Asymptotic Optimal Noisy Neural Coding

Mark D. McDonnell [a,*], Nigel G. Stocks [b], Charles Pearce [c], Derek Abbott [a]

[a] *School of Electrical and Electronic Engineering,*

&

Centre for Biomedical Engineering,

The University of Adelaide, SA 5005, Australia

[b] *School of Engineering,*

The University of Warwick, Coventry CV4 7AL, UK

[c] *School of Applied Mathematics,*

The University of Adelaide, SA 5005, Australia

Abstract

Populations of neurons manage to operate effectively despite very noisy operating conditions. Indeed, recent work has shown that in a population of model neurons, optimal performance occurs for nonzero noise. This mechanism, known as suprathreshold stochastic resonance, requires all neurons to have the same threshold. We remove this restriction, and present an analysis of the asymptotic optimal encoding of a random stimulus by a population size approaching infinity. We find that below a certain SNR, the asymptotic optimal solution coincides with suprathreshold stochastic resonance and derive an approximation for the optimal threshold values for SNRs in the region above this point.

Key words: Neural coding, noise, stochastic resonance, suprathreshold stochastic

resonance, quantization, mutual information, Fisher information

1 Introduction

Recently, there has been much interest in the use of information theory for analysis of neural coding [4]. It has been shown that despite operating with very Signal to Noise Ratios (SNRs), the brain is able to efficiently encode information [1], implying that encoding of information is performed in an optimal manner, given these ambient noise conditions. There has also been much interest in Stochastic Resonance (SR) in neurons [9]. SR occurs when the optimal output, usually expressed as the SNR, of a system occurs for nonzero input or internal noise [5]. In particular, it has been shown that the mutual information through a population of identical parallel noisy neurons is optimized by a certain non zero noise intensity. This phenomenon is a form of SR known as Suprathreshold Stochastic Resonance (SSR) [11,12]. Here, we link SSR with population coding, in particular, the asymptotic behavior of population codes, which occurs when the number of neurons is very large [13]. Specifically, we find the optimal encoding for the same system in which SSR occurs, for a population size approaching infinity, by using a Fisher information approximation to the mutual information. We show that below a certain SNR that SSR is the optimal encoding.

* Corresponding author.
Email address: mmcdonne@eleceng.adelaide.edu.au (Mark D. McDonnell).

2 Problem formulation

We consider a population of neurons as an information channel consisting of N individual neurons receiving the same random input signal, x. The channel output, y, is the sum of all the individual neuron outputs. We use a very simple discrete time neuron model which simply emits an output spike when an input signal sample crosses a threshold. This model provides ease of analysis, and yet encapsulates the main nonlinearity in well known neuron models. The input to each neuron is subject to *iid* additive noise, and the output from each is unity if the input signal plus the noise is greater than the threshold, θ_n, of that neuron and zero otherwise. Thus the overall output is a discrete encoding of the input signal signal, taking on integer values from 0 to N.

If the input signal has probability density function $P_x(x)$, then the mutual information between the input and output is given by

$$I(x,y) = -\sum_{n=0}^{N} P_y(n)\log_2 P_y(n) - \left(-\int_{-\infty}^{\infty} P_x(x)\sum_{n=0}^{N} P(n|x)\log_2 P(n|x)dx\right),$$

where $P_y(n) = \int_{-\infty}^{\infty} P(n|x)P_x(x)dx$ [11].

Therefore for given signal and noise distributions, the mutual information depends entirely on the transition probabilities, $P(n|x)$, and the only free variables are the population size, N, and the threshold values of the neurons.

For the specific case of N identical neurons with threshold values equal to the signal mean it has been shown that the mutual information has a maximum for a nonzero noise intensity [11]. For large N, this is a realistic neuronal situation, since experimental evidence suggests that sensory neurons threshold values can adapt to an input signal mean [3]. In this case, for the optimal value of noise intensity, the mutual information is of the order of $0.5 \log_2 N$ bits per sample, which is about half the noiseless channel capacity. It has also been shown that the optimal noise standard deviation converges for large N to $\sigma_r \simeq \sqrt{1 - \frac{2}{\pi}}\sigma_x$ [6] which is a SNR of about 4.4 dB.

In order to determine whether this realistic neural configuration is optimal, we need to find the threshold values that optimize the mutual information for a range of SNRs. Let $\hat{P}_n$ be the probability of neuron n being triggered by signal value x so that $\hat{P}_n = \int_{\theta_n - x}^{\infty} R(\eta) d\eta = 1 - F_R(\theta_n - x)$, where F_R is the cumulative distribution function of the noise and $n = 1, .., N$. Assuming $\hat{P}_n$ has been calculated for desired values of x, we have previously derived a convenient way of recursively calculating the probabilities $P(n|x)$ for a population of size N [7]. The problem of finding the threshold settings that maximize the mutual information is a nonlinear optimization problem which can be expressed as

$$
\begin{aligned}
\text{Find:} \quad & \max_{\{\theta_n\}} I(x, y) = f(P(n|x)), \\
\text{where} \quad & P(n|x) \text{ is a function of} \\
\hat{P}_n \quad & = \int_{\theta_n - x}^{\infty} R(\eta) d\eta = 1 - F_R(\theta_n - x), \\
\text{subject to:} \; & \theta_n \in \mathbf{R}, \;\; (n = 1, .., N).
\end{aligned} \tag{1}
$$

3 Results

We use both Gaussian signal and noise, each with zero mean, and variance σ_x^2 and σ_r^2 respectively. Using an algorithm based on deterministic annealing [10], we can solve (1) numerically [8]. For example, the optimal thresholds for $N = 127$ are shown in Fig. 1. Our results show several interesting features, including a bifurcational structure where as the SNR decreases, more and more thresholds coincide to the same values. In particular, for very low SNRs, the optimal solution is indeed the realistic neural situation of all thresholds equal to the same value.

For the asymptotic case of large N, we can make use of a Fisher information approximation to the mutual information to find an approximate solution to the optimal thresholds in the region near the first bifurcation. The Fisher information is given by

$$J(x) = \sum_{n=0}^{N} \frac{\left(\frac{dP(n|x)}{dx}\right)^2}{P(n|x)}.$$

It can be shown that for large N the mutual information between x and y can be expressed in terms of the Fisher information as [2]

$$I(x,y) = H(x) - 0.5 \int_x P_x(x) \log_2 \frac{2\pi e}{J(x)} dx. \tag{2}$$

This expression is only valid if the distribution of y is close to Gaussian, which we have found to be true for large noise. It can be shown that for large N, $J(x)$ is the sum of the Fisher information for each neuron,

$$J(x) = \sum_{n=1}^{N} F_n(x) = \sum_{n=1}^{N} \left(\frac{d\hat{P}_n}{dx}\right)^2 \frac{1}{\hat{P}_n(1-\hat{P}_n)}, \tag{3}$$

with, in this case, $\hat{P}_n = 0.5 + 0.5\text{erf}\left((x-\theta_n)/\sqrt{2\sigma_r^2}\right)$. We now make use of the fact that our numerical results show that for SNRs larger than the first bifurcation point the optimal solution is to have half the thresholds equal to one value, $\theta > 0$, and the other half equal to $-\theta$. Then we get from (2)

$$I_1(x,y) = \log_2 \sigma_x^2 + 0.5 \int_x P_x(x) \log_2 \frac{N}{2} (F_1(x) + F_2(x)) dx, \tag{4}$$

where $F_1(x)$ and $F_2(x)$ are given by (3) with $\theta_n = \pm\theta$. We can find an approximation to the optimal thresholds in this region, by fitting a Gaussian to the error function squared, that is, we let $\mathrm{erf}^2(x) \simeq 1 - \exp(-x^2/c^2)$ and obtain

$$F_n(x) = \frac{2}{\pi\sigma_r^2} \exp\left(-\frac{(x-\theta_n)^2}{\sigma_r^2}\right) \frac{1}{\left(1 - \mathrm{erf}^2\left(\frac{x-\theta_n}{\sqrt{2}\sigma_r}\right)\right)} \simeq \frac{2}{\pi\sigma_r^2} \exp\left(\frac{-(x-\theta_n)^2}{2s^2}\right),$$

where $s^2 = (c^2/(2c^2-1))\sigma_r^2$. Our problem now becomes

$$\max_{\theta} \int_x P_x(x) \log_2 \left(\exp\left(\frac{-(x-\theta)^2}{2s^2}\right) + \exp\left(\frac{-(x+\theta)^2}{2s^2}\right)\right) dx. \tag{5}$$

This reduces to

$$\max_{\theta} \quad -\frac{\sigma_x^2}{2s^2} - \frac{\theta^2}{2s^2} + \int_x P_x(x) \ln\left(\cosh\left(\frac{x\theta}{s^2}\right)\right) dx.$$

Differentiating with respect to θ and setting to zero, followed by Taylor expanding tanh gives

$$-\frac{\theta}{s^2} + \frac{1}{s^2} \int_x x P_x(x) \tanh\left(\frac{x\theta}{s^2}\right) = -\frac{\theta}{s^2} + \frac{\theta}{s^4} M_2 - \frac{\theta^3}{3s^8} M_4 + O(\theta^5) + ... = 0,$$

where M_i is the i-th moment of $P_x(x)$. This expansion is valid iff $|\frac{x\theta}{s^2}| < \frac{\pi}{2}$, which is true for small θ. Dividing throughout by θ and taking the limit as $\theta \longrightarrow 0$ gives the value of σ_r where the first bifurcation occurs as $\sigma_r = \frac{\sqrt{2c^2-1}}{c}\sqrt{M_2} \simeq 0.907\sigma_x$, with $c^2 = 0.84965$. For $\theta > 0$ we have the maximizing θ for a Gaussian signal as

$$\theta \simeq \sqrt{\left(\frac{s^2}{\sigma_x^2} - \frac{s^4}{\sigma_x^4}\right)} = \frac{s}{\sigma_x}\sqrt{\left(1 - \frac{s^2}{\sigma_x^2}\right)}, \quad \sigma_r \leq 0.907\sigma_x. \tag{6}$$

This expression is shown plotted in Fig. 2, as well as the numerical solution of the maximization of (4). It is clear that as the SNR becomes larger, the optimal value of θ increases, and the approximation gets less accurate.

4 Conclusions

We have shown that below a certain SNR, the optimal encoding of a Gaussian input signal by a population of a large number of neurons, subject to *iid* Gaussian noise, occurs when all neurons have threshold values set to the signal mean. The maximum SNR at which this is the case is about 0.85 dB. This corresponds well with experimental knowledge of sensory neuronal behavior, where it is known that the ambient SNR is of the order of 0 dB [1] (i.e. $\sigma_r = \sigma_x$), and that a single stimulus can provide a response in a large number of neurons.

Acknowledgments

This work was funded by the Australian Research Council and the Leverhulme Trust, grant number F/00 215/J.

References

[1] W. Bialek, M. DeWeese, F. Rieke, D. Warland, Bits and brains: Information flow in the nervous system, Physica A 200 (1993) 581–593.

[2] N. Brunel, J.P. Nadal, Mutual information, Fisher information and population coding, Neural Computation 10 (1998) 1731–1757.

[3] M. DeWeese, A. Zador, Asymmetric dynamics in optimal variance adaptation, Neural Computation 10 (1998) 1179–1202.

[4] A.G. Dimitrov, J.P. Miller, Z. Aldworth, T. Gedeon, Non-uniform quantization

of neural spike sequences through an information distortion measure, Neurocomputing 38-40, (2001) 175–181.

[5] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic Resonance, Reviews of Modern Physics 70 (1998) 223-287.

[6] T. Hoch, G. Wenning, K. Obermayer, Adaptation using local information for maximizing the global cost, Neurocomputing 52-54 (2003) 541-546.

[7] M. D. McDonnell, C. E. M. Pearce, D. Abbott, An analysis of noise enhanced information transmission in an array of comparators, Microelectronics Journal 33 (2002) 1079–1089.

[8] M. D. McDonnell, N.G. Stocks, D. Abbott, C. E. M. Pearce, Singular solutions and suprathreshold stochastic resonance in optimal coding, Submitted to Physical Review Letters (2004).

[9] H.E. Plesser, T. Geisel, Signal processing by means of noise, Neurocomputing 38-40 (2001) 307–312.

[10] K. Rose, Deterministic annealing for clustering, compression, classification, regression and related optimization problems, Proceedings of the IEEE 86 (1998) 2210–2239.

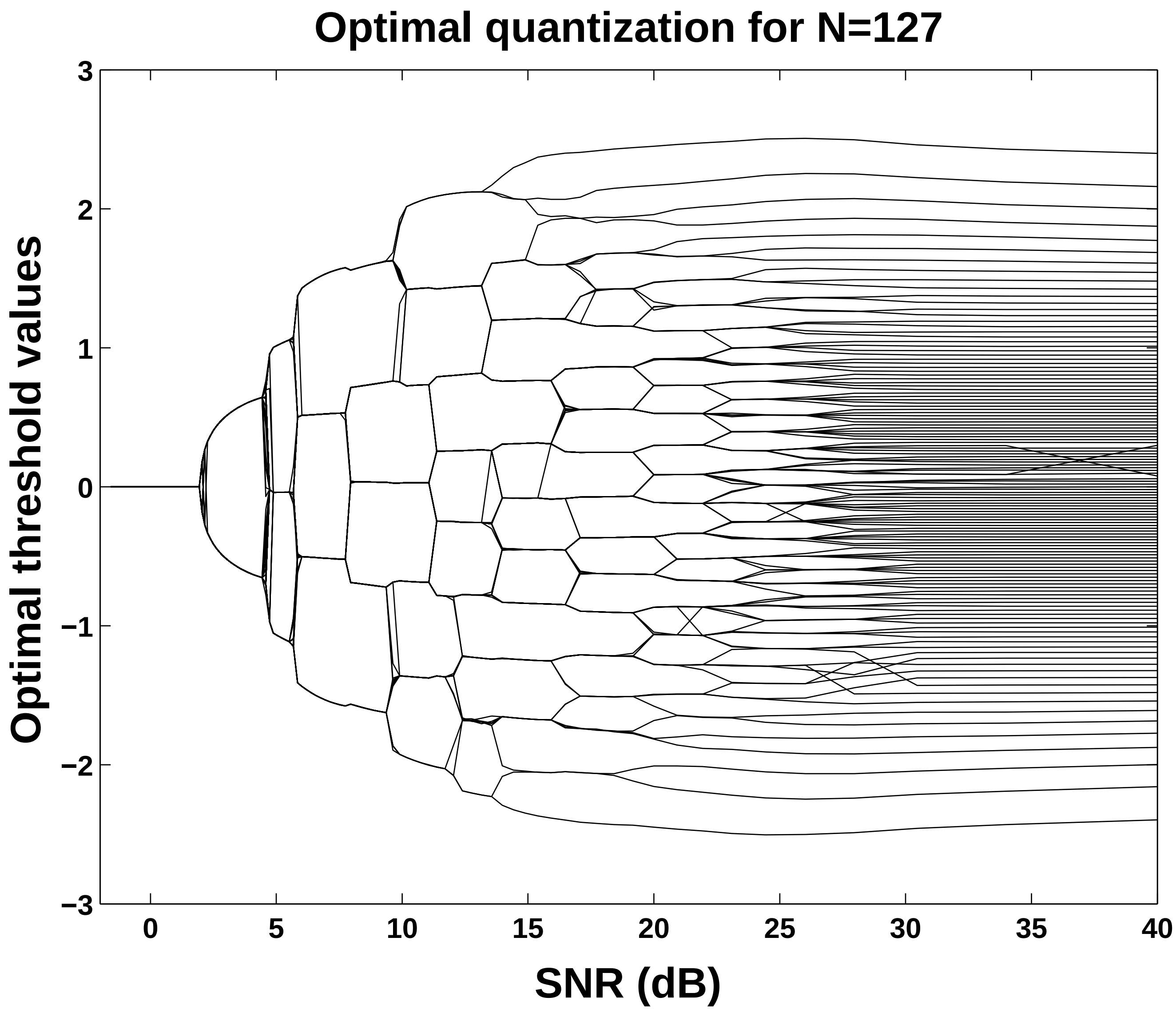
[11] N.G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, Physical Review Letters 84 (2000) 2310–2313.

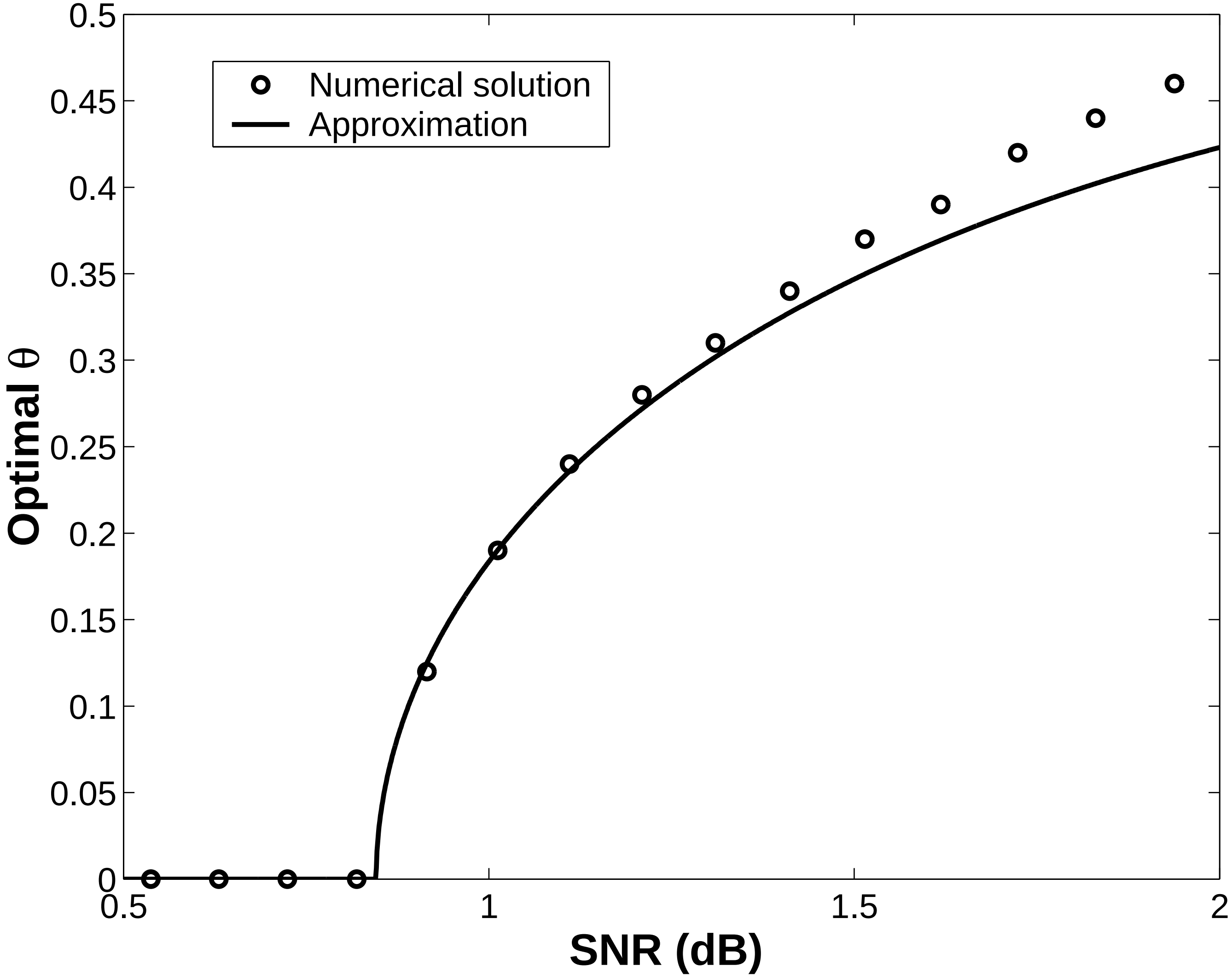
[12] N.G. Stocks, R. Mannella, Generic noise enhanced coding in neuronal arrays, Physical Review E 64 (2001) 030902(R).

[13] S. Wu, S. Amari, H. Nakahara, Asymptotic behaviors of population codes, Neurocomputing 44-46 (2002) 697-702.

Fig. 1. Plot of optimal thresholds against SNR for $N = 127$ and Gaussian signal and noise.

Fig. 2. Optimal thresholds for Gaussian signal and noise and large N, where $N/2$ thresholds are θ and the other half are $-\theta$. It can be seen that below a SNR of about 0.85 dB the optimal thresholds are all zero. Our approximation to the optimal θ can be seen to become less accurate as θ increases.


Optimal quantization for N=127
Optimal threshold values
SNR (dB)
3
2
1
0
−1
−2
−3
0
5
10
15
20
25
30
35
40


Numerical solution
Approximation
Optimal θ
SNR (dB)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.5
1
1.5
2