Optimal Noisy Neural Coding

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

[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 optimal encoding of a random stimulus by such a system. The objective is to maximize the mutual information between stimulus and response by optimally setting each neuron's threshold. We find that for low SNR's, the optimal solution is for all thresholds to be equal to the signal mean, which is a realistic neuronal situation.

Key words: Neural coding, noise, stochastic resonance, quantization, mutual

information

1 Introduction

Recently, there has been much interest in the use of information theory for analysis of neural coding [5]. It has been shown that despite operating with very low signal to noise ratios, 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 of a system occurs for nonzero input or internal noise [6]. 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]. In this paper we investigate the optimality of this encoding, by determining the optimal threshold values of each neuron in such a population.

2 Model

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

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

simple neuron model consisting of a single threshold, since this provides ease of analysis, and yet encapsulates the main nonlinearity in well known neuron models such as the leaky-integrate and fire model and the Fitzhugh-Nagumo model. The input to each neuron is subject to *iid* additive noise, η_n ($n = 1, .., N$), and the output from each, y_n, 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 [11]

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

where

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

Therefore for given signal and noise distribution, 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. Let $\hat{P}_n$ be the probability of neuron n being triggered by signal value x. Then

$$\hat{P}_n = \int_{\theta_n - x}^{\infty} R(\eta)d\eta = 1 - F_R(\theta_n - x), \tag{1}$$

where F_R is the cumulative distribution function of the noise and $n = 1, .., N$.

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]. This is a realistic neuronal situation, since experimental evidence suggests that sensory neurons threshold values can adapt to an input signal mean [4]. For large N, with 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 $\sigma_r \simeq \sqrt{1-\frac{2}{\pi}}\sigma_x$ [7] which is a signal to noise ratio of about $4.4dB$.

3 Optimal quantization

In order to determine whether the realistic neuronal situation of large N and all thresholds equal to the signal mean is optimal, we wish to determine the threshold values that optimize the mutual information for a range of signal to noise ratios. In general, it is difficult to find an analytical expression for $P(n|x)$ and we will rely on numerics. Given a noise density and threshold value, $\hat{P}_n$ can be calculated exactly for any value of x from (1). Assuming $\hat{P}_n$ has been calculated for desired values of x, a convenient way of numerically calculating the probabilities $P(n|x)$ for an array of size N is as follows. Let $T_{n,k}$ denote the probability that n of the devices $1,\ldots,k$ are "on," given x. Then let $T_{0,1} = 1-\hat{P}_1$ and $T_{1,1} = \hat{P}_1$ and we have the recursive formula

$$T_{n,k+1} = \hat{P}_{k+1}T_{n-1,k} + (1-\hat{P}_{k+1})T_{n,k},$$

where $k = 1,..,N-1$, $n = 0,..,k+1$, $T_{-1,k} = T_{k+1,k} = 0$ and we have $P(n|x)$ given by $T_{n,N}$ [8].

The problem of finding the threshold settings that maximize the mutual information is a nonlinear optimization problem which can be expressed as

$$\text{Find:} \quad \max_{\{\theta_n\}} I(x,y) = f(P(n|x)),$$

$$\begin{aligned} &\text{where} \quad P(n|x) \text{ is a function of} \\ &\hat{P}_n = \int_{\theta_n - x}^{\infty} R(\eta) d\eta = 1 - F_R(\theta_n - x), \\ &\text{subject to:}\, \theta_n \in \mathbf{R}, \ (n = 1, .., N). \end{aligned} \tag{2}$$

Using an algorithm based on deterministic annealing [10], we can solve (2) numerically. The optimal thresholds for Gaussian signal and noise, with $N = 127$ are shown in Fig. 1. Our results show several interesting features, including a bifurcational structure where as the signal to noise ratio decreases, more and more thresholds coincide to the same values. In particular, for very low signal to noise ratios, the optimal solution is indeed the realistic neuron situation of all thresholds equal.

4 Optimal SSR for large N

We make use of Fisher information [3], which for our system 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 a lower bound on the mutual information between x and y can be expressed in terms of the Fisher information and the entropy of an efficient estimator for x. In the limit of large N, the entropy of an efficient estimator approaches the entropy of the input signal, and if the distribution of this estimator becomes close to Gaussian then we have [2]

$$I(x, y) = H(x) - 0.5 \int_x P_x(x) \log_2 \frac{2\pi e}{J(x)} dx. \tag{3}$$

For large noise, we can use this approximation to find the SNR at which the final bifurcation occurs for large N, below which the SSR configuration is optimal. We make use of the fact that our numerical results show that for signal to noise ratios higher than this bifurcation point the optimal solution is to have half the thresholds equal to one value, θ, and the other half equal to $-\theta$.

To begin, the Fisher information for each neuron is

$$F_n(x) = \left(\frac{d\hat{P}_n}{dx}\right)^2 \frac{1}{\hat{P}_n(1-\hat{P}_n)}. \tag{4}$$

where for Gaussian noise with variance σ_r^2 we have

$$\hat{P}_n = 0.5 + 0.5\text{erf}\left(\frac{x-\theta_n}{\sqrt{2\sigma_r^2}}\right). \tag{5}$$

Let N be large and even, and $N/2$ of the thresholds be ϵ and the other $N/2$ be $-\epsilon$, where $\epsilon > 0$. Then the mutual information using the Fisher information approximation of (3) is

$$I_1(x,y) = \log_2 \sigma_x^2 + 0.5 \int\limits_x P_x(x) \log_2 \frac{N}{2} (F_1(x) + F_2(x))dx, \tag{6}$$

where $F_1(x)$ and $F_2(x)$ are given by (4), with $\theta_n = \pm\epsilon$. The optimal ϵ can be found numerically by varying ϵ for each value of σ_r and calculation $I_1(x,y)$. These are plotted in Fig. 2, for a Gaussian signal. We find that the bifurcation point occurs at $\sigma_r = 0.907\sigma_x$, which is a signal to noise ratio of $0.847dB$.

5 Conclusions

We have shown that below a certain signal to noise ratio, the optimal encoding of a Gaussian input signal by a population of N neurons subject to *iid* Gaussian noise occurs when all neurons have threshold values set to the signal mean. Furthermore, for large N we have shown that the maximum signal to noise ratio at which this is the case is about $0.85dB$. For higher signal to noise ratios, the optimal threshold settings is for a certain fraction of the thresholds to be unique, and as signal to noise ratio increases further, bifurcations occur that cause more and more unique optimal threshold values.

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] T.M. Cover, J.A. Thomas, Elements of Information Theory (Wiley Series in Telecommunications, 2nd Edition, Wiley, New York, 1991.)

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

[5] 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.

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

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

[8] 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.

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

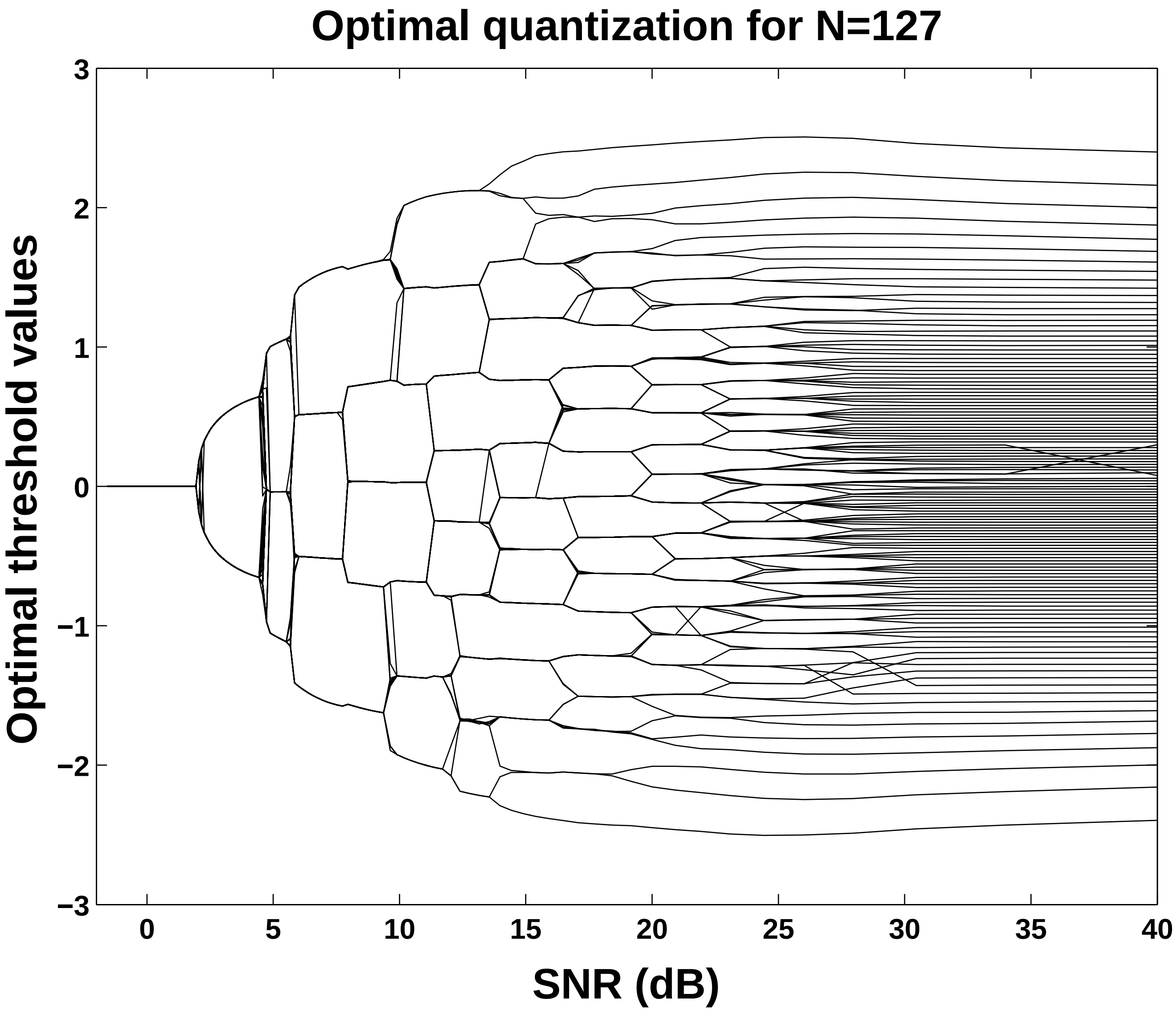
[10] K. Rose, Deterministic annealing for clustering, compression, classification, regression and related optimization problems, Proceedings of the IEEE 86 (1998) 2210–2239.

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

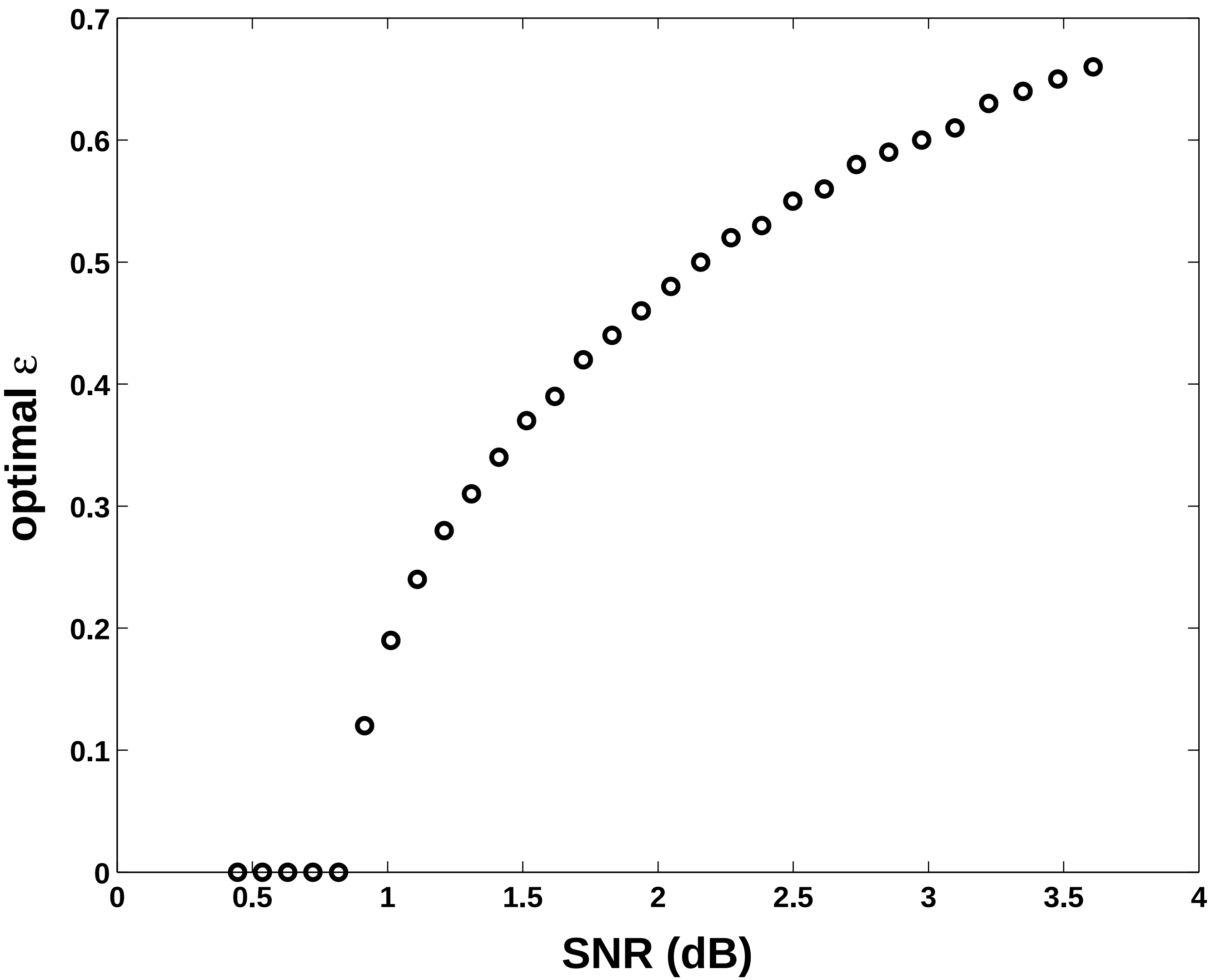
[12] N.G. Stocks, R. Mannella, Generic noise enhanced coding in neuronal arrays, Physical Review E 64 (2001) 030902(R).

Fig. 1. Plot of optimal thresholds against signal to noise ratio 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 ϵ and the other half are $-\epsilon$. It can be seen that below a signal to noise ratio of about $0.85dB$ the optimal thresholds are all zero.



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


optimal ε
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5
4
SNR (dB)