

Information transmission in parallel threshold arrays: Suprathreshold stochastic resonance

N. G. Stocks

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

(Received 22 September 2000; published 28 March 2001)

The information transmitted through a parallel summing array of noisy threshold elements with a common threshold is considered. In particular, using theoretical and numerical analysis, a recently reported [N. G. Stocks, Phys. Rev. Lett. **84**, 2310 (2000)] form of stochastic resonance, termed suprathreshold stochastic resonance (SSR), is studied in detail. SSR is observed to occur in arrays with two or more elements and, unlike stochastic resonance (SR) in a single element, gives rise to noise-induced information gains that occur independent of the setting of the threshold or the size of the signal. The transmitted information is maximized when all thresholds are set to coincide with the signal mean. In this situation, and for large arrays, the noise can enhance performance up to approximately half the theoretical noiseless channel capacity. The theory is tested against digital simulation.

DOI: 10.1103/PhysRevE.63.041114

PACS number(s): 05.40.-a, 02.50.-r, 05.45.-a

I. INTRODUCTION

The study of stochastic resonance (SR) [2] in threshold based systems has received considerable attention in recent years [2–5]. In such systems it is well known that, via the SR effect, the addition of noise can lead to an enhancement of the system's response to subthreshold signals. Initial studies used sinusoidal input signals and the output signal-to-noise ratio (SNR) to characterize the SR effect—the effect manifests itself as a noise induced maximum in the SNR. More recently, SR has been extended to include aperiodic broadband signals [6] and information theoretic measures, such as the average mutual information, have been introduced to characterize the dynamics [3,7–11]. Both broadband signals and information theory are employed in this study.

The great majority of previous studies [2] have focused on single element SR systems. Attention has recently turned to the study of networks of SR elements. A large number of network configurations and connectivity have been studied and include, globally coupled networks [12–14], randomly connected networks [15] and linear chains [16–18]. In comparison, parallel arrays of threshold elements, in which N SR elements are placed in parallel and their outputs summed at a common summing point, have received relatively little attention with only a handful of studies [19,6,20–22]. Parallel arrays (ensembles) are of considerable importance in many signal processing applications. For example, they can be used to model DIMUS sonar arrays (in the on target position) [23] and, for regularly spaced thresholds, Flash analogue-to-digital converters (ADC's) [24]. Additionally, parallel arrays have recently been used to model ensembles of sensory neurons [6,21,22]. Consequently, the study of internal system noise and SR effects in these systems are of importance to a number of signal processing and neurophysiological applications.

Stochastic resonance is commonly understood to be the enhancement, by noise, of the response of a system to a *weak* signal. By weak, one normally means with reference to an appropriate scale. This scale can either be taken as the (internal/external) noise intensity or, in a single threshold system such as a simple comparator, as the threshold level.

However, when determining whether SR is in principle observable, it is the size of the signal compared to the threshold level that is the important quantity. Normally, SR is only observed if the signal is smaller than the threshold level, i.e., it is subthreshold. For larger, suprathreshold signals, the SR effect disappears [3,25,7]. This has led to the common belief that SR type effects can only be observed for predominantly subthreshold signals. However, as will be demonstrated, this is only true for single element threshold systems. It is demonstrated here that parallel arrays can display another form of SR—termed suprathreshold stochastic resonance (SSR)—that occurs when the signal is predominantly suprathreshold. Additionally, this new form of SR leads to significantly greater signal enhancements than can be obtained using subthreshold signal levels. The paper is organized as follows: Sec. II introduces the model array and in Sec. III the calculation of the transmitted information and the theory of the SSR effect are presented. Section IV discusses the digital simulation and in Sec. V the theory and digital simulation results are compared and discussed. Finally, in Sec. VI, conclusions are drawn.

II. MODEL

A summing array of N threshold devices (Fig. 1) is considered. Each threshold device is subject to the same input signal $x(t)$ but independent noise, $\eta_i(t)$. The devices are modeled as Heaviside functions, the outputs, $y_i(t)$, being given by the response function,

$$y_i(t) = \begin{cases} 1 & \text{if } x(t) + \eta_i(t) > \theta_i, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where θ_i are the threshold levels and $i = 1, 2, \dots, N$. The response of the array is obtained by summing the individual responses of each device. Consequently, $y(t)$ represents the number of devices that are triggered at any instant of time. The signal is taken to be aperiodic and broadband. Although the initial theoretical discussion in Sec. III assumes no specific form for the signal distribution—all results presented are for Gaussian signal and noise.

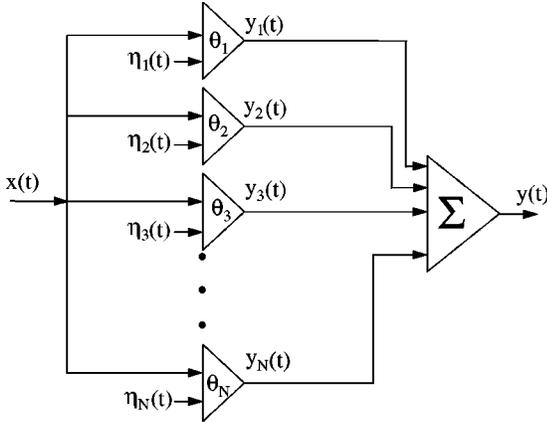


FIG. 1. A summing array of N threshold devices. Each device is subject to the same signal (a Gaussianly distributed signal with zero mean and standard deviation σ_x) but independent Gaussian noise.

By a suitable choice of threshold settings, the array can be used to model a number of applications arising in engineering and neurophysiology. For example, placing the thresholds levels regularly across the signal space results in a uniform quantization scheme, identical to that found in Flash analogue-to-digital converters. This case has been studied previously [1,26]. However, in this study, only the situation where all the threshold levels are set to same value, θ , will be considered. Arrays of this form have recently been considered in connection with neuronal ensembles [6,21,22] and are also applicable to digital-multibeam-steering (DIMUS) arrays used in passive sonar [23].

III. AVERAGE TRANSMITTED (MUTUAL) INFORMATION

A. Theoretical preliminaries

An information theoretic measure—the average mutual information—will be used to quantify the amount of information transmitted through the array. Although, traditionally, the SNR is often used to characterize SR, in practice this has only a limited utility for nonlinear systems subject to broadband excitation. The SNR only provides a meaningful measure under the assumption that the dynamics are approximately linear and the noise is Gaussian. If these assumptions are valid then the SNR can be related to the information flow through the system [27]. Consequently, measuring the SNR is equivalent to measuring the transmitted information. However, for weak noise, the dynamics of these arrays are highly nonlinear and no simple relation between the SNR and the transmitted information exist. It is in this regime that the SNR fails to give a meaningful characterization of the response. For example, passing a signal through a deterministic (noiseless), but noninvertible transfer function yields an infinite output SNR but, by construction, information is lost about the signal, i.e., no unique function exists that maps the response back into the signal. Other linear signal processing techniques, such as cross correlation, also suffer from similar deficiencies.

The average mutual (or transmitted) information, I , for the array shown in Fig. 1 (which in information theory is re-

garded as a semi-continuous channel) can be written [28]

$$\begin{aligned}
 I &= H(y) - H(y|x) \\
 &= - \sum_{n=0}^N P_y(n) \log_2 P_y(n) \\
 &\quad - \left(- \int_{-\infty}^{\infty} dx P_x(x) \sum_{n=0}^N P(n|x) \log_2 P(n|x) \right). \quad (2)
 \end{aligned}$$

$H(y)$ is the information content (or entropy) of $y(t)$ and $H(y|x)$ can be interpreted as the amount of encoded information lost in the transmission of the signal. $P_y(n)$ is the probability of the output $y(t)$ being numerically equal to n and $P(n|x)$ is the conditional probability density of the output being in state n given knowledge of the signal value, x . $P_x(x)$ is the probability density function (pdf) of the signal. The logarithms are taken to base 2 so I is measured in bits.

Equation (2) represents the appropriate definition of information for a channel that has a continuous input signal but a discrete output [28]. It should be noted that this definition has no explicit time dependence and, therefore, does not strictly treat the signal as a stochastic process. For this reason, I does not represent an information flow (measured in bits/s) but the average amount of information, measured in bits, that a measurement of the output yields about the instantaneous input signal value.

It will be assumed initially that the noise has an arbitrary pdf, $P_\eta(\eta)$, with a standard deviation, $\sqrt{\langle \eta^2 \rangle - \langle \eta \rangle^2} = \sigma_\eta$. If all information is lost in transmission $H(y|x) = H(y)$ (which occurs as $\sigma_\eta \rightarrow \infty$) and hence $I = 0$. Alternatively, if all encoded information is transmitted ($\sigma_\eta = 0$) $H(y|x) = 0$ and $I = H(y)$. Given it is straightforward to show [27] that for any nonzero σ_η , $H(y|x) < H(y)$, it would seem to follow that maximum information transfer occurs when there is no internal noise. However, this is not necessarily the case because internal noise also serves to increase $H(y)$. Consequently, the maximization of I by internal noise is a balance between additional useful information generated by the noise and the increased loss in information transmitted through the array with increasing σ_η . It is this ability of noise to maximize the transmitted information that is termed SR.

B. Calculation of I

The calculation of the transmitted information is straightforward when all thresholds are set equal to an arbitrary value θ . To proceed, it is first useful to simplify the formula for I . The conditional probability, $P(n|x)$, is easily obtained by noting that for any given signal value, x , each device acts independently under the influence of its own noise source. Consequently, the probability that n devices are triggered is given by the binomial distribution and hence $P(n|x) = C_n^N P_{1|x}^n P_{0|x}^{N-n}$, where $P_{1|x}$ is the conditional probability of a device being in state 1 (triggered) which is given by the cumulative distribution $P_{1|x} = \int_{\theta-x}^{\infty} P_\eta(\eta) d\eta$. C_n^N is the binomial coefficient and $P_{0|x} = 1 - P_{1|x}$. Using this result and

noting that $\sum_0^N n P(n|x) = N P_{1|x}$ and $\sum_0^N P(n|x) = 1$, the second summation on the right-hand side (RHS) of Eq. (2) can be carried out to yield

$$\begin{aligned} & \sum_{n=0}^N P(n|x) \log_2 P(n|x) \\ &= \sum_{n=0}^N P_y(n) \log_2 C_n^N + N(P_{1|x} \log_2 P_{1|x} + P_{0|x} \log_2 P_{0|x}). \end{aligned} \quad (3)$$

Noting that $P_y(n) = \int_{-\infty}^{\infty} P(n|x) P_x(x) dx$, the transmitted information can now be written in the simplified form

$$\begin{aligned} I &= - \sum_{n=0}^N P_y(n) \log_2 P'(n) - \left(-N \int_{-\infty}^{\infty} dx P_x(x) \right. \\ & \quad \left. \times (P_{1|x} \log_2 P_{1|x} + P_{0|x} \log_2 P_{0|x}) \right), \quad (4) \\ P'(n) &= \int_{-\infty}^{\infty} dx P_x(x) P_{1|x}^n P_{0|x}^{N-n}, \end{aligned}$$

where $P_y(n) = C_n^N P'(n)$.

Given the signal and noise distributions, $P_x(x)$ and $P_\eta(\eta)$, Eq. (4) can be used to calculate I numerically.

C. Suprathreshold SR (SSR)

The specific case when $\theta = \langle x \rangle$, where $\langle x \rangle$ is the first moment of $P_x(x)$, is now considered. It will be demonstrated that under this condition a new form of SR, not observable in a single threshold device, can be anticipated.

Although, in general, $P_x(x)$ and $P_\eta(\eta)$ must be specified to enable I to be obtained, this is not the case when these distributions are of the same form and $\theta = \langle x \rangle$. Under these conditions, the integrals and summations in Eq. (4) can be solved analytically. However, it should be stressed that the following analysis only applies under the assumption that $P_x(x)$ and $P_\eta(\eta)$ are identical, i.e., all moments, except the first, are the same. This condition applies, for the Gaussian signal and noise to be considered, when the signal and noise variances are equal.

Assuming identical signal and noise distributions (a part from a nonzero signal mean) the signal distribution can be written as $P_x(x) = P_\eta(x - \langle x \rangle)$. Taking $P_\eta(\eta)$ to be an even function of η and, without loss of generality, zero mean, it follows that, $P_{1|x} = \int_{\theta-x}^{\infty} P_x(\langle x \rangle - s) ds$. Introducing the change of variable, $y = P_{1|x}$, yields $dy/dx = P_x(\langle x \rangle - \theta + x)$ and setting $\theta = \langle x \rangle$, gives for $P'(n)$ in Eq. (4),

$$P'(n) = \int_0^1 dy y^n (1-y)^{N-n}. \quad (5)$$

The expression for $P'(n)$ is in the form of a beta function [29], the integral can now be solved to give $P'(n) = n!(N-n)!/(N+1)!$. This in turn yields $P_y(n) = 1/(N+1)$ which is independent of n . This result states that all output states, n ,

are occupied with equally probability; this is the well known condition for maximum entropy [27]. Hence, this leads to the interesting result that $H(y)$ is maximized when the signal and noise distribution are matched. This does not imply, however, that I is also maximized.

To finalize the calculation, the second integral on the RHS of Eq. (4) can be solved using the same procedure to give $-N/2 \log_e 2$. Combing this result with the result for $P_y(n)$ and substituting into Eq. (4) finally yields the exact result

$$I = \log(N+1) - \frac{N}{2 \log_e 2} - \frac{1}{N+1} \sum_2^N (N+1-2n) \log_2 n. \quad (6)$$

It is interesting to note that this results states that I is dependent only on the number of elements in the array and is, therefore, independent on the exact form of the signal and noise distributions. Equation (6) was, however, derived under the assumption that the signal and noise distributions were the same. If they are not, then no simple change of variable could be found that removed the functional dependence of the signal and noise PDFs.

Of particular interest is how I scales with N . This can be determined by approximating the summation in Eq. (6) using the formula $\sum_2^N f(n) = \int_1^{N+1} f(n) dn - 1/2(f(1) + f(N+1))$. This yields

$$I = \frac{1}{2} \log_2(N+1) - \frac{1}{2 \log_e 2} + O(1/N). \quad (7)$$

Consequently, for large N , $I \approx 1/2 \log_2 N$.

It is now straightforward to establish that this result implies that a noise induced maximum must occur in I . As already discussed, in the limit $\sigma_\eta \rightarrow \infty$, I must tend to zero. In the limit $\sigma_\eta \rightarrow 0$, $I \rightarrow H(y)$. Therefore, a noise induced maximum must occur if $H(y)$ (evaluated at $\sigma_\eta = 0$) is smaller than the finite noise result given in Eq. (7). It is straightforward to show that this must be the case when N is sufficiently large. In the absence of noise all the devices switch in unison and hence $P_y(n) = 0$ for $n = 1, 2 \dots N-1$, $P_y(0) = \int_{-\infty}^{\theta} P_x(x) dx$ and $P_y(N) = 1 - P_y(0)$. This implies

$$I(\sigma_\eta = 0) = -P_y(0) \log_2 P_y(0) - P_y(N) \log_2 P_y(N). \quad (8)$$

$I(\sigma_\eta = 0)$ is maximized when $P_y(0) = P_y(N) = 1/2$. This occurs when $\theta = \langle x \rangle$ [as long as $P_x(x)$ is even functioned about its mean] and gives $I(\sigma_\eta = 0) = 1$ bit. In general, $I(\sigma_\eta = 0) \leq 1$ for arbitrary θ and $P_x(x)$. Consequently, in the absence of noise, the maximum information the array can transmit is only 1 bit; hence from Eq. (7) it is easy to infer that for sufficiently large N a noise induced maximum must occur. Unlike conventional SR, this SR effect is to be anticipated when all thresholds are set suprathreshold (in the sense that deterministic signal induced threshold crossings are maximized) with respect to the signal and will, therefore, be referred to as suprathreshold SR (SSR).

D. Gaussian signal and noise

Although much of the above discussion has placed no restrictions on the signal and noise pdfs, Gaussian signal and noise are used for the results presented in Sec. V. Here the calculation of I is considered for this particular case. Taking $P_x(x) = (1/\sqrt{2\pi\sigma_x^2})\exp(-(x-\langle x \rangle)^2/2\sigma_x^2)$ and $P_\eta(\eta) = (1/\sqrt{2\pi\sigma_\eta^2})\exp(-\eta^2/2\sigma_\eta^2)$ gives $P_{1|x} = \frac{1}{2}\text{erfc}(-u)$ where $u = (x - \theta)/\sqrt{2\sigma_\eta^2}$. Substituting these quantities into Eq. (4) and performing a change of variable to u gives

$$I = - \sum_{n=0}^N P_y(n) \log_2 P'(n) + N \int_{-\infty}^{\infty} du \frac{\sigma}{\sqrt{\pi}} \times \exp\left(-\left(\sigma u - \frac{(\theta - \langle x \rangle)}{\sqrt{2\sigma_x}}\right)^2\right) \left(\frac{1}{2} \text{erfc}(-u) \log_2 \frac{1}{2} \times \text{erfc}(-u) + \left(1 - \frac{1}{2} \text{erfc}(-u)\right) \log_2 \left(1 - \frac{1}{2} \text{erfc}(-u)\right)\right), \quad (9)$$

where

$$P'(n) = \int_{-\infty}^{\infty} du \frac{\sigma}{\sqrt{\pi}} \exp\left[-\left(\sigma u - \frac{(\theta - \langle x \rangle)}{\sqrt{2\sigma_x}}\right)^2\right] \times \left(\frac{1}{2} \text{erfc}(-u)\right)^n \left(1 - \frac{1}{2} \text{erfc}(-u)\right)^{N-n},$$

and $\sigma = \sigma_\eta/\sigma_x$, $P_y(n) = C_n^N P'(n)$. Equation (9) can be solved numerically to obtain I . In the particular case $\theta = \langle x \rangle$ it can be seen that I depends only on the parameter σ —where σ^2 is interpreted as the inverse signal-to-internal-noise ratio.

E. $N=1$ and $\theta = \langle x \rangle$

The results obtained in Sec. III C indicate that SSR is to be anticipated if N is sufficiently large—thus suggesting that SSR cannot be observed in a single threshold system. This is consistent with previous studies [3,25] that have demonstrated that SR effects are removed if the threshold is lowered to suprathreshold levels. However, this can be proved under general conditions and for completeness a proof is included here. The only assumption that will be made is that the signal and noise PDF's are even functions of their state variable. In this case setting $\theta = \langle x \rangle$ results in $P_y(0) = P_y(1) = 1/2$ (independent of σ_η) and hence $H(y) = 1$. Therefore, $I = 1 - H(y|x)$ for arbitrary σ_η , where

$$H(y|x) = - \int_{-\infty}^{\infty} dx P_x(x) [P_{1|x} \log_2 P_{1|x} + (1 - P_{1|x}) \log_2 (1 - P_{1|x})]. \quad (10)$$

Differentiating I with respect to σ_η gives

$$\frac{dI}{d\sigma_\eta} = \int_{-\infty}^{\infty} dx P_x(x) \log_2 \left(\frac{P_{1|x}}{1 - P_{1|x}}\right) \frac{dP_{1|x}}{d\sigma_\eta}. \quad (11)$$

Given that $P_{1|x} - 1/2$ is always an odd function about $x = \langle x \rangle$, it follows that $dP_{1|x}/d\sigma_\eta$ will also be an odd function about $x = \langle x \rangle$ with a negative sign for $x > \langle x \rangle$. Splitting the integral in Eq. (11) into the ranges $-\infty \rightarrow \langle x \rangle$ and $\langle x \rangle \rightarrow \infty$ (that is $P_{1|x}$ between 0 and 1/2 and then 1/2 and 1) it can be seen that, the factor $\log_2 [P_{1|x}/(1 - P_{1|x})] < 0$ for $x < \langle x \rangle$ and positive otherwise. The integrand is therefore negative for all x and it follows that $dI/d\sigma_\eta < 0$ for all σ_η . Therefore SR can not occur.

IV. SIMULATIONS

To confirm the validity of the theory developed in Sec. III digital simulations were also undertaken. Due to the simplicity of the array, the algorithm to generate $y(t)$ was straightforward but some care was required when calculating the transmitted information. The transmitted information was obtained by constructing the distributions $P_y(n), P_x(x), P(n|x)$ and using Eq. (2) directly. To obtain good statistics (approximately 1% error), signal lengths in excess of 10^5 independent samples were used to obtain the distributions. The ‘‘bin’’ size used to construct $P_x(x)$ was also found to critically affect the results. Because x is treated as a continuous variable and y as a discrete one, it is necessary that x is discretized into many more bins than the number of states occupied by y (which is simply given by N). In practice, to produce convergent results it was necessary to discretize x with a resolution better than $1/100N$. Consequently, increasing the number of elements not only increases the time required to compute y , but also requires an increase in the signal length to compensate for the finer resolution required when discretizing x . This was found to limit the number of devices that could be simulated to < 100 .

V. RESULTS AND DISCUSSION

In the previous section, theoretical arguments were forwarded that suggest a new form of SR is to be anticipated when all the thresholds are set to coincide with the dc signal level. In this section a comparison between the theory and the results of the digital simulation are presented. All results presented are for Gaussian signal and noise with $\langle x \rangle = 0$ and $\langle \eta \rangle = 0$.

A. $N=1$

To place the multielement results in context, the results for $N=1$ are first reviewed. Figure 2 shows the dependence of the transmitted information on noise intensity for a single element. The threshold value, θ , is expressed in units of the standard deviation of the signal, σ_x . Clearly, an SR effect (i.e., a noise is induced maximum) is only observed when $\theta = 2.83$. Lowering the threshold below this level significantly improves the transmitted information but removes the beneficial role of the noise. This behavior has been reported previously in a number of different threshold systems [25,3]. Lowering the threshold removes SR effects by making the signal ‘‘more suprathreshold.’’ The suprathreshold nature of the signal is clearly observed by considering the behavior of I when $\sigma = 0$. Without noise, signal information can only be

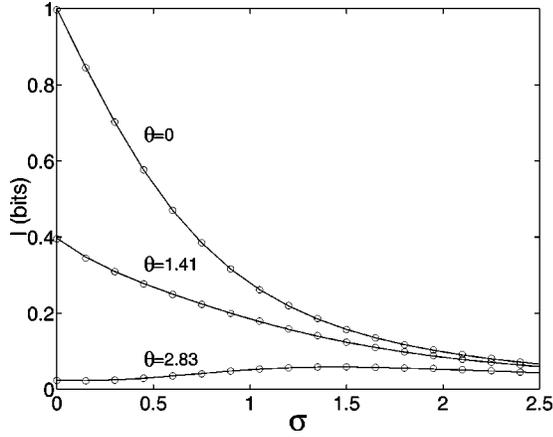


FIG. 2. Plot of transmitted information against $\sigma = \sigma_\eta/\sigma_x$ for $N=1$ and various θ . The data points are the results of the digital simulation and the solid lines were obtained by numerically evaluating Eq. (9).

transmitted through deterministic threshold crossings—as the threshold is lowered these increase and hence the information increases. As discussed in Sec. III E $H(y)$ (and hence I) is maximized by setting $\theta = \langle x \rangle$; this condition maximizes the deterministic threshold crossings and leads to a transfer of 1-bit of information.

For a single device, it can be seen that the noise-induced enhancements are relatively weak. Consequently, they are only observed if the contribution to the information from deterministic threshold crossings do not dominate—it was observed that this requires $\theta \geq 2.45$. In this case the signal spends at least 99.3% of its time below the threshold and hence can be termed *predominantly* subthreshold. It is interesting to note that, due to symmetry, the same SR effects are also observed if $\theta \leq -2.45$, i.e., when the signal spends 99.3% of its time above the threshold. This indicates that the fashion of using the term suprathreshold to distinguish the dynamics from the subthreshold case is somewhat ambiguous. Clearly what is meant here is to distinguish between the cases of deterministic and non-deterministic threshold crossings. For this reason, the term *predominantly suprathreshold* will be used to indicate the situation where deterministic threshold crossings dominate the transmitted information in an individual device.

B. $\theta = \langle x \rangle$

Here it is demonstrated that the situation is quite different if there is more than one device. Figure 3 shows the results for all $\theta = \langle x \rangle$ and various N . It is immediately observed that SR type behavior is manifest for all $N > 1$. As N increases, the maximum value attained by I also increases. These results are in keeping with those presented in Sec. III and confirm the existence of SSR. It is noted that SSR occurs even if there are only two elements; this was not anticipated because Eq. (7) is only valid when $\sigma = \sigma_\eta/\sigma_x = 1$ and hence does not predict the value of the maximum. The maximum is seen to shift to higher noise intensities as N increases but, for reasons discussed below, cannot pass $\sigma = 1$.

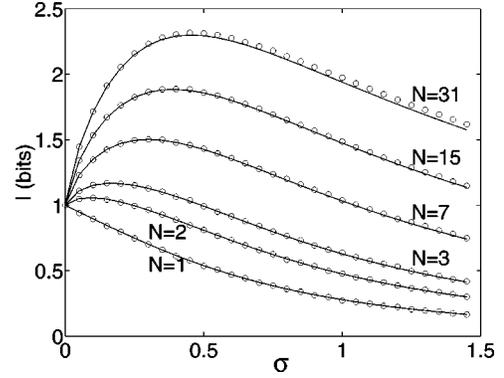


FIG. 3. Plot of transmitted information against $\sigma = \sigma_\eta/\sigma_x$ for various N and $\theta=0$. The data points are the results of the digital simulation and the solid lines were obtained by numerically evaluating Eq. (9).

The results in Fig. 3 were obtained with $\sigma_x = 1$, however, the theoretical analysis in Sec. III indicates that the information is dependent only on the ratio σ_η/σ_x and not on σ_x alone. Hence, the same set of curves can be reproduced for arbitrary σ_x —as long as the noise is scaled accordingly, i.e., SSR occurs for arbitrary signal strengths. This is, of course, not true of SR in a single element where SR effects disappear for sufficiently large suprathreshold signal amplitudes.

The mechanism giving rise to SSR is quite different to that of classical SR and is not connected to a previously reported form of suprathreshold SR [30]. In the absence of noise, all devices switch in unison and consequently the array acts like a single bit ADC ($I=1$). The fact that all devices switch in unison in response to the signal implies that they also carry identical information about the signal. Consequently, no additional information is obtained by having more than one device. Ideally, one would wish that the devices carry at least a degree of independent signal information. This degree of independence is facilitated by the addition of noise. At any instant of time, finite noise results in a distribution of thresholds that, in turn, leads to the signal being “sampled” at N randomly spaced points across the signal space. In effect, the noise allows additional bits of information (output states) of the system to be accessed, resulting in an increase in the output entropy $H(y)$. Although, the information content of each individual device is reduced because of the noise (all of which individually follow the $N=1$ curve), the sum total from all the devices results in a net gain in information. Consequently, SSR largely arises due to the initial increase in $H(y)$.

This can clearly be seen in Fig. 4, which shows the contributions of $H(y)$ and $H(y|x)$ to I . For all values of N , except $N=1$, $H(y)$ rapidly increases with noise intensity, reaches a maximum, and then decreases. $H(y)$ reaches its maximum value of $\log_2(N+1)$ when all output states are equally probable; as discussed in Sec. III, this occurs when the signal and noise PDF’s are exactly matched which requires $\sigma=1$. In contrast, the conditional entropy, $H(y|x)$, always increases monotonically with increasing noise intensity but, initially, at a slower rate. It increases asymptotically

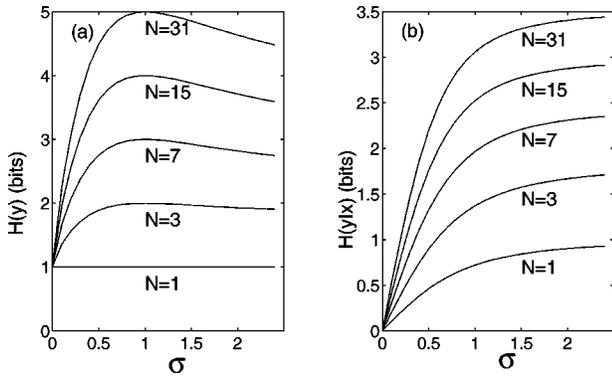


FIG. 4. Plot of (a) $H(y)$ and (b) $H(y|x)$ against σ for various N and $\theta=0$. The curves were obtained from Eq. (9).

to the entropy given by $-\sum_0^N P_y(n) \log_2 P_y(n)$ where $P_y(n) = C_n^N / 2^N$. The monotonic behavior of $H(y|x)$ and the fact that $H(y)$ always reaches its maximum when $\sigma=1$, implies that the maximum in I must always occur for $\sigma \leq 1$. This conclusion is valid independent of the value of N but, as discussed below, is dependent on the threshold level being set equal to the signal mean. Consequently, although as N increases the maximum in the information shifts to higher noise intensities it cannot exceed unity.

These results tend to suggest that at large N Eq. (6) gives a reasonable estimate for the maximum information attainable which is approximately equal to half the maximum output entropy. In the absence of noise, the maximum output entropy of $\log_2(N+1)$ also represents the channel capacity of the array and, hence, these results indicate that SSR can lead to information gains which approach half the noiseless channel capacity. This is an interesting result when compared to the performance of a single device. In Fig. 2 the results for $\theta=2.83$ indicate that SR effects are relatively weak—the maximum noise-enhanced information gain is only about 0.05 bits—compared to a noiseless channel capacity of 1 bit. Thus, noise only helps to recover at most 5% of the potential channel capacity of a single device compared to a possible 50% for an array.

In Sec. III it was predicted that for large N and $\sigma=1$, I should scale approximately as $1/2 \log_2(N)$. This is tested in Fig. 5. The solid line was calculated using Eq. (6) and the circular data points are the results from the digital simulation. Good agreement is observed thus confirming the validity of Eq. (6). The dashed line was obtained from the approximation derived in Eq. (7). Despite slightly underestimating the exact result this approximation clearly predicts the correct scaling at large N , thus confirming the $1/2 \log_2(N)$ scaling. However, it is important to note this scaling was derived for the specific case $\sigma=1$. For other values of σ the scaling breaks down; this is demonstrated by the simulation results for $\sigma=0.2$ (crosses)—the dot-dashed line is a fit to the data. The fit predicts a scaling of $I \sim 0.29 \log_2(N)$ and, therefore, generally the scaling is seen to depend on noise intensity as $I \sim f(\sigma) \log_2(N)$ where $f(1) = 1/2$.

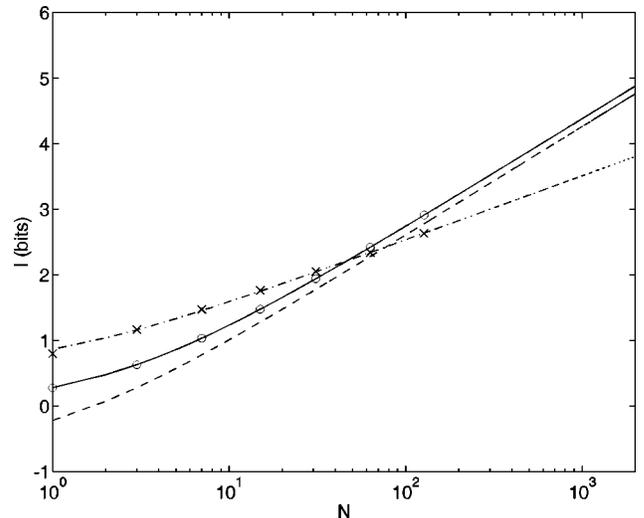


FIG. 5. Scaling of I with N . The crosses and circles are data points from the digital simulation. The solid line was obtained from Eq. (6), the dashed line is the approximation Eq. (7) and the dot-dashed line is a fit to the data.

C. θ dependence

All the results discussed so far were obtained by setting all thresholds equal to the mean of the signal, the dependence on θ is now discussed. Figures 6 and 7 show the effects of varying θ ; Fig. 6 shows the transmitted information and Fig. 7 the contributions from $H(y)$ and $H(y|x)$. In these figures the number of elements was held fixed ($N=63$). Considering Fig. 6, it can be seen that, similar to the $N=1$ case (Fig. 2), the information is maximized when $\theta = \langle x \rangle = 0$. However, SR type behavior is now observed independent of the threshold value and is significantly enhanced—even for the sub-threshold signals. Additionally, as θ is increased the maximum in I is shifted to higher noise values and can easily surpass $\sigma=1$. The reason for this can be seen in Fig. 6(a);

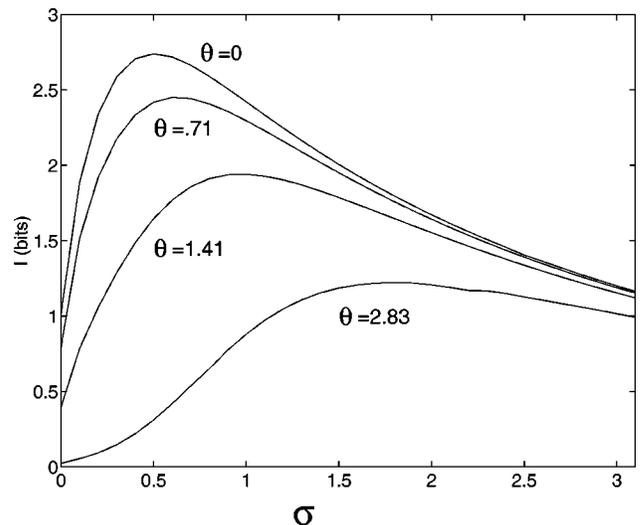


FIG. 6. Plot of transmitted information against σ for $N=64$ and various θ . The curves were obtained from Eq. (9).

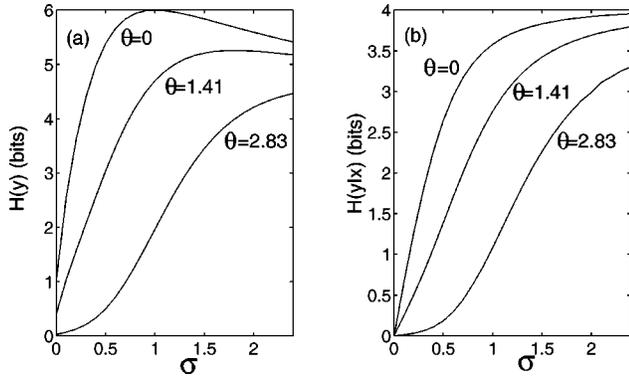


FIG. 7. Plot of (a) $H(y)$ and (b) $H(y|x)$ against σ for various $N=64$ and various θ . The curves were obtained from Eq. (9).

the position of the maximum in $H(y)$ is seen to strongly depend on the threshold setting. This is, of course, an obvious result—as the thresholds are moved away from the signal mean, larger noise intensities are required to distribute the threshold levels over the signal space.

The results in Fig. 6 raise an interesting point; as the thresholds are increased away from the signal mean there is a smooth change from suprathreshold to subthreshold signals. However, there is no obvious change over between SSR and conventional SR. This implies these two effects cannot be distinguished. Obviously when the signal is subthreshold the mechanism giving rise to SSR does not “switch off”—the noise still enables additional output states of the system to be accessed thus resulting in an improvement of the transmitted information. In this respect the term SSR is a bit of a misnomer, the SSR mechanism does enable noise to enhance the detection of suprathreshold signals but not exclusively so. Comparing the subthreshold results ($\theta=2.83$) in Figs. 2 and 6 shows that the subthreshold SR effect is significantly enhanced by increasing N . This is because both SSR and conventional SR effects are now contributing to the information gain.

D. Signal power and signal-to-noise ratio SNR

Traditionally, SR effects have been characterized by quantities such as the output signal power and SNR. It is, therefore, of some interest to consider these quantities for this system. The broadband nature of the signals means that Fourier techniques cannot be employed to experimentally measure either the signal power or the SNR; they can, however, be easily obtained from a time domain analysis. The response of the array, $y(t)$, can be split into a signal contribution, $\langle y(t) \rangle$ and a noise contribution, $n(t)$ where $y(t) = \langle y(t) \rangle + n(t)$. The brackets $\langle \dots \rangle$ denote an ensemble average over the noise and hence, in the limit $N \rightarrow \infty$, $y(t) \rightarrow \langle y(t) \rangle$. The signal power, S_p , and noise power, N_p are then given by

$$S_p = \overline{(\langle y(t) \rangle - \langle y(t) \rangle)^2}, \quad (12)$$

$$N_p = \overline{(n(t) - \langle n(t) \rangle)^2}, \quad (13)$$

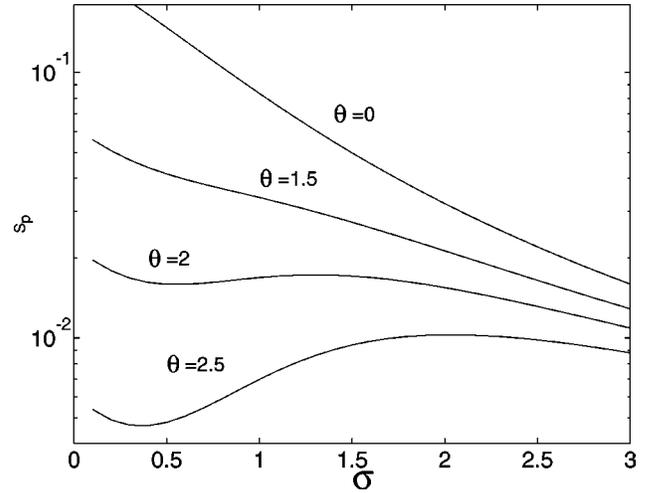


FIG. 8. Plot of the signal power, S_p , against σ for $N=64$ and various θ . The curves were obtained from Eq. (13).

where bars denote time averages. The SNR is defined in the conventional manner as the ratio S_p/N_p . Equations (12) and (13) enable the SNR to be obtained from simulation data in the time domain. To proceed further with the calculation one notes that the signals and noise used in this study are ergodic, hence the time averages can be replaced with averages over the signal distribution. For the case $\theta = \langle x \rangle$, the statistics of the response follow a Binomial distribution and hence

$$S_p = N^2 \int_{-\infty}^{\infty} P_x(x) (P_{1|x} - \bar{y})^2 dx, \quad (14)$$

$$N_p = N \int_{-\infty}^{\infty} P_x(x) P_{1|x} (1 - P_{1|x}) dx, \quad (15)$$

where $\bar{y} = \int_{-\infty}^{\infty} P_x(x) P_{1|x} dx$. Equations (14) and (15) can now be used to calculate the SNR numerically.

Before discussing the results it is worth noting that the definition of the SNR differs slightly from that typically used to characterize SR for sinusoidal signals. For a sinusoidal signal S_p yields the *total* integrated signal output power, i.e., the power in the fundamental plus all harmonics. This differs from the conventional definition used for SR which normally only considers the power in the fundamental. However, for broadband signals it is not possible to distinguish individual spectral components and hence the modified definition is necessary. N_p is essentially unchanged from the standard definition and is simply equal to total integrated power of the spectral background (as opposed to the power at the forcing frequency.) It should be noted that the SNR defined in this manner (i.e., total signal power over total noise power) is the one conventionally used in engineering when treating broadband signals [27].

Figure 8 shows the dependence of S_p on noise intensity for several values of θ and Fig. 9 shows the SNR. Considering the signal power first, it is clear that for predominantly suprathreshold signals (i.e., $\theta=0$ and 1.5) no maximum is observed. Increasing the threshold value, however, does lead to a noise-induced maximum when $\theta > 2$. These results seem

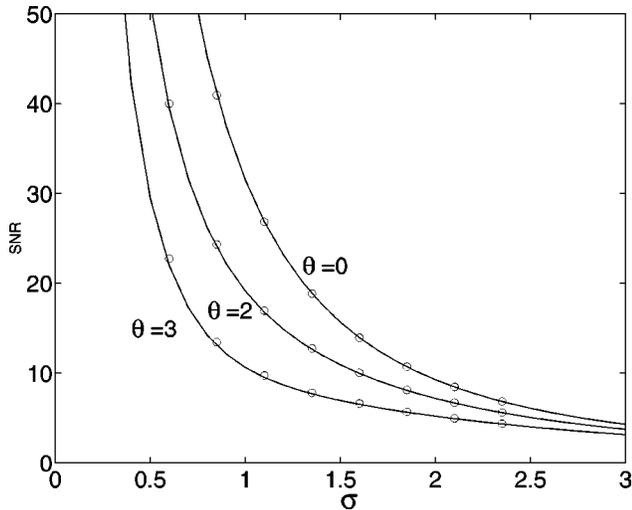


FIG. 9. Plot of the output signal-to-noise ratio against σ for $N=64$ and various θ . The data points are results from the digital simulation and the curves were obtained from Eqs. (13) and (14).

to indicate, therefore, that a subthreshold SR effect occurs but there is no evidence that the SSR mechanism leads to an enhancement of the signal power. In principle, this raises question about whether SSR is truly an SR effect—this is, of course, purely a matter of definition—but it does illustrate that SSR has to be defined as an improvement in the transmitted information rather than signal power. The reason why SSR improves the transmitted information but not the signal power is straightforward to understand. For $\theta = \langle x \rangle = 0$ and $\sigma = 0$ the output of the array switches between its maximum range 0 and N —thus the signal power is maximized. Adding noise can only result in this range being reduced—thus lowering the signal power. However, in the noiseless case all that is known about the signal is its polarity—the whole array switches to N when the signal is above zero and to 0 when below—thus all but 1-bit (assuming both output states are equally probable) of signal information is lost. The problem with signal power as a measure of system performance is that it does not take into account the form of the response but only quantifies the size of the signal-induced effect. This may have some use in the context of signal detection (provided one does not need to know much about the signal being detected) but is not so useful when the “quality” of signal detection is important.

The results for the SNR (Fig. 9) are more surprising. No form of SR effect, SSR or subthreshold SR, could be detected regardless of the positioning of the threshold. That SSR effects do not occur can easily be inferred from the results in Fig. 8, but the fact that conventional SR effects are also not observed is unexpected. This result has been confirmed both numerically [solid line obtained using Eqs. (14) and (15)] and by digital simulation using Eqs. (12) and (13) (circular data points). The reason that SR does not occur even for predominantly subthreshold signals is due to the distributed nature (in signal space) of the signal. Gaussian signals, although appearing to decay relatively quickly due to

the $\exp(-x^2)$ factor, do not decay “quickly enough” for SR effects to occur. The peak-signal-to-threshold distance is not well defined, as it is in the case of a sinusoid, and hence a smearing of the SR effect takes place. This is also the reason why the curves of signal power in Fig. 8 show an initial decrease with increasing noise intensity. Although a Gaussian distribution is not a good representation of real signals, for signals such as speech and music the situation is worse. These distributions tend to have long tails that are typically characterized by exponential decays and hence have an even slower decay than a Gaussian [31]. Consequently, SR effects in the SNR are not generally to be anticipated in threshold based systems when complex signals are used.

These results help to set the SSR effect in context. Based on the SNR it seems that, at least for Gaussianly distributed signals, no form of SR effect is observed; but in the transmitted information both SSR and conventional SR occur. Consequently, the fact that SSR is defined as an improvement in the transmitted information is not so restrictive as at first might appear.

VI. CONCLUSIONS

The phenomenology of a new form of stochastic resonance, termed SSR, that occurs in parallel threshold arrays has been discussed in detail. SSR is found to differ from SR in a single threshold element in a number of important ways. First, SSR can be observed with signals of arbitrary magnitude; there is not, therefore, the restriction that the signal must be weak (subthreshold) for SR effects to be observed—as is the case in a single threshold element. Second, the SSR effect is maximized (maximum noise-induced information gains) when the threshold level is set to coincide with the dc-signal level. In a single element, this condition removes the beneficial role of the noise. Third, SSR can result in large information gains—for large arrays the noise can recovery of up to 50% of the noiseless channel capacity. This appears to be significantly greater than is attainable in a single element. Indeed, these information gains appear to significant. In a recent study [1] it was found that, for signals comparable in size to the internal noise, arrays designed around the SSR effect could outperform those designed around conventional quantization techniques. This therefore seems to indicate that SSR may well have signal processing applications for the detection of weak signals. As discussed in the introduction, the array studied is a good model of a DIMUS sonar array in the on target position. Although the noise sources at the hydrophones are predominantly from external sources, due to the physical separation of the hydrophones they are nevertheless largely independent. Consequently, the results presented suggest that these noise sources are actually necessary for high fidelity signal detection and coding and should be taken into account when designing both the array and the signal-processing stages.

Finally, it is of some interest to discuss the results in the context of neuronal ensembles. It has recently been demonstrated that the SSR effect also occurs in an array of FitzHugh-Nagumo neurons [32]. Therefore, the results pre-

sented for the simple threshold network should have some applicability to real neuronal ensembles. It has been established that the transinformation is maximized when all thresholds are set to coincide with the dc-signal level. This has an interesting analogy to the well established phenomenon of dc adaptation in sensory neurons. One well known example is in light and dark adaptation of the human eye

[33]. The threshold levels adapt, via chemical changes in the cones and rods, to the ambient mean light level. This enables the eye to operate over a wide range of light intensities covering a 10^9 fold change in energy flux. Therefore, it may be possible that dc adaptation takes place, not only to increase the dynamic range over which the eye can operate, but also to enhance signal encoding via the SSR effect.

-
- [1] N. G. Stocks, Phys. Rev. Lett. **84**, 2310 (2000).
- [2] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998); *Proceedings of the International Workshop on Fluctuations in Physics and Biology: Stochastic Resonance, Signal Processing and Related Phenomena*, edited by R. Mannella and P. V. E. McClintock, [Nuovo Cimento Soc. Ital. Fis., D **17D**, 653 (1995)].
- [3] A. R. Bulsara and A. Zador, Phys. Rev. E **54**, R2185 (1996).
- [4] P. Jung, Phys. Lett. B **207**, 93 (1995).
- [5] Z. Gingl, L. B. Kiss, and F. Moss, Europhys. Lett. **29**, 191 (1995).
- [6] J. J. Collins, C. C. Chow, T. T. Imhoff, Nature (London) **376**, 236 (1995).
- [7] J. E. Levin, and J. P. Miller, Nature (London) **380**, 165 (1996).
- [8] C. Heneghan, C. C. Chow, J. J. Collins, T. T. Imhoff, S. B. Lowen, and M. C. Teich, Phys. Rev. E **54**, R2228 (1996).
- [9] F. Chapeau-Blondeau, Phys. Rev. E **55**, 2016 (1997).
- [10] G. Deco and B. Schurmann, Physica D **17**, 276 (1998).
- [11] A. Neiman, B. Shulgin, V. Anishchenko, W. Ebeling, L. SchimanskyGeier, J. Freund, Phys. Rev. Lett. **76**, 4299 (1996).
- [12] A. R. Bulsara, and G. Schmera, Phys. Rev. E **47**, 3734 (1993).
- [13] M. E. Inchiosa and A. R. Bulsara, Phys. Rev. E **52**, 327 (1995); **53**, R2021 (1996).
- [14] W. J. Rappel and A. Karma, Phys. Rev. Lett. **77**, 3256 (1996).
- [15] J. Pham, K. Pakdaman, and J. F. Vibert, Phys. Rev. E **58**, 3610 (1998).
- [16] J. F. Lindner, B. K. Meadows, W. L. Ditto, M. E. Inchiosa, and A. R. Bulsara, Phys. Rev. Lett. **75**, 3 (1995); J. F. Lindner, S. Chandramouli, A. R. Bulsara, M. Locher, and W. L. Ditto, *ibid.* **81**, 5048 (1998).
- [17] M. Locher, G. A. Johnson, and E. R. Hunt, Phys. Rev. Lett. **77**, 23 (1996); **77**, 4698 (1996).
- [18] F. Marchesoni, L. Gammaitoni, and A. R. Bulsara, Phys. Rev. Lett. **76**, 2609 (1996).
- [19] E. Pantazelou, F. Moss, and D. Chialvo, *Noise in Physical Systems and 1/f Noise* (AIP, New York, 1993), p. 549.
- [20] A. Neiman, L. SchimanskyGeier, and F. Moss, Phys. Rev. E **56**, R9 (1997).
- [21] X. Pei, L. Wilkens, and F. Moss, Phys. Rev. Lett. **77**, 4679 (1996).
- [22] D. R. Chialvo, A. Longtin, and J. Müller-Gerking, Phys. Rev. E **55**, 1798 (1997).
- [23] V.C. Anderson, J. Acoust. Soc. Am. **32**, 867 (1960); P. Rudnick, *ibid.* **32**, 871 (1960); W. R. Remley, *ibid.* **39**, 702 (1966).
- [24] V. K. Madisetti and D. B. Williams, *The Digital Signal Processing Handbook* (CRC Press, Boca Raton, FL, 1998).
- [25] D. DeWeese, and W. Bialek, Nuovo Cimento D **17**, 733 (1995).
- [26] L. Gammaitoni, Phys. Rev. E **52**, 4691 (1995); Phys. Lett. A **208**, 315 (1995).
- [27] M. F. Reza, *An Introduction to Information Theory* (Dover, New York, 1994).
- [28] A. Feinstein, *Foundations of Information Theory* (McGraw-Hill, New York, 1958).
- [29] A. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).
- [30] F. Apostolico, L. Gammaitoni, F. Marchesoni, and S. Santucci, Phys. Rev. E **55**, 36 (1997).
- [31] J. Dunlop and B. G. Smith, *Telecommunications Engineering* (Van Nostrand Reinhold, England, 1986).
- [32] N. G. Stocks and R. Mannella, (unpublished).
- [33] G. Somjen, *Neurophysiology—The Essentials* (Willaims & Wilkins, Baltimore, 1983).