

2009 Special Issue

Error-probability noise benefits in threshold neural signal detection

Ashok Patel, Bart Kosko*

Department of Electrical Engineering, Signal and Image Processing Institute, University of Southern California, Los Angeles, CA 90089-2564, USA

ARTICLE INFO

Article history:

Received 6 May 2009

Received in revised form 7 June 2009

Accepted 25 June 2009

Keywords:

Optimal noise intensity

Stochastic resonance

Threshold neurons

Signal detection

Error probability

Array detection

ABSTRACT

Five new theorems and a stochastic learning algorithm show that noise can benefit threshold neural signal detection by reducing the probability of detection error. The first theorem gives a necessary and sufficient condition for such a noise benefit when a threshold neuron performs discrete binary signal detection in the presence of additive scale-family noise. The theorem allows the user to find the optimal noise probability density for several closed-form noise types that include generalized Gaussian noise. The second theorem gives a noise-benefit condition for more general threshold signal detection when the signals have continuous probability densities. The third and fourth theorems reduce this noise benefit to a weighted-derivative comparison of signal probability densities at the detection threshold when the signal densities are continuously differentiable and when the noise is symmetric and comes from a scale family. The fifth theorem shows how collective noise benefits can occur in a parallel array of threshold neurons even when an individual threshold neuron does not itself produce a noise benefit. The stochastic gradient-ascent learning algorithm can find the optimal noise value for noise probability densities that do not have a closed form.

© 2009 Elsevier Ltd. All rights reserved.

1. Neural noise benefits: Total and partial SR

Stochastic resonance (SR) occurs when a small amount of noise improves nonlinear signal processing (Amblard, Zozor, McDonnell, & Stocks, 2007; Chapeau-Blondeau & Rousseau, 2004; Gammaitoni, 1995; Kay, 2000; Kosko, 2006; Levy & Baxter, 2002; McDonnell, Stocks, Pearce, & Abbott, 2006, 2008; Moss, Ward, & Sannita, 2004; Patel & Kosko, 2009b; Rousseau & Chapeau-Blondeau, 2005b; Saha & Anand, 2003; Stocks, 2001). SR occurs in many types of subthreshold and suprathreshold neural signal detection (Bulsara, Jacobs, Zhou, Moss, & Kiss, 1991; Deco & Schürmann, 1998; Hänggi, 2002; Hoch, Wenning, & Obermayer, 2003; Li, Hou, & Xin, 2005; Mitaim & Kosko, 1998, 2004; Moss et al., 2004; Patel & Kosko, 2005, 2008; Sasaki et al., 2008; Stacey & Durand, 2000; Stocks, Appligham, & Morse, 2002; Stocks & Mannella, 2001; Wang & Wang, 1997; Wiesenfeld & Moss, 1995). Biological noise can arise from an internal source such as thermal noise (Faisal, Selen, & Wolpert, 2008; Manwani & Koch, 1999) or ion channel noise (Schneidman, Freedman, & Segev, 1998; White, Rubinstein, & Kay, 2000). Or it can arise from an external source such as synaptic transmission (Levy & Baxter, 2002; Markram & Tsodyks, 1996). We focus on noise-enhanced signal detection in threshold neurons where a user can control only the noise variance or dispersion (Läer et al., 2001; Pantazelou, Dames, Moss, Douglass,

& Wilkens, 1995; Rao, Wolf, & Arkin, 2002). We measure detection performance with the probability of correct decision $P_{CD} = 1 - P_e$ when P_e is the probability of error (Patel & Kosko, 2009a).

We classify SR noise benefits as either total SR or partial SR. The SR effect is *total* if adding independent noise in the received signal reduces the error probability. Then the plot of detection probability versus noise intensity increases monotonically in some noise-intensity interval starting from zero. The SR effect is *partial* when the detection performance increases in some noise-intensity interval away from zero. Total SR ensures that adding small amounts of noise gives a better detection performance than not adding noise. Partial SR ensures only that there exists a noise intensity interval where the detection performance increases as the noise intensity increases. The same system can exhibit both total and partial SR. We derive conditions that screen for total or partial SR noise benefits in almost all suboptimal simple threshold detectors because the SR conditions apply to such a wide range of signal probability density functions (pdfs) and noise pdfs. Learnings laws can then search for the optimal noise intensity in systems that pass the screening conditions. Section 5 presents one such stochastic learning law.

We have already proven necessary and sufficient “*forbidden interval*” conditions on the noise mean or location for total SR in mutual-information-based threshold detection of discrete weak binary signals (Kosko & Mitaim, 2003, 2004): SR occurs if and only if the noise mean or location parameter μ obeys $\mu \notin (\theta - A, \theta + A)$ for threshold θ where $-A < A < \theta$ for bipolar subthreshold signal $\pm A$. More general forbidden interval theorems apply to many stochastic neuron models with Brownian or even Levy (jump)

* Corresponding author.

E-mail address: kosko@usc.edu (B. Kosko).

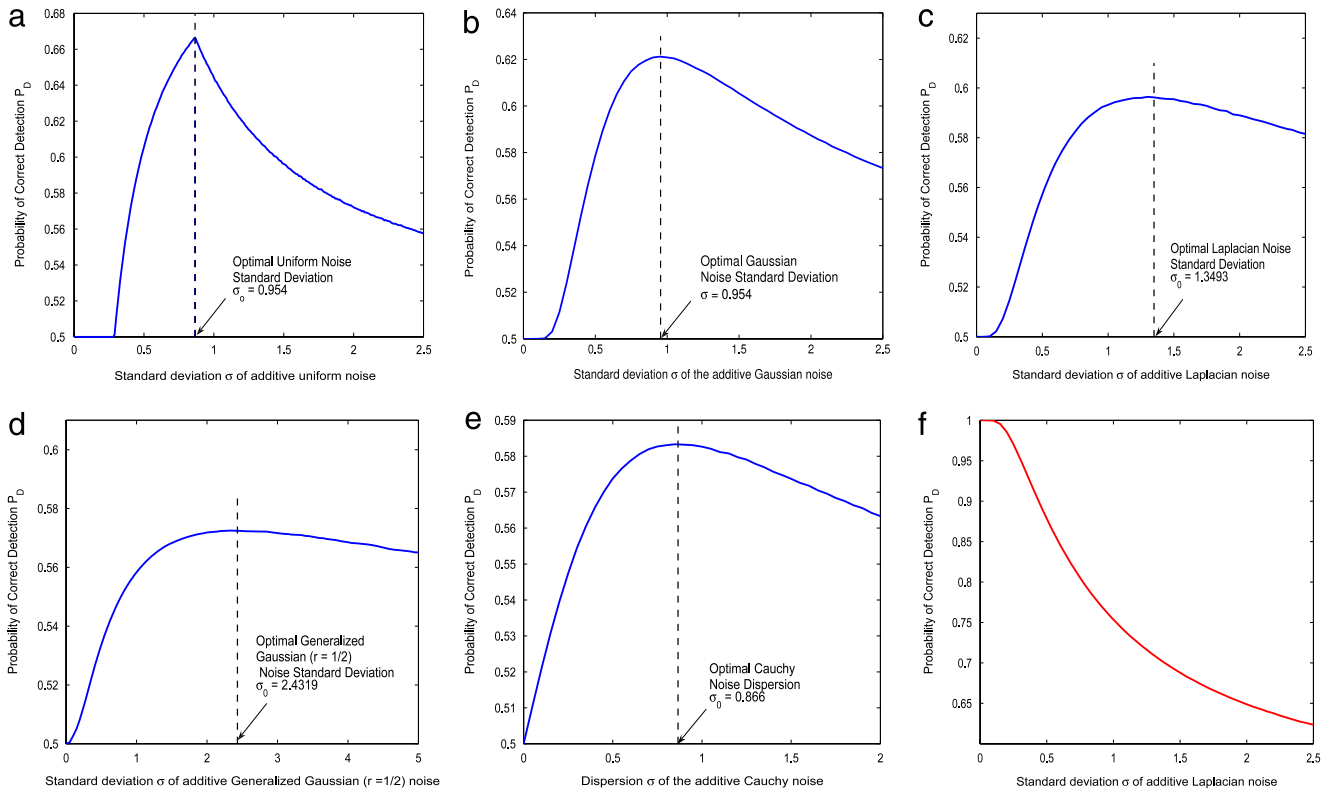


Fig. 1. Stochastic resonance (SR) noise benefits in binary neural signal detection for five out of six types of additive noise. The discrete signal X can take subthreshold values $s_0 = -1$ or $s_1 = 0$ with equal probability ($p_0 = P(s_0) = P(s_1) = p_1$) and $\theta = 1.5$ is the detection threshold. We decide $X = s_1$ if the observation $W = X + N > \theta$. Else $X = s_0$. There is a noise benefit in (a)–(e) because the zero-mean uniform, Gaussian, Laplacian, generalized Gaussian ($r = \frac{1}{2}$), and zero-location Cauchy noise satisfy condition (5) of Theorem 1. The SR effect is partial in (a) because condition (6) does not hold while the SR effect in (b)–(e) is total in each case because condition (6) holds. The dashed vertical lines show that the maximum SR effect occurs at the theoretically predicted optimal noise intensities. There is no noise benefit in (f) for Laplacian noise because its mean μ lies in the forbidden interval in accord with Corollary 1: $\mu = 1 \in (0.5, 1.5) = (\theta - s_1, \theta - s_0)$.

noise (Patel & Kosko, 2005, 2008). Corollary 1 gives a forbidden-interval necessary condition for SR in error-probability detection. But we did find necessary and sufficient conditions for both total and partial SR noise benefits in error-probability-based threshold signal detection when the noise has a scale-family distribution.

Theorem 1 gives the a simple necessary and sufficient SR condition for a noise benefit in threshold detection of discrete binary signals. This result appears in Section 3. The condition also determines whether the SR effect is total or partial if the noise density belongs to a *scale family*. Scale-family densities include many common densities such as the normal and uniform but not the Poisson. The condition implies that SR occurs in simple threshold detection of discrete binary signals only if the mean or location of additive location-scale family noise does not fall in an open forbidden interval. The uniform, Gaussian, Laplacian, and generalized Gaussian ($r = \frac{1}{2}$) noise in Fig. 1(a)–(e) produce a noise benefit because they satisfy condition (5) of Theorem 1. But the Laplacian noise in Fig. 1(f) violates this forbidden-interval condition of Corollary 1 and so there is no noise benefit. Section 4 shows that the SR condition of Theorem 1 also allows us to find the optimal noise dispersion that maximizes the detection probability for a given closed-form scale-family noise pdf. Section 5 shows that an adaptive gradient-ascent learning algorithm can find this optimal intensity from sample data even when the noise pdf does not have a closed form as with many thick-tailed noise pdfs.

Total SR can never occur in an optimal threshold system if we add only independent noise in the received signal. Kay and coworkers showed that the optimal independent additive SR noise is just a constant that minimizes the detection error probability of a given detection scheme (Kay, Michels, Chen, & Varshney, 2006). So total SR can never occur if the detection threshold

location is optimal even when the overall detection scheme is suboptimal. But we show that partial SR can still occur in a single-threshold suboptimal system even if the detection threshold is optimal. Rousseau and Chapeau-Blondeau found earlier that what we call partial SR occurs in some special cases of optimal threshold detection (Rousseau & Chapeau-Blondeau, 2005b). Fig. 3 shows such a partial SR effect for the important but special case of an optimal threshold. Our result still holds in the general case of non-optimal thresholds. The suboptimality of the signal detection remains only a necessary condition for total SR noise benefits based on error probability.

Theorem 2 in Section 6 presents a related necessary and sufficient condition for a noise benefit in a more general case of threshold detectors when the signals have continuous pdfs and when the additive independent noise has a pdf from a scale family. Then Theorem 3 gives a necessary and sufficient condition for total SR with zero-mean discrete bipolar noise. Corollary 2 gives a necessary and sufficient condition for partial SR with zero-mean discrete bipolar noise when there is no total SR in Theorem 3. Theorems 3 and 4 each gives a necessary and sufficient condition for total SR when the additive noise is zero-mean discrete bipolar or when it comes from a finite-mean symmetric scale family. These two theorems compare weighted derivatives of continuously differentiable signal pdfs at the detection threshold to determine the total SR effect. Theorem 5 shows when noise produces a collective SR effect in parallel arrays of threshold neurons even when an individual threshold neuron does not produce an SR effect. The next section describes a general problem of threshold-based neural signal detection and defines the two SR effects based on error probability.

2. Binary signal detection based on error-probability

We now cast the problem of threshold-based neural signal detection as a statistical hypothesis test. So consider the binary hypothesis test where a neuron decides between $H_0 : f_X(x, H_0) = f_0(x)$ and $H_1 : f_X(x, H_1) = f_1(x)$ using a single noisy observation of $X + N$ and a detection threshold θ where noise N is independent of X . Here $X \in R$ is the original signal input to the threshold neuron and f_i is its pdf under the hypothesis H_i for $i = 0$ or 1 . So we use the classical McCulloch–Pitts threshold neuron (McCulloch & Pitts, 1943) where the neuron’s output Y has the form

$$Y = \begin{cases} 1 & \text{(accept } H_1) \text{ if } X + N > \theta \\ 0 & \text{(accept } H_0) \text{ else.} \end{cases} \quad (1)$$

This simple threshold neuron model has numerous applications (Auer, Burgsteiner, & Maass, 2008; Beiu, Member, Quintana, & Avedillo, 2003; Caticha, Palo Tejada, Lancet, & Domany, 2002; Freund & Schapire, 1999; Minsky & Papert, 1988).

The probability of correct decision $P_{CD}(\sigma) = 1 - P_e(\sigma)$ measures detection performance. Suppose that $p_0 = P(H_0)$ and $p_1 = P(H_1) = 1 - p_0$ are the prior probabilities of the respective hypotheses H_0 and H_1 . Let $\alpha(\sigma)$ and $\beta(\sigma)$ be the respective Type-I and Type-II error probabilities when the intensity of the additive noise N is σ :

$$\alpha(\sigma) = P(\text{reject } H_0 | H_0 \text{ is true at noise intensity } \sigma) \quad (2)$$

$$\beta(\sigma) = P(\text{accept } H_0 | H_1 \text{ is true at noise intensity } \sigma). \quad (3)$$

Then define the probability of error as the usual probability-weighted sum of decision errors (Proakis & Salehi, 2008)

$$P_e(\sigma) = p_0\alpha(\sigma) + p_1\beta(\sigma). \quad (4)$$

We assume that the additive noise N is a scale-family noise with pdf $f_N(\sigma, n)$ where σ is the noise intensity (standard deviation or dispersion): $f_N(\sigma, n) = \frac{1}{\sigma} f(\frac{n}{\sigma})$ where f is the standard pdf for the family (Casella & Berger, 2001). Then the noise cumulative distribution function (CDF) is $F_N(\sigma, n) = F(\frac{n}{\sigma})$ where F is the standard CDF for the family.

We next define SR effects in neural signal detection based on error probability. A binary signal detection or hypothesis testing system exhibits the SR effect in the noise intensity interval (a, b) for $0 \leq a < b < \infty$ iff $P_{CD}(\sigma_1) < P_{CD}(\sigma_2)$ for any two noise intensities σ_1 and σ_2 such that $a \leq \sigma_1 < \sigma_2 \leq b$. The SR effect is *total* if $a = 0$ and *partial* if $a \neq 0$. We say that the SR effect occurs at the noise intensity σ iff the SR effect occurs in some noise intensity interval (a, b) and $\sigma \in (a, b)$.

3. Noise benefits in threshold detection of discrete binary signals

We first consider the binary signal detection problem where the signal X is a binary discrete random variable with the two values s_0 and s_1 so that $s_0 < s_1$ and that $P(X = s_0) = p_0$ and $P(X = s_1) = p_1$. Then Theorem 1 gives a necessary and sufficient condition for an SR effect in the threshold neuron model (1) for discrete binary signal detection if the additive noise comes from an absolutely continuous scale-family distribution.

Theorem 1. Suppose that the additive continuous noise N has scale-family pdf $f_N(\sigma, n)$ and that the threshold neuron model is (1). Suppose that signal X is a binary discrete random variable with the two values s_0 and s_1 so that $P(X = s_0) = p_0$ and $P(X = s_1) = p_1$. Then the SR noise benefit occurs in a given noise intensity interval (a, b) if and only if

$$p_0(\theta - s_0)f_N(\sigma, \theta - s_0) < p_1(\theta - s_1)f_N(\sigma, \theta - s_1) \quad (5)$$

for almost every noise intensity $\sigma \in (a, b)$. The SR effect is total if

$$\lim_{\sigma \downarrow 0} p_0(\theta - s_0)f_N(\sigma, \theta - s_0) < \lim_{\sigma \downarrow 0} p_1(\theta - s_1)f_N(\sigma, \theta - s_1). \quad (6)$$

Proof. The signal X is a binary discrete random variable with the two values s_0 and s_1 . So the Type-I and Type-II error probabilities (2) and (3) become

$$\alpha(\sigma) = 1 - F_N(\sigma, \theta - s_0) \quad (7)$$

$$\beta(\sigma) = F_N(\sigma, \theta - s_1) \quad (8)$$

where F_N is the absolutely continuous CDF of the additive noise random variable N . Then the error probability $P_e(\sigma) = p_0\alpha(\sigma) + p_1\beta(\sigma)$ is an absolutely continuous function of σ in any closed interval $[c, d] \subset R^+$ where $c > 0$. Then the above definition of SR effects and the fundamental theorem of calculus (Folland, 1999) imply that the SR effect occurs in the noise intensity interval (a, b) if and only if $\frac{dP_e(\sigma)}{d\sigma} < 0$ for almost all $\sigma \in (a, b)$. So the SR effect occurs in the noise intensity interval (a, b) if and only if

$$0 < -p_1 \frac{\partial F_N(\sigma, \theta - s_1)}{\partial \sigma} - p_0 \frac{\partial [1 - F_N(\sigma, \theta - s_0)]}{\partial \sigma} \quad (9)$$

for almost all $\sigma \in (a, b)$. Rewrite (9) as

$$0 < -p_1 \frac{\partial F(\frac{\theta - s_1}{\sigma})}{\partial \sigma} - p_0 \frac{\partial [1 - F(\frac{\theta - s_0}{\sigma})]}{\partial \sigma} \quad (10)$$

where F is the standard scale-family CDF of the additive noise N . Then (10) gives

$$\begin{aligned} 0 < p_1 \frac{(\theta - s_1)}{\sigma} f\left(\frac{\theta - s_1}{\sigma}\right) - p_0 \frac{(\theta - s_0)}{\sigma} f\left(\frac{\theta - s_0}{\sigma}\right) \\ = p_1(\theta - s_1)f_N(\sigma, \theta - s_1) - p_0(\theta - s_0)f_N(\sigma, \theta - s_0) \end{aligned} \quad (11)$$

because the additive noise N has scale-family pdf $f_N(\sigma, n) = \frac{1}{\sigma} f(\frac{n}{\sigma})$ and because the noise scale σ is always positive. Inequality (5) now follows from (11). The definition of a limit implies that condition (5) holds for all $\sigma \in (0, b)$ for some $b > 0$ if (6) holds. So (6) is a sufficient condition for the total SR effect in the simple threshold detection of discrete binary random signals. \square

Theorem 1 lets users screen for total or partial SR noise benefits in discrete binary signal detection for a wide range of noise pdfs. This screening test can prevent a fruitless search for nonexistent noise benefits in many signal-noise contexts. The inequality (5) leads to the simple stochastic learning law in Section 5 that can find the optimal noise dispersion when a noise benefit exists. The learning algorithm does not require a closed-form noise pdf.

Inequality (5) differs from similar inequalities in standard detection theory for likelihood ratio tests. It specifically resembles but differs from the maximum a posteriori (MAP) likelihood ratio test in detection theory (Proakis & Salehi, 2008):

$$\begin{cases} \text{Reject } H_0 & \text{if } p_0 f_N(\sigma, z - s_0) < p_1 f_N(\sigma, z - s_1) \\ \text{Else accept } H_0. \end{cases} \quad (12)$$

The MAP rule (12) minimizes the detection-error probability in optimal signal detection. But it requires the noisy observation z of the received signal $Z = X + N$ whereas inequality (5) does not. Inequality (5) also contains the differences $\theta - s_0$ and $\theta - s_1$. So Theorem 1 gives a general way to detect SR noise benefits in suboptimal detection.

Theorem 1 implies a forbidden-interval necessary condition if the noise pdf comes from a location-scale family.

Corollary 1. Suppose that the additive noise N has location-scale family pdf $f_N(\sigma, n)$. Then the SR noise benefit effect occurs only if the noise mean or location μ obeys the forbidden-interval condition $\mu \notin (\theta - s_1, \theta - s_0)$.

