

A hierarchy of phase transitions in optimal neuronal coding: from binary to M -ary discrete optimal codes

A. Nikitin^a, N. G. Stocks^a and R. P. Morse^b

^aSchool of Engineering, University of Warwick, Coventry CV4 7AL, UK;

^bSchool of Life and Health Sciences, University of Aston, Birmingham B4 7ET, UK

ABSTRACT

We have investigated how optimal coding for neural systems changes with the time available for decoding. Optimization was in terms of maximizing information transmission. We have estimated the parameters for Poisson neurons that optimize Shannon transinformation with the assumption of rate coding. We observed a hierarchy of phase transitions from binary coding, for small decoding times, toward discrete (M -ary) coding with two, three and more quantization levels for larger decoding times. We postulate that the presence of subpopulations with specific neural characteristics could be a signature of an optimal population coding scheme and we use the mammalian auditory system as an example.

Keywords: Mutual information, optimal neuronal coding, phase transition

1. INTRODUCTION

In this paper we present preliminary results that investigate how optimal coding for neural systems changes with the time available for decoding. Our study was motivated by two contrasting findings. First, in a theoretical study based on estimation theory, Bethge et al.^{1,2} have studied the quality of information transmission by calculating the mean-square error for various decoding times. They predict a single phase transition going from binary coding for small decoding times to analogue encoding for large decoding times. Second, many anatomical and physiological studies, suggest that neural fibres can be grouped into subpopulations; for example mammalian inner hair cells (the receptors for hearing^{3,4}) can be divided into two or three subpopulations on the basis of size, mitochondrial content, synaptic morphology, spontaneous discharge rate, and threshold.⁵⁻⁹ The number and role of the subpopulations is still unclear. However, based on the findings from Bethge et al.,^{1,2} might we expect a single population if short observation times are relevant (thus implementing binary coding) or a homogenous distribution of neural parameters if longer observation times are relevant (thus implementing analogue coding)? Given that, in general, this is not what is observed we reinvestigate optimal coding but this time using Shannon information¹⁰ as the measure to be optimised. By introducing a parameterised tuning function, we estimate the parameters for Poisson neurons that optimize Shannon transinformation with the assumption of rate coding. Unlike Bethge et al.,^{1,2} we show that the tuning functions display a hierarchy of phase transitions from binary coding, for small decoding times, toward discrete (M -ary) coding with two, three and more quantization levels for larger decoding times. We therefore postulate that subpopulations, for example in the ear, could form part of an optimal coding system.

2. MODEL

The problem investigated in this paper is in principle generic. We are interested in how neural parameters should be tuned to optimally transfer information through a population of fibres. Should the parameters of each neuron be approximately the same, should there be a homogenous distribution of parameters or should parameters form subpopulations? To give a firmer focus to this question we consider the specific case of coding and transmission of signals in the mammalian auditory system, however, we stress that the conclusion are in principle quite general.

Further author information: (Send correspondence to A.N.)

A.N.: E-mail: a.nikitin@warwick.ac.uk, Telephone: +44 (0) 2476 523 141

N.G.S.: E-mail: n.g.stocks@warwick.ac.uk, Telephone: +44 (0) 2476 522 857

R.P.M.: E-mail: r.p.morse@aston.ac.uk

Noise and Fluctuations in Biological, Biophysical, and Biomedical Systems, edited by Sergey M. Bezrukov,
Proc. of SPIE Vol. 6602, 66020H, (2007) · 0277-786X/07/\$18 · doi: 10.1117/12.724410

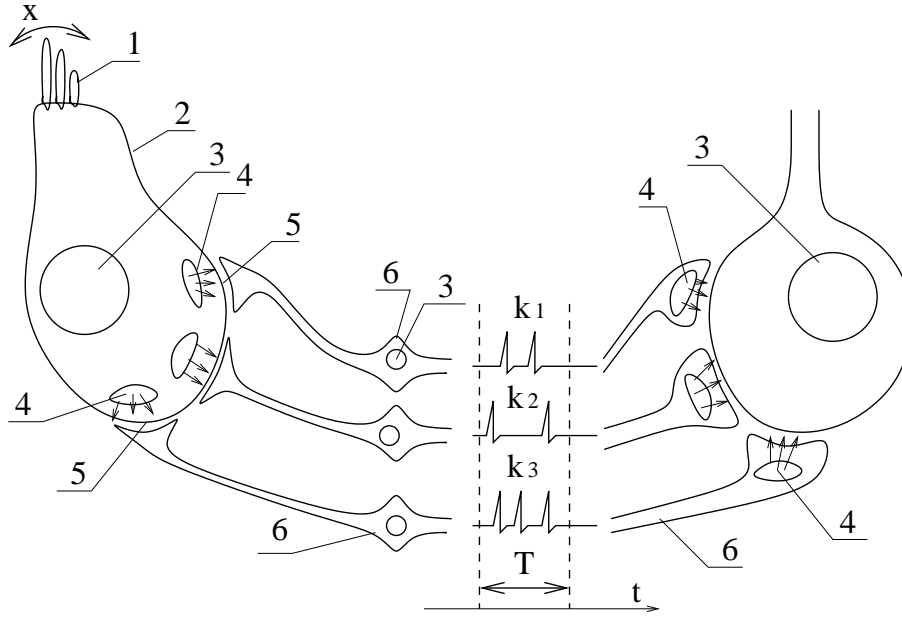


Figure 1. The structure of the mammalian auditory system. The signal, x , stimulates the sensory cell through the displacement of the stereocilia. The stereocilia displacement opens ion channels that cause depolarization of the membrane potential. Depolarization leads to the opening of calcium ion channels that in turn mediate the release of neurotransmitter into the synaptic cleft. Neurotransmitter stimulates the afferent neurons to produce action potentials. The postsynaptic neuron (seen on the right side of the figure) integrates over all incoming spikes within a time window of length T . Numbers correspond to; 1 – the stereocilia, 2 – the IHC, 3 – the nucleus, 4 – the immediate store of neurotransmitter, 5 – the cleft, 6 – afferent neuron.

It is well known that a typical mammalian inner hair cell is innervated by 10-20 fibres (see Fig. 1) that transmit signals to the brain³ and furthermore that these fibres can be grouped into subpopulations that possess common neural characteristics. Advanced models of hair cell transduction exist that take into account dynamics of the ion channels, the membrane potential of the hair cell body, quantal release of neurotransmitter into the synaptic cleft and refractoriness of the fibres.¹¹ However, such models have many parameters and hence are unsuitable for studies of optimal coding. Here we assume that the system can be modelled as a population of Poisson neurons (Fig. 2 b). Poisson neurons (i.e. neurons whose firing statistics are governed by the Poisson distribution) are known to be a good model for many types of sensory neuron.^{12,13}

The Poisson neuron can be described by the conditional probability,^{1,13}

$$P[k|\mu(x)] = \frac{[\mu(x)]^k}{k!} \exp(-\mu(x)), \quad \mu(x) = Tf(x), \quad (1)$$

where T is a time interval available for decoding and the number of observed spikes within the time interval T is k i.e. the Poisson neuron is a random map transforming the signal value x (which is modelled as a continuous random variable) into the discrete random variable k (see Fig. 2 a). In Eq. (1) the firing rate f is a function of the stimulus level x and $0 \leq f(x) \leq 1$. Therefore, $\mu(x) = Tf(x)$, corresponds to the average number of spikes observed in time T . This enables us to introduce the maximum average number of spikes N as, $N = \max_x Tf(x)$, and given $\max_x f(x) = 1$ this yields $N \equiv T$. If one is interested in optimizing information flow through a Poisson neuron the only parameters that can be varied are T and the tuning function $f(x)$. More precisely, by the introduction of a parameterised tuning function, we aim to calculate the optimal functional form for the tuning function as a function of T . It should be stated that a population of Poisson neurons whose outputs are pooled can be modelled by a single Poisson neuron (see appendix A.) and hence we can optimise Eq. (1) without having to consider a population directly.

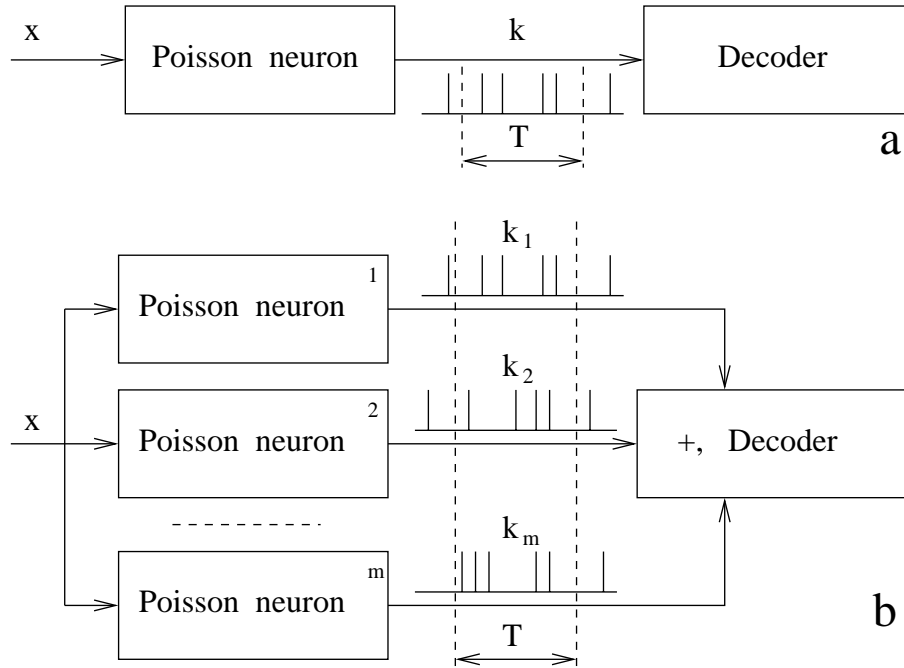


Figure 2. (a) The signal, x , stimulates the Poisson neuron the response of which, k , is analyzed by the decoder. (b) The array of the Poisson neurons stimulated by the common signal x . The responses of the Poisson neurons, k_i , are summed and analyzed by the decoder. The single Poisson neuron in (a) can be used to model an arbitrary population of Poisson neurons by choosing an appropriate tuning function. Hence without loss of generality we focus on the optimisation of model Eq. (1)

We now proceed to introduce the optimisation problem. Without loss of generality the stimulus distribution is chosen to be uniform,

$$P_x(x) = \begin{cases} 0 & : x < 0, \\ 1 & : x \in [0 : 1], \\ 0 & : x > 1. \end{cases} \quad (2)$$

We assume that $f(x)$ is a non-decreasing function, $df(x)/dx \geq 0$, with the boundary conditions: $f(0) = 0$ and $f(1) = 1$.

To estimate the quality of signal transmission, the mutual information is used,

$$I = H_k - H_{k|x}, \quad (3)$$

where the entropies are

$$H_{k|x} = - \sum_{k=0}^{\infty} \int_0^1 P_x(x) P[k|\mu(x)] \log_2 P[k|\mu(x)] dx \quad (4)$$

and

$$H_k = - \sum_{k=0}^{\infty} P_k(k) \log_2 P_k(k). \quad (5)$$

Here the probability $P_k(k)$ is

$$P_k(k) = \int_0^1 P_x(x) P[k|\mu(x)] dx. \quad (6)$$

To calculate the optimal tuning function we choose the piecewise linear approximation of the tuning function,

$$f(x) = \begin{cases} 0 & : x \in [0 : x_0), \\ \alpha_0(x - x_0) + f_0 & : x \in [x_0 : x_1), \\ \alpha_1(x - x_1) + f_1 & : x \in [x_1 : x_2), \\ \dots & : \dots, \\ \alpha_M(x - x_M) + f_M & : x \in [x_M : x_{M+1}), \\ 1 & : x \in [x_{M+1} : 1), \end{cases} \quad (7)$$

where

$$\alpha_i = \frac{f_{i+1} - f_i}{x_{i+1} - x_i}, \quad i = 0, 1, 2, \dots, M, \quad (8)$$

α_i can be transformed into a continuous function $\alpha(x) = df(x)/dx$ in the limit $M \rightarrow \infty$.

For the piecewise linear tuning function the condition entropy can be found as

$$H_{k|x} = - \sum_{k=0}^{\infty} \sum_{i=0}^M \int_{x_i}^{x_{i+1}} P_x(x) P[k|\mu(x)] \log_2 P[k|\mu(x)] dx - \sum_{k=0}^{\infty} \int_0^{x_0} P_x(x) P[k|\mu(x)] \log_2 P[k|\mu(x)] dx - \sum_{k=0}^{\infty} \int_{x_{M+1}}^1 P_x(x) P[k|\mu(x)] \log_2 P[k|\mu(x)] dx. \quad (9)$$

The probability, $P_k(k)$, can be easily calculated for the the piecewise linear tuning function,

$$\begin{aligned} P_k(k) &= \sum_{i=0}^M \int_{x_i}^{x_{i+1}} P_x(x) P[k|\mu(x)] dx + \int_0^{x_0} P_x(x) P[k|\mu(x)] dx + \int_{x_{M+1}}^1 P_x(x) P[k|\mu(x)] dx \\ &= \frac{1}{T k!} \sum_{i=0}^M \frac{1}{\alpha_i} F_i(k) + \frac{(1 - x_{M+1}) T^k}{k! \exp(T)} + \begin{cases} x_0 & : k = 0, \\ 0 & : k > 0. \end{cases} \end{aligned} \quad (10)$$

Where the function $F_i(k)$ is introduced as,

$$F_i(k) = \int_{Y_i}^{Y_{i+1}} y^k \exp(-y) dy = \Gamma_{Y_i, Y_{i+1}}(k + 1). \quad (11)$$

And the parameter y is,

$$y = T[\alpha_i(x - x_i) + f_i], \quad Y_i = T f_i. \quad (12)$$

Here $\Gamma_{Y_i, Y_{i+1}}(k)$ is the Gamma function.¹⁴

The functions, $F_i(k)$, can be calculated by using the following recurrent expressions,

$$F_i(k + 1) = (k + 1) F_i(k) - Y_{i+1}^{k+1} \exp(-Y_{i+1}) + Y_i^{k+1} \exp(-Y_i), \quad (13)$$

and

$$F_i(0) = \exp(-Y_i) - \exp(-Y_{i+1}). \quad (14)$$

Then the conditional entropy is,

$$\begin{aligned} H_{k|x} &= \frac{1}{T \log(2)} \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^M \frac{1}{\alpha_i} [k L_i(k) - F_i(k + 1) - \log(k!) F_i(k)] \\ &\quad - \frac{1 - x_{M+1}}{\log(2)} \left(T \log(T) - T - \exp(-T) \sum_{k=0}^{\infty} T^k \frac{\log(k!)}{k!} \right). \end{aligned} \quad (15)$$

Where the function $L_i(k)$ is introduced as,

$$L_i(k) = \int_{Y_i}^{Y_{i+1}} y^k \exp(-y) \log(y) dy. \quad (16)$$

The functions, $L_i(k)$, can be found by using the following recurrent expressions,

$$L_i(0) = \exp(-Y_i) \log(Y_i) - \exp(-Y_{i+1}) \log(Y_{i+1}) + E_1(Y_i) - E_1(Y_{i+1}), \quad (17)$$

where $E_1(Y_i)$ is the Exponential Integral,¹⁴

$$E_1(Y_i) = \int_{Y_i}^{\infty} \frac{\exp(-y)}{y} dy, \quad (18)$$

and

$$L_i(k+1) = Y_i^{k+1} \exp(-Y_i) \log(Y_i) - Y_{i+1}^{k+1} \exp(-Y_{i+1}) \log(Y_{i+1}) + (k+1)L_i(k) + F_i(k). \quad (19)$$

We use a simple but highly robust method of maximizing the mutual information. For a given T we calculate the mutual information using Eqs. (3), (5), (10) and (15) for an initial vector of the parameters (the vector holds the components f_i and x_i). A component of the vector is modified (with a small step) and a new value of the mutual information is calculated. Similarly, other components of the vector are modified and the mutual information is calculated for every case. Then we choose the vector corresponding to the maximum of the mutual information. We repeat this process until a local maximum of the mutual information is attained for a given numerical error. By changing the initial vector and seeking other local maxima we can find a global maximum of the mutual information and hence the vector of parameters f_i and x_i that optimise the mutual information. This procedure then gives the optimal tuning function.

3. RESULTS

Fig. 3 shows the main results of the study. The insets display the optimal tuning functions for four different values of N ($N = 2, 7, 15, 22$). The main point to note is that the optimal tuning functions contain step-like discontinuities. The presence of these discontinuities means that the firing rate can only take on discrete values - the number of discrete values depends on the value of N . For small N ($N < 3$) only two firing rates are observed; for values of $x < \theta_1$ $f = 0$ and for larger values of x , $f = 1$, where we introduce θ_1 to denote the position of the step in the tuning function. Consequently, these results demonstrate that the input signal is effectively partitioned (quantized) into two regions - each region corresponding to a different firing rate. This is similar to binary coding, although note that unlike normal binary coding the two output states are stochastic. This form of binary coding has been predicted previously^{1,2} using estimation theory. As N is increased (by increasing the decoding time T) the number of steps in the optimal transfer function also increases; for $N = 7$ two steps (denoted by θ_1, θ_2) are observed - these steps partition the signal into three regions associated with firing rates of 0, $\phi_1, 1$ and hence this is a tertiary coding scheme. For $N = 15$, three steps, $\theta_1, \theta_2, \theta_3$ partition the signal into four firing rates 0, $\phi_1, \phi_2, 1$ - hence this is a 4-ary (quaternary) coding scheme - and for $N = 22$ a 5-ary coding scheme is optimal. Due to excessive computational time it was not possible to find the optimal tuning function for values greater than $N = 25$; however, we anticipate that, in general, an M -ary code will be optimal with M increasing with increasing N . Fig. 3 shows how the partition boundaries, θ_i , vary as N is increased. Clearly, new boundaries form at existing boundaries via bifurcations (or phase transitions in the language of statistical mechanics), as far as we can tell, these bifurcations are continuous and hence are akin to second order phase transitions of the optimal transfer function.

The binary encoding is a simple regime and it allows us to obtain the analytical solution for the optimal dependence of the partition boundary, θ_1 , on the parameter T . We assume that the tuning function is the one-step function, $f(x) = \sigma(x - \theta_1)$, with location of the partition boundary (threshold) at θ_1 . The conditional distributions for the step tuning function are

$$P[k|0] = \begin{cases} 1 & : k = 0, \\ 0 & : k > 0, \end{cases} \quad (20)$$

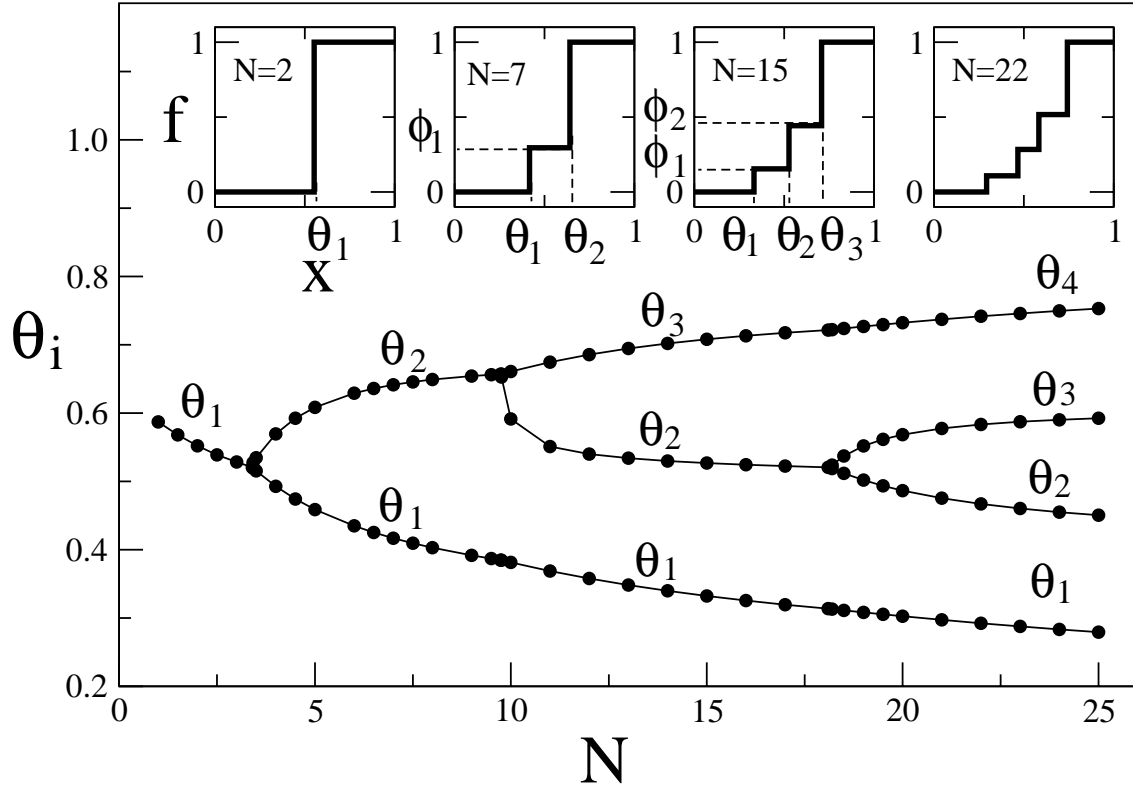


Figure 3. The partition boundaries θ_i against the number of spikes N . Optimal tuning functions, $f(x)$, for $N = 2, 7, 15, 22$ are shown as insets. Note that the optimal tuning functions contain steps that give rise to a discrete optimal code.

and

$$P[k|T] = \begin{cases} \exp(-T) & : k = 0, \\ \frac{T^k}{k!} \exp(-T) & : k > 0. \end{cases} \quad (21)$$

Then the output distribution is

$$P_k(k) = \theta_1 P[k|0] + (1 - \theta_1) P[k|T] = \begin{cases} \theta_1 + (1 - \theta_1) \exp(-T) & : k = 0, \\ (1 - \theta_1) \frac{T^k}{k!} \exp(-T) & : k > 0. \end{cases} \quad (22)$$

The entropies can be written as,

$$H_k = [1 - (1 - \theta_1)(1 - \exp(-T))] \log_2 [1 - (1 - \theta_1)(1 - \exp(-T))] - (1 - \theta_1) \sum_{k=1}^{\infty} \left(\frac{T^k}{k!} \exp(-T) \right) \log_2 (1 - \theta_1) - (1 - \theta_1) \sum_{k=1}^{\infty} \left(\frac{T^k}{k!} \exp(-T) \right) \log_2 \left(\frac{T^k}{k!} \exp(-T) \right), \quad (23)$$

and

$$H_{k|x} = -(1 - \theta_1) \sum_{k=1}^{\infty} \left(\frac{T^k}{k!} \exp(-T) \right) \log_2 \left(\frac{T^k}{k!} \exp(-T) \right). \quad (24)$$

The maximum of the mutual information can be found from the extremal condition,

$$\frac{dI}{d\theta_1} = \frac{d}{d\theta_1} (H_k - H_{k|x}) = 0, \quad (25)$$

or

$$\frac{dI}{d\theta_1} = -(1 - \exp(-T)) \log_2 \left(\frac{1 - (1 - \theta_1)(1 - \exp(-T))}{1 - \theta_1} \right) + \frac{T \exp(-T)}{\ln(2)} = 0 \quad (26)$$

Then the threshold θ_1 corresponding to the maximum mutual information is

$$\theta_1 = 1 - \left(\exp \left(T \frac{\exp(-T)}{1 - \exp(-T)} \right) + 1 - \exp(-T) \right)^{-1}. \quad (27)$$

It is easy to see the solution (27) differs from the solution of Bethge et al.,^{1,2}

$$\theta_1^B = 1 - \frac{3 - \sqrt{8 \exp(-T) + 1}}{4(1 - \exp(-T))}, \quad (28)$$

which was obtained when the mean-square error between stimulus x and the mean square estimator, $\hat{x}(k) = E[x|k] = \int_0^1 x P_{x|k}[x|k] dx$, is minimized. Though minimization of the mean-square error and the maximization of mutual information point to binary encoding for weak values of times T , both methods of optimization give different quantities of the thresholds, θ_1 and θ_1^B . For example, in limit $T \rightarrow 0$ the optimal thresholds are $\theta_1 = 1 - \exp(-1) \simeq 0.63$ and $\theta_1^B = 2/3 \simeq 0.67$.

The optimal multi-step form of the tuning function (see Fig. 3) has been found for the model consisting of a single nerve fibre. An estimation of the realistic length of the decoding time T done by Bethge et al.² indicates $1 < T < 5$ and hence suggests binary coding. Such a system cannot transmit information of more than 1 bit. To transmit more information the time window T must become longer, however, this method is not effective because the information rate, $R = I/T$, will not increase. Alternatively, the number of fibres must become bigger. As it is shown in Appendix A the Poisson neuron with a multi-step tuning function can be replaced by an array of Poisson neurons with a single-step tuning functions (see Fig. 2). Then by keeping T short we can increase N by increasing the number of Poisson neurons. By allowing each neuron of a subpopulation to have the same threshold (i.e. each subpopulation contains identical binary neurons) an optimal M -ary coding scheme can be implemented by having different thresholds (i.e. partition boundaries) between subpopulations. In this picture an M -ary code is implemented by $M - 1$ subpopulations. This method enables optimal coding and high information rates to be obtained for small T .

Could this explain why auditory neurons that innervate hair cells can be classified into three subpopulations? This question cannot be answered with any certainty, but we note that the presence of three subpopulation would suggest that quaternary coding ($M = 4$) could be implemented. From our results we note that a quaternary code is optimal when, approximately, $10 < N < 18$. Taking the average firing rate of an auditory neuron to be 100Hz and assuming that there are 10 auditory neurons innervating a single hair cell gives the total spike rate as 1000 spike/s per hair cell. The time required on average to receive 15 spikes (in the region optimal for quaternary coding) is therefore 15ms during which time each fibre produces just 1.5 spikes. This time scale is consistent with the time scales of auditory perception and hence this does not rule out the possibility of quaternary coding.

4. CONCLUSION

In this paper we attempt to find an answer to the question concerning a role of fibre subpopulations in optimal coding. We show that, in general, discrete M -ary codes are optimal for the coding of information using Poisson neurons. Furthermore, the results indicate that the presence of three subpopulations of fibres in the ear could be indicative of a quaternary coding scheme. A simple analysis confirms that quaternary coding is not inconsistent with known neural characteristics of auditory neurons and sound perception.

Finally, we note that the results obtained in this paper are qualitatively similar to results of Rose¹⁵ and McDonnell et al.¹⁶ although they considered systems that are quite different to Poisson neurons. They also observed phase transition like sequences that resulted in M -ary codes. However, these studies did not consider decoding time but varied other parameters, for example, the internal noise intensity. Nevertheless, this leads to the conclusion that optimal quantization leading to discrete M -ary codes is probably a generic phenomenon and occurs independent of the type of noise or detail of the system.

APPENDIX A. EQUIVALENCE BETWEEN A SINGLE AND AN ARRAY OF POISSON NEURONS

Consider two Poisson neurons,

$$P_{k_1|x}[k_1|\mu_1(x)] = \frac{[\mu_1(x)]^{k_1}}{k_1!} \exp(-\mu_1(x)), \quad (29)$$

and

$$P_{k_2|x}[k_2|\mu_2(x)] = \frac{[\mu_2(x)]^{k_2}}{k_2!} \exp(-\mu_2(x)), \quad (30)$$

and there is a cortical neuron collecting spikes of the Poisson neurons, $P_{k_1, k_2|x}[k_1, k_2|\mu_1(x), \mu_2(x)]$. If the cortical neuron estimates the sum of the spikes, $k = k_1 + k_2$, in the time window T , then the distribution of the spikes satisfies the equation,

$$\begin{aligned} P_{k|x}[k|\mu_1(x), \mu_2(x)] &= \sum_{k_2=0}^{\infty} P_{k_1|x}[k-k_2|\mu_1(x)] P_{k_2|x}[k_2|\mu_2(x)] \\ &= \sum_{k_2=0}^k P_{k_1|x}[k-k_2|\mu_1(x)] P_{k_2|x}[k_2|\mu_2(x)], \end{aligned} \quad (31)$$

where the condition probability is used,

$$P_{k_1|x}[k-k_2|\mu_1(x)] = \begin{cases} P_{k_1|x}[k-k_2|\mu_1(x)] & : k \geq k_2, \\ 0 & : k < k_2. \end{cases} \quad (32)$$

Then by substitution Eqs. (29) and (30) in to Eq. (31) we obtain,

$$\begin{aligned} P_{k|x}[k|\mu_1(x), \mu_2(x)] &= \sum_{k_2=0}^k \frac{[\mu_1(x)]^{k-k_2}}{(k-k_2)!} \exp(-\mu_1(x)) \frac{[\mu_2(x)]^{k_2}}{k_2!} \exp(-\mu_2(x)) \\ &= [\mu_1(x)]^k \exp(-[\mu_1(x) + \mu_2(x)]) \sum_{k_2=0}^k \left[\frac{\mu_2(x)}{\mu_1(x)} \right]^{k_2} \frac{1}{(k-k_2)! k_2!} \\ &= [\mu_1(x)]^k \exp(-[\mu_1(x) + \mu_2(x)]) \left[\frac{\mu_2(x)}{\mu_1(x)} + 1 \right]^k \frac{1}{k!} \\ &= \frac{[\mu_1(x) + \mu_2(x)]^k}{k!} \exp(-[\mu_1(x) + \mu_2(x)]) \end{aligned} \quad (33)$$

Now, by using the principle of mathematical induction¹⁷ it is easy to prove that if the cortical neuron estimates the sum of the spikes emitted by m Poisson neurons, $k = \sum_{i=1}^m k_i$, in the time window T , then the distribution of the spikes satisfies the equation,

$$P_{k|x} \left[k \left| \sum_{i=1}^m \mu_i(x) \right. \right] = \frac{[\sum_{i=1}^m \mu_i(x)]^k}{k!} \exp \left(- \sum_{i=1}^m \mu_i(x) \right), \quad (34)$$

or

$$P_{k|x}[k|\mu_{\text{eff}}(x)] = \frac{[\mu_{\text{eff}}(x)]^k}{k!} \exp(-\mu_{\text{eff}}(x)), \quad (35)$$

where $T = \sum_{i=1}^m \mu_{i, \text{max}}$, and

$$\mu_{\text{eff}}(x) = \sum_{i=1}^m \mu_i(x). \quad (36)$$

In the limit $m \rightarrow \infty$ the multi-step function $\mu_{\text{eff}}(x)$ can be rewritten in the continuous form

$$\mu_{\text{eff}}(x) = \int_0^1 \mu_{z,\text{max}}(z) P_z(z) dz, \quad (37)$$

where $P_z(z)$ is the distribution of the thresholds of single-step functions, $\mu_{z,\text{max}}(z)$ is the density of the magnitudes of the step functions.

Eqs. (35), (36) and (37) show that the array of Poisson neurons with the cortical neuron summing the spikes can be replaced by an effective Poisson neuron with the smooth continuous tuning function or multi-step tuning function.

ACKNOWLEDGMENTS

This work was funded by the EPSRC (grant GR/R35650/01 and EP/D05/1894/1(P)).

REFERENCES

1. M. Bethge, D. Rotermund, and K. Pawelzik, "Second order phase transition in neural rate coding: binary encoding is optimal for rapid signal transmission," *Phys. Rev. Lett.* **90**, p. 088104, 2003.
2. M. Bethge, D. Rotermund, and K. Pawelzik, "Optimal neural rate coding leads to bimodal firing rate distributions," *Network: Computation in Neural Systems* **14**, pp. 303–319, 2003.
3. P. Dallos, A. N. Popper, and R. R. Fay, *The Cochlea*, Springer, New York, 1996.
4. W. L. Gulick, *Hearing*, Oxford University Press, New York, 1971.
5. M. C. Liberman, "Morphological differences among radial afferent fibers in the cat cochlea: an electron-microscopic study of serial section," *Hearing Research* **3**, pp. 45–64, 1980.
6. M. C. Liberman, "Single-neuron labeling in the cat auditory nerve," *Science* **216**, pp. 1239–1241, 1982.
7. H. Spoendlin, "Degeneration behavior of the cochlear nerve," *Arch. Klin. Exp. Ohren Nasen Kehlkopfheilkd* **200**, pp. 275–291, 1971.
8. T. Kawase and M. Liberman, "Spatial organization of the auditory nerve according to spontaneous discharge rate," *The Journal of Comparative Neurology* **319**, pp. 312–318, 1992.
9. J. Tsuji and M. Liberman, "Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections," *J. Comp. Neurol.* **381**, p. 188202, 1997.
10. C. E. Shannon and W. Weaver, *The mathematical theory of communication*, University of Illinois Press, Urbana, 1949.
11. C. Summer, E. A. Lopez-Poveda, L. P. O'Mard, and R. Meddis, "A revised model of the inner-hair cell and auditory-nerve complex," *J. Acoust. Soc. Am.* **111**, pp. 2178–72188, 2002.
12. P. Dayan and L. F. Abbot, *Theoretical Neuroscience*, MIT Press, Cambridge, 2001.
13. W. Gerstner and W. Kistler, *Spiking Neuron Models*, Cambridge University Press, Cambridge, 2002.
14. M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, Dover Publications, Inc., New York, 1965.
15. K. Rose, "A mapping approach to rate-distortion computation and analysis," *IEEE Transaction on information theory* **40**, pp. 1939–1952, 1994.
16. M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, "Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance," *Physics Letters A* **352**, pp. 183–189, 2006.
17. H. L. Royden, *Real Analysis*, Macmillan Publishing Company, New York, 1988.