



Encoding efficiency of suprathreshold stochastic resonance on stimulus-specific information



Fabing Duan ^{a,*}, François Chapeau-Blondeau ^b, Derek Abbott ^c

^a Institute of Complexity Science, Qingdao University, Qingdao 266071, China

^b Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France

^c Centre for Biomedical Engineering (CBME) and School of Electrical & Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Article history:

Received 30 May 2015

Received in revised form 23 September 2015

Accepted 24 September 2015

Available online 1 October 2015

Communicated by C.R. Doering

Keywords:

Stimulus-specific information

Suprathreshold stochastic resonance

Encoding efficiency

Neuron population

ABSTRACT

In this paper, we evaluate the encoding efficiency of suprathreshold stochastic resonance (SSR) based on a local information-theoretic measure of stimulus-specific information (SSI), which is the average specific information of responses associated with a particular stimulus. The theoretical and numerical analyses of SSIs reveal that noise can improve neuronal coding efficiency for a large population of neurons, which leads to produce increased information-rich responses. The SSI measure, in contrast to the global measure of average mutual information, can characterize the noise benefits in finer detail for describing the enhancement of neuronal encoding efficiency of a particular stimulus, which may be of general utility in the design and implementation of a SSR coding scheme.

© 2015 Elsevier B.V. All rights reserved.

It is now a well-known fact that noise can sometimes improve, without degrading, the responses of certain nonlinear systems. This viewpoint is primarily motivated by the phenomenon of stochastic resonance (SR) [1–3], where a suitable amount of noise brings an optimized system response characterized by various measures, such as spectral amplification [4], correlation coefficient [5], signal-to-noise ratio [6,7], and mutual information [8]. In its original form, SR often applies to a noise-enhanced subthreshold signal [1–5]. However, in a parallel summing network, a new form of SR termed suprathreshold stochastic resonance (SSR) results in the maximum input-output mutual information at a non-zero level of noise intensity, even if the input signal is predominantly above the threshold [8]. The information content in each individual subsystem is monotonically decreased by the addition of noise, but the summed outputs from all subsystems yield a net gain in information [8–17].

The case of SSR is now well established as an important paradigm that suggests neuronal noise can possibly have a beneficial role in sensory systems [9–11]. This paradigm is based on the facts that there are large numbers of interconnected neurons in the nervous system of animals and humans with variations in structure, function and size, and noise permeates every level of

the nervous system, from the perception of sensory signals to the generation of motor responses [18,19]. The foregoing nervous system features indicate that the potential exploitation of SSR in a neuronal population stands as an interesting question in neuroscience, relevant for instance to sensory neurons [13,20], cochlear implants [10,11,14], motion detection [15], and stochastic pooling sensor networks [21–23]. These research results show that SSR does appear to serve as an efficient coding strategy of information transformation—providing a possible explanation of the role of noise in human sensory processes [8,10,11,13,16].

Mutual information is often used to calculate the information gain or the reduction of uncertainty of the neuronal responses [24]. However, it cannot address which particular stimuli or responses are significant in information transmission [25,26]. Therefore, based on the specific information [25], Butts proposes a new measure of stimulus-specific information (SSI) defined as the average reduction in the uncertainty of one observation given a particular stimulus [26]. This information bearing measure of SSI can be calculated without prior knowledge about the coding scheme, and is also robust to nonlinearities in the system [26]. Recent studies of SSI give rise to a number of interesting results: The effect of variability on SSI illustrates that the best encoded stimulus with the maximum SSI can change systematically from the high-slope region of tuning curve for low noise to the peak of the tuning curve for high noise [27]. The neuronal encoding of sound frequency in the auditory cortex shows that the maximum SSI is always at the

* Corresponding author.

E-mail addresses: fabing.duan@gmail.com (F. Duan), chapeau@univ-angers.fr (F. Chapeau-Blondeau), derek.abbott@adelaide.edu.au (D. Abbott).

best frequency and never in the tuning curve tails [28]. In finite neural populations, the shape of the marginal SSI can converge toward that of the Fisher information as the population size increases, and predict the best encoded stimulus precisely [29]. In this paper, we are particularly interested in the noise benefits in enhancing the encoding efficiency of the stimulus in a population of neurons. The obtained theoretical and numerical results show that the SSIs not only straightforwardly explicate how much information a neuronal population provides about a particular stimulus, but also illustrate how effectively each stimulus is enhanced by the optimal noise level to transmit more information via the mechanism of SSR. The SSI measure is particularly useful in making an analytical observation of the positive role of noise, and offers an interesting insight into the application of SSR to neuronal coding schemes.

1. Information measures

Consider a neural system with an ensemble of stimuli X and whose behavior can be classified in a set of responses Y , the mutual information between the ensemble of stimuli X and the set of responses Y is given by

$$I(X, Y) = H(X) - H(X|Y), \quad (1)$$

where the information entropy of the stimulus ensemble [24]

$$H(X) = - \sum_{x \in X} P_x(x) \log_2 P_x(x), \quad (2)$$

and the average conditional entropy [24]

$$\begin{aligned} H(X|Y) &= \sum_{y \in Y} P_y(y) H(X|y) \\ &= \sum_{y \in Y} P_y(y) \left[- \sum_{x \in X} P(x|y) \log_2 P(x|y) \right]. \end{aligned} \quad (3)$$

The conditional entropy associated with a particular response y is defined as $H(X|y)$ [24], and lowercase characters x and y represent an individual observation within ensembles.

The mutual information can be used to quantify the information provided by an entire response ensemble about an entire stimulus ensemble, but it is often of interest to know which particular stimuli are effectively encoded by the system, and which particular responses communicate information about the stimuli [25–29]. Then, DeWeese and Meister [25] propose the specific information measure

$$I_{sp}(y) = H(X) - H(X|y), \quad (4)$$

which is an appropriate representation of the degree to which a given response y contributes to the overall mutual information in Eq. (1) [25,26]. However, due to the asymmetry of stimulus and response with respect to causality, the specific information of a particular stimulus $I_{sp}(x) = H(Y) - H(Y|x)$ fails to select the effective encoded stimuli [26,27]. The largest value of $I_{sp}(x)$ corresponds to those stimuli that have few responses associated with them, without regard to whether these responses are informative or not [26]. Therefore, Butts defines a new information theoretic measure of SSI as

$$I_{ssi}(x) = \sum_{y \in Y} P(y|x) I_{sp}(y), \quad (5)$$

which explicitly represents the average specific information of the response ensemble Y that occurs when a particular stimulus x is present [26]. It is also noted that the average SSI over the entire ensemble of stimulus yields the mutual information, as follows

$$I(X, Y) = \sum_{x \in X} P_x(x) I_{ssi}(x). \quad (6)$$

In Eq. (6), it is interesting to note that the term $P_x(x) I_{ssi}(x)$ represents the average informative contribution of each stimulus to the mutual information. Therefore, we can define the encoding efficiency of each stimulus as

$$E_{ssi}(x) = P_x(x) I_{ssi}(x), \quad (7)$$

which can be considered as a useful metric to characterize the positive role of noise in the enhancement of the encoding efficiency of neuronal information transmission.

2. Binary threshold SSR model

Consider a typical SSR model consisting of N binary threshold neurons, and each neuron is subject to the same continuous stimulus signal $x(t)$ but independent noise components $\eta_i(t)$ [8]. The output y_i is given by the neuronal response function

$$y_i = \begin{cases} 1 & \text{if } x + \eta_i \geq \theta, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where θ is the threshold level of the neuronal population. The overall response is $y = \sum_{i=1}^N y_i$ with probability mass function $P_y(n)$ as y being equal to n for $n = 0, 1, \dots, N$ [8,9]. Furthermore, we assume that the noise distribution is f_η and the cumulative distribution function is F_η , then the transition probability

$$P(1|x) = P(x + \eta > \theta|x) = 1 - F_\eta(\theta - x) \quad (9)$$

is the conditional probability of neuron responses being in state 1. Consequently, the conditional probability of neuron responses being in state 0 can be written as $P(0|x) = 1 - P(1|x)$ [8,9]. Noting that, for any given stimulus value x , each device acts independently under the influence of its own noise η_i , thus the probability that n neurons are triggered accords with the binomial distribution

$$P(n|x) = C_n^N P^n(1|x) P^{N-n}(0|x) \quad (10)$$

with the binomial coefficient C_n^N [8]. Therefore, the probability mass function $P_y(n)$ can be calculated as

$$P_y(n) = \int P(n|x) P_x(x) dx = C_n^N B(n), \quad (11)$$

$$B(n) = \int P^n(1|x) P^{N-n}(0|x) P_x(x) dx. \quad (12)$$

Using Bayes' theorem, the conditional probability for a particular stimulus x knowing the response $y = n$ is given by

$$P(x|n) = \frac{P(n|x) P_x(x)}{P_y(n)}. \quad (13)$$

Then, the specific information $I_{sp}(n)$ in Eq. (4) can be written as

$$\begin{aligned} I_{sp}(n) &= H(X) - H(X|n) \\ &= - \int P_x(x) \log P_x(x) dx + \int P(x|n) \log_2 P(x|n) dx, \end{aligned} \quad (14)$$

with the differential entropy of stimulus signal $H(X)$ and the conditional entropy $H(X|n)$. Based on Eqs. (10) and (14), the SSI of a particular stimulus value x can be expressed as

$$I_{ssi}(x) = \sum_{n=0}^N P(n|x) I_{sp}(n). \quad (15)$$

Next, we specifically consider the generalized Gaussian stimulus signal $x(t)$ with its distribution

$$P_x(x) = \frac{c_1}{\sigma_x} \exp\left(-c_2 \left|\frac{x-\mu}{\sigma_x}\right|^{\frac{2}{1+\beta}}\right), \quad (16)$$

with the mean μ , the signal standard deviation σ_x and the exponent $\beta \geq -1$ [30]. Here, $c_1 = \frac{1}{1+\beta} \Gamma^{\frac{1}{2}}(\frac{3(1+\beta)}{2}) / \Gamma^{\frac{3}{2}}(\frac{1+\beta}{2})$, $c_2 = [\Gamma(\frac{3(1+\beta)}{2}) / \Gamma(\frac{1+\beta}{2})]^{\frac{1}{1+\beta}}$ and Gamma function $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. As the exponent β varies, we can conveniently consider different stimulus distributions, including Laplacian ($\beta = 1$), Gaussian ($\beta = 0$) and uniform ($\beta = -1$) densities [30]. For the stimulus density of Eq. (16), the differential entropy

$$H(X) = \log_2 \frac{\sigma_x}{c_1} + \frac{1+\beta}{2} \log_2 e. \quad (17)$$

We assume the Gaussian noise density $f_\eta(\eta) = \exp[-\eta^2/(2\sigma_\eta^2)] / \sqrt{2\pi\sigma_\eta^2}$ with the noise level σ_η , and Eq. (9) becomes

$$P(1|x) = \frac{1}{2} \operatorname{erfc}[-(x-\theta) / \sqrt{2\sigma_\eta^2}], \quad (18)$$

with the complementary error function

$$\operatorname{erfc}(x) = 2 \int_x^\infty \exp(-z^2) dz / \sqrt{\pi}.$$

Making the change of variables of $\tau = (x-\theta) / \sqrt{2\sigma_\eta^2}$ and $\sigma = \sigma_\eta / \sigma_x$ and taking the threshold $\theta = \mu$, Eq. (12) can be rewritten as

$$B(n) = \int \sqrt{2} c_1 \sigma \exp\left(-c_2 \left|\sqrt{2} \tau \sigma\right|^{\frac{2}{1+\beta}}\right) \left(\frac{1}{2} \operatorname{erfc}(-\tau)\right)^n \times \left(1 - \frac{1}{2} \operatorname{erfc}(-\tau)\right)^{N-n} d\tau, \quad (19)$$

and the conditional entropy is given by

$$H(X|n) = - \int P(x|n) \log_2 P(x|n) dx = \log_2 B(n) + \log_2 \frac{\sigma_x}{c_1} - \frac{D(n)}{B(n)}, \quad (20)$$

with

$$D(n) = \int \sqrt{2} c_1 \sigma \exp\left(-c_2 \left|\sqrt{2} \tau \sigma\right|^{\frac{2}{1+\beta}}\right) \left(\frac{1}{2} \operatorname{erfc}(-\tau)\right)^n \times \left(1 - \frac{1}{2} \operatorname{erfc}(-\tau)\right)^{N-n} \log_2 \left[\exp\left(-c_2 \left|\sqrt{2} \tau \sigma\right|^{\frac{2}{1+\beta}}\right) \left(\frac{1}{2} \operatorname{erfc}(-\tau)\right)^n \left(1 - \frac{1}{2} \operatorname{erfc}(-\tau)\right)^{N-n}\right] d\tau. \quad (21)$$

Using Eqs. (17) and (20), the specific information $I_{sp}(n)$ can be computed as

$$I_{sp}(n) = H(X) - H(X|n) = \frac{1+\beta}{2} \log_2 e - \log_2 B(n) + \frac{D(n)}{B(n)}. \quad (22)$$

Substituting Eqs. (18) and (22) into Eq. (15) and noting the standardized random variable of $x_n = (x-\mu) / \sigma_x$, the SSI can be calculated as

$$I_{ssi}(x_n) = \sum_{n=0}^N C_n^N \left(\frac{1}{2} \operatorname{erfc}\left(\frac{-x_n}{\sqrt{2}\sigma}\right)\right)^n \left(1 - \frac{1}{2} \operatorname{erfc}\left(\frac{-x_n}{\sqrt{2}\sigma}\right)\right)^{N-n} I_{sp}(n), \quad (23)$$

which is a single variable function of the ratio of σ for a given value of x_n . We note that the random variable x_n has the standardized density

$$P_{x_n}(x_n) = \sigma_x P_x(\sigma_x x_n + \mu) = c_1 \exp\left(-c_2 |x_n|^{\frac{2}{1+\beta}}\right), \quad (24)$$

with zero mean and unity variance. Then, the mutual information can be computed as

$$I(X, Y) = \int P_{x_n}(x_n) I_{ssi}(x_n) dx_n, \quad (25)$$

which depends only on the parameter σ . In the considered case of $\theta = \mu$, then $1/\sigma^2 = \sigma_x^2 / \sigma_\eta^2$ can be interpreted as the input signal-to-noise ratio [9].

For instance, we choose the stimulus density in Eq. (16) with the exponent $\beta = 1$ and the threshold $\theta = \mu$. An illustrative SSR effect is shown in Fig. 1(A) that, for a single neuron with $N = 1$, the mutual information I monotonically decreases as the parameter σ increases. However, as the population size $N \geq 2$, the noise-enhanced behavior of mutual information appears [8,9]. For a given noise level of $\sigma = 0.34$, the effect of population size N on the SSI I_{ssi} is shown in Fig. 1(B). When $N = 1$, the injection of internal noise $\eta_1(t)$ makes each value of x_n trigger the neuron with response states 1 possible, but with a lower constant value of $I_{ssi}(x_n) = 0.56$ bits. When the population size N increases, the maximum SSI I_{ssi} appears at $x_n = 0$ (i.e. $x = \theta$), and the collective actions of $\eta_i(t)$ are concentrated in the regions on both sides of $x_n = 0$ with more informative contents of I_{ssi} , as indicated by the curves in Fig. 1(B) for $N = 5$ and $N = 31$. Furthermore, from Fig. 1(A) and for $N = 31$, we select four values of $\sigma = 0, 0.1, 0.34$ and 1.0 with the corresponding mutual information $I(X, Y) = 1$ bits (up triangle), 1.94 bits (circle), 2.33 bits (square) and 1.85 bits (down triangle), respectively. Without noise ($\sigma = 0$), signal transmissions in all neurons can only be achieved through deterministic threshold crossing and leads to a transfer of 1-bit of information. Correspondingly, only $I_{ssi}(0) = 1$ bits exists (see the black point in Fig. 1(C)), and the mutual information $I(X, Y) = 1$ bits at $x_n = 0$ ($x = \theta$). At the other three values of σ , the obtained SSIs are also shown in Fig. 1(C). Although the $I_{ssi}(0) = 5.34$ bits at $\sigma = 0.1$, the area under the corresponding I_{ssi} terms are multiplied by $P_{x_n}(x_n)$ to obtain the mutual information $I(X, Y) = 1.94$ bits in Eq. (25). While, at the optimal $\sigma = 0.34$, the largest $I_{ssi}(0)$ is reduced to 3.63 bits, but many values of $I_{ssi}(x_n)$ in the neighborhood of $x_n = 0$ are enlarged. Based on Eq. (25), the corresponding I_{ssi} terms are multiplied by the weighted probabilities $P_{x_n}(x_n)$ to yield the largest mutual information $I(X, Y) = 2.33$ bits. Of course, as σ increases to 1.0, too much noise leads to a flatter distribution of I_{ssi} in an extended region around $x_n = 0$, as shown in Fig. 1(C). Therefore, the corresponding mutual information $I(X, Y)$ degrades to 1.85 bits.

More interestingly, from Eq. (25), the encoding efficiency in Eq. (7) can be reexpressed as

$$E_{ssi}(x_n) = P_{x_n}(x_n) I_{ssi}(x_n), \quad (26)$$

which is actually the encoding efficiency density with the units of bits per stimulus amplitude (bits/volt) for a continuous stimulus. For instance, E_{ssi} is shown in Fig. 1(D) as a function of σ for different stimulus amplitude $|x_n|$ and $N = 31$. At $\sigma = 0$, only $I_{ssi}(0)$ bits exists and $E_{ssi}(0)$ in Eq. (26) tends to infinity, then the addition of more noise leads to a monotonic decrease of the encoding efficiency. However, for $|x_n| \neq 0$, it is illustrated in Fig. 1(D) that there is an optimal noise level of σ that corresponds to the best encoding efficiency, which is a new form of SR based on the measure of encoding efficiency E_{ssi} . It is seen in Fig. 1(D) that the positive role of noise in the enhancement of encoding efficiency over almost the entire region of stimulus (except for $x_n = 0$) can be observed explicitly.

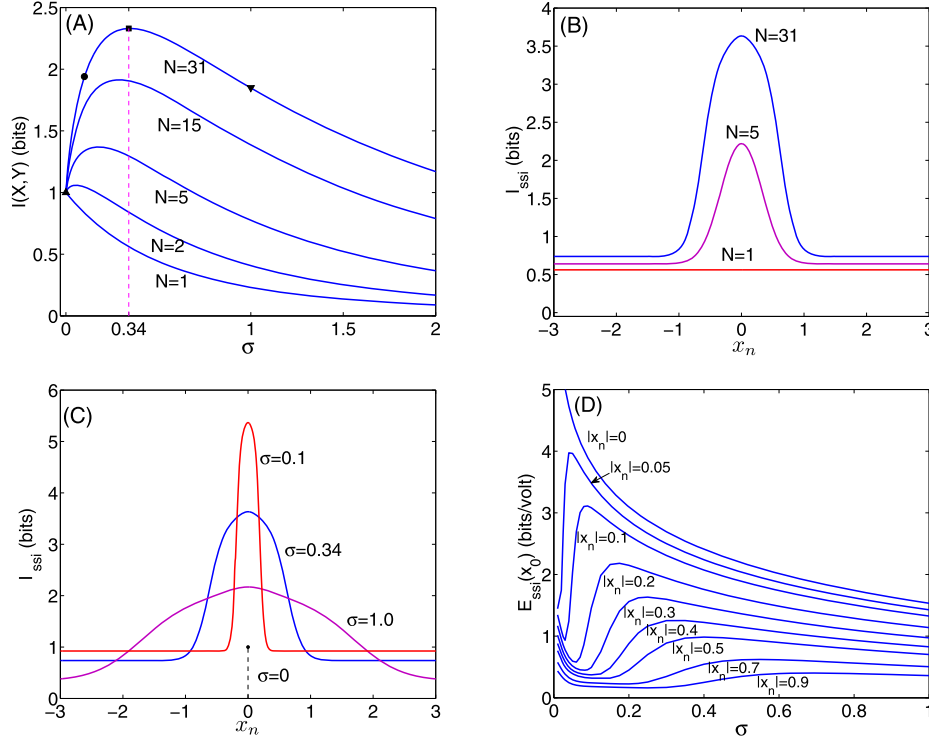


Fig. 1. (A) Mutual information $I(X, Y)$ in Eq. (25) versus $\sigma = \sigma_n/\sigma_x$ for $N = 1, 2, 5, 15$ and 31 ; (B) I_{ssi} in Eq. (23) for $N = 1, 5, 31$ at $\sigma = 0.34$; (C) I_{ssi} in Eq. (23) at $\sigma = 0, 0.1, 0.34$ and 1.0 for $N = 31$; (D) For $N = 31$, $E_{\text{ssi}}(x_n)$ in Eq. (26) versus σ at $|x_n| = 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.7$ and 0.9 . Here, the stimulus density in Eq. (16) is with the exponent $\beta = 1$ and the threshold $\theta = \mu$.

3. Leaky integrate-and-fire neuron population

We further consider a population of leaky integrate-and-fire neurons subjected to a common stimulus and mutually independent internal noise. Despite its simplicity, the integrate-and-fire model is still an extremely useful description of neuronal activity [2,11,14,16,17,31,32]. The membrane voltage $V_i(t)$ of each neuron obeys

$$\tau_m \frac{dV_i(t)}{dt} = -[V_i(t) - V_r] + x(t) + \eta_i(t), \quad (27)$$

until it exceeds the threshold θ , and τ_m is the membrane time constant. Then, at this moment, a spike is generated, and $V_i(t)$ is reset to a post-discharge hyperpolarization V_r with an absolute refractory period τ_{abs} . After the time duration τ_{abs} , the membrane voltage then steps into a relative refractory period τ_{rel} , which can be modeled by multiplying the noisy input by the factor of $1 - \exp(-t/\tau_{\text{rel}})$ [14]. The spike trains from all neurons are then summed as the collective action potential $y_s(t)$, and the final response $y(t)$ can be obtained by the convolution

$$y(t) = \int y_s(\tau) f(t - \tau) d\tau, \quad (28)$$

where the kernel function $f(t)$ is given by

$$f(t) = \begin{cases} 1 - \exp(-t/\tau_m), & t < T_p, \\ (1 - \exp(-T_p/\tau_m)) \exp(-(t - T_p)/\tau_m), & t \geq T_p, \end{cases} \quad (29)$$

with the time width T_p [14]. This convolution operation is equivalent to a low-pass filter that provides an estimate of the input stimulus, and the response $y(t)$ can be viewed as the membrane voltage of the neuronal population [14].

This considered neuron model in Eq. (27) with the membrane voltage $y(t)$ of a neuronal population is a continuous channel, then I_{ssi} in Eq. (5) can be rewritten as

$$I_{\text{ssi}}(x) = \int P(y|x) \times \int [P(x|y) \log_2 P(x|y) - P_x(x) \log_2 P_x(x)] dx dy, \quad (30)$$

and the mutual information can be computed as

$$I(X, Y) = \int I_{\text{ssi}}(x) P_x(x) dx. \quad (31)$$

Here, we extend the information measure of I_{ssi} to the continuous random variables, and I_{ssi} remains finite if the distribution of random variable exists and the integral is Riemann integrable.

In practical simulation, we divide the ranges of variation of $x(t)$ and $y(t)$ with intervals Δx and Δy , respectively. For the i th interval of $x(t)$ and the j th interval of $y(t)$, the statistical number of pairs (x_i, y_j) is k_{ij} , and then the joint probability density is approximated as $P(x_i, y_j) \approx k_{ij}/(K \Delta x \Delta y)$ with $K = \sum_i \sum_j k_{ij}$. Similarly, the probability density of x_i is $P_x(x_i) \approx \sum_j k_{ij}/(K \Delta x)$, the conditional probabilities $P(y_j|x_i) \approx k_{ij}/(\sum_j k_{ij} \Delta y)$ and $P(x_i|y_j) \approx k_{ij}/(\sum_i k_{ij} \Delta x)$ [33,34]. In this way, Eqs. (30) and (31) can be numerically calculated as

$$I_{\text{ssi}}(x_i) \approx \sum_j \frac{k_{ij}}{\sum_j k_{ij}} \left[\sum_i \left(\frac{k_{ij}}{\sum_i k_{ij}} \log_2 \frac{k_{ij}}{\sum_i k_{ij}} \right) - \sum_i \left(\frac{\sum_j k_{ij}}{K} \log_2 \frac{\sum_j k_{ij}}{K} \right) \right], \quad (32)$$

$$I(X, Y) \approx \sum_i I_{\text{ssi}}(x_i) P_x(x_i) \Delta x = \sum_i I_{\text{ssi}}(x_i) \frac{\sum_j k_{ij}}{K}. \quad (33)$$

This estimation method of the informative quantities is based on two-dimension histogram of the observations of X and Y . In Refs. [33,34], it is proved that, as the total number K of division intervals is large enough, the bias and the variance of $I(X, Y)$ are

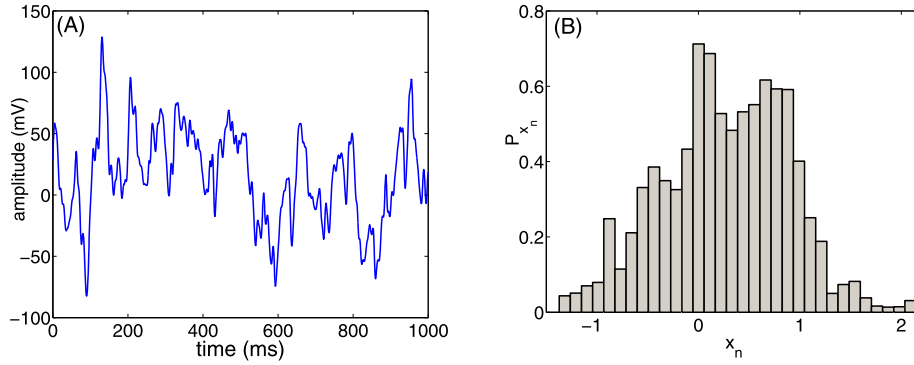


Fig. 2. (A) Gaussian distributed stimulus $x(t)$ with a correlation time of 20 ms, the mean $\mu = 15$ mV and the duration of 1000 ms; (B) Empirical probability distribution P_{x_n} of the scaled stimulus $x_n(t) = x(t)/\sigma_x$.

with the orders of K^{-1} and K^{-2} , respectively. In the following numerical experiments and for the observation sampling length L , the total number K of division intervals is selected as the adjacent integer to $\sqrt[3]{L} + 10$, which is large enough for the estimation accuracy of the mutual information [33,34]. We note that, for a finite number of observations and a very large number of K , k_{ij} might be zero. In this case, $\log_2 k_{ij}$ in Eq. (32) is assumed to be zero for proceeding numerical calculation, otherwise I_{SSI} will be infinite. Moreover, the direct calculations of the SSI will produce a biased estimate, due to the limited number of trials and under-sampling of the probability distributions [28,31,32]. We can simply compute the SSI as above, but with all stimulus-response combinations randomly shuffled. In this case, the true SSI is zero theoretically, but the calculation will produce the biased value $I_{\text{SSI}}^{\text{shuf}}$. In this way, the calculation of $I_{\text{SSI}}^{\text{shuf}}$ will be carried out M times, and the average value is $\langle I_{\text{SSI}}^{\text{shuf}} \rangle = I_{\text{SSI}}^{\text{shuf}}/M$ [28,31,32]. Subtracting $\langle I_{\text{SSI}}^{\text{shuf}} \rangle$ from the raw SSI calculated above, the unbiased SSI is obtained as [28,31,32]

$$I_{\text{SSI}}^{\text{un}} = I_{\text{SSI}} - \langle I_{\text{SSI}}^{\text{shuf}} \rangle. \quad (34)$$

The trial results presented in the following were obtained using stimuli that are Gaussian distributed and exponentially correlated with a correlation time of 20 ms, the mean $\mu = 15$ mV and the duration of 1000 ms [5]. In experiments, the sampling time is $\Delta t = 0.1$ ms, the membrane time $\tau_m = 3$ ms, the resetting potential $V_r = -65$ mV, the firing threshold $\theta = -50$ mV, the absolute refractory period $\tau_{\text{abs}} = 1.5$ ms and the relative refractory period $\tau_{\text{rel}} = 2$ ms. In order to sum the population spikes, the convolution function $f(t)$ in Eq. (29) is with the time width $T_p = 2$ ms. In absence of noise, the standard deviation $\sigma_x = 60$ mV of $V_i(t)$ is assumed to be the signal strength of the output of one neuron. Furthermore, the internal Gaussian noise components η_i in each neuron are with the common standard deviation σ_η . A representative stimulus is shown in Fig. 2(A), and the empirical probability distribution P_{x_n} of the scaled stimulus $x_n(t) = x(t)/\sigma_x$ is plotted in Fig. 2(B).

Upon increasing the scaled ratio $\sigma = \sigma_\eta/\sigma_x$, the evolutions of the mutual information $I(X, Y)$ are shown in Fig. 3(A) for different population sizes N . It is seen that the typical SSR effect appears, and the enhancement of $I(X, Y)$ is manifested by an optimal nonzero level of internal noise, not in a single neuron, but in more than two neurons. The SSIs $I_{\text{SSI}}^{\text{un}}$ are shown in Fig. 3(B) for $N = 1, 4$ and 256 at $\sigma = 5.8$. It is obviously shown in Fig. 3(B) that, for the population size $N = 256$, almost all regions of the stimulus are well encoded, and the SSIs marked by green circles are higher than that corresponding to population sizes $N = 1$ (blue triangles) and 4 (magenta squares). For a given population size $N = 256$, we illustrate the SSIs $I_{\text{SSI}}^{\text{un}}$ for three levels of $\sigma = 0, 5.8$ and 30 in Fig. 3(C). It is visible that, in comparison to $I_{\text{SSI}}^{\text{un}}$ (blue

stars) of no noise ($\sigma = 0$), the stimulus, especially the parts below the mean value $\mu = 15$ mV ($x_n \leq \mu/\sigma_s = 0.25$), is effectively encoded with larger $I_{\text{SSI}}^{\text{un}}$ (green circles) with the help of the optimal noise intensity of $\sigma = 5.8$. While, for a very large noise level $\sigma = 30$, $I_{\text{SSI}}^{\text{un}}$ (red pluses) decrease over entire regions, as shown in Fig. 3(C). Finally, for different values of the scaled stimulus x_n , Fig. 3(D) shows the non-monotonic behavior of the encoding efficiency $E_{\text{SSI}}(x_n) = P_{x_n}(x_n)I_{\text{SSI}}^{\text{un}}(x_n)$ versus the noise level σ . It is interesting to note that the noise-enhanced effects also appear, and the positive role of noise is clearly demonstrated with improved encoding efficiency of the considered region of x_n in the integrate-and-fire neuron population.

Next, we note that neurons have adaptive capabilities, and for instance, the response threshold might be tuned for maximizing the mutual information by a neuron itself. In Fig. 4(A), the mutual information $I(X, Y)$ versus the threshold θ is plotted for the absence of noise ($\sigma = 0$) and a single neuron ($N = 1$). It is seen in Fig. 4(A) that, as θ decreases, the mutual information $I(X, Y)$ becomes larger. While, according to electrophysiological results, a neuron typically fires an action potential when its membrane potential reaches a threshold value that is larger than the resetting potential [31]. Since the resetting potential takes $V_r = -65$ mV, we here just assume that the threshold θ can be reduced to -60 mV. Then, for $\sigma = 0$ and $\theta = -60$ mV, $I_{\text{SSI}}^{\text{un}}$ of a single neuron are illustrated in Fig. 4(B) by red squares. For comparison, at $\sigma = 5.8$, the corresponding $I_{\text{SSI}}^{\text{un}}$ are also plotted by green circles for the population size $N = 256$ neurons and the given threshold $\theta = -50$ mV. It is shown in Fig. 4(B) that, with a non-zero noise level σ and a larger population size N , $I_{\text{SSI}}^{\text{un}}$ are effectively enhanced for the region of $x_n < 0.5$. Meanwhile, we also observe that, at the largest value of $x_n = 2.15$, the corresponding $I_{\text{SSI}}^{\text{un}}$ of a single neuron is remarkably large. Considering the distribution P_{x_n} in Fig. 2(B), the average mutual information $I(X, Y) = 1.80$ bits that is obtained for an optimal non-zero noise level and a larger population size is superior to $I(X, Y) = 1.39$ bits of a non-noisy neuron by adjusting its threshold. This result shows that a number of neurons with the same fixed threshold can benefit from the mechanism of SSR, and has a better encoding efficiency than a single non-noise neuron with a tunable threshold. However, a population of neurons might optimize the threshold θ_n of the n th neuron as the noise level σ varies, and the average mutual information will be maximized by the optimally distributed thresholds. Moreover, realistic models of neurodynamics must ultimately encompass multiple interacting modules, and neural information processing is the integration of excitatory and inhibitory synaptic contributions [31]. Therefore, the resetting potential V_r is also a tunable parameter. All of these forms of plasticity, and many others, are important elements of the adaptability of neural systems. It is interesting to further investigate the

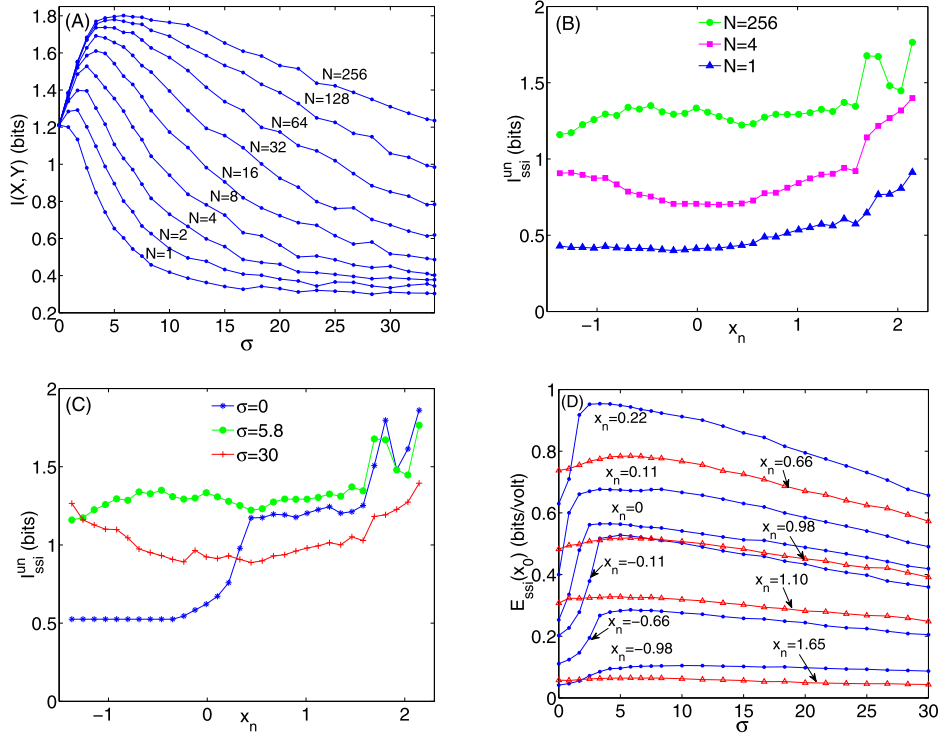


Fig. 3. (A) Mutual information $I(X, Y)$ versus the noise level σ for population sizes $N = 1, 2, 4, 8, 16, 32, 64, 128$ and 256 ; (B) I_{SSI}^{un} in Eq. (34) for $N = 1, 4$ and 256 at $\sigma = 5.8$; (C) I_{SSI}^{un} at three noise levels of $\sigma = 0, 5.8$ and 30 for $N = 256$; (D) $E_{SSI}(x_n)$ versus σ for different values of $x_n = -0.98, -0.66, -0.11, 0, 0.11, 0.22, 0.66, 0.98, 1.10$ and 1.65 . The leaky integrate-and-fire neuron is with parameters $\tau_m = 3$ ms, $V_r = -65$ mV, $\theta = -50$ mV, $T_p = 2$ ms, $\tau_{abs} = 1.5$ ms and $\tau_{rel} = 2$ ms. The total number of trials is 100 and the sampling time $\Delta t = 0.1$ ms. The calculation of I_{SSI}^{shuf} is carried out $M = 20$ times.

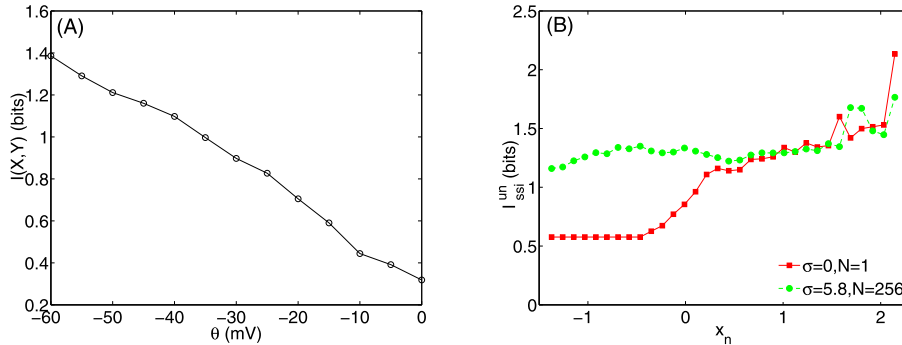


Fig. 4. (A) In the absence of noise ($\sigma = 0$), the mutual information $I(X, Y)$ versus the threshold θ for a single neuron ($N = 1$); (B) I_{SSI}^{un} versus x_n (red squares) for $\sigma = 0$, $N = 1$ and $\theta = -60$ mV. For comparison, for $\sigma = 5.8$, $N = 256$ and the given $\theta = -50$ mV, I_{SSI}^{un} (green circles) versus x_n are also illustrated.

encoding efficiency of a number of neurons with tunable neural parameters, and clarify whether the neuronal noise can still play a positive beneficial role for adaptive capability of neurons or not.

4. Conclusion

In this paper, we utilize a local information-theoretic measure, namely SSI, as a stimulus-specific decomposition of the mutual information, to evaluate the noise-enhanced encoding efficiency of a population of neurons within the context of SSR. For the analytical SSR model of binary threshold neurons, the theoretical expression of SSI is deduced, and a clear picture of noise benefits in neural encoding efficiency is illustrated. Furthermore, the SSR effect in the integrate-and-fire neuron population is investigated using the SSI measure, and the numerical calculation method for differential SSI is given. The obtained results of SSIs and encoding efficiency fully show the positive role of noise components for coding within

a neuronal population. A maximal number of input stimuli are efficiently encoded at the output due to the assistance of an optimal nonzero amount of added noise. We argue that this local SSI information measure may be generally applicable in analyzing neural encoding strategies via the mechanism of SSR. However, we assume a bundle of homogeneous neurons with a fixed response threshold, and then our results may be only applicable to these small populations with the identical characteristics. An open question is to further demonstrate the positive role of neuronal noise in improving the encoding efficiency of real neuronal populations, in view of the fact that neurons have the adaptive capability to optimize their threshold or employ the inhibitory connections.

Acknowledgements

This work is sponsored by the National Natural Science Foundation of China (No. 61573202), and the Science and Technology Development Program of Shandong Province (No. 2014GGX101031).

References

- [1] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of stochastic resonance, *J. Phys. A, Math. Gen.* 14 (1981) L453–L457.
- [2] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance, *Rev. Mod. Phys.* 70 (1998) 233–287.
- [3] K. Wiesenfeld, F. Moss, Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs, *Nature* 373 (1995) 33–36.
- [4] P. Jung, P. Hänggi, Amplification of small signal via stochastic resonance, *Phys. Rev. A* 44 (1991) 8032–8042.
- [5] J.J. Collins, C.C. Chow, T.T. Imhoff, Stochastic resonance without tuning, *Nature* 376 (1995) 236–238.
- [6] F. Chapeau-Blondeau, X. Godivier, Theory of stochastic resonance in signal transmission by static nonlinear systems, *Phys. Rev. E* 55 (1997) 1478–1495.
- [7] F. Duan, F. Chapeau-Blondeau, D. Abbott, Exploring weak-periodic-signal stochastic resonance in locally optimal processors with a Fisher information metric, *Signal Process.* 92 (2012) 3049–3055.
- [8] N.G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, *Phys. Rev. Lett.* 84 (2000) 2310–2313.
- [9] N.G. Stocks, Information transmission in parallel threshold arrays: suprathreshold stochastic resonance, *Phys. Rev. E* 63 (2001) 041114.
- [10] M.D. McDonnell, N.G. Stocks, C.E.M. Pearce, D. Abbott, *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization*, Cambridge University Press, 2008.
- [11] X. Gao, D.B. Grayden, M.D. McDonnell, Stochastic information transfer from cochlear implant electrodes to auditory nerve fibers, *Phys. Rev. E* 90 (2014) 022722.
- [12] D. Rousseau, F. Duan, F. Chapeau-Blondeau, Suprathreshold stochastic resonance and noise-enhanced Fisher information in arrays of threshold devices, *Phys. Rev. E* 68 (2003) 031107.
- [13] N.G. Stocks, R. Mannella, Generic noise-enhanced coding in neuronal arrays, *Phys. Rev. E* 64 (2001) 030902(R).
- [14] N.G. Stocks, D. Allingham, R.P. Morse, The application of suprathreshold stochastic resonance to cochlear implant coding, *Fluct. Noise Lett.* 2 (2002) L169–L181.
- [15] G.P. Harmer, B.R. Davis, D. Abbott, A review of stochastic resonance: circuits and measurement, *IEEE Trans. Instrum. Meas.* 51 (2002) 299–309.
- [16] R.P. Morse, N.G. Stocks, Enhanced cochlear implant coding using multiplicative noise, in: N.G. Stocks, D. Abbott, R.P. Morse (Eds.), *Fluctuations and Noise in Biological, Biophysical and Biomedical Systems III*, in: Proc. SPIE, vol. 5841, 2005, pp. 23–30.
- [17] T. Hoch, G. Wenning, K. Obermayer, Optimal noise-aided signal transmission through populations of neurons, *Phys. Rev. E* 68 (2003) 011911.
- [18] R.Q. Quiroga, S. Panzeri, Extracting information from neuronal populations: information theory and decoding approaches, *Neuroscience* 10 (2009) 173–185.
- [19] A.A. Faisal, L.P.J. Selen, D.M. Wolpert, Noise in the nervous system, *Neuroscience* 9 (2008) 292–303.
- [20] S. Durrant, Y. Kang, N. Stocks, J. Feng, Suprathreshold stochastic resonance in neural processing tuned by correlation, *Phys. Rev. E* 84 (2011) 011923.
- [21] S. Zozor, P.O. Amblard, C. Duchêne, On pooling networks and fluctuation in suboptimal detection framework, *Fluct. Noise Lett.* 7 (2007) L39–L60.
- [22] M.D. McDonnell, P.O. Amblard, N.G. Stocks, Stochastic pooling networks, *J. Stat. Mech. Theory Exp.* 1 (2009) P01012.
- [23] M.D. McDonnell, F. Li, P.O. Amblard, A.J. Grant, Optimal sensor selection for noisy binary detection in stochastic pooling networks, *Phys. Rev. E* 88 (2013) 022118.
- [24] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
- [25] M.R. DeWeese, M. Meister, How to measure the information gained from one symbol, *Netw. Comput. Neural Syst.* 10 (1999) 325–340.
- [26] D.A. Butts, How much information is associated with a particular stimulus?, *Netw. Comput. Neural Syst.* 14 (2003) 177–187.
- [27] D.A. Butts, M.S. Goldman, Tuning curves, neuronal variability, and sensory coding, *PLoS Biol.* 4 (2006) 0040092.
- [28] N. Montgomery, M. Wehr, Auditory cortical neurons convey maximal stimulus-specific information at their best frequency, *J. Neurosci.* 30 (2010) 13362–13366.
- [29] S. Yarrow, E. Challis, P. Seriès, Fisher and Shannon information in finite neural populations, *Neural Comput.* 24 (2012) 1740–1780.
- [30] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice Hall, Englewood Cliffs, NJ, 1998.
- [31] P. Dayan, L.F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, MIT Press, Cambridge, MA, 2001.
- [32] P. Wallisch, M. Lusignan, M. Benayoun, T.I. Baker, A.S. Dickey, N.G. Hatsopoulos, *MATLAB for Neuroscientists: An Introduction to Scientific Computing in MATLAB*, Elsevier Inc., 2009.
- [33] R. Moddemeijer, On estimation of entropy and mutual information of continuous distributions, *Signal Process.* 16 (1989) 233–248.
- [34] R. Moddemeijer, A statistic to estimate the variance of the histogram-based mutual information estimator based on dependent pairs of observations, *Signal Process.* 75 (1999) 51–63.