Suprathreshold Stochastic Resonance in Multilevel Threshold Systems

Article *in* Physical Review Letters · April 2000

DOI: 10.1103/PhysRevLett.84.2310 · Source: PubMed

1 author:

Nigel G Stocks
University of Warwick

Suprathreshold Stochastic Resonance in Multilevel Threshold Systems

N. G. Stocks
School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom
(Received 18 June 1999)

A new form of stochastic resonance (SR) that occurs in multilevel threshold signal detectors is reported. In contrast to classical SR, which extends the dynamic range of threshold detectors to subthreshold signal levels, this new form of SR extends the dynamic range to suprathreshold signal strengths. The effect is most dominant, and can outperform networks based on standard engineering design, when all thresholds adapt to the dc level of the signal. This has an interesting analogy to dc adaptation in neurons. The possible connection between these two effects is discussed.

PACS numbers: 05.40.Ca, 02.50.Ey

The study of stochastic resonance (SR) in threshold based systems has received considerable attention in recent years [1–3]. However, the great majority of these studies have focused on systems with a single threshold. In this Letter a multilevel parallel network of threshold devices is considered. Such networks are of considerable importance in many signal processing applications. For example, they are the basis on which Flash analog-to-digital converters (ADCs) [4] work. Additionally, multilevel networks have recently been used to model ensembles of sensory neurons [5–7]. Consequently, the study of SR effects in these systems is of some importance.

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 be taken either 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 observed only if the signal is smaller than the threshold level, i.e., it is subthreshold. For large, suprathreshold signals, the SR effect disappears [2,3]. This has led to the common belief that SR type effects can be observed only for predominantly subthreshold signals (although residual SR effects are known to occur for marginally suprathreshold signals [8]). However, as we will see, this is true only for single threshold systems. It is demonstrated here that multithreshold networks can display another form of SR that occurs when the signal is predominantly suprathreshold.

In this study a summing network of N threshold devices (Fig. 1) is considered. Each threshold device is subject to the same input signal $x(t)$ but independent Gaussian noise, $\eta_i(t)$, with a common standard deviation, σ_η. The devices are modeled as Heaviside functions, and the individual outputs, $y_i(t)$, are equal to unity if $x(t) + \eta_i(t) > \theta_i$ and zero otherwise. The θ_i are the threshold levels and $i = 1, 2, \ldots, N$. The response of the network is obtained by summing the individual responses of each device. Consequently, the output $y(t)$ represents the number of devices that are triggered at any instant of time.

The network is similar in spirit to those previously studied in the context of neuronal ensembles [5–7]. However, two significant differences exist. First, each threshold level is independently adjustable. This feature is incorporated into the model primarily because it is well known that the transmitted information can be significantly improved if a distribution of thresholds is employed. Second, simple nondynamical threshold elements are incorporated into the network instead of “excitable units” such as the FitzHugh-Nagumo model. This simplification helps to mitigate the increase in complexity caused by having independently adjustable thresholds and makes the model more amenable to theoretical analysis.

An information theoretic measure—the average mutual information—will be used to quantify the amount of information transmitted through the network. The average mutual (or transmitted) information, I, for the network shown in Fig. 1 can be written [9]

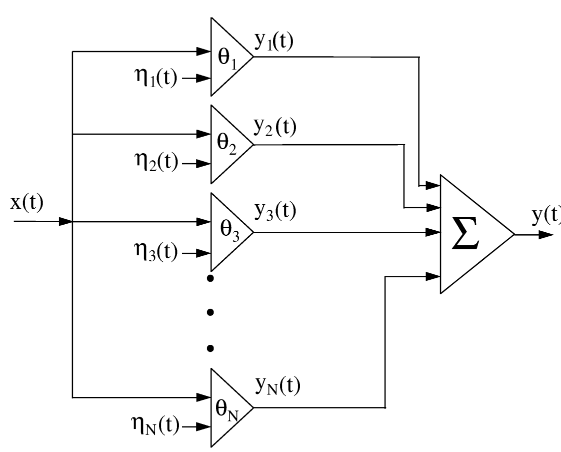

FIG. 1. A summing network of N threshold devices. Each device is subject to the same signal but independent Gaussian noise.

 0031-9007/00/84(11)/2310(4)$15.00

$$I = H(y) - H(y \mid x) = -\sum_{n=0}^{N} P_y(n) \log_2 P_y(n) - \left(-\int_{-\infty}^{\infty} dx \sum_{n=0}^{N} P(n \mid x) P_x(x) \log_2 P(n \mid x)\right). \quad (1)$$

$H(y)$ is the information content (or entropy) of $y(t)$ and $H(y \mid x)$ can be interpreted as the amount of encoded information lost in the transmission of the signal. $P_y(n)$ is the probability that $y(t) = n$ and $P(n \mid x)$ is the conditional probability given knowledge of the signal value, x. The logarithms are taken to base 2 so I is measured in bits.

Initially, the case similar to that studied in [5,7], where all the threshold levels have the same value, will be studied. However, in addition to the subthreshold signal enhancements previously reported [5], we will see that, by an appropriate choice of threshold level, noise can also optimize the detection of suprathreshold signals. Figure 2 shows results for all thresholds set equal to the mean of the signal and various N. In this situation the signal is strongly suprathreshold, yet SR type behavior (i.e., a noise induced maximum) is still observed for all $N > 1$. As N increases, the maximum value attained by I also increases. In complete contrast to classical SR, this effect is at its most pronounced when the deterministic (signal induced) threshold crossings are maximized; for this reason it will be termed suprathreshold stochastic resonance (SSR). SSR effects do not diminish with increasing signal magnitude—as long as the noise is scaled accordingly. Consequently, SSR can be observed with signals of any magnitude without having to modify the threshold levels. It should be stressed that the mechanism giving rise to SSR is quite different from that of classical SR. In the absence of noise, all devices switch in unison, and consequently the network acts like a single bit ADC ($I = 1$). Finite noise results in a distribution of thresholds that gives access to more bits of information; effectively the noise is accessing more degrees of freedom (output states) of the system and hence generating information. In addition to the improvement of the transmitted information the noise also acts to linearize the system response. This is a common effect and occurs in nearly all nonlinear systems [10]. Indeed, the dynamics of a similar network was interpreted in terms of a noise induced linearization (NIL) mechanism [7]. However, while NIL nearly always occurs with increasing noise, SR does not. The advantage of using an information theoretic measure instead of a linear signal processing technique, such as cross correlation, is that it is robust to a deterministic nonlinear deformation of the signal. Consequently, it is possible to quantify the dynamics of these networks, and hence SR, largely independently of the linearity of the response.

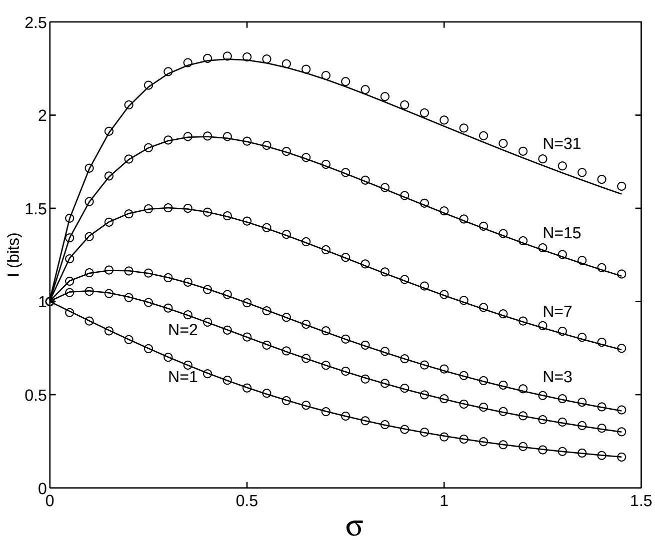

FIG. 2. Transmitted information using a Gaussian signal source with a standard deviation σ_x. $\sigma = \sigma_\eta/\sigma_x$ and all $\theta_i = 0$. The data points are the results of a digital simulation of the network, and the solid lines were obtained by numerically evaluating Eq. (2).

The existence of the SSR effect can be confirmed by direct calculation of I. When all threshold values are equal to θ it is straightforward to show that

$$I = -\sum_{n=0}^{N} P_y(n) \log_2 P'(n) - \left(-N \int_{-\infty}^{\infty} dx\, P_x(x) \times (P_{1|x} \log_2 P_{1|x} + P_{0|x} \log_2 P_{0|x})\right), \quad (2)$$

and $P'(n) = \int_{-\infty}^{\infty} dx\, P_x(x) P_{1|x}^n P_{0|x}^{N-n}$ where $P_y(n) = C_n^N P'(n)$, C_n^N is the binomial coefficient. The signal distribution is taken as Gaussian, $P_x(x) = (1/\sqrt{2\pi\sigma_x^2}) \exp(-x^2/2\sigma_x)$, and $P_{1|x} = 1/2 \mathrm{erfc}[(\theta - x)/\sqrt{2\sigma_\eta^2}]$, is the conditional probability of $y_i = 1$ for a given x and similarly $P_{0|x} = 1 - P_{1|x}$ is the probability of a zero given x. Erfc is the complimentary error function. The integrals and summation in Eq. (2) were calculated numerically (solid lines in Fig. 2). Good agreement between simulation and theory is observed confirming the existence of the SSR effect.

In practice, setting the threshold levels to a common value seems like a rather inefficient use of N devices. In general, it is standard engineering practice to employ a distribution of thresholds to help maximize the transmitted information. The most commonly used distribution is regular uniform quantization. Flash ADCs are designed on this principle, but, of course, it is also the principle behind the operation of most ADCs. For this reason, the situation when the threshold levels are uniformly distributed between the limits ± 1 with separation $\Delta = 2/(N + 1)$ will be considered. In addition, a random uniformly distributed signal between limits $\pm L$ will be used instead of a Gaussian source. This is because, in the absence of noise, a uniform quantization scheme is optimal (i.e., I reaches its theoretical maximum) for a uniformly distributed signal.

Consequently, the network is at its most efficient when a uniform signal is employed, and hence serves as a useful benchmark for comparative studies. This network configuration will be termed the ADC network.

Figure 3 shows the behavior of I against σ for various signal strengths and $N = 64$. Figure 3(a) shows results up to a signal strength $L = 1$ and Fig. 3(b) for signal strengths $L > 1$. In Fig. 3(a), an SR effect is observed when $L = \Delta/2$. At this signal strength the signal falls just between the middle two levels, and hence it is just subthreshold. Therefore, adding noise results in the observation of a classical SR effect. This is the principle behind *dithering* in ADCs. The effect of the dithering is to improve the dynamic range of the system, allowing signals smaller than the quantization level to be detected. The connection between dithering and SR has been pointed out previously [11]. It can be seen that further increases in L improve I—as one would expect—but the noise has only a detrimental effect. The results for $L = 1$ represent the maximum attainable performance for this system. In the absence of noise, the maximum number of bits of information achievable for an N level system is given by $\log_2(N + 1)$, which in this case gives $I = 6.02$. However, in practice such a high information rate is rarely achieved.

Figure 3(b) illustrates what happens if L is increased beyond unity. In this situation the signal is clipping and consequently, at zero noise level, I falls off with increasing L. However, for $L \geq 4$, a noise induced maximum is seen to occur—this coincides with the onset of the SSR effect. Indeed, for all $L > 4$, the maximum I is always attained at finite noise. In the limit $L \to \infty$ the system reduces to that considered in Fig. 2. These results help to set the SSR effect in context. The ability of the network to encode large, suprathreshold, signals is improved by the addition of noise. In other words, the dynamic range of the ADC network is improved for large signal amplitudes. Therefore, SSR can be thought of as the large signal complement of classical SR. Classical SR extends the dynamic range of the system to subthreshold signal levels, whereas SSR improves it at larger signal intensities.

This can be illustrated more generally. The integrals in Eq. (2) can be solved explicitly in the case $\sigma = 1$ from which it can be determined that I scales for large N as $1/2 \log_2(N)$. Given that the maximum is reasonably approximated by the $I(\sigma = 1)$ value, then the maximum I attainable for a given L can be obtained and is shown in Fig. 4. The solid line represents the theoretical performance of the ADC network in the absence of noise. The dashed lines indicate situations, at small and large L, where noise enhances the transmission of information. For L between $\Delta/2$ and 4 the maximum I is attained at

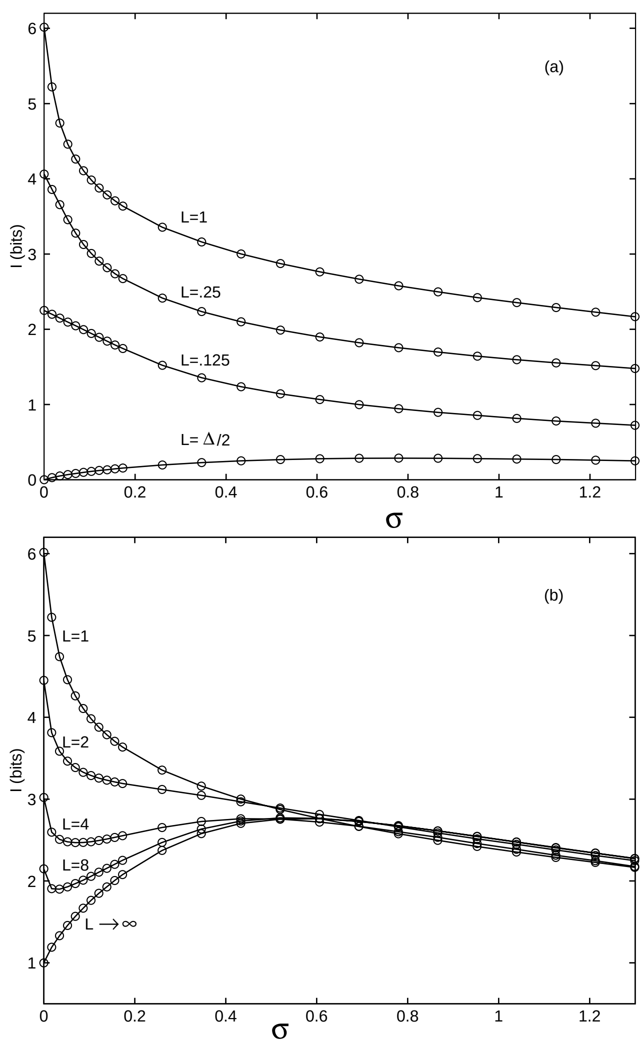

FIG. 3. Plots of transmitted information against $\sigma = \sigma_\eta/\sigma_x$ for $N = 64$ and various signal strengths L: (a) $L \leq 1$, and (b) $L \geq 1$. The data points are from digital simulation and the solid lines are guides to the eye.

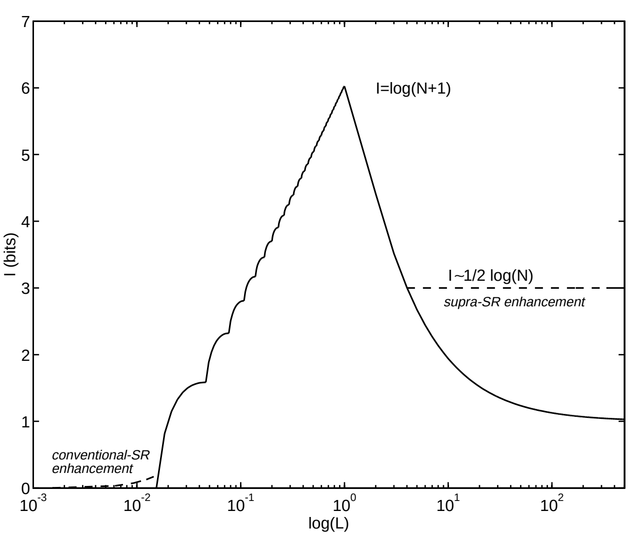

FIG. 4. Plot of transmitted information against $\log_{10}(L)$ for $N = 64$. The solid line is the theoretical I calculated in the absence of noise. The dashed lines indicate the maximum value of I that can be obtained when an optimal amount of noise is added to the system.

zero noise intensity. For $L < \Delta/2$, classical SR can be used to improve I, while for larger values ($L > 4$) SSR can be employed. In the SSR regime a value of σ can always be found, regardless of the size of L, such that $I \approx 1/2 \log_2(N)$ (while in the absence of noise $I \to 1$ as $L \to \infty$). This implies that the dynamic range of the ADC network can be enhanced up to arbitrarily large L.

The results of Fig. 3 show that the transmitted information is generally higher in the absence of noise. This is hardly a surprising result. However, noise forms an irreducible part of any signal detection system—it is this that often imposes a fundamental lower limit on the size of signal that the system can detect. In this regime the results presented yield some interesting conclusions. Consider again Figs. 3(a) and 3(b) when $\sigma \approx 1$. It can be seen from Fig. 3(b) that at this noise level optimal signal transmission occurs provided $L > 1$. Whereas when $L < 1$, as considered in Fig. 3(a), the performance of the network is reduced. These results therefore demonstrate that no benefit is to be gained by employing a distribution of threshold levels when trying to detect signals similar in magnitude to the internal noise. It is simpler to set all threshold levels to coincide with the dc component of the signal (this normally maximizes the deterministic threshold crossings for an even functioned signal distribution) and rely on the SSR effect to enhance I. Indeed, if one adopts the standard engineering practice of setting $L < 1$ the network does not achieve the same level of performance.

Additionally, it should be borne in mind that the performance of the ADC networks never achieve that displayed by the $L = 1$ curve. This is due, in part, to the fact that setting the dynamic range of the ADC with such precision requires *a priori* knowledge of the signal distribution. In practice the distribution is not usually known with great accuracy. For example, the dynamic range of an ADC is normally set to allow for possible large excursions in the signal and, hence, system performance much closer to the $L = 0.25$ curve in Fig. 3(a) is typical. If this is the case there is still an argument for designing the system using the SSR effect if $\sigma > 0.1$. This is the noise intensity at which the performance of the system falls below 3 bits which, in turn, is the number of bits of information that can be obtained using the SSR effect. This implies that, for the $N = 64$ case studied here, SSR is probably useful for signal detection when the signal-to-(internal)noise ratio (SNR) of the system is less than 20 dB.

In most signal detection/encoding systems SNRs are usually well in excess of 20 dB. However, one application where this is not the case is sensory neurons. In a recent set of papers the basic characteristics of sensory neurons were summarized [12]. One of the most remarkable features of sensory neurons seems to be that they have only a SNR of order unity (0 dB). This therefore leads to the following question: Could SSR be employed in neuronal ensembles to enhance the transmitted information? Indeed, evidence does exist to suggest that this may be a possibility. The dc adaptive capabilities of sensory neurons is well established. One well known example is in light and dark adaptation of the human eye [13]. 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. While the network studied in this Letter is obviously a very crude model of a real biological network, neurons are still predominantly threshold based systems. Consequently, the question as to how they should distribute their threshold levels to maximize information flow still applies. The summed response of these networks is largely governed by the statistics of the "firing" of all the devices rather than the exact shape or dynamics of each firing event. Therefore, the results presented for the simple threshold network should have some applicability to real neuronal ensembles. Given that sensory neurons have an SNR of order 0 dB, the above results suggest that adapting all thresholds to the dc signal component should give better information flow than trying to redistribute their levels to accommodate changes in the signal. Additionally, this strategy also simplifies the mechanism required to achieve the adaptation process. Consequently, SSR does seem to offer a possible explanation for dc adaptation in sensory neurons.

[1] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 23 (1998); in *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. **17D**, 653 (1995)].
[2] A. R. Bulsara and A. Zador, Phys. Rev. E. **54**, R2185 (1996).
[3] P. Jung, Phys. Lett. A **207**, 93 (1995).
[4] V. K. Madisetti and D. B. Williams, *The Digital Signal Processing Handbook* (CRC Press, Boca Raton, 1998).
[5] J. J. Collins, C. C. Chow, and T. T. Imhoff, Nature (London) **376**, 236 (1995).
[6] X. Pei, L. Wilkens, and F. Moss, Phys. Rev. Lett. **77**, 4679 (1996).
[7] D. R. Chialvo, A. Longtin, and J. Muller-Gerking, Phys. Rev. E **55**, 1798 (1997).
[8] F. Apostolico *et al.,* Phys. Rev. E **55**, 36 (1997).
[9] A. Feinstein, *Foundations of Information Theory* (McGraw-Hill, New York, 1958).
[10] M. I. Dykman *et al.,* Phys. Lett. A **193**, 61 (1994); N. G. Stocks *et al.,* in *Fluctuations and Order—The New Synthesis,* edited by M. Millonas (Springer, New York, 1996).
[11] L. Gammaitoni, Phys. Rev. E **52**, 4691 (1995).
[12] W. Bialek *et al.,* Physica (Amsterdam) **200A**, 581 (1994); M. DeWeese, Netw. Comput. Neural Syst. **7**, 325 (1996); D. DeWeese and W. Bialek, Nuovo Cimento Soc. Ital. Fis. **17D**, 733 (1995).
[13] G. Somjen, *Neurophysiology—The Essentials* (Williams & Wilkins, Baltimore, 1983).