

Research

Open Access

The mechanism for stochastic resonance enhancement of mammalian auditory information processing

Dawei Hong¹, Joseph V Martin*² and William M Saidel²

Address: ¹Department of Computer Science, Rutgers University, Camden, New Jersey, USA and ²Department of Biology, Rutgers University, Camden, New Jersey, USA

Email: Dawei Hong - dhong@camden.rutgers.edu; Joseph V Martin* - jomartin@camden.rutgers.edu; William M Saidel - saidel@camden.rutgers.edu

* Corresponding author

Published: 01 December 2006

Received: 19 May 2006

Theoretical Biology and Medical Modelling 2006, **3**:39 doi:10.1186/1742-4682-3-39

Accepted: 01 December 2006

This article is available from: <http://www.tbiomed.com/content/3/1/39>

© 2006 Hong et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: In a mammalian auditory system, when intrinsic noise is added to a subthreshold signal, not only can the resulting noisy signal be detected, but also the information carried by the signal can be completely recovered. Such a phenomenon is called stochastic resonance (SR). Current analysis of SR commonly employs the energies of the subthreshold signal and intrinsic noise. However, it is difficult to explain SR when the energy addition of the signal and noise is not enough to lift the subthreshold signal over the threshold. Therefore, information modulation has been hypothesized to play a role in some forms of SR in sensory systems. Information modulation, however, seems an unlikely mechanism for mammalian audition, since it requires significant *a priori* knowledge of the characteristics of the signal.

Results: We propose that the analysis of SR cannot rely solely on the energies of a subthreshold signal and intrinsic noise or on information modulation. We note that a mammalian auditory system expends energy in the processing of a noisy signal. A part of the expended energy may therefore deposit into the recovered signal, lifting it over threshold. We propose a model that in a rigorous mathematical manner expresses this new theoretical viewpoint on SR in the mammalian auditory system and provide a physiological rationale for the model.

Conclusion: Our result indicates that the mammalian auditory system may be more active than previously described in the literature. As previously recognized, when intrinsic noise is used to generate a noisy signal, the energy carried by the noise is added to the original subthreshold signal. Furthermore, our model predicts that the system itself should deposit additional energy into the recovered signal. The additional energy is used in the processing of the noisy signal to recover the original subthreshold signal.

Background

Stochastic resonance (SR) is a phenomenon resulting from the interactions between stochastic processes and many physical systems [1-4]. In the early 1990s, Moss and colleagues [5] pointed out the importance of SR phenom-

ena in biological sensory systems. Subsequently, Moss developed a more general theory (see reviews in [6,7]). We will use the term "SR" for stochastic resonance in biological sensory systems [6]. As a stochastic phenomenon, SR consists of three ingredients: a threshold, a subthresh-

old signal (the original signal), and intrinsic noise. The original signal is insufficient to reach threshold and stimulate the appropriate sensory system unless it interacts with some intrinsic noise. Such an interaction generates a "noisy signal". When the derived noisy signal exceeds threshold in a sensory system, a sequence of action potentials (the spike train) is produced by the first stages of the system. Subsequent neural processes use these spikes to recover the information contained within the original signal. For a biological sensory system, SR enhances sensory information processing, particularly near the system's threshold.

As summarized in a recently published review [7], a core idea of Moss' theory on SR is that "(t)he role of noise is to sample the stimulus. This means that the larger amplitude excursions of the noise cross the threshold and provide a sample of the subthreshold signal's amplitude at a given instant in time. For good information transmission, the sampling rate should be greater than the stimulus frequency." (p. 269). As noise takes samples (in amplitude) from a subthreshold signal at a series of instants of time, a noisy signal is created. This process can be formulated as follows. An input to the mammalian auditory system, which we will call the original signal, is commonly modeled by a mathematical curve, a function $h(t): t \in [0,1] \mapsto \mathbb{R}$. Here, h is supposed, at least, to be continuous; t represents time; the time period is normalized as $[0,1]$; and $h(t)$ stands for the amplitude of the signal at time instant t . The information carried by h is encoded in both amplitude and frequency. Noise is commonly modeled by a random variable, which in mathematical terms is a measurable function $e(t): t \in [0,1] \mapsto \mathbb{R}$ where $e(t)$ is the amplitude of the noise at time instant t . The noise in a mammalian auditory system is intrinsic. That is, the physiological evidence suggests that noise is generated by the system internally. For the mammalian auditory system, we can set a baseline such that the intensity is zero. Since amplitude is measured by intensity against the baseline, we can let $h(t) > 0$ and $e(t) > 0$ for all $t \in [0,1]$. The resulting noisy signal is represented by

$$f(t) = h(t) + e(t) \quad (1)$$

which usually is quite irregular. As previously adopted in the literature on SR in sensory systems [8], (1) indicates that noise is additive with the original signal.

Thus, in the original formulation of Moss' theory [8], the energies of signal and noise were not considered. The theory simply required a mechanism by which addition of the raw data of an original signal $h(t)$ and noise $e(t)$ would eventually enhance mammalian auditory information processing.

Energy Addition and Information Modulation

If the core idea of Moss' theory [8] is valid for mammalian auditory information processing (as we strongly believe), one has to accept that a mammalian auditory system is capable of recovering an original signal $h(t)$ from the noisy signal $f(t)$ expressed in (1). Guided by Occam's razor, we expect the mechanism of (1) to be generally applicable in the mammalian auditory system. As a first step, it is natural to analyze the energies carried by the signal $h(t)$ and noise $e(t)$. Indeed, in many cases, e.g. [9], the energy addition of the signal and noise is sufficient to explain SR. Moss and his coworkers categorized such SR as Type E (for energy). However, SR has also been observed when the energy addition of the signal and noise is not sufficient to explain the enhancement in sensory perception [10]. Moss *et al.* used the concept of information modulation to explain this observation and categorized such SR as Type I (for information). In other words, the occurrence of Type I SR relies on characteristics of the signal other than energy. Still, the distinction between Types E and I SR has the disadvantage of requiring evolution of multiple mechanisms for SR in the mammalian auditory system, which would seem less likely than evolution of a single unitary mechanism.

At this point, it is instructional to consider the historical progression of research on signal processing in the latter half of the twentieth century. (We refer the reader to section 1 of [11] for a summary of this history.) Filtering out noise from a noisy signal $f(t)$ as expressed in (1) is a major concern of the community of signal processing, where this task is termed "de-noising". Early researchers developed a substantial number of algorithms for de-noising. However, most of the de-noising algorithms were mathematically proven to be optimal when the characteristics of original signal $h(t)$ could be known to the algorithm in advance. De-noising was thought to require information modulation. In 1994, Donoho and Johnstone [12] dramatically changed the modern understanding of de-noising by proposing wavelet shrinkage. Importantly, wavelet-based algorithms do not require *a priori* knowledge of the characteristics of the signal (see below) and can be implemented more efficiently than earlier methods such as the fast Fourier transform (FFT).

Wavelet Shrinkage

Since our proposed model employs a recent improvement on analysis of wavelet shrinkage, we will mention some details related to this algorithm. Recall that an original signal is modeled by a function $h(t): t \in [0,1] \mapsto \mathbb{R}^+$. In the mammalian auditory system, $h(t)$ necessarily has a certain degree of "smoothness". In the literature on signal processing, this is formulated as a requirement that $h(t)$ belongs to a Hölder class. Recall that a Hölder class $\Lambda^\alpha(M)$

is a family of functions, which is determined by two parameters α and M as follows: Let $\mathbb{R}^{[0,1]}$ denote the set of all functions defined on $[0,1]$. For $0 < \alpha \leq 1$, $\Lambda^\alpha(M) \underline{\text{def}} \{h \in \mathbb{R}^{[0,1]}: (\forall x_1, x_2 \in [0,1]), |h(x_1) - h(x_2)| \leq M|x_1 - x_2|^\alpha\}$. For $1 < \alpha$, $\Lambda^\alpha(M) \underline{\text{def}} \{h \in \mathbb{R}^{[0,1]}: (\forall x \in [0,1])|h'(x)| \leq M, h^{,\alpha}$ exists, and $(\forall x_1, x_2 \in [0,1])|h^{,\alpha}(x_1) - h^{,\alpha}(x_2)| \leq M|x_1 - x_2|^{\alpha - 1}\}$. It is straightforward to see that the concept of Hölder class contains information modulation. For example, a sine wave belongs to a Hölder class with $1 < \alpha$; however, the higher the frequency of the wave is, the larger the M must be. Before the advent of wavelet shrinkage, proposed de-noising algorithms required α and M as part of their inputs. Unlike the earlier algorithms, wavelet shrinkage only requires that $h(t)$ belongs to a Hölder class, without further knowledge of α and M . Therefore, wavelet shrinkage provides a universal solution for de-noising. Strikingly, it was mathematically proven that the recovery of a signal by wavelet shrinkage is as good as that obtained by earlier algorithms requiring specific knowledge of α and M [12]. Therefore, based on wavelet shrinkage we can propose a model that universally explains SR, including both Types E and I in the same model.

To realize the new model, we must first overcome a mathematical difficulty. Throughout the rest of this paper, we will always denote by $\tilde{h}(t)$ the signal recovered from a noisy signal as expressed in (1). In signal processing, the performance of a de-noising algorithm is mainly judged by the closeness between the recovered signal $\tilde{h}(t)$ and original signal $h(t)$, and this closeness is measured in terms of L_2 norm $\|\tilde{h}(t) - h(t)\|_2 \underline{\text{def}} \sqrt{\int_0^1 (\tilde{h}(t) - h(t))^2 dt}$. For SR in the mammalian auditory system, however, when $h(t)$ has few sharp transients which are lost in \tilde{h} , one may still have $\|\tilde{h}(t) - h(t)\|_2 \approx 0$. Of greater concern, for a given original signal $h(t)$, the recovered signal $\tilde{h}(t)$ is random. This is because the noise $e(t)$ is random, and hence, the noisy signal $f(t) = h(t) + e(t)$ is random. While in signal processing, the performance of a de-noising algorithm such as wavelet shrinkage (see [12]), is judged by $E[\|\tilde{h}(t) - h(t)\|_2]$, the average closeness between $\tilde{h}(t)$ and $h(t)$, it would clearly be unacceptable to claim that SR enhances mammalian auditory information processing *on the average*. Fortunately, the performance of wavelet shrinkage can be judged by $\sup_{1 \leq t \leq 1} |\tilde{h}(t) - h(t)|$ with very

high probability [13]. That is, the signal recovered by wavelet shrinkage is almost surely (with probability 1) close to an original signal, even when examined in a point-wise fashion. In the next section, we will propose a model for SR based on this new result of Hong and Birget [13]. Since in mammalian hearing any part of $h(t)$ may contain crucial information, a necessary condition to recover an original signal $h(t)$ from the noisy signal $f(t) = h(t) + e(t)$ is that the noisy signal be detectable. The proposed model will show that in mammalian hearing, SR occurs *if and only if the noisy signal is detectable*. In addition, we will demonstrate that the model explains both so-called Types E and I SR in a unitary mechanism.

In the final section, we will indicate how observed physiological structures and functions in mammalian auditory system are consistent with and suggest the proposed model.

Results and Discussion

The proposed model

Recall that in SR the role of noise $e(t)$ is to sample an original signal $h(t)$ generating a noisy signal $f(t) = h(t) + e(t)$; and that the sampling rate needs to be greater than the frequency of $h(t)$ [7]. Mathematically, the sampling of the original signal by noise indicates that SR has a discrete nature. A mammalian auditory system can therefore be viewed as a "device" with the following characteristics. Let n , a large positive integer, denote the sampling rate.

Input: At time instants $t = \frac{i}{n}, i = 1, 2, \dots, n$, an original subthreshold signal $h(t), t \in [0,1]$, is sampled by a noise $e(t)$. This results in the noisy samples $f(\frac{i}{n}) = h(\frac{i}{n}) + e(\frac{i}{n})$.

Output: A recovered signal $\tilde{h}(t)$ obtained by processing the noisy samples $f(\frac{i}{n}), i = 1, 2, \dots, n$.

Since the noise $e(t)$ is intrinsic, i.e., generated within the mammalian auditory system, the intensity is clearly always bounded. That is, the random variable $e(t)$ is bounded. We assume there are two constants $0 \leq a < b$ such that $e(t) \in [a,b]$. The criterion for the closeness between the recovered signal $\tilde{h}(t)$ and original signal $h(t)$ is

the smallness of $\sup_{t \in [0,1]} |\tilde{h}(t) - h(t)|$ (2)

where the meaning of the standard mathematical notation "sup" is as follows. Consider all upper bounds for $P(\cdot)$ where (\cdot) stands for an expression and P is a predicate which the expression must satisfy. Then, $\sup_P(\cdot)$ is the smallest possible of all the upper bounds.

The procedure to process the noisy samples follows from the notion of wavelet shrinkage in signal processing. It consists of two linear transforms and one non-linear thresholding. That is, we model a mammalian auditory system as a non-linear system.

First, a linear transform is carried out to decompose the noisy samples in the cochlea. For simplicity, in accordance with (1) we denote the noisy samples by f . Auditory information carried by the original signal $h(t)$ is encoded by the changes in both amplitude and frequency. Hence, retrieval of the information from $h(t)$ requires its decomposition according to both amplitude and frequency. Now, $h(t)$ is mixed with a noise $e(t)$, generating $f(t)$; and the function of the auditory system is to process the noisy samples f . Thus, a decomposition of f is necessary at the very beginning of the procedure. The principle for such a decomposition is as follows: f is viewed as an element in a function space, usually $L^2[0,1]$; and then, with the choice of a basis of $L^2[0,1]$, it finds the projections of f on each component in the basis. Thus, the mathematical quality of the decomposition is determined by the basis. Technically, during mammalian auditory information processing, the noisy samples f are decomposed to allow recovery of $h(t)$. Any basis that is chosen for the decomposition must be sensitive in detecting changes in both amplitude and frequency at the same time. It is mathematically proven that a wavelet basis is the best choice for such a decomposition. While there are many wavelet bases, from the Haar to the Daubechies, we do not specify a particular wavelet basis in the proposed model, except that it is required to be orthonormal.

It must be noted that once a wavelet basis is chosen, the linear transform is constant in the following sense. Recall that in a standard way, a linear transform can be represented by a matrix and vice versa. The matrix representing this linear transform is constant if all entries in the matrix are constants. From a viewpoint of physiology, this indicates that once a mammalian auditory system is developed, it may decompose signals to filter out noise in a fixed manner.

Since the first linear transform decomposes the noisy samples, it is necessary to filter out the noise right after this

transform. A non-linear thresholding is applied immediately as the second step in the procedure. It also must be noted that the threshold here is again a constant if the sampling rate n is regarded as fixed. The output from the second step is the decomposed noisy samples with the noise filtered out. Thus, the third (final) step of the procedure is to re-compose the filtered output of the second step. It is carried out by a linear transform, which again is constant in the sense mentioned above for the first step.

Mathematically, we describe the three steps as follows. Two related $n \times n$ orthonormal matrices V and \tilde{V} respectively for a discrete wavelet transform (DWT) and its inverse are used for the first and third steps, respectively.

$$V = \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{pmatrix}$$

and

$$\tilde{V} = \begin{pmatrix} \widetilde{v}_{11} & \widetilde{v}_{12} & \dots & \widetilde{v}_{1n} \\ \widetilde{v}_{21} & \widetilde{v}_{22} & \dots & \widetilde{v}_{2n} \\ \dots & \dots & \dots & \dots \\ \widetilde{v}_{n1} & \widetilde{v}_{n2} & \dots & \widetilde{v}_{nn} \end{pmatrix}$$

For a mammalian auditory system, the two matrices are fixed, i.e. v_{ij} and \widetilde{v}_{ij} are fixed during development, and they are used to process any noisy signal entering the system.

A threshold for the second step is defined as

$$\lambda_{n,\delta} = c \cdot (b - a) \cdot \left(1 + 2\sqrt{1 + \delta}\right) \ln 2 \cdot \sqrt{\frac{\log_2 n}{n}}$$

where $c > 0$ is a parameter determined according to which wavelet basis is used, and $\delta > 0$ is a parameter related to the accuracy of the auditory information processing. The threshold $\lambda_{n,\delta}$ is different from the threshold s used by the spike train. Notice that for an auditory system, $\lambda_{n,\delta}$ is fixed (recall that $[a, b]$ is the range of the intrinsic noise) and is used to process any noisy signal entering the auditory system.

In what follows, we will use some simplified notations.

We let h_i denote $h(\frac{i}{n})$, $i = 1, 2, \dots, n$; and let $h = (h_1 \ h_2 \ \dots \ h_n)^T$ where $(\cdot)^T$ stands for the transposition of a vector (\cdot) .

We apply the same notation to e_i , f_i , and \tilde{h}_i . Now, we mathematically formulate the three steps mentioned above.

Step 1 Discrete wavelet transform (DWT)

$$\sqrt{n} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \Leftarrow \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}$$

where $(\eta_1 \eta_2 \dots \eta_n)^T$ is the decomposed noisy samples.

Step 2 Thresholding

$$\begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix} \Leftarrow \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

where $(\zeta_1 \zeta_2 \dots \zeta_n)^T$ is the result of filtering out the noise from the decomposed noisy samples, which is obtained by

$$\zeta_i = \begin{cases} \eta_i & \text{if } |\eta_i| \geq \lambda_{n,\delta} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i = 1, \dots, n$$

Step 3 Inverse of DWT

$$\begin{pmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \vdots \\ \tilde{h}_n \end{pmatrix} \Leftarrow \begin{pmatrix} \widetilde{v}_{11} & \widetilde{v}_{12} & \dots & \widetilde{v}_{1n} \\ \widetilde{v}_{21} & \widetilde{v}_{22} & \dots & \widetilde{v}_{2n} \\ \dots & \dots & \dots & \dots \\ \widetilde{v}_{n1} & \widetilde{v}_{n2} & \dots & \widetilde{v}_{nn} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix}$$

We refer the reader to chapters 4 and 6 of [14] for details on DWT, thresholding, and the inverse of DWT. This work also describes the nature of simple constant matrices V and \tilde{V} , once a wavelet basis is chosen and the sampling n is given. Importantly, the three steps can be carried out within time of an order of n , i.e., the amount of time needed to process the noisy samples is only proportional to the number n of noisy samples. As the sampling rate n can be thought as fixed in a mammalian auditory system, the three steps can process an auditory signal as a stream. In other words, as a linear function of n , the processing rate of the proposed model is as rapid as conceivably possible. In contrast, the processing rate of a FFT is $n \times \log n$, so that the delay in processing with an increase in n would preclude online processing.

Mathematical analysis

We set a baseline for the mammalian auditory system. In terms of hearing, this baseline is understood as "absolute silence". Mathematically, the baseline is represented by a constant 0. All signals, including noise, are measured against this baseline and are evaluated in terms of pressure. In this way, $h(t)$, $e(t)$, $f(t)$ and $\tilde{h}(t)$ take positive values. Setting up such a baseline serves the following two purposes:

1. The baseline is fixed, yielding a metric system. For any given mammalian auditory system, our analysis of SR depends on this fixed metric system.
2. While our analysis of SR is *not* concerned with energy, we will be able to compute energy based on the metric system, so as to show how the proposed model naturally covers all types of SR in the mammalian auditory system.

We can assume that the threshold for a stimulus to be detected by the system is $s > 0$. Here, s is a constant against the baseline, i.e., the threshold is fixed (see Figure 1). Recall that within time interval $[0,1]$, noise $e(t)$ samples an original subthreshold signal $h(t)$, generating the noisy samples $f_i = h_i + e_i$ where $f_i = f(\frac{i}{n})$, $h_i = h(\frac{i}{n})$, $e_i = e(\frac{i}{n})$, $i = 1, 2, \dots, n$. Since in mammalian auditory information processing, any h_i may contain a critical part of the information carried by h , a necessary condition for SR to occur is

$$f_i = h_i + e_i \geq s \text{ for all } 1 \leq i \leq n \quad (3)$$

That is, all noisy samples must be detectable. However, the detectability of the samples does not at all automatically imply that the information carried by $h(t)$ is retrievable, since at the moment of acquisition, the samples are a mixture of signal $h(t)$ and noise $e(t)$.

Our goal is to use the proposed model to prove that the necessary condition expressed in (3) is also a sufficient condition for SR to occur. In precise mathematical terms, if all the noisy samples are detectable, then the information carried by $h(t)$ can be retrieved almost surely.

Recall that an original signal is represented by a function $h(t): t \in [0,1] \mapsto \mathbb{R}^+$ in a Hölder class $\Lambda^\alpha(M)$. It is straightforward to see that for all $t_1, t_2 \in [0,1]$

$$|h(t_1) - h(t_2)| \leq M|t_1 - t_2|^\alpha$$

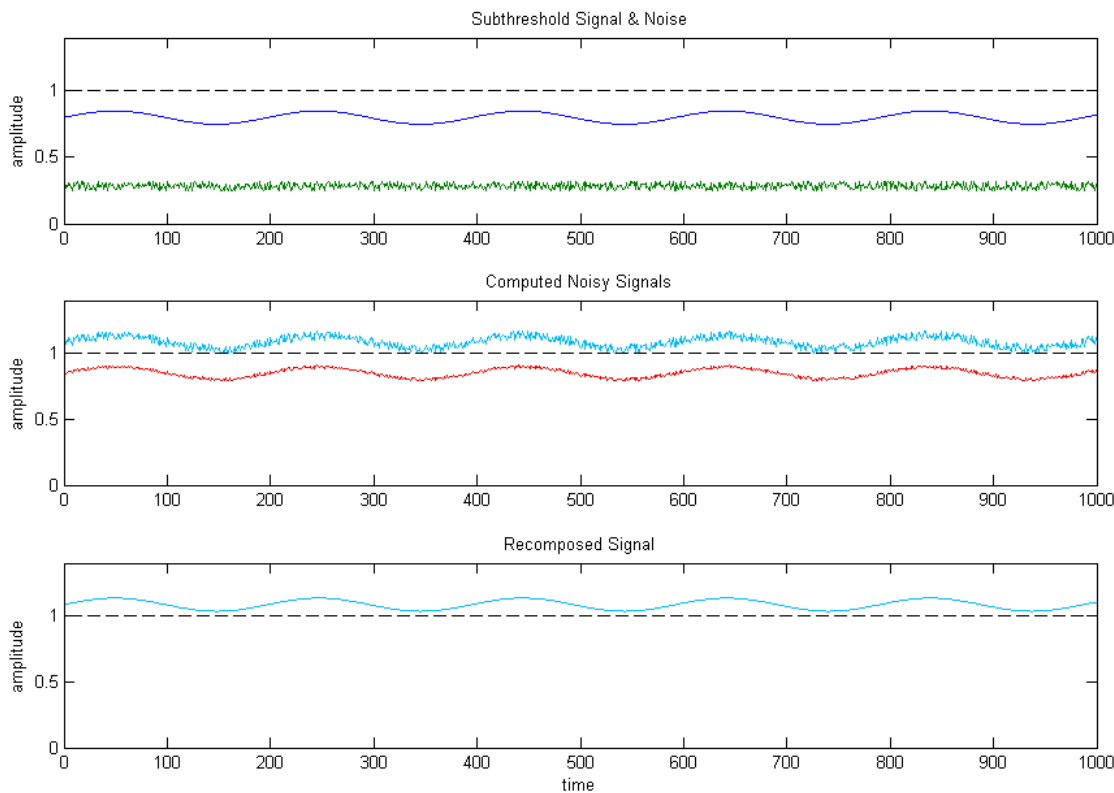


Figure 1

Comparison of energy addition and geometric translation of a signal by noise. In each panel the amplitude is plotted as a function of time and threshold is indicated by a horizontal dotted line. Top Panel: Intrinsic noise ε is plotted on the lowest (green) trace. The sinusoidal wave represents the subthreshold signal h (dark blue). Middle panel: The traces indicate the interaction of h and noise to obtain a noisy signal by energy addition (red) or by geometric translation (light blue). The noisy signal obtained by energy addition (red) does not reach threshold, so the SR would traditionally be classified as Type I. However, if the original subthreshold signal h is translated by the intrinsic noise ε with its mean m , the noisy signal obtained by geometric translation (light blue) is entirely above threshold. Lower Panel: As a result of Steps 1,2 and 3, the "denoised" signal is recovered entirely above threshold.

which implies that for all $1 \leq i \leq n$ and for all $t \in \left[\frac{i-1}{n}, \frac{i}{n} \right]$

$$|h(t) - h_i| \leq M \left(\frac{1}{n} \right)^\alpha \tag{4}$$

In physiological terms, (4) indicates that a mammalian auditory system uses its intrinsic noise to sample the original signal at a high rate so that information loss between two consecutive samples is negligible. On the other hand, (4) implies that the capability of a mammalian auditory system is limited; it has to lose some information between two consecutive noisy samples. Mathematically, with (4)

we need to focus only on $\tilde{h} = (\tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_n)$, the recovery from the noisy samples obtained in Step 3.

In signal processing, the criterion to judge the quality of the recovery \tilde{h} is the average squared error $E[Q_{avg}(n)] \stackrel{\text{def}}{=} E \left[\frac{1}{n} \sum_{i=1}^n (h_i - \tilde{h}_i)^2 \right]$, yet, as discussed in the previous section, this criterion is not acceptable for mammalian auditory information processing.

It was mathematically proven that for a Hölder class $\Lambda^\alpha(M)$ the best that any de-noising algorithm can achieve is

$$E[Q_{\text{avg}}(n)] \leq C \cdot \left(\frac{\log_2 n}{n}\right)^{\frac{2\alpha}{1+2\alpha}} \tag{5}$$

where C is a constant (see section 2 of [11]). One would expect that we can sharpen (5) by considering the probabilistic behavior of

$$Q_{\text{max}}(n) \stackrel{\text{def}}{=} \max_{1 \leq i \leq n} \{ |h_i - \tilde{h}_i| \} \tag{6}$$

For the moment, suppose that we can prove that

$$Q_{\text{max}}(n) \leq C \cdot \left(\frac{\log_2 n}{n}\right)^{\frac{2\alpha}{1+2\alpha}} \text{ almost surely. This will not}$$

violate (5); but will strengthen it so that we can use (6) in the analysis of SR. However, there is a technical flaw in (6), which is an issue commonly overlooked in the current analysis on SR.

First, (5) and (6) can be used in signal processing since noise is always considered as a random variable with zero mean. However, in a mammalian auditory system, noise is a random variable with positive mean. For a mammalian auditory system, the baseline can logically be set at absolute silence and mathematically fixed at 0. When noise is used to sample an original signal, then it must be measured above the baseline, and hence, must have a positive mean. Accordingly, noise with a positive mean was used in [9]. Thus, we would expect $\tilde{h}_i > h_i$ which in mathematical terms can be expressed as

$$\tilde{h}_i = h_i + \mathcal{W}_i$$

Obviously, \mathcal{W}_i must be a constant; for otherwise, \tilde{h}_i , the recovery from the noisy samples, will be skewed, which may cause a severe loss of information carried by h_i (the original signal). On the other hand, \mathcal{W}_i are from the noise. Recall that we assumed that the noise is represented by a bounded random variable $e(t)$ with $0 \leq a \leq e(t) \leq b$ where $a < b$ are constants. Without loss of generality, we let the mean of this random variable be $m = \frac{a+b}{2} > 0$. The

best scenario that one can expect is $\mathcal{W}_i = m$ almost surely. For the moment, suppose this can be proven. Then, we can rewrite (6) as

$$\widetilde{Q}_{\text{max}}(n) \stackrel{\text{def}}{=} \max_{1 \leq i \leq n} \{ |h_i + m - \tilde{h}_i| \}$$

Hong and Birget [13] showed that with the threshold $\lambda_{n,\delta}$ by Steps 1, 2 and 3 we have, for all $n \geq 512$

$$\Pr \left\{ \widetilde{Q}_{\text{max}}(n) \leq (c_1 + c_2 \delta) \left(\frac{\log_2 n}{n}\right)^{\frac{\alpha}{1+2\alpha}} \right\} \geq 1 - \frac{9}{n^{1+\delta}} \tag{7}$$

where c_1 and c_2 depend only on $(b - a)$, M , and α . (We note the following. In [13] the mean m of the random variable $e(t)$ was supposed to be zero; however, with a trivial modification, all proofs can be applied when $m > 0$.) Since n is

large and $\delta > 0$, we have $1 - \frac{9}{n^{1+\delta}}$ and $\left(\frac{\log_2 n}{n}\right)^{\frac{\alpha}{1+2\alpha}}$

which are extremely close to 1 and 0, respectively. Thus, (7) indicates that the error $\widetilde{Q}_{\text{max}}(n)$ is almost surely close to 0. A key step in proving (7) was to apply a deep result in measure concentration [15], a recently developed field in probability.

Summarizing all discussed thus far in this subsection, with the proposed model we can conclude that a mammalian auditory system processes an original subthreshold signal $h(t) < s$ as follows. At time instants $t = \frac{i}{n}$, $i = 1, 2, \dots, n$, the intrinsic noise $e(t)$ with mean $m > 0$ is employed to sample the original signal, generating the detectable noisy samples $f(\frac{i}{n}) = h(\frac{i}{n}) + e(\frac{i}{n}) \geq s$. Then, by the Step 1, 2 and 3, the system recovers the noisy samples, obtaining $\tilde{h}(\frac{i}{n})$. (7) indicates that almost surely

$$\left| \left(h\left(\frac{i}{n}\right) + m \right) - \tilde{h}\left(\frac{i}{n}\right) \right| \approx 0 \text{ for all } i = 1, 2, \dots, n \tag{8}$$

This means that the system amplifies the original subthreshold signal by simply translating it up with the mean m of the intrinsic noise. Figure 1 illustrates this. Some remarks need to made.

- Our analysis of SR takes an approach that differs substantially from the current view of SR as applied to sensory physiology. However, our approach does follow from the core idea by Moss [8] that noise enhances hearing by sampling the subthreshold signal. Using recent deep results in signal processing ([15,13]), our analysis further provides a strong statement that a **necessary and sufficient condi-**

tion for SR to occur in a mammalian auditory system is that all the samples by the noise are detectable.

- Our model and analysis do not involve energy and information modulation (as was also apparent in Moss' original description of SR in sensory systems [8]). However, we formulate this idea in a rigorous and concrete way: **using noise to sample an original subthreshold signal, a mammalian auditory system processes the noisy samples to translate the original signal up (in amplitude) by the mean of the noise.**

- A new insight that our model and analysis adds to SR is as follows: **When a mammalian auditory system processes the noisy samples, it may deposit energy into the recovered signal, and this added energy is expended in the recovery process.** As a consequence, our result suggests that information modulation is not a likely mechanism for SR, as discussed below.

Recall that in the analysis of a mammalian auditory system above, all signals and noise are evaluated in terms of their pressure against a fixed baseline. Thus, we can compute energies of signals and noise. The energy carried by an original subthreshold signal is

$$E_{\text{signal}} = \int_0^1 h(t)^2 dt$$

Since the sample rate n is large, we can interpolate $\tilde{h}_i, i = 1, 2, \dots, n$, by segments to have a function $\tilde{h}(t), t \in [0, 1]$. This is equivalent to taking the Haar wavelet as the basis, and thus, will not violate our analysis presented above. Our analysis above showed $\tilde{h}(t) \equiv h(t) + m$ almost surely. Hence, we can write the energy carried by the recovered signal as

$$E_{\text{recovery}} = \int_0^1 \tilde{h}(t)^2 dt = \int_0^1 (h(t) + m)^2 dt$$

and hence,

$$\begin{aligned} E_{\text{recovery}} &= \int_0^1 h(t)^2 + 2m \int_0^1 h(t) dt + m^2 \\ &= E_{\text{signal}} + 2m \int_0^1 h(t) dt + m^2 \end{aligned}$$

Recall that the energy carried by a random variable equals the deviation of the random variable. Thus, we can see that the energy carried by the intrinsic noise as a random variable is

$$E_{\text{noise}} = \lambda m^2 \text{ for } 0 < \lambda < 1$$

where for a given noise λ is a constant. Therefore,

$$E_{\text{recovery}} = [E_{\text{signal}} + E_{\text{noise}}] = 2m \int_0^1 h(t) dt + (1 - \lambda)m^2 > 0 \tag{9}$$

since $h(t) > 0$ and $m > 0$. In addition to the energies of the original signal and intrinsic noise, (9) indicates that as the noisy samples are processed, the auditory system itself adds an extra energy of amount $2m \int_0^1 h(t) dt + (1 - \lambda)m^2$ to the recovered signal. The extra energy allows SR to occur even if $[E_{\text{signal}} + E_{\text{noise}}]$ is not sufficient to reach threshold. Indeed, if one explained SR by energy addition, then it would be necessary that

$$[E_{\text{signal}} + E_{\text{noise}}] \geq s^2$$

i.e., the added energy is at least more than a constant signal with intensity equal to the threshold s . Thus, when

$$[E_{\text{signal}} + E_{\text{noise}}] < s^2 \tag{10}$$

then energy addition can no longer be used to explain SR. Moss and his coworkers called SR under condition (10) Type I, and asserted that it would require information modulation. However, an infinite number of signals h can be shown to satisfy the necessary and sufficient condition for SR to occur as suggested by our proposed model

$$h(t) + m \geq s \text{ all } t \in [0, 1]$$

and yet have the property that

$$[E_{\text{single}} + E_{\text{noise}}] < s^2$$

Here, we present one example, as summarized in Figure 1. Suppose that the noise $e(t)$ is characterized as $0.25 \leq e(t) \leq 0.32$ and $m = 0.285$; and the threshold $s = 1.0$. Consider an original signal $h(t)$ such that its intensity is within $[0.75, 0.85]$ and its average intensity is 0.8. Then the energy of $e(t)$ is no more than $0.285^2 = 0.08$, and the energy of h is no more than $0.8^2 = 0.64$. Then $0.08 + 0.64 < 1$ implies that the energy addition of $e(t)$ and $h(t)$ is not sufficient to explain how SR can enhance the reception of $h(t)$. However, since $h(t) + e(t) \geq (0.75 + 0.25) = 1.0$, the necessary and sufficient condition for our proposed model is satisfied and SR will occur (even without invoking information modulation).

Physiological analysis

Our proposed model points out that the mammalian auditory system needs only to be capable of performing Steps 1, 2 and 3 to process a noisy signal. The mammalian auditory system has a long history of neuroanatomical, physiological and psychophysical analysis (cf. [16-23])

with which to draw parallels to the steps of this model. Moreover, SR phenomena have been clearly documented in this system [4,7], thus providing the impetus for modeling. Since our model was inspired by an analysis of SR in the mammalian auditory system, we find it important to consider how the steps of the processing might be performed. Since (1) indicates that noise is added directly to subthreshold signals, this process is likely to occur in the inner ear, the origin of the neural aspects of the auditory system. Processing of an auditory signal involves both transduction by hair cells and synaptic integration by innervating spiral ganglion neurons. Outer hair cells (OHC) insert energy into the signal as they modulate the stiffness of the tectorial membrane. Changes in stiffness of the tectorial membrane modulate transduction by the inner hair cells (IHC) and enhance signal transduction at near-threshold amplitudes. Evidence for this statement is implicit in the degradation of transduction capabilities by IHC's when OHC's are immobilized [24].

The IHC's drive spiking of spiral ganglion cells, but spiral ganglion axons also convey spikes in the absence of a signal [25]. Thus, the central output of the spiral ganglion appears to include added noise, as in (1). A spectrally complex signal is transduced by the spatially-organized frequency-based array formed at the cochlea and by the IHC's [23]. The spiral ganglion cells convey that information in their spike trains. These axons terminate in a spatially-organized pattern in the cochlear nuclei [23], thereby preserving the array derived at the cochlea.

The spiking activity within the orderly array of spiral ganglion cells and their central terminations in the cochlear nuclei can therefore be seen to represent the matrix of Step 1 and the thresholding operation of Step 2. Both frequency and amplitude information are simultaneously represented in the output of the array of spiral ganglion cells (eg, [23,26]). The orderly spatial mapping of the cochlea and cochlear nuclei is preserved in the serial pathway that includes the midbrain inferior collicular nucleus, the thalamic medial geniculate nucleus, and primary auditory cortex (e.g., [27-29]).

The neuroanatomical array maintains the signal representation to the cortex. Thus, the extensive representation of the cochlear array continued throughout the auditory system embodies the first two steps of our model.

It is currently difficult to precisely localize the anatomical site of occurrence of Step 3, the recovery of the signal. Step 3 is likely to occur sometime after primary auditory cortex, in which the array is also preserved [30-33]. Linguistic recognition in humans and animal recognition of species-specific vocalizations occurs beyond primary acoustic cortex [12,34], indicating that the reconstruction of a signal

must also occur in higher order cortical areas involved in auditory function. Since the mammalian auditory system is capable of the concurrent recovery of frequency and amplitude information in short time segments [35,36], this, too, suggests a relationship between performance in the mammalian auditory system and our model based on wavelet analysis.

Hopfield [37,38] suggested that, as a consequence of evolution, interconnections built among a large number of simple neurons will form a stable network; and these networks compute [39]. Among the computational abilities of Hopfield networks are thresholding and linear transform (cf. [39]), both of which are required for our model. Hopfield networks may play the role of subsystems for DWT and its inverse. Thus, the mammalian auditory system, either at the stage of hair cell and spiral ganglion response integration (for Steps 1 and 2) or more centrally, in, for example, the auditory association cortex (for Step 3), may be considered as containing multiple Hopfield networks, and capable of the computations necessary for our model.

Central to our model's representation of SR is the realization that the system must add energy to the input (the initial signal and the noise) to exceed a perceptual threshold. To obey the first law of thermodynamics, the auditory system itself must therefore intrinsically add some energy to the noisy signal, and this extra energy is expended during the processing of information in Steps 1, 2 and 3. Three types of evidence suggest that this requirement is met experimentally. The first is the demonstration in mammalian hearing that a significant loss of threshold occurs with the loss of outer hair cell function [24]. Thus, one source of intrinsic energy might be embedded in the role of the OHCs. A second type of evidence is reflected in the physiology of eighth nerve afferents to the brain from the cochlea in the absence of a stimulus. Many studies of spiral ganglionic axons reveal classes of axons with different spontaneous activity (SA): one with Gaussian-like SA, one with bursting SA, and one with little SA. Spontaneous activity in an axon reflects an intrinsic property of the system that correlates with the sensitivity at an axon's characteristic frequency. Even in kittens raised in the absence of obvious sound stimuli, 8th nerve axons of these animals carry spontaneous activity [40]. Thirdly, in experiments with implanted cochlear electrodes in deaf people, Zeng et al. [9] showed that the addition of noise (i.e., extra energy) to a defined signal enhanced the perceptual sensitivity when near threshold levels.

In summary, the auditory system of mammals contains the necessary elements for using SR to process acoustic information according to the requirements and steps of our proposed model.

Conclusion

We present a new theoretical viewpoint for the analysis of SR in the mammalian auditory system. Most strikingly, the analysis indicates that the mechanism for reception of auditory sensation is necessarily more active than previously considered.

Although energy-requiring aspects of cochlear function have been described previously [24], the current analysis indicates that the addition of energy is a key feature of auditory receptor function. The new model suggests that the effect of noise is to carry out a geometric translation, "lifting" the original signal by the mean of the noise and creating a noisy signal which is above threshold and discernable (see Figure 1). The result of this geometric translation is more than the energy addition of the original (subthreshold) signal and intrinsic noise.

The model shows that the mechanism underlying the geometric translation does not need to be very complex. The function of the mammalian auditory system can be modeled very simply in three steps by a DWT, followed by thresholding and the inverse of DWT. Wavelet analysis is considered a useful model of the auditory system because of the capability to concurrently represent temporal and intensity information in short time segments. Furthermore, the parameters used in the DWT, thresholding and inverse DWT are invariant and the processing can therefore proceed instantaneously. Since the parameters are invariant, they are components of the phenotype and therefore would be subject to natural selection. The mammalian auditory system, optimized by evolution, appears to have evolved unique specializations to take advantage of the phenomenon of SR to enhance sensory perception. The auditory system should be considered as an active, not passive, receptor.

Authors' contributions

DH carried out the mathematical derivations and the drafting and review of the manuscript. JVM and WMS participated in the analysis of the model and the revision of the manuscript

Acknowledgements

The project was partially supported by NSF CNS 0310793 (DH) and the Rutgers University Academic Excellence Fund (JVM). This is a publication of the Rutgers University Center for Computational and Integrative Biology.

References

- Bennzi R, Sutera R, Vulpiani A: **The mechanism of stochastic resonance.** *J Phys A* 1981, **14**:L453-457.
- Casado-Pascual J, Guez-Ordóñez J, Morillo M: **Stochastic resonance: theory and numerics.** *Chaos* 2005, **15**:1-12.
- Nicolis C: **Stochastic aspects of climatic transitions.** *Tellus* 1982, **24**:1-9.
- Moss F, Chiou-Tan F, Klinke R: **Will there be noise in their ears?** *Nat Med* 1996, **2**:860-862.
- Douglass J, Wilkens L, Pantazelou E, Moss F: **Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance.** *Nature* 1993, **365**:337-340.
- Moss F: **Stochastic resonance: from the ice ages to the monkey's ear.** In *Contemporary Problems in Statistical Physics* Edited by: Weiss GH. Society for Industrial and Applied Mathematics; 1994:16-27.
- Moss F, Ward L, Sannita W: **Stochastic resonance and sensory information processing: a tutorial and review of application.** *Clin Neurophysiol* 2004, **115**:267-281.
- Greenwood P, Ward L, Wefelmeyer W: **Statistical analysis of stochastic resonance in a simple setting.** *Phys Rev E* 1999, **60**:4687-4695.
- Zeng FG, Fu QJ, Morse R: **Human hearing enhanced by noise.** *Brain Res* 2000, **869**:251-255.
- Ward L, Desai S, Rootman D, Tata M, Moss F: **Noise can help as well as hinder seeing and hearing.** *Bull Am Phys Soc* 2001, **46**:N23.002.
- Donoho D, Johnstone I, Kerkyacharian G, Picard D: **Wavelet shrinkage: Asymptopia?** *J R Stat Soc Ser B* 1995, **57**:301-369.
- Davis M, Johnsruide I: **Hierarchical processing in spoken language comprehension.** *J Neuroscience* 2003, **23**:3423-3431.
- Hong D, Birget JC: **Deviation bounds for wavelet shrinkage.** *IEEE Trans Inf Theory* 2003, **49**:1851-1858.
- Vidakovic B: *Statistical Modeling by Wavelets* New York: John Wiley; 1999.
- Talagrand M: **Concentration of measure and isoperimetric inequalities in product spaces.** *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques* 1995, **81**:73-205.
- de No RL: *The Primary Acoustic Nuclei* New York, NY: Raven Press; 1981.
- Scott S, Johnsruide I: **The neuroanatomical and functional organization of speech perception.** *Trends Neurosci* 2003, **26**:100-107.
- Yost W, Popper A, Fay R, (Eds): *Human psychophysics Volume 3.* New York: Springer-Verlag; 1993. [Springer handbook of auditory research]
- Webster D, Popper A, Fay R, (Eds): *The mammalian auditory pathway: neuroanatomy Volume 1.* New York, Springer-Verlag; 1992. Springer handbook of auditory research
- Oertel D, Fay R, Popper A, (Eds): *Integrative functions in the mammalian auditory pathway Volume 15.* New York: Springer-Verlag; 2002. [Springer handbook of auditory research]
- Keidel W, Neff W, (Eds): *Auditory system: clinical and special topics Volume 5.* Issue part 3 New York: Springer-Verlag; 1976. [Handbook of sensory physiology]
- Keidel W, Neff W, (Eds): *Auditory system: physiology (CNS), behavioral studies, psychoacoustics Volume 5.* Issue part 2 New York: Springer-Verlag; 1975. [Handbook of sensory physiology]
- Keidel W, Neff W, (Eds): *Auditory system: anatomy, physiology (ear) Volume 5.* Issue part 1 New York: Springer-Verlag; 1974. [Handbook of sensory physiology]
- Lieberman M, Gao JA, He D, Wu SP, Jia SP, Zuo J: **Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier.** *Nature* 2002, **419**:300-304.
- Lieberman M: **The cochlear frequency map for the cat: labeling auditory-nerve fibers of known characteristic frequency.** *J Acoust Soc Am* 1982, **72**:1441-1449.
- Javel E: **Basic response properties of auditory nerve fibers.** In *Neurobiology of Hearing: The Cochlea* Edited by: Altschuler R, Hoffman D, Bobbin R. New York: Raven Press; 1986:213-245.
- Ehret G: **The auditory midbrain, a "Shunting Yard" of acoustical information processing.** In *The Central Auditory System* Edited by: Ehret G, Romand R. New York: Oxford University Press; 1997:259-316.
- de Ribaupierre F: **Acoustical information processing in the auditory thalamus and cerebral cortex.** In *The Central Auditory System* Edited by: Ehret G, Romand R. New York: Oxford University Press; 1997:317-397.
- Schreiner C: **Order and disorder in auditory cortical maps.** *Current Opinion in Neurobiology* 1995, **4**:489-496.
- Merzenich M, Brugge J: **Representation of the cochlear partition on the superior temporal plane of the Macaque monkey.** *Brain Research* 1973, **50**:275-296.

31. Rauschecker J, Tian B, Pons T, Mishkin M: **Serial and parallel processing in rhesus monkey auditory cortex.** *J Comp Neurol* 1997, **382**:89-103.
32. Recanzone G, Guard D, Phan M: **Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey.** *J Neurophysiol* 2000, **83**:2315-2331.
33. Bilecen D, Scheffler K, Schmid N, Tschopp K, Seelig J: **Tonotopic organization of the human auditory cortex as detected by BOLD-FMRI.** *Hearing Research* 1998, **126**:19-27.
34. Gil-da Costa R, Martin A, Lopes M, Munoz M, Fritz J, Braun A: **Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque.** *Nature Neuroscience* 2006, **9**:1064-1070.
35. Langner G, Schreiner C: **Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms.** *J Neurophysiol* 1988, **60**:1799-1822.
36. Cooper N, Rhode W: **Nonlinear mechanics at the apex of the guinea-pig cochlea.** *Hear Res* 1995, **82**:225-243.
37. Hopfield J: **Neural networks and physical systems with emergent collective computational abilities.** *Proc Natl Acad Sci USA* 1982, **79**:2554-2558.
38. Hopfield J: **Neurons with graded response have collective computational properties like those of two-state neurons.** *Proc Natl Acad Sci USA* 1984, **81**:3088-3092.
39. Haykin S: *Neural networks: a comprehensive foundation* 2nd edition. Upper Saddle River, New Jersey: Prentice-Hall International; 1999.
40. Liberman M: **Auditory-nerve response from cats raised in a low-noise chamber.** *J Acoust Soc Am* 1978, **63**:442-455.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

