

Enhanced information transmission with signal-dependent noise in an array of nonlinear elements

A. Nikitin and N. G. Stocks

School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom

R. P. Morse

School of Life and Health Sciences, University of Aston, Birmingham B4 7ET, United Kingdom

(Received 17 August 2006; published 26 February 2007)

We have investigated information transmission in an array of threshold units that have signal-dependent noise and a common input signal. We demonstrate a phenomenon similar to stochastic resonance and suprathreshold stochastic resonance with additive noise and show that information transmission can be enhanced by a nonzero level of noise. By comparing system performance to one with additive noise we also demonstrate that the information transmission of weak signals is significantly better with signal-dependent noise. Indeed, information rates are not compromised even for arbitrary small input signals. Furthermore, by an appropriate selection of parameters, we observe that the information can be made to be (almost) independent of the level of the noise, thus providing a robust method of transmitting information in the presence of noise. These results could imply that the ability of hair cells to code and transmit sensory information in biological sensory systems is not limited by the level of signal-dependent noise.

DOI: [10.1103/PhysRevE.75.021121](https://doi.org/10.1103/PhysRevE.75.021121)

PACS number(s): 05.40.-a, 87.17.-d

I. INTRODUCTION

In recent years there has been a significant increase in interest in the interplay between signals and noise in nonlinear systems. This activity has largely been motivated by studies of stochastic resonance (SR) [1–4] but also by a desire to understand stochastic aspects of neural coding [5–9]. In particular, the study of signal coding in parallel arrays (populations) of nonlinear devices (neurons) has received considerable attention; in such arrays, a new form of SR—termed suprathreshold stochastic resonance (SSR)—has been discovered [10–13]. In a similar vein to SR, SSR can lead to an improvement in information transmission when internal noise is added to the system. However, SSR, which can only occur in arrays of nonlinear devices, has a number of advantages over conventional SR. First, it occurs for all signal levels—it does not require that the signal be subthreshold. Consequently SSR can be used to improve information transmission for a broader class of signal than standard SR. Second, it provides an optimal method of enhancing information when the signal to be detected is comparable (or smaller) than the residual internal noise [14,15]. In contrast, for SR in a single device, greater information flow is usually obtained by simply increasing the input signal (or lowering the threshold if possible) rather than setting the signal to be subthreshold and utilizing SR. For these reasons, the potential exploitation of SSR in technological applications is arguably greater than for conventional SR and SSR may well be relevant to optimal neural coding. Possible applications are novel digital-to-analog converters [10,11,16–18], sonar arrays [11,19], and cochlear implants [20].

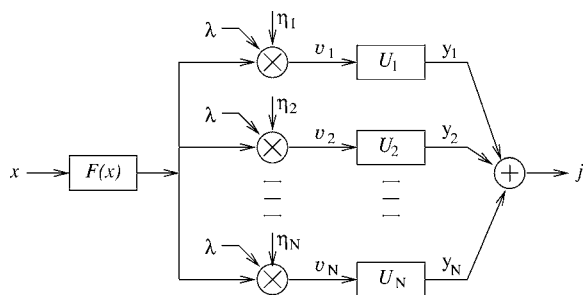
Although SSR has now been studied in a wide variety of different contexts [10–18,20–25], all these studies have been undertaken assuming that the noise is additive. However, it is well established that in neural systems the noise may enter in a signal-dependent or multiplicative fashion [26–30]. Signal-

dependent noise (SDN), for example, is characteristic of the propagation or initiation of nerve impulses at synapses through the quantal release of neurotransmitter. Furukawa *et al.* [31,32] have shown experimentally that the variation of the neurotransmitter release into a synaptic cleft is proportional to the intensity of the stimulus. Consequently, the variance of the fluctuations in the postsynaptic potentials is signal dependent and, therefore, they cannot be modeled using an additive noise process. This provides motivation to investigate the role of SDN in the transmission of information in nonlinear systems. Furthermore, although there are a number of studies of SR that consider SR with signal-dependent or multiplicative noise [33–39], few have considered an array of such devices (a notable exception is the study by Lindner and Schimansky-Geier [34]) and, as yet, there have been no studies of SSR with SDN. This paper addresses these issues—specifically, we consider SSR effects with SDN.

II. MODEL

The model we study is loosely based on that proposed by Furukawa *et al.* [31] for the modeling of synaptic transmission in hair cell transduction. However, the model has been abstracted and simplified to enable theoretical analysis to be undertaken and relate to other studies in SSR. The abstraction of the model from that in [31] means that we do not wish this model to be viewed as a realistic model for synaptic transmission, but as a model for investigating the role of SDN in information transmission in a simple nonlinear array. Nevertheless, the results and conclusion may well have some relevance to hair cell transduction.

We have modeled an array of N nerve fibers by an array of N simple threshold units (level-crossing detectors), Fig. 1. Each threshold unit was subject to the same input signal x , which for generality can be considered to be transformed by


 FIG. 1. A summing array of N modeled nerve fibers.

the linear or nonlinear function $F(x)$. The output y of each unit is given by the Heaviside function

$$y_i = \begin{cases} 1: v_i \geq U_i, \\ 0: v_i < U_i, \end{cases} \quad (1)$$

where U_i is the threshold of the unit and $i=1, \dots, N$, and the input to each threshold unit is given by

$$v_i = \lambda F(x) + D \eta_i F(x). \quad (2)$$

Here, η_i has a Gaussian noise distribution with zero average mean and unit dispersion, D is the coefficient of proportionality (the “noise intensity”), and λ is a parameter of the model that dictates the level of the deterministic signal component $F(x)$. The term $D \eta_i F(x)$ is clearly a fluctuational term whose intensity varies with signal strength and which is additive to the deterministic signal component. Through an appropriate choice of threshold settings and noise intensities, the array of threshold units can model a number of applications. For simplicity, however, we considered only the case where the thresholds were identical ($U_1 = U_2 = \dots = U$) and all units had the same noise intensity. This corresponds to the configuration implemented in the original studies of additive SSR. Moreover, the results were for a Gaussian signal, such that the probability density for the input signal was given by

$$P_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (3)$$

where σ is the intensity of the signal.

Because the noises η_i were mutually independent—i.e., $\langle \eta_i \eta_j \rangle = 0$ if $i \neq j$ —the variables v_i were also mutually independent. The probability that exactly j units were in state 1 for a given value of x was therefore given by the binomial distribution

$$\text{Prob}\left\{\sum_{i=1}^N y_i = j \mid x\right\} = C_j^N q_x^j (1 - q_x)^{N-j},$$

$$C_j^N = \frac{N!}{j! (N-j)!}. \quad (4)$$

Given that the noise intensity for each unit was identical, $q_x = \text{Prob}\{y_i = 1 \mid x\}$ was identical for all units too.

The overall output of the system was taken to be the sum of the outputs of the individual units. The system can therefore be taken to map the instantaneous amplitude of the sig-

nal, x , onto j , the number of units in state 1 such that $j = \sum_{i=1}^N y_i$.

The structure of Eq. (2) is similar to additive noise models previously studied. For comparison we will also show some results for the additive noise model (5),

$$v_i = \lambda F(x) + D \eta_i. \quad (5)$$

In these comparisons models (2) and (5) have identical parameters, and hence the behaviors of the models can be simply compared.

A. Linear model

We analyzed two versions of the model characterized by different expressions for v_i : a linear model and one with rectification. In the linear model, we take the function $F(x)$ to be linear for simplicity such that $F(x) = x$ and, therefore,

$$v_i = \lambda x + D \eta_i x. \quad (6)$$

With this condition, the conditional probability density of v_i was Gaussian,

$$P_{v_i|x}(v_i|x) = \frac{1}{\sqrt{2\pi D^2 x^2}} \exp\left(-\frac{(v_i - \lambda x)^2}{2D^2 x^2}\right), \quad (7)$$

but with a mean of λx and a standard deviation of $D^2 x^2$ that are signal dependent. The probability that $y_i = 1$ for a given signal level x can be calculated as

$$q_x = \text{Prob}\{y_i = 1 \mid x\} = \text{Prob}\{v_i > U \mid x\} = \int_U^\infty P_{v_i|x}(v_i|x) dv_i$$

$$= \frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{U - \lambda x}{\sqrt{2D^2 x^2}}\right). \quad (8)$$

B. Model with rectification

The second model we considered included a half-wave rectification of the signal. The rectification is described by

$$v_i = \lambda R(x) + D \eta_i R(x), \quad (9)$$

where $R(x)$ is the rectification function,

$$R(x) = \begin{cases} x: x > 0, \\ 0: x \leq 0, \end{cases} \quad (10)$$

Again, the motivation for studying a rectified signal comes from the hair cell model [31]. The stereocilia (“hairs”) of hair cells half-wave rectify signals due to the ion channel opening probabilities being directionally sensitive; the signal is only transduced when the stereocilia are displaced in a specific direction (e.g., on the positive half cycle of the signal). We therefore wish to quantify the effect of this rectification on information transmission.

The variable v_i is not Gaussian for this model over all values of x . When $x > 0$, the conditional probability density is given by Eq. (7)—that is,

$$P_{v_i|x}(v_i|x) = \frac{1}{\sqrt{2\pi D^2 x^2}} \exp\left(-\frac{(v_i - \lambda x)^2}{2D^2 x^2}\right), \quad (11)$$

as before, but when $x \leq 0$, we have

$$P_{v_i|x}(v_i|x) = \delta(v_i), \quad (12)$$

where $\delta(v_i)$ is the delta function. The probability that $y_i=1$ for a given positive value of x is therefore

$$q_x = \text{Prob}\{y_i = 1|x > 0\} = \text{Prob}\{v_i \geq U|x > 0\} \\ = \int_U^\infty P_{v_i|x}(v_i|x) dv_i = \frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{U - \lambda x}{\sqrt{2Dx}}\right), \quad (13)$$

and for negative x , $x \leq 0$, it is

$$q_x = \text{Prob}\{y_i = 1|x \leq 0\} = \int_U^\infty \delta(v_i) dv_i = \begin{cases} 0: U > 0, \\ 1: U \leq 0. \end{cases} \quad (14)$$

III. MUTUAL INFORMATION

The mutual information between the instantaneous level of the input signal x and the number of units in state 1, j , is given by Shannon and Weaver [40] to be

$$I = H(j) - H(j|x), \quad (15)$$

where $H(j)$ denotes output entropy and $H(j|x)$ denotes the output entropy conditional on the input defined, respectively, by

$$H(j) = - \sum_{j=0}^N Q_j \log_2 Q_j \quad (16)$$

and

$$H(j|x) = - \int_{-\infty}^{\infty} dx P_x(x) \sum_{j=0}^N B_j \log_2 B_j, \quad (17)$$

where $B_j = \text{Prob}\{\sum_{i=1}^N y_i = j|x\}$, and

$$Q_j = \text{Prob}\left\{\sum_{i=1}^N y_i = j\right\} = \int_{-\infty}^{\infty} dx P_x(x) B_j \\ = C_j^N \int_{-\infty}^{\infty} dx P_x(x) q_x^j (1 - q_x)^{N-j} \quad (18)$$

and

$$H(j|x) = - \int_{-\infty}^{\infty} dx P_x(x) \sum_{j=0}^N C_j^N q_x^j (1 - q_x)^{N-j} \\ \times [\log_2 C_j^N + j \log_2 q_x + (N - j) \log_2 (1 - q_x)]. \quad (19)$$

The mutual information can be calculated by using Eqs. (15)–(19) in conjunction with Eq. (8) or Eqs. (13) and (14) for the linear and rectified cases, respectively.

IV. RESULTS

A. Linear model

In this section we present results of the “linear” model, by which we mean results obtained from computer simula-

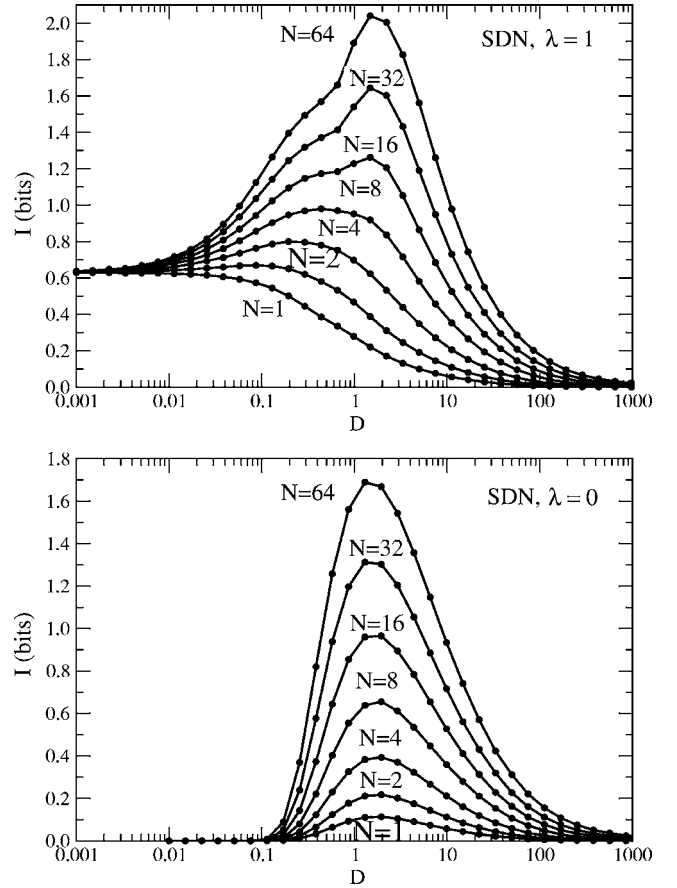


FIG. 2. The mutual information I as a function of the noise intensity D for an array of threshold units for the linear model (6). The results were obtained for a common threshold $U=1$ and a common Gaussian input signal x with a standard deviation $\sigma=1$ and for various numbers of threshold units from $N=1$ to $N=64$. In the upper panel $\lambda=1$ and in the lower panel $\lambda=0$.

tion of Eqs. (1) and (6). The results presented in Fig. 2 correspond to $\lambda=1$ (upper panel) and $\lambda=0$ (lower panel). This figure shows that a nonmonotonic dependence of the mutual information I on the noise intensity D is observed and, hence, that the mutual information is optimized by a nonzero level of noise. The phenomenon is therefore similar to stochastic resonance for additive noise [1–4].

Furthermore, the upper panel of Fig. 2 displays an effect similar to SSR [10]. For very weak noise ($D=0.001$) the information is largely independent of the noise intensity but yet the transmitted information is nonzero (just above 0.6 bits)—this is because in the absence of noise the signal can be suprathreshold and hence some signal information is transmitted. For $N=1$ the effect of increasing the noise is to reduce information transmission, while for $N > 1$, an initial increase in the information is observed. A distinguishing characteristic of SSR is that a nonmonotonic dependence on noise intensity is only observed for multiple elements; hence, we can conclude that SSR appears to occur for SDN.

The results in the lower panel differ in that there is no information transmission at low noise intensity. This is simple to understand; for $\lambda=0$, we have $v_i = D\eta_i x$ and hence there is no deterministic signal component present in v_i .

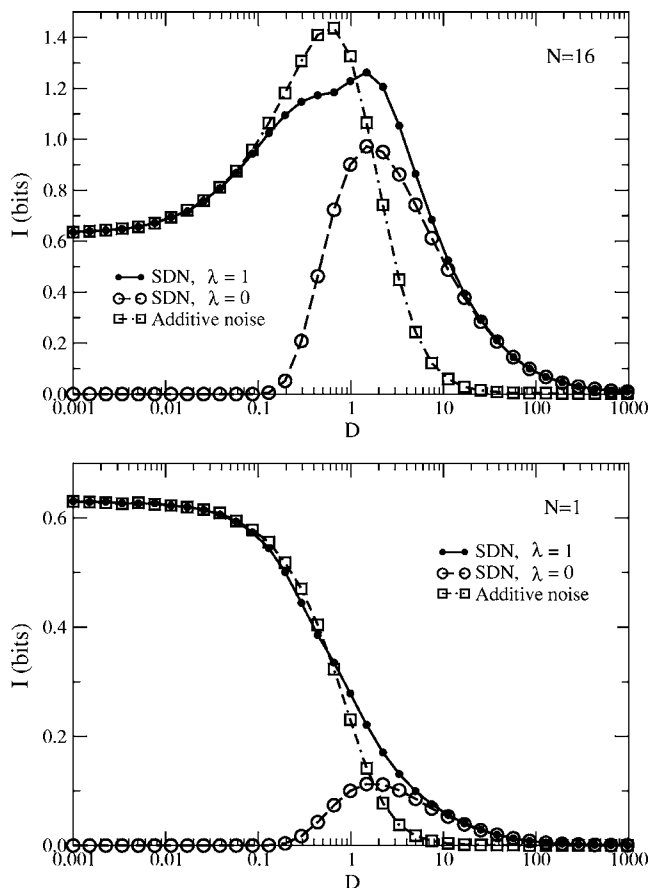


FIG. 3. Comparison of signal-dependent noise (SDN) (with $\lambda=0$ and $\lambda=1$) to additive noise. The mutual information I is shown as a function of the noise intensity D for the linear model (6). The results were obtained for a common threshold $U=1$ and a common Gaussian input signal x with a standard deviation $\sigma=1$. Results with $\lambda=1$ are shown by the curve with black dots and for $\lambda=0$ by the circles. The results of the additive noise model (5) for $\lambda=1$ are shown by the squares. The number of threshold units was $N=1$ (lower panel) and $N=16$ (upper panel).

Thus, v_i can only exceed the threshold for nonzero D . Consequently, the behavior here is similar to that of conventional SR when the signal is predominantly subthreshold. At larger noise intensities the results in the upper and lower panels are similar as one would expect (because the noise term dominates the expression for v_i). Finally, we note that the information increases as the number of devices increase — again this is to be expected [10].

Figure 3 shows a comparison between the cases for $\lambda=0$ (circles), $\lambda=1$ (solid dots), and the additive noise model (5) (squares) for $N=16$ (upper panel) and for $N=1$, the single-element case (lower panel). It can be seen that in the limit of weak noise intensity ($D \ll \lambda$) the mutual information of the linear model coincides with that for the model with additive noise. To some extent this is to be anticipated because Eqs. (6) and (5) reduce to the same model in the limit $D \rightarrow 0$. In this limit the mutual information can be calculated analytically as

$$I = -A \log_2 A - (1-A) \log_2 (1-A),$$

where

$$A = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{U}{\sqrt{2\lambda}} \right).$$

However, it is some surprise that the agreement between these models persists up to $D=0.2$. This close agreement is observed independent of the number of elements. These results demonstrate that, qualitatively, there is little difference in the SR and SSR effects between additive and SDN.

Figure 3 also shows (as was observed in Fig. 2) that in limit of high noise intensity ($D \gg \lambda$) the dependence of the mutual information on noise intensity is largely independent of λ . In this limit, Eqs. (19) and (16) can be calculated analytically. For a finite threshold ($-\infty < U < +\infty$), and in the limit of large values of D , the instantaneous level of the input signal and the threshold level have no effect on the output state of each threshold unit and the probability that the output is in state 1, q_x , is given by the probability that the Gaussian noise exceeds 0; in the limit of large D , q_x is therefore 0.5. The probability that j units are in state 1 is therefore independent of x and is given by

$$\operatorname{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \right\} = \frac{C_j^N}{2^N}, \quad (20)$$

From this independence between j and x , it follows that

$$\operatorname{Prob} \left\{ \sum_{i=1}^N y_i = j \right\} = \operatorname{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \right\} \quad (21)$$

and the output entropy and output entropy conditions on the input are therefore identical. From Eq. (21), it follows that the mutual information in the limit of large D is zero; this analytic result is confirmed by the simulated results in Fig. 2. While this analysis may seem somewhat trivial, it is included to contrast with the results obtained with signal rectification in the next section.

Figure 4 demonstrates the effect of varying the threshold. The top panel is for $\lambda=1$, the middle panel for $\lambda=0$, and the bottom panel for the additive noise model. Increasing the threshold is well known to decrease the maximum information transmitted by the SR and SSR effects when the noise is additive [11]. This can clearly be seen in the bottom panel; the maximum information is reduced from a value of 1.4 bits ($U=1$) to about 0.3 bits ($U=4$). However, the reduction in the information with increasing threshold is less prevalent in the top panel and does not occur at all in the middle panel. Consequently it would appear that SR and SSR are more robust to changes in threshold when the noise is signal dependent. To understand this behavior we note that the following equivalences are to be expected when $\lambda=0$:

$$v_i = D\tilde{x}\eta_i = D(ax)\eta_i = (D\alpha)x\eta_i = \tilde{D}x\eta_i. \quad (22)$$

This demonstrates that a scaling of the signal is equivalent to a rescaling of the noise. Given that for $\lambda=0$ the mutual information only depends on the ratio between the threshold and the signal strength, this implies that the information must also scale equivalently with the threshold level. Hence changing the threshold simply results in a rescaling of the noise. This scaling is observed in the middle panel of Fig. 4

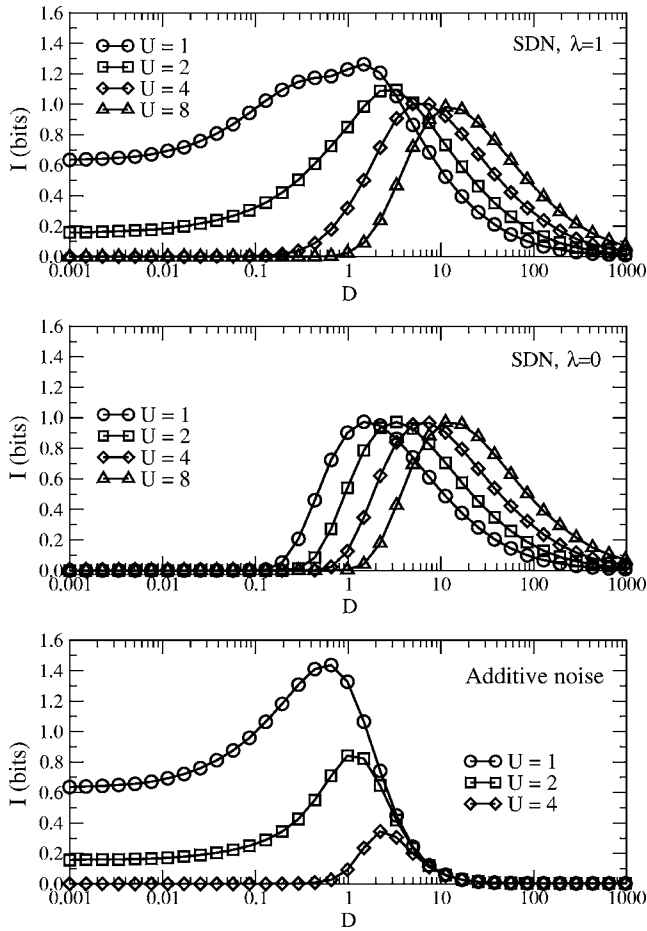


FIG. 4. The mutual information I as a function of the noise intensity D for various thresholds U with signal-dependent noise (6) (upper and middle panels) and for additive noise (5) (lower panel). The results were obtained for $N=16$ and a common Gaussian input signal x with a standard deviation $\sigma=1$. Other parameters were $\lambda=1$ (upper and lower panels) and $\lambda=0$ (middle panel). The curves with circles, squares, diamonds, and triangles correspond to $U=1$, $U=2$, $U=4$, and $U=8$, respectively.

which shows that the curves of mutual information against noise intensity have identical forms for different values of U but are shifted along the noise intensity axis; e.g., increasing the threshold by a factor of 2 causes the optimum mutual information to occur when the noise intensity is twice as large. These scalings do not hold for nonzero λ , but nevertheless, the top panel indicates that the maximum in the information curves is only weakly dependent on threshold.

The robustness of the maximal mutual information to changes in threshold gives SDN a significant advantage compared with additive noise when detecting signals that are very weak—i.e., signals that are well below threshold. The scaling arguments for $\lambda=0$ suggest that, with SDN, the information transmitted is not compromised regardless of how weak the signal is. In fact, information transmission rates of 1 bit can, in principle, be achieved for arbitrary small signals. In contrast, with additive noise the information tends to zero as the signal strength is reduced. In practice, the information transmitted by a system with SDN will not be limited by the

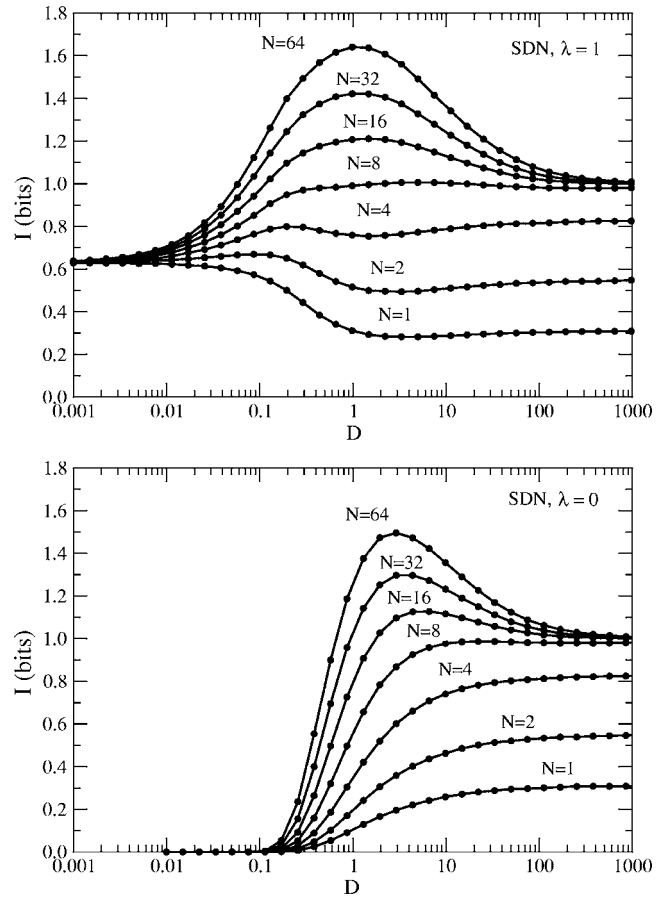


FIG. 5. The mutual information I as a function of the noise intensity D for an array of threshold units with rectification (9). The results were obtained for a common threshold $U=1$ and a common Gaussian input signal x with a standard deviation $\sigma=1$ and for various numbers of threshold units from $N=1$ to $N=64$. The upper panel is for $\lambda=1$ and the lower panel for $\lambda=0$.

size of the input signal but by the size of the residual *additive* noise that is present in all physical systems.

B. Model with rectification

In this section we present results of the “rectified” model. The results shown are for identical parameters to those in the previous section but Eq. (6) has been replaced by Eq. (9).

Figure 5 shows the results for $\lambda=1$ (top panel) and $\lambda=0$ (bottom panel). These contrast with those for the linear model shown in Fig. 2 in the following ways: first, a local maximum existed (lower panel of Fig. 5) only for a sufficiently large value N , much like for suprathreshold stochastic resonance [10,12,13,20,41]. Second, the local maxima for different values of N occurred at different noise intensities; the location of the maximum shifted further to the region of weak noise intensity for larger values N . (For the linear model the locations of the maxima of the curves are almost identical for different values of N .) Third, the mutual information asymptotically approached 1 from above or below for sufficiently large N and D . Indeed, this is the most notable difference between the linear and rectified models. For the

linear model the information tends to zero as the noise intensity tends to infinity—as would be expected from results based on additive noise. It is therefore of some surprise that, at large values of the SDN, the information approaches an asymptotic limit that is nonzero. The value of this asymptote increases as N is increased and saturates at 1 bit for $N > 8$.

To understand this result we have considered the information in the limit of large noise intensity. For the threshold U less than or equal to zero, the probability $q_x=0.5$ if $x > 0$ and $q_x=1$ if $x \leq 0$ at high noise intensity and the behavior of the models with and without rectification is therefore the same—i.e., $\lim_{D \rightarrow +\infty} I = 0$. But when the threshold U is positive and finite ($0 < U < +\infty$), the probability $q_x=0$ if $x \leq 0$ and $q_x=1/2$ if $x > 0$. The conditional probability is therefore given by

$$\text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \leq 0 \right\} = \begin{cases} 0 & : j \neq 0, \\ 1 & : j = 0, \end{cases}$$

$$\text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x > 0 \right\} = \frac{C_j^N}{2^N}. \quad (23)$$

Therefore, in contrast to the model without rectification, the conditional probability of the model with rectification is dependent on x . From Eq. (23) it follows that

$$\begin{aligned} & \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid j \neq 0 \right\} \\ &= \int_{-\infty}^{\infty} dx P_x(x) \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \right\} \\ &= \int_{-\infty}^0 dx P_x(x) \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \leq 0 \right\} \\ & \quad + \int_0^{\infty} dx P_x(x) \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x > 0 \right\} \\ &= \frac{C_j^N}{2^N} \int_0^{\infty} P_x(x) dx, \text{Prob} \left\{ \sum_{i=1}^N y_i = 0 \right\} \\ &= \int_{-\infty}^0 P_x(x) dx. \end{aligned} \quad (24)$$

Since the probability distribution of the Gaussian input signal is symmetric—i.e., $P_x(x) = P_x(-x)$ and $\int_{-\infty}^0 P_x(x) dx = \int_0^{\infty} P_x(x) dx = 1/2$ —the previous expression can be rewritten as

$$\text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid j \neq 0 \right\} = \frac{C_j^N}{2^{N+1}},$$

$$\text{Prob} \left\{ \sum_{i=1}^N y_i = 0 \right\} = \frac{1}{2}. \quad (25)$$

By using Eqs. (19), (16), and (25) the following expression for the information entropies is obtained:

$$\begin{aligned} H(j|x) &= - \int_{-\infty}^0 dx P_x(x) \sum_{j=0}^N \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x \leq 0 \right\} \\ & \quad \times \log_2 P \left\{ \sum_{i=1}^N y_i = j \mid x \leq 0 \right\} \\ & \quad - \int_0^{\infty} dx P_x(x) \sum_{j=0}^N \text{Prob} \left\{ \sum_{i=1}^N y_i = j \mid x > 0 \right\} \\ & \quad \times \log_2 P \left\{ \sum_{i=1}^N y_i = j \mid x > 0 \right\} = \frac{N}{2} - \frac{1}{2^{N+1}} \sum_{j=1}^N C_j^N \log_2 C_j^N \end{aligned} \quad (26)$$

and

$$H(j) = 1 + \frac{N}{2} - \frac{N+1}{2^{N+1}} - \frac{1}{2^{N+1}} \sum_{j=1}^N C_j^N \log_2 C_j^N, \quad (27)$$

The mutual information in the limit of large D is therefore

$$\lim_{D \rightarrow +\infty} I = 1 - \frac{N+1}{2^{N+1}}. \quad (28)$$

The last term of Eq. (28) decreases with increasing N , and for large N , the mutual information can be approximated by $I=1$. Confirmation of this analytic result is shown in Fig. 5. The asymptotic result, Eq. (28), is in good agreement with the numerical results observed in Fig. 5.

The results shown in Fig. 6 are similar to those obtained without rectification (Fig. 3)—with the exception that the curves approach nonzero asymptotic values at large values of SDN. We again see that results for $\lambda=1$ coincide with those for $\lambda=0$ at large noise intensity and, in limit of weak noise, the $\lambda=1$ results coincide with the additive noise results. The results for $\lambda=1$ bridge those of the additive and $\lambda=0$ results.

We note that the model with rectification and for $\lambda=0$ has similar scaling equivalences to the linear model,

$$\tilde{v}_i = DR(\tilde{x}) \eta_i = DR(ax) \eta_i = D\alpha R(x) \eta_i = \tilde{D}R(x) \eta_i, \quad (29)$$

and hence there is a simple scaling behavior between the variance of the noise and the common threshold level U . This is confirmed in Fig. 7 (middle panel), the curves of mutual information against noise intensity have identical forms for different values of U but are shifted along the noise intensity axis. Such equivalences do not hold for the additive noise model (bottom panel) or for nonzero λ (top panel). However, for nonzero λ the scaling is observed to hold approximately at large values of the SDN. By comparing the top and middle panels of Fig. 7 to the bottom panel (additive noise) it can again be observed that for weak signals (i.e., large U) SDN substantially outperforms additive noise. However, with rectification this improved performance is

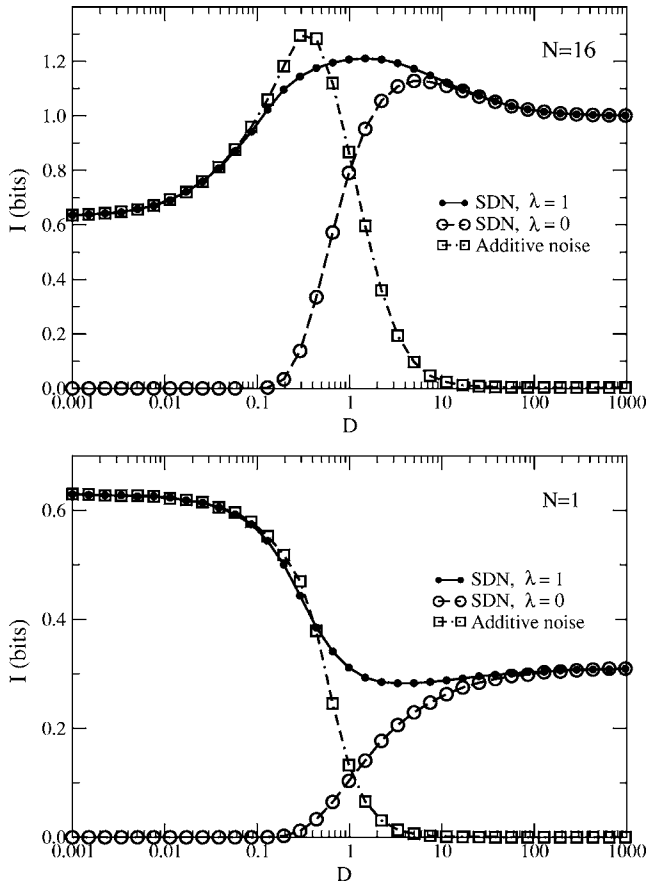


FIG. 6. Comparison of signal-dependent noise (with $\lambda=0$ and $\lambda=1$) to additive noise. The mutual information I is shown as a function of the noise intensity D for the rectified model (9). The results were obtained for a common threshold $U=1$ and a common Gaussian input signal x with a standard deviation $\sigma=1$. Results with $\lambda=1$ are shown by the curve with black dots and for $\lambda=0$ by the circles; the results of the additive noise model (5) for $\lambda=1$ are shown by the squares. The number of threshold units was $N=1$ (lower panel) and $N=16$ (upper panel).

even more striking (compared with the nonrectified case) at large noise intensities; weak signals and large SDN do not limit the ability of the system to transmit information.

We make one final observation regarding the effect of rectification and SDN. In principle, with an appropriate choice of threshold and setting λ to be nonzero, it is possible to obtain information transmission rates that only depend weakly on the level of noise. This is particularly noticeable in Fig. 5 (e.g., top panel and $N=4$) and Fig. 6 (solid dotted curves). In principle, and unlike additive noise, this means that SDN need not limit the ability of the system to transmit information and that the information rates are robust to changes in the noise level. This may have implications for the transduction and transmission of information by hair cells—on which our models are based. Hair cells both rectify signals and also, via synaptic transmission, introduce a form of SDN. It is possible, therefore, that sources of SDN do not limit the ability of hair cells to transit information.

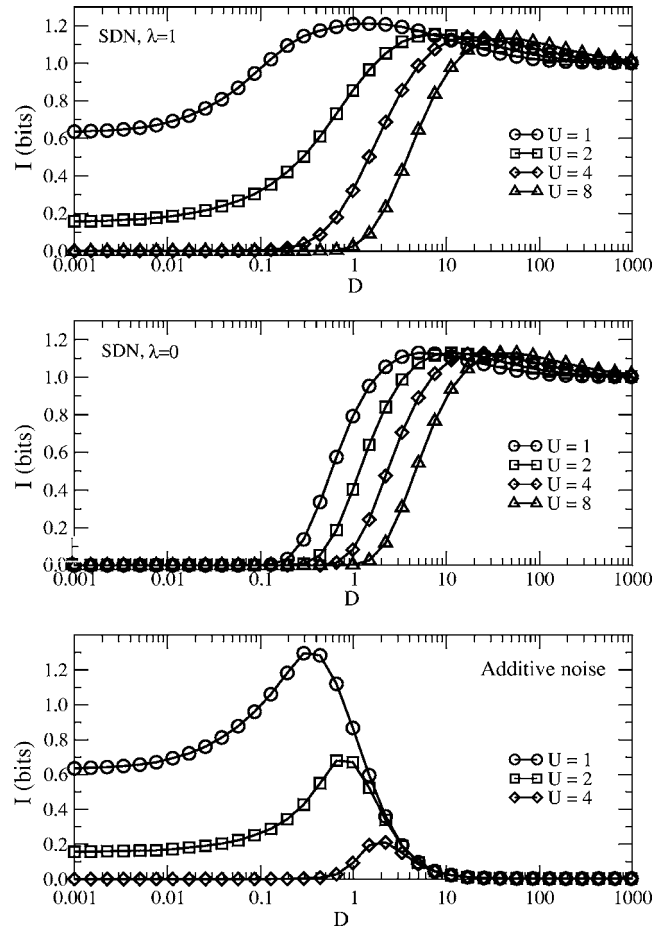


FIG. 7. The mutual information I as a function of the noise intensity D for various thresholds U with the rectified model (9) (upper and middle panels) and for additive noise (5) (lower panel). The results were obtained for $N=16$, various thresholds U , and a common Gaussian input signal x with a standard deviation $\sigma=1$. Other parameters were $\lambda=1$ (upper and lower panels) and $\lambda=0$ (middle panel). The curves with circles, squares, diamonds, and triangles correspond to $U=1$, $U=2$, $U=4$, and $U=8$, respectively.

V. CONCLUSION

In this paper we have investigated a simple neural-inspired model that consisted of an array of threshold units with signal-dependent noise.

The three main findings are as follows. First, we have shown a relationship between the mutual information and noise intensity that is similar to the phenomenon of stochastic resonance and suprathreshold stochastic resonance that have previously been observed with additive noise [10,11]. In particular, we have shown that a nonzero level of signal-dependent noise can lead to increased information transfer in threshold systems.

Second, we found that, for weak signals, the information transmission with signal-dependent noise is more effective than with additive noise. Furthermore, the performance of the system with signal-dependent noise is not compromised by high thresholds or, equivalently, weak signals; approxi-

mately the same level of information can be transmitted *independent* of the size of the signal. In principle, signals of arbitrary small amplitude can be enhanced by signal-dependent noise.

Finally, we observed that if the signal is first rectified, then the transmitted information can be made largely independent of the level of the signal-dependent noise. We speculate that, in principle, hair cells could potentially exploit

these result to provide robust signal coding in biological sensory systems.

ACKNOWLEDGMENT

The authors thank gratefully acknowledge support from EPSRC Grant No. EP/C523334/1.

-
- [1] K. Wiesenfeld and F. Moss, *Nature (London)* **373**, 33 (1995).
 [2] A. Bulsara and L. Gamaitoni, *Phys. Today* **49** (3), 39 (1996).
 [3] L. Gamaitoni, P. Hanggi, P. Jung, and F. Marchesoni, *Rev. Mod. Phys.* **70**, 223 (1998).
 [4] V. S. Anishchenko, A. B. Neiman, F. Moss, and L. Schimansky-Geier, *Phys. Usp.* **169**, 7 (1999) [*Phys. Usp.* **42**, 7 (1999)].
 [5] A. Longtin, A. Bulsara, and F. Moss, *Phys. Rev. Lett.* **67**, 656 (1991).
 [6] J. K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, *Nature (London)* **365**, 337 (1993).
 [7] J. J. Collins, C. C. Chow, and T. T. Imhoff, *Nature (London)* **376**, 236 (1995).
 [8] J. J. Collins, C. C. Chow, A. C. Capela, and T. T. Imhoff, *Phys. Rev. E* **54**, 5575 (1996).
 [9] D. R. Chialvo, A. Longtin, and J. Müller-Gerking, *Phys. Rev. E* **55**, 1798 (1997).
 [10] N. G. Stocks, *Phys. Rev. Lett.* **84**, 2310 (2000).
 [11] N. G. Stocks, *Phys. Rev. E* **63**, 041114 (2001).
 [12] N. G. Stocks and R. Mannella, in *Future Directions for Intelligent Systems and Information Sciences*, edited by N. Kasabov (Physica-Verlag, Heidelberg, 2000), pp. 236–247.
 [13] N. G. Stocks and R. Mannella, *Phys. Rev. E* **64**, 030902(R) (2001).
 [14] N. G. Stocks, in *Proceedings of the 16th International Conference on Noise in Physical Systems and Its Fluctuations*, edited by G. Bosman (World Scientific, Singapore, 2001), pp. 594–597.
 [15] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Phys. Lett. A* (to be published).
 [16] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Fluct. Noise Lett.* **5**, L457 (2005).
 [17] D. Rousseau and F. Chapeau-Blondeau, *Signal Process.* **85**, 571 (2005).
 [18] M. D. McDonnell, D. Abbott, and C. E. M. Pearce, *Microelectron. Eng.* **33**, 1079 (2002).
 [19] V. C. Anderson, *J. Acoust. Soc. Am.* **32**, 867 (1960).
 [20] N. G. Stocks, D. Allingham, and R. P. Morse, *Fluct. Noise Lett.* **2**, L169 (2002).
 [21] D. Rousseau, F. Duan, and F. Chapeau-Blondeau, *Phys. Rev. E* **68**, 3110701 (2003).
 [22] D. Rousseau and F. Chapeau-Blondeau, *Phys. Lett. A* **321**, 280 (2004).
 [23] Y. Wang and L. Wu, *Fluct. Noise Lett.* **5**, L435 (2005).
 [24] T. Hoch, G. Wenning, and K. Obermayer, *Neurocomputing* **52-54**, 541 (2003).
 [25] N. G. Stocks, *Phys. Lett. A* **279**, 308 (2001).
 [26] A. Manwani and C. Koch, *Neural Comput.* **11**, 1797 (1999).
 [27] E. F. Evans, in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer, Berlin, 1975), pp. 1–108.
 [28] P. Lánský and L. Sacerdote, *Phys. Lett. A* **285**, 132 (2001).
 [29] P. E. Greenwood and P. Lánský, *Biol. Cybern.* **92**, 199 (2005).
 [30] W. J. McGill and M. C. Teich, *J. Math. Psychol.* **39**, 146 (1995).
 [31] T. Furukawa, Y. Hayashida, and S. Matsuura, *J. Physiol. (London)* **276**, 211 (1978).
 [32] T. Furukawa, M. Kuno, and S. Matsuura, *J. Physiol. (London)* **322**, 181 (1982).
 [33] M. I. Dykman, D. G. Luchinsky, P. V. E. McClintock, N. D. Stein, and N. G. Stocks, *Phys. Rev. A* **46**, R1713 (1992).
 [34] B. Lindner and L. Schimansky-Geier, *Phys. Rev. Lett.* **86**, 2934 (2000).
 [35] L. Gamaitoni, F. Marchesoni, E. Menichella-Saetta, and S. Santucci, *Phys. Rev. E* **49**, 4878 (1994).
 [36] A. V. Barzykin and K. Seki, *Europhys. Lett.* **40**, 117 (1997).
 [37] A. V. Barzykin, K. Seki, and F. Shibata, *Phys. Rev. E* **57**, 6555 (1998).
 [38] V. Berdichevsky and M. Gitterman, *Phys. Rev. E* **60**, 1494 (1999).
 [39] C. C. Chow, T. T. Imhoff, and J. J. Collins, *Chaos* **8**, 616 (1998).
 [40] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
 [41] A. Nikitin and N. G. Stocks, *Proc. SPIE* **5467**, 202 (2004).