LEVY NOISE BENEFITS IN NEURAL SIGNAL DETECTION

Ashok Patel and Bart Kosko

University of Southern California Viterbi School of Engineering, Department of Electrical Engineering Los Angeles, CA 90089-2564, USA

ABSTRACT

We use the Ito calculus to prove that a general type of 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. Levy noise generalizes Brownian motion and includes several important jump and impulsive random processes often found in neural and financial-engineering models. A global Lipschitz condition implies 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. Simulation results show that the same 'stochastic resonance' noise benefit occurs for at least some impulsive or infinite-variance (stable) Levy noise processes.

Index Terms— Levy noise, stochastic resonance, neural signal detection, mutual information, jump diffusion

1. NOISE BENEFITS IN NEURAL SIGNAL DETECTION

Small amounts of noise can enhance some forms of nonlinear signal processing while too much noise degrades the signal processing [6, 14, 19, 20, 26, 30]. Such stochastic resonance (SR) noise benefits arise in many physical and biological systems [12, 18, 22, 31]. The Figure 2 panels show the characteristic nonmonotonic signature of SR when the noise additively perturbs a subthreshold stochastic neuron model for the three white Levy noise types whose samples appear in Figure 1.

We generalize the recent 'forbidden interval' theorems [16, 17, 21, 24, 25] to the case of Levy noise with finite second moments. Those theorems state that simple threshold neurons will have an SR noise benefit if and only if the mean noise level does not fall in a threshold-related interval. The theorem below shows that such an SR noise benefit will occur if the additive white Levy noise process has a scaled drift velocity that falls within a threshold-based interval. This holds for general feedback neuron models that include common signal functions such as logistic sigmoids or Gaussians. The result requires a finite second moment. The last SR plot in Figure 2 shows that a noise benefit still occurs in the more general infinite-variance case of at least some types of α -stable Levy noise.

Levy noise has advantages over standard Gaussian noise in neuron models. Adding white Levy noise more accurately describes how the neuron's membrane potential evolves than do simpler diffusion models because the more general Levy model includes not only pure-diffusion and pure-jump models but jump-diffusion models as well [28]. 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. This requires at least that the number of impinging synapses is large and that they have small membrane effects due to the small coupling coefficient or the synaptic weights [10]. So the Gaussian noise assumption is more accurate for signal inputs from dendritic trees. Fewer inputs come from synapses near the trigger zone and they have large amplitudes due to their voltage sensitive sodium channels [11].

2. LEVY NOISE AND CONTINUOUS NEURON MODELS

We consider the noisy dynamical neuron models of the general form

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

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

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 , g is a static (usually threshold) transformation function, and y(t) is the neuron's output. The neuron feeds its activation or membrane potential signal x(t) back to itself through -x(t) + f(x(t)) and emits the thresholded signal y(t) as output. The neuronal signal function f(x) can be quite general. n(t) is the additive white Levy noise with intensity scale κ . We next describe Levy processes and the neuronal signal function f(x).

Levy processes [29] form a wide class of random processes that include Brownian motion, α -stable motion, compound Poisson processes, generalized inverse Gaussian processes, and generalized Hyperbolic processes. Levy processes can account for the impulsiveness or discreteness of the signal. Researchers have used Levy processes to model diverse phenomena such as financial data, network traffic, acoustic signals, and storage processes [1, 4, 23]. A Levy process is a realvalued adapted stochastic process with stationary and independent increments in a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{0 \le t \le \infty})$ where \mathcal{F}_0 contains all the *P*-null sets of \mathcal{F} and (\mathcal{F}_t) is right continuous. So a real-valued adapted process $L_t = (L(t), t \ge 0)$ with L(0) = 0 (a.s) is a Levy process if

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

Any Levy process has a specific (infinitely divisible) form for its characteristic function. The Levy-Khintchine formula gives the characteristic function of L(t) as

$$E(e^{iuL(t)}) = e^{t\eta(u)} \text{ for all } 0 \le t, u \in \mathbf{R}$$
(3)

with the characteristic exponent or Levy exponent

$$\eta(u) = iu\mu - \frac{1}{2}u^2\sigma^2 u + \int_{\mathbf{R}_0} \left(e^{iuy} - 1 - \frac{iuy}{1+y^2}\right)\nu(dy) \quad (4)$$

for some $\mu \in \mathbf{R}$, $\sigma \geq 0$, and measure ν on Borel subsets of $\mathbf{R}_0 = \mathbf{R} \setminus \{0\}$. ν is a Levy measure and has the property that $\int_{\mathbf{R}_0} \min\{1, |y|^2\} \nu(dy) < \infty$.

A Levy process has a drift, a Brownian motion, and a jump component. The Levy-Khintchine triplet (μ, σ, ν) completely determines these components. The Levy measure ν defines the behavior of the jump component of L(t) and determines the frequency and magnitudes of jumps. Jumps of any size in Borel set B form a compound Poisson process with rate $\int_B \nu(dy)$ and jump density $\hat{\nu}(dy) / \int_B \nu(dy)$ if the closure \overline{B} does not contain zero. If $\nu = 0$ then L(t) is a continuous Brownian motion for $t \ge 0$ because then equation (3) takes the form $E(e^{iuL_t}) = e^{t[iu\mu - \frac{1}{2}u^2\sigma^2]}$ —the characteristic function of a Gaussian random variable with mean μt and variance $t\sigma^2$. If $\nu(\mathbf{R}) < \infty$ then L_t is a compound Poisson process. If $\nu(\mathbf{R}) = \infty$ then L_t is a purely discontinuous jump process and has an infinite number of small jumps in any time interval of positive length. Figure 2 shows one-dimensional samples of such Levy processes. We assume one-dimensional Levy processes with a *finite* second moment $(E|L_t|^2 < \infty)$. A finite second moment assumption excludes the thick-tailed family of pure-jump infinite-variance α -stable processes (including the $\alpha = 0.5$ Levy stable case) [23]. But this assumption does not imply that the Levy measure is finite ($\nu(\mathbf{R}) < \infty$). Normal inverse Gaussian NIG $(\alpha, \beta, \gamma, \mu)$ distributions are examples of semi-thick-tailed pure-jump Levy processes that have infinite Levy measure and possess moments of all order [27].

We can rewrite (1)-(2) in more general form as the Ito stochastic differential equation [3]

$$dX_t = b(X_{t^-})dt + c(X_{t^-})dL_t$$
(5)

$$Y_t = g(X_t) \tag{6}$$

for initial condition $X_{t_0} = X_0$ where $b(X_{t^-}) = -X_{t^-} + f(X_{t^-}) + S_t$ is a Lipschitz continuous drift term, $c(X_{t^-}) = a$ is a constant Levy diffusion term, and dL_t is a white Levy noise with noise scale κ . Neuronal signal function f has the general form that includes most common signal functions:



Fig. 1. Sample paths of 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] (so $\nu(dy) = (3/0.4)dy$ for $y \in [-0.2, 0.2]$ and zero else), (c) Normal Inverse Gaussian (NIG) process with α arameter set ($\alpha = 20$; $\beta = 0$, $\gamma = 0.1$; $\mu = 0$), (d) α -stable process with $\alpha = 1.9$ and scale $\kappa = 0.15$ ($\mu = 0$, $\sigma = 0$, and $\nu(dy)$ is of the form $\frac{k}{|y|^{1+\alpha}}dy$).

• *Logistic*. The logistic signal function [15] is sigmoidal and strictly increasing

$$f(x) = \frac{1}{1 + e^{-cx}}$$
 (7)

for scaling constant c > 0. 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, 7, 13, 15]:

$$f(x) = 2 \tanh x \tag{8}$$

• *Linear Threshold*. This linear-threshold signal has the form [15]:

$$f(x) = \begin{cases} cx & |cx| < 1 \\ 1 & cx > 1 \\ -1 & cx < -1 \end{cases}$$
(9)

for constant c > 0.

• *Exponential*. This signal function is asymmetric and has the form [15]

$$f(x) = \begin{cases} 1 - \exp\{-cx\} & \text{if } x > 0\\ 0 & \text{else} \end{cases}$$
(10)

for constant c > 0.

• *Gaussian*. The Gaussian or 'radial basis' signal function [15] differs from the other signal functions above because it is nonmonotonic:

$$f(x) = \exp\{-cx^2\} \tag{11}$$

for constant c > 0.

The above neuron models can have one-to-three fixed points depending on the input signal and the model parameters. The input signal is subtreshold in the sense that switching it from s_1 to s_2 or vice versa does not change the output Y_t of (6). There exist θ_1 and θ_2 such that the input S is subtreshold 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 (9) 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. So the input is subtreshold because switching it from s_1 to s_2 or vice versa does not change the output Y_t .

3. LEVY SR IN CONTINUOUS NEURONS

We now prove that Levy noise can benefit the noisy continuous neurons (5)-(6) and (7)-(11) with subthreshold 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 Y denote the output signal Y_t for a sufficiently large randomly chosen time $t \leq T$. We use mutual information to measure the noise benefits because the input signal is random [5, 21]. The Shannon mutual information of the discrete input random variable S and the output random variable Y is the difference between its unconditional and conditional entropy [9]: I(S, Y) = H(S) - H(S|Y). Jensen's inequality implies that $I(S, Y) \ge 0$ [9]. 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. This implies that the system exhibits the SR noise benefit if I(S,Y) > 0, and if $I(S,Y) \to 0$ when noise parameters σ, ν $\rightarrow 0.$

We prove Levy SR theorem with the stochastic calculus and a limiting argument. This avoids trying to solve for the process X_t in (5). The proof structure follows that of the 'forbidden interval' theorems [16, 17, 24]. The proof strategy is that *what goes down must go up*. The proof assumes that the nonnegative mutual information is positive for some level of input-output correlation. Then the mutual information goes to zero as the noise variance goes to zero. Hence the mutual information must increase as the noise variance or standard deviation increases from zero—and thus a noise benefit must occur. The theorem requires the following technical lemma. The proof is lengthy and we omit it for space reasons.

Lemma: Let $b : \mathbf{R} \to \mathbf{R}$ and $c : \mathbf{R} \to \mathbf{R}$ be measurable functions that satisfy global Lipschitz conditions

$$||b(x_1) - b(x_2)|| \leq K_1 ||x_1 - x_2||$$
(12)

$$||c(x_1) - c(x_2)|| \leq K_2 ||x_1 - x_2||$$
(13)

for all $x_1, x_2 \in \mathbf{R}$. Suppose

$$dX_t = b(X_{t^-})dt + c(X_{t^-})dL_t$$
(14)

$$d\hat{X}_t = b(\hat{X}_t)dt. \tag{15}$$

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

$$P[\sup_{0 \le t \le T} ||X_{t-}\hat{X}_t||^2 > \varepsilon] \to 0 \quad \text{as} \quad \sigma, \nu \to 0.$$
 (16)

The lemma holds for any continuous neuron model that has a neuronal signal function f(x) of the form (7)-(11) because its respective drift term $b(X_{t-})$ in (5) is globally Lipschitz in accord with (12). Conditions (12)-(13) ensure a unique solution of (5).

We now state the theorem that gives a sufficient interval condition for the SR effect in the continuous neuron models (5)-(6) with neuronal signal functions (7)-(11). Lack of a necessary condition broadens rather than lessens the scope of potential SR effects in the neurons. The proof of the theorem is similar to the proof of Theorem 1 in [25] and we here omit it.

Theorem: Suppose that continuous neuron models (5)-(6) with constant Levy diffusion term $c(X_{t^-}) = a$ and (7)-(11) have additive white Levy noise with drift velocity μ and that the input signal $S(t) \in \{s_1, s_2\}$ is subthreshold: $\theta_1 \leq s_1 < s_2 \leq \theta_2$. Suppose that there is some statistical dependence between the input random variable S and the output random variable Y so that I(S, Y) > 0. Then the neuron models exhibit the nonmonotone SR effect in the sense that $I(S, Y) \to 0$ as the Levy noise parameters $\sigma, \nu \to 0$ if $\theta_1 - s_1 \leq a\mu \leq \theta_2 - s_2$.

Figure 2 (a)-(b) show simulation instances of the theorem for finite-variance jump-diffusion and pure-jump additive white Levy noise. Small amounts of additive Levy noise in a bistable potential neuron model produces 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 effect in Figure 2 (c) lies outside the scope of the theorem because it occurs for infinite-variance α -stable noise. We considered signal-*independent* Levy noise $(c(X_{t^-}) = \text{constant})$ in equation (5) but the lemma and a similar theorem also hold for signal dependent noise when $c(X_{t^-})$ is bounded. We also have theoretical and simulation evidence that the FitzHugh-Nagumo (FHN) spiking neuron model [8, 10] shows the SR effect for additive white Levy noise.

4. CONCLUSION

We have shown that a general type of 'forbidden interval' SR theorem holds for a wide range of Levy processes and a wide range of feedback stochastic neuron models. The underlying lemma requires that the Levy process have a finite second moment and thus it does not apply to the infinite-variance α -stable SR that Fig. 2 (c) demonstrates. An open research question is whether a more general Levy result holds that drops the finite-variance assumption and yet accounts for α -stable Levy SR results.



Fig. 2. Mutual information Levy noise benefits in continuous neurons. Additive white Levy noise dL_t increases the mutual information of the bistable potential neuron (1)-(2) and (8) for the subthreshold input signal $s_1 = -0.3$, $s_2 = 0.4$, and a = 1. The types of Levy noise dL_t are (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 [23]). The dashed vertical lines show the total min-max deviations of the mutual information in 100 simulation trials.

5. REFERENCES

- H. Ahn and R.E. Feldman, "Optimal Filtering of Gaussian Signals in the Presence of Levy Noise," *SIAM Journal of Applied Mathematics*, vol. 60, No. 1, pp. 359-369, 1999.
- [2] S. Amari, "Neural Theory of Association and Concept Formation," *Biological Cybernetics*, vol. 26, pp. 175-185, 1977.
- [3] D. Applebaum, Levy Processes and Stochastic Calculus, Cambridge University Press, 2004.
- [4] O. Barndoff-Nielson and N. Shephard, "Non-Gaussian Ornstein-Uhlenbeck-based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, Series B* vol. 63, pp. 167-241, 2001.
- [5] A.R. Bulsara and A. Zador, "Threshold Detection of Wideband Signals: A Noise Induced Maximum in the Mutual Information," *Physi*cal Review E, vol. 54, no. 3, R2185-R2188, 1996.
- [6] F. Chapeau-Blondeau and D. Rousseau, "Noise-Enhanced Performance for an Optimal Bayesian Estimator", *IEEE Transactions on Signal Processing*, vol. 52, pp. 1327-1334, 2004.
- [7] M.A. Cohen and S. Grossberg, "Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks," *IEEE Transactions on System, Man, and Cybernetics*, vol. SMC-13, pp. 815-826, 1983.
- [8] J.J. Collins, C.C. Chow, A.C. Capela, and T.T. Imhoff, "Aperiodic Stochastic Resonance in Excitable Systems," *Physical Review E*, vol. 52, No. 4, pp. R3321-R3324, 1995.
- [9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [10] W. Gerstner and W.M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.
- [11] M.T. Giraudo and L. Sacerdote, "Jump-Diffusion Processes as Models of Neuronal Activity," *BioSystems*, vol. 40, pp. 75-82, 1997.
- [12] 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, 2002.
- [13] J.J. Hopfield, "Neural Networks with Graded Response Have Collective Computational Properties Like Those of Two-state Neurons," *Proceeding of National Academy of Science*, vol. 81, pp. 3088-3092, 1984.
- [14] S. Kay, "Can Detectability be Improved by Adding Noise?," *IEEE Signal Processing Letters*, vol. 7, pp. 8-10, 2000.
- [15] B. Kosko, Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Englewood Cliffs, NJ: Prentice-Hall, 1992.

- [16] B. Kosko and S. Mitaim, "Stochastic Resonance in Noisy Threshold Neurons," *Neural Networks*, vol. 14, no. 6-7, pp. 755-761, 2003.
- [17] B. Kosko and S. Mitaim, "Robust Stochastic Resonance for Simple Threshold Neurons," *Physical Review E*, vol. 70, 031911-1 – 031911-10, 2004.
- [18] I.Y. Lee, X. Liu, B. Kosko, and C. Zhou, "Nanosignal Processing: Stochastic Resonance in Carbon Nanotubes that Detect Subthreshold Signals," *Nano Letters*, vol. 3, pp. 1683-1686, 2003.
- [19] M.D. McDonnell, N. Stocks, C.E.M. Pearce, and D. Abbott, "Quantization in the Presence of Large Amplitude Threshold Noise," *Fluctuation Noise Letters*, vol. 5, No. 3, L457-L468, 2005.
- [20] 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, 1998.
- [21] 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. 1526-1540, 2004.
- [22] F. Moss, L.M. Ward, and W.G. Sannita, "Stochastic Resonance and Sensory Information Processing: A Tutorial and Review of Applications," *Clinical Neurophysiology*, vol. 115, No. 2, pp. 267-281, 2004.
- [23] C.L. Nikias and M. Shao, Signal Processing with Alpha-Stable Distributions and Applications, John Wiley and Sons, New York, 1995.
- [24] A. Patel and B. Kosko, "Stochastic Resonance in Noisy Spiking Retinal and Sensory Neuron Models," *Neural Networks*, vol. 18, pp. 467-478, 2005.
- [25] A. Patel and B. Kosko, "Mutual-Information Noise Benefits in Brownian Models of Continuous and Spiking Neuron Models," *International Joint Conference on Neural Networks*, pp. 2347-2354, July 2006.
- [26] D. Rousseau and F. Chapeau-Blondeau, "Constructive Role of Noise in Signal Detection From Parallel Array of Quantizers," *Signal Processing*, vol. 85, No. 3, pp. 571-580, 2005.
- [27] T. Rydberg, "The Normal Inverse Gaussian Levy Processes: Simulation and Approximation," *Communications in Statistics: Stochastic Models*, vol. 13, pp. 887-910, 1997.
- [28] 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, 2003.
- [29] K-I. Sato, Levy Processes and Infinitely Divisible Distributions, Cambridge University Press, 1999.
- [30] N.G. Stocks, "Information Transmission in Parallel Threshold Arrays: Suprathreshold Stochastic Resonance," *Physical Review E*, vol. 63, 041114-1 – 041114-9, 2001.
- [31] 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, 1995.