_j^i(z) - c_j^i(y)| \leq K_1 \|z - y\| \quad (81)$$

$$|c_j^i(z)|^2 \leq H_j^i \quad (82)$$

for all z and $y \in \mathbf{R}^d$, and for $i = 1, \dots, d$ and $j = 1, \dots, m$. Suppose

$$dX_t = b(X_{t-})dt + c(X_{t-})dL_t \quad (83)$$

$$d\hat{X}_t = b(\hat{X}_t)dt \quad (84)$$

where dL_t is a Levy noise with $\mu = \mathbf{0}$ and finite second moments. Then for every $T \in \mathbf{R}^+$ and for every $\varepsilon > 0$

$$E \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \text{ for all } j = 1, \dots, m \quad (85)$$

and hence

$$P \left[\sup_{0 \leq t \leq T} \|X_t - \hat{X}_t\|^2 > \varepsilon \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \text{ for all } j = 1, \dots, m \quad (86)$$

because mean square convergence implies convergence in probability.

Proof: Define the function \tilde{b}_r^i such that

- i) $\tilde{b}_r^i(x) = b^i(x)$ for $\|x\| \leq r$;
- ii) $\tilde{b}_r^i(x) = 0$ for $\|x\| \geq 2r$;
- iii) $\tilde{b}_r^i(x) = ((2r - \|x\|)/r)b^i(rx/\|x\|)$ for $r \leq \|x\| \leq 2r$.

We then show that the function \tilde{b}_r^i is globally Lipschitz: $|\tilde{b}_r^i(x) - \tilde{b}_r^i(y)| \leq C_r^i \|x - y\|$ for all $x, y \in \mathbf{R}^d$.

Consider the function $\tilde{b}_r^i(x)$. Write

$$\tilde{b}_r^i(x) = \begin{cases} b^i(x), & \text{if } \|x\| \leq r \\ f(x)g^i(x), & \text{if } r \leq \|x\| \leq 2r \end{cases} \quad (87)$$

where

$$f(x) = \left(\frac{2r - \|x\|}{r} \right) \quad \text{and} \quad g^i(x) = b^i \left(\frac{rx}{\|x\|} \right). \quad (88)$$

The definition of \tilde{b}_r^i implies that it is Lipschitz continuous on the region $D_1 = \{\|x\| \leq r\}$

$$\|\tilde{b}_r^i(x) - \tilde{b}_r^i(y)\| \leq C_r \|x - y\|, \quad \text{for all } x, y \in D_1. \quad (89)$$

We first show that $\tilde{b}_r^i(x)$ is Lipschitz continuous on the region $D_2 = \{r \leq \|x\| \leq 2r\}$. For $x, y \in D_2 = \{r \leq \|x\| \leq 2r\}$

$$\begin{aligned} |f(x) - f(y)| &= \frac{\| \|y\| - \|x\| \|}{r} \quad \text{by definition of } f \\ &\leq \frac{\|x - y\|}{r} \end{aligned} \quad (90)$$

and

$$\begin{aligned} |g^i(x) - g^i(y)| &= \left| b^i \left(\frac{rx}{\|x\|} \right) - b^i \left(\frac{ry}{\|y\|} \right) \right| \\ &\quad \text{by definition of } g^i \end{aligned} \quad (91)$$

$$\leq C_r \left\| \frac{rx}{\|x\|} - \frac{ry}{\|y\|} \right\|$$

because $\frac{rs}{\|s\|} \in D_1$ for all $s \in \mathbf{R}^d$

and b^i is Lipschitz continuous on D_1 (92)

$$\leq \frac{C_r}{2} \|x - y\| \text{ because } r \geq \|x\|, \|y\| \geq 2r. \quad (93)$$

Hence

$$\begin{aligned} & |\tilde{b}_r^i(x) - \tilde{b}_r^i(y)| \\ &= |f(x)g^i(x) - f(y)g^i(y)| \end{aligned} \quad (94)$$

$$\leq |f(x)g^i(x) - f(z)g^i(z)| + |f(z)g^i(z) - f(y)g^i(y)| \quad (95)$$

$$= |f(x)g^i(x) - f(x)g^i(y)| + |f(x)g^i(y) - f(y)g^i(y)| \quad (96)$$

by choosing z on the line segment between 0 and y such that $\|z\| = \|x\|$

$$= |f(x)||g^i(x) - g^i(y)| + |g^i(y)||f(x) - f(y)| \quad (97)$$

$$\leq \|f\|_{\infty,2} \|g^i(x) - g^i(y)\| + \|g\|_{\infty,2} \|f(x) - f(y)\| \quad (98)$$

where we define $\|v\|_{\infty,i} = \sup\{\|v(s)\| : s \in D_i\}$

$$\leq \|f\|_{\infty,2} \frac{C_r}{2} \|x - y\| + \|g\|_{\infty,2} \frac{\|x - y\|}{r} \quad (99)$$

$$\leq C'_r \|x - y\| \quad \text{where } C'_r = \|f\|_{\infty,2} \frac{C_r}{2} + \frac{\|g\|_{\infty,2}}{r}. \quad (100)$$

So $\tilde{b}^r(x)$ is Lipschitz continuous on D_2 .

We next show that $\tilde{b}^r(x)$ is Lipschitz continuous on D_1 and D_2 . Choose $x \in D_1, y \in D_2$, and a point z of ∂D_1 on the line segment between x and y . Then

$$|\tilde{b}_r^i(x) - \tilde{b}_r^i(y)| \leq |\tilde{b}_r^i(x) - \tilde{b}_r^i(z)| + |\tilde{b}_r^i(z) - \tilde{b}_r^i(y)| \quad (101)$$

$$\leq C_r \|x - z\| + C'_r \|z - y\| \quad (102)$$

$$\leq C'_r \|x - y\| \quad (103)$$

because $C'_r \geq C_r$ and $\|x - z\| + \|z - y\| = \|x - y\|$. So $\tilde{b}^r(x)$ is Lipschitz continuous with coefficient C'_r on $\|x\| \leq 2r$. Choose $x \in (D_1 \cup D_2), y \in (D_1 \cup D_2)^c$ and a point z of $\partial(D_1 \cup D_2)^c$ on the line segment between x and y . Then

$$|\tilde{b}_r^i(x) - \tilde{b}_r^i(y)| \leq |\tilde{b}_r^i(x) - \tilde{b}_r^i(z)| + |\tilde{b}_r^i(z) - \tilde{b}_r^i(y)| \quad (104)$$

$$\leq C_r \|x - z\| + 0 \quad (105)$$

$$\leq C'_r \|x - z\| + C'_r \|z - y\| \quad (106)$$

$$= C'_r \|x - y\|. \quad (107)$$

Then (89), (100), (103), and (107) show that $\tilde{b}_r^i(x)$ is Lipschitz continuous with coefficient C'_r on \mathbf{R}^d .

Consider next the SDE

$$d\tilde{X}_t = \tilde{b}_r(\tilde{X}_t)dt + c(\tilde{X}_t)dL_t. \quad (108)$$

Lemma 1 holds for (108) and so we can write

$$E \left[\sup_{0 \leq t \leq \tilde{T}_r} \|\tilde{X}_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \quad (109)$$

for all $j = 1, \dots, m$ where $\tilde{T}_r = \inf\{t \geq 0 : \|\tilde{X}_t\| \geq r\}$ and we choose r such that $\|\hat{X}_t\| < r$ for all t . Now define $T_r = \inf\{t \geq 0 : \|X_t\| \geq r\}$ and $\tau_r = \inf\{t : \|X_t\| \geq r \text{ or } \|\hat{X}_t\| \geq r\} = \min\{\tilde{T}_r, T_r\}$. Then X_t and \hat{X}_t satisfy (83) on $[0, \tau_r]$. Note that \tilde{T}_r and T_r are stopping times and thus τ_r is

also a stopping time. So arguments similar to those of the proof of Lemma 1 ((68)–(76) with appropriate modifications) give

$$E \left[\sup_{0 \leq u \leq \min\{t, \tau_r\}} \|X_u - \hat{X}_u\|^2 \right] \leq Q' \int_0^t E \left[\sup_{0 \leq u \leq \min\{s, \tau_r\}} \|X_u - \hat{X}_u\|^2 \right] ds. \quad (110)$$

Then

$$E \left[\sup_{0 \leq u \leq \min\{t, \tau_r\}} \|X_u - \hat{X}_u\|^2 \right] \leq 0 \quad (111)$$

by Gronwall's inequality. Hence $X_t = \hat{X}_t$ holds almost surely on $[0, \tau_r]$. This result and (109) give

$$E \left[\sup_{0 \leq t \leq \tau_r} \|X_t - \hat{X}_t\|^2 \right] \rightarrow 0 \text{ as } \sigma^j \rightarrow 0 \text{ and } \nu^j \rightarrow 0 \quad (112)$$

for all $j = 1, \dots, m$.

We need to show only that $\tau_r \rightarrow \infty$ almost surely as $r \rightarrow \infty$ to prove (86). Let $X_{\min\{T, \tau_r\}}$ be the value of the process X_t at time $\min\{T, \tau_r\}$. Note first that $\|X_{\min\{T, \tau_r\}}\|^2 = (I_{\{\tau_r > T\}} \|X_T\|^2 + I_{\{\tau_r \leq T\}} \|X_{\tau_r}\|^2)$. Then

$$\begin{aligned} E[\|X_{\min\{T, \tau_r\}}\|^2] &= E \left[(I_{\{\tau_r > T\}} \|X_T\|^2) \right] \\ &\quad + E \left[(I_{\{\tau_r \leq T\}} \|X_{\tau_r}\|^2) \right] \\ &= E \left[(I_{\{\tau_r > T\}} \|X_T\|^2) \right] + P(\tau_r \leq T) r^2 \end{aligned}$$

because $\|X_{\tau_r}\| = r$. Therefore

$$P(\tau_r \leq T) \leq \frac{E[\|X_{\min\{T, \tau_r\}}\|^2]}{r^2}. \quad (113)$$

Applying Itô's lemma [3] to $\|X_{\min\{T, \tau_r\}}\|^2$ gives

$$E[\|X_{\min\{T, \tau_r\}}\|^2] \leq A'' + Q'' \int_0^T E[\|X_{\min\{s, \tau_r\}}\|^2] ds. \quad (114)$$

Thus

$$E[\|X_{\min\{T, \tau_r\}}\|^2] \leq A'' e^{Q'' T} \quad (115)$$

by Gronwall's inequality where A'' and Q'' do not depend on r because we do not use the Lipschitz condition in the derivation of (114). Then (113) and (115) imply that

$$P(\tau_r \leq T) \leq \frac{E[\|X_{\min\{T, \tau_r\}}\|^2]}{r^2} \rightarrow 0 \text{ as } r \rightarrow \infty. \quad (116)$$

Thus $\tau_r \rightarrow \infty$ almost surely as $r \rightarrow \infty$. This implies the claim (85). Q.E.D.

REFERENCES

- [1] H. Ahn and R. E. Feldman, "Optimal filtering of Gaussian signals in the presence of Levy noise," *SIAM J. Appl. Math.*, vol. 60, no. 1, pp. 359–369, Nov. 1999.
- [2] S. Amari, "Neural theory of association and concept formation," *Biol. Cybern.*, vol. 26, pp. 175–185, May 1977.
- [3] D. Applebaum, *Levy Processes and Stochastic Calculus*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[4] O. Barndorff-Nielsen and N. Shephard, "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics," *J. Roy. Statist. Soc. B*, vol. 63, no. 2, pp. 167–241, 2001.

[5] S. M. Bezrukov and I. Vodyanov, "Noise-induced enhancement of signal transduction across voltage-dependent ion channels," *Nature*, vol. 378, no. 6555, pp. 362–364, Nov. 1995.

[6] A. R. Bulsara, E. W. Jacobs, T. Zhou, F. Moss, and L. Kiss, "Stochastic resonance in a single neuron model: Theory and analog simulation," *J. Theor. Biol.*, vol. 152, no. 4, pp. 531–555, Oct. 1991.

[7] A. R. Bulsara, A. J. Maren, and G. Schnera, "Single effective neuron: Dendritic coupling effect and stochastic resonance," *Biol. Cybern.*, vol. 70, no. 2, pp. 145–156, Dec. 1993.

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

[9] J. Cao and J. Lu, "Adaptive synchronization of neural networks with and without time-varying delays," *Chaos*, vol. 16, pp. 013133-1–013133-6, Mar. 2006.

[10] J. Cao, Z. Wang, and Y. Sun, "Synchronization in an array of linearly stochastically coupled networks with time delays," *Physica A*, vol. 385, pp. 718–728, Nov. 2007.

[11] F. Chapeau-Blondeau, X. Godivier, and N. Chambet, "Stochastic resonance in a neuron model that transmits spike trains," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 53, no. 1, pp. 1273–1275, Jan. 1996.

[12] F. Chapeau-Blondeau and D. Rousseau, "Noise-enhanced performance for an optimal Bayesian estimator," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1327–1334, May 2004.

[13] F. Chapeau-Blondeau, S. Blanchard, and D. Rousseau, "Noise-enhanced fisher information in parallel arrays of sensors with saturation," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 74, pp. 0311021–03110210, Sep. 2006.

[14] M. Chatterjee and M. E. Robert, "Noise enhances modulation sensitivity in cochlear implant listeners: Stochastic resonance in a prosthetic sensory system?," *J. Assoc. Res. Otolaryngol.*, vol. 2, no. 2, pp. 159–171, Aug. 2001.

[15] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-13, no. 5, pp. 815–826, Sep./Oct. 1983.

[16] J. J. Collins, C. C. Chow, A. C. Capela, and T. T. Imhoff, "Aperiodic stochastic resonance in excitable systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 52, no. 4, pp. R3321–R3324, Oct. 1995.

[17] J. J. Collins, C. C. Chow, A. C. Capela, and T. T. Imhoff, "Aperiodic stochastic resonance," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 54, no. 5, pp. 5575–5584, Nov. 1996.

[18] J. J. Collins, T. T. Imhoff, and P. Grigg, "Noise enhanced information transmission in rat SA1 cutaneous mechanoreceptors via aperiodic stochastic resonance," *J. Neurophysiol.*, vol. 76, pp. 642–645, Jul. 1996.

[19] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*. London, U.K.: Chapman and Hall/CRC Press, 2003.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[21] G. Deco and B. Schurmann, "Stochastic resonance in the mutual information between input and output spike trains of noisy central neurons," *Physica D*, vol. 117, pp. 276–282, Jun. 1998.

[22] T. Ditzinger, M. Stadler, D. Struber, and J. A. Kelso, "Noise improves three-dimensional perception: Stochastic resonance and other impacts of noise to the perception of autostereograms," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, pp. 2566–2575, Aug. 2000.

[23] J. K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, "Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance," *Nature*, vol. 365, pp. 337–340, Sep. 1993.

[24] R. Durrett, *Stochastic Calculus: A Practical Introduction*. Boca Raton, FL: CRC Press, 1996.

[25] R. Durrett, *Probability: Theory and Examples*. Belmont, CA: Duxbury Press, 2004.

[26] R. A. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophys. J.*, vol. 1, pp. 445–466, Jul. 1961.

[27] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 52, no. 5, pp. 4691–4698, Nov. 1995.

[28] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[29] M. T. Giraud and L. Sacerdote, "Jump-diffusion processes as models of neuronal activity," *BioSystems*, vol. 40, no. 1-2, pp. 75–82, 1997.

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

[31] B. S. Gutkin and G. B. Ermentrout, "Dynamics of membrane excitability determine interspike interval variability: A link between spike generation mechanism and cortical spike train statistics," *Neural Comput.*, vol. 10, no. 5, pp. 1047–1065, Jul. 1998.

[32] P. Hanggi, "Stochastic resonance in biology: How noise can enhance detection of weak signals and help improve biological information processing," *Chemphyschem*, vol. 3, pp. 285–290, Mar. 2002.

[33] A. Hanssen and T. A. Oigard, "The normal inverse Gaussian distribution: A versatile model for heavy-tailed stochastic processes," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2001, pp. 3985–3988.

[34] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, pp. 500–540, Aug. 1952.

[35] N. Hohn and A. N. Burkitt, "Shot noise in leaky integrate-and-fire neuron," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 63, pp. 031902-1–031902-11, Mar. 2001.

[36] F. C. Hoppensteadt and E. M. Izhikevich, *Weakly Connected Neural Networks*. New York: Springer-Verlag, pp. 111–142.

[37] J. J. Hopfield, "Neural networks with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci.*, vol. 81, pp. 3088–3092, May 1984.

[38] H. Huang and J. Cao, "Exponential stability analysis of uncertain stochastic neural networks with multiple delays," *Nonlinear Anal.: Real World Appl.*, vol. 8, pp. 646–653, Apr. 2007.

[39] M. E. Inchiosa, J. W. C. Robinson, and A. R. Bulsara, "Information-theoretic stochastic resonance in noise-floor limited systems: The case for adding noise," *Phys. Rev. Lett.*, vol. 85, pp. 3369–3372, Oct. 2000.

[40] F. Jaramillo and K. Wiesenfeld, "Mechano-electrical transduction assisted by Brownian motion: A role for noise in the auditory system," *Nature Neurosci.*, vol. 1, pp. 384–388, 1998.

[41] P. Jung, A. Cornell-Bell, F. Moss, S. Kadar, J. Wang, and K. Showalter, "Noise sustained waves in subexcitable media: From chemical waves to brain waves," *Chaos*, vol. 8, no. 3, pp. 567–575, Sep. 1998.

[42] A. Kalemnova, B. Schmid, and R. Werner, "The normal inverse Gaussian distribution for synthetic CDO pricing," *J. Financial Derivatives*, vol. 14, pp. 80–93, 2007.

[43] G. G. Karapetyan, "Application of stochastic resonance in gravitational-wave interferometer," *Phys. Rev. D, Part. Fields*, vol. 73, pp. 122003-1–122003-9, Jun. 2006.

[44] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer-Verlag, 1988.

[45] S. Kay, "Can detectability be improved by adding noise?," *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.

[46] E. R. Kandel, J. H. Schwartz, and M. J. Thomas, *Principles of Neuroscience*, 4th revised ed. New York: McGraw-Hill, 2000, pp. 224–224.

[47] A. F. Kohn, "Dendritic transformations on random synaptic inputs as measured from neuron's spike train-modeling and simulation," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 1, pp. 44–54, Jan. 1989.

[48] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

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

[50] B. Kosko and S. Mitaim, "Stochastic resonance in noisy threshold neurons," *Neural Netw.*, vol. 14, no. 6-7, pp. 755–761, Jun. 2003.

[51] B. Kosko and S. Mitaim, "Robust stochastic resonance for simple threshold neurons," in *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, Sep. 27, 2004, vol. 70, pp. 031911-1–031911-10.

[52] I. Y. Lee, X. Liu, C. Zhou, and B. Kosko, "Noise-enhanced detection of subthreshold signals with carbon nanotubes," *IEEE Trans. Nanotechnol.*, vol. 5, no. 6, pp. 613–627, Nov. 2006.

[53] J. E. Levin and J. P. Miller, "Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance," *Nature*, vol. 380, pp. 165–168, Mar. 1996.

[54] B. Lindner, A. Longtin, and A. R. Bulsara, "Analytic expressions for rate and CV of a type I neuron driven by white Gaussian noise," *Neural Comput.*, vol. 15, pp. 1761–1788, Aug. 2003.

[55] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier, "Effects of noise in excitable systems," *Phys. Rep.*, vol. 392, pp. 321–424, Mar. 2004.

- [56] A. Longtin, "Autonomous stochastic resonance in bursting neurons," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 55, no. 1, pp. 868–876, Jan. 1997.
- [57] X. Mao, *Stochastic Differential Equations and their Applications*. Chichester, U.K.: Horwood, 1997.
- [58] M. D. McDonnell, N. Stocks, C. E. M. Pearce, and D. Abbott, "Quantization in the presence of large amplitude threshold noise," *Fluctuation Noise Lett.*, vol. 5, no. 3, pp. L457–L468, Sep. 2005.
- [59] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
- [60] S. Mitaim and B. Kosko, "Adaptive stochastic resonance," *Proc. IEEE*, vol. 86, Special Issue on Intelligent Signal Processing, no. 11, pp. 2152–2183, Nov. 1998.
- [61] S. Mitaim and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1526–1540, Nov. 2004.
- [62] F. Moss, L. M. Ward, and W. G. Sannita, "Stochastic resonance and sensory information processing: A tutorial and review of applications," *Clin. Neurophysiol.*, vol. 115, no. 2, pp. 267–281, Feb. 2004.
- [63] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*. New York: Wiley, 1995.
- [64] A. Pacut, "Separation of deterministic and stochastic neurotransmission," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2001, vol. 1, pp. 55–60.
- [65] A. Patel and B. Kosko, "Stochastic resonance in noisy spiking retinal and sensory neuron models," *Neural Netw.*, vol. 18, pp. 467–478, Aug. 2005.
- [66] A. Patel and B. Kosko, "Mutual-information noise benefits in Brownian models of continuous and spiking neuron models," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2006, pp. 2347–2354.
- [67] O. V. Poliannikov, Y. Bao, and H. Krim, "Levy processes for image modelling," in *Proc. IEEE Signal Process. Workshop Higher-Order Statist.*, Caesarea, Israel, Jun. 14–16, 1999, pp. 233–236.
- [68] S. I. Resnick, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer-Verlag, 2007.
- [69] D. Rousseau, J. Rojas Varela, and F. Chapeau-Blondeau, "Stochastic resonance for nonlinear sensors with saturation," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 67, no. 2, pp. 0211021–02110216, Feb. 2003.
- [70] D. Rousseau and F. Chapeau-Blondeau, "Neuronal signal transduction aided by noise at threshold and at saturation," *Neural Process. Lett.*, vol. 20, no. 2, pp. 71–83, Nov. 2004.
- [71] D. Rousseau and F. Chapeau-Blondeau, "Constructive role of noise in signal detection from parallel array of quantizers," *Signal Process.*, vol. 85, no. 3, pp. 571–580, Mar. 2005.
- [72] D. Rousseau, G. V. Anand, and F. Chapeau-Blondeau, "Noise-enhanced nonlinear detector to improve signal detection in non-Gaussian noise," *Signal Process.*, vol. 86, no. 11, pp. 3456–3465, Nov. 2006.
- [73] T. Rydberg, "The normal inverse Gaussian Levy processes: Simulation and approximation," *Commun. Statist.: Stochastic Models*, vol. 13, pp. 887–910, 1997.
- [74] L. Sacerdote and R. Sirovich, "Multimodality of the interspike interval distribution in a simple jump-diffusion model," *Scientiae Mathematicae Japonicae*, vol. 58, no. 2, pp. 307–322, Sep. 2003.
- [75] Y. Sakumara and K. Aihara, "Stochastic resonance and incidence detection in single neuron," *Neural Process. Lett.*, vol. 16, pp. 235–242, Dec. 2002.
- [76] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. London, U.K.: Chapman & Hall/CRC Press, 1994.
- [77] K-I. Sato, *Levy Processes and Infinitely Divisible Distributions*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [78] W. Schoutens, *Levy Processes in Finance: Pricing Financial Derivatives*. New York: Wiley, 2003.
- [79] W. Schoutens, "Exotic options under Levy models: An overview," *J. Comput. Appl. Math.*, vol. 189, no. 1, pp. 526–538, May 2006.
- [80] R. Segev, M. Benveniste, E. Hulata, N. Cohen, A. Palevski, E. Kapon, Y. Shapira, and E. Ben-Jacob, "Long term behavior of lithographically prepared in vitro neuronal networks," *Phys. Rev. Lett.*, vol. 88, no. 11, pp. 118102-1–118102-4, Mar. 2002.
- [81] M. F. Shlesinger, G. M. Zaslavskii, and U. Frisch, Eds., "Levy flights and related topics in physics," in *Lecture Notes in Physics*, ser. 450. Berlin, Germany: Springer-Verlag.
- [82] O. Sotolongo-Costa, J. C. Antoranz, A. Posadas, F. Vidal, and A. Vazquez, "Lévy flights and earthquakes," *Geophys. Res. Lett.*, vol. 27, pp. 1965–1967, Jul. 2000.
- [83] W. C. Stacey and D. M. Durand, "Stochastic resonance improves signal detection in hippocampal CA1 neurons," *J. Neurophysiol.*, vol. 83, pp. 1394–1402, Sep. 2001.
- [84] N. G. Stocks, "Information transmission in parallel threshold arrays: Suprathreshold stochastic resonance," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 63, no. 4, pp. 041114-1–041114-9, Mar. 2001.
- [85] N. G. Stocks, D. Appligham, and R. P. Morse, "The application of suprathreshold stochastic resonance to cochlear implant coding," *Fluctuation Noise Lett.*, vol. 2, no. 3, pp. L169–L181, Sep. 2002.
- [86] Y. Sun and J. Cao, "Adaptive lag synchronization of unknown chaotic delayed neural networks with noise perturbations," *Phys. Lett. A, Gen. Phys.*, vol. 364, pp. 277–285, Apr. 2007.
- [87] Z. Wang, Y. Liu, M. Li, and X. Liu, "Stability analysis for stochastic Cohen-Grossberg neural networks with mixed time delays," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 814–820, May 2006.
- [88] Z. Wang, Y. Liu, K. Fraser, and X. Liu, "Stochastic stability of uncertain Hopfield neural networks with discrete and distributed delays," *Phys. Lett. A, Gen. Phys.*, vol. 354, no. 4, pp. 288–297, Jun. 2006.
- [89] 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, Jan. 1995.
- [90] W. Yu and J. Cao, "Synchronization control of stochastic delayed neural network," *Physica A*, vol. 373, pp. 252–260, Jan. 2007.
- [91] F. G. Zeng, Q. J. Fu, and R. Morse, "Human hearing enhanced by noise," *Brain Res.*, vol. 869, no. 1–2, pp. 251–255, Jun. 2000.



Ashok Patel received the B.E. and M.E. degrees in electrical engineering from Gujarat University, Ahmedabad, India, and the M.A. degree in applied mathematics and the M.S. degree in electrical engineering from University of Southern California, Los Angeles, where he is currently working towards the Ph.D. degree at the Department of Electrical Engineering.

His current research interests are in the areas of nonlinear and statistical signal processing.



Bart Kosko holds degrees in philosophy, economics, mathematics, electrical engineering, and law.

Currently he is a Professor of Electrical Engineering at the University of Southern California (USC), Los Angeles. He is a licensed attorney and former director of USC's Signal and Image Processing Institute. He is author of the trade books *Noise, Heaven in a Chip*, and *Fuzzy Thinking*, as well as the novel *Nanotime*. He has authored the textbooks *Fuzzy Engineering* and *Neural Networks and Fuzzy Systems*, edited the volume *Neural Networks for Signal Processing*, and co-edited the volume *Intelligent Signal Processing*. He is author of the trade books *Fuzzy Thinking, Heaven in a Chip*, and the novel *Nanotime*. His most recent book is *Noise*.

Dr. Kosko is an elected governor of the International Neural Network Society. He has chaired or cochaired several international conferences on neural and fuzzy systems conferences and sits on the editorial board of several scientific journals.