What Does a Neuron Talk About?

Stefan D. Wilke and Christian W. Eurich

Institut für Theoretische Physik, Universität Bremen
Kufsteiner Str., D-28359 Bremen, Germany
(swilke,eurich)@physik.uni-bremen.de

Abstract. We study the coding accuracy of a population of stochastically spiking neurons that respond to different features of a stimulus. By using Fisher information as a measure of the encoding error, it can be shown that narrow tuning functions in one of the encoded dimensions increase the coding accuracy for this dimension as long as the active sub-population is large enough. This can be achieved by neurons that are broadly tuned in the other dimensions. If one or more stimulus features encoded by the neural population are unknown, the relative widths of the tuning curves in the remaining dimensions are a measure of the corresponding relative accuracies. This feature allows a quantitative description of the kind of information conveyed by the neural population.

1. Introduction

Most neurons in biological systems respond to many different stimulus features. A so-called visual neuron, for example, may respond to the location, shape, movement pattern, color, and contrast of an object in the visual field. This coding is not uniform in several respects. First, the tuning width of a single neuron is generally not the same for all dimensions representing the stimulus features. Second, neurons in a population differ with respect to the location and the width of their tuning curves. On the other hand, a general feature of biological systems is the existence of physiological classes of neurons with similar tuning properties (e. g., [1, 2]). It is assumed that the classes process different aspects of the incoming stimuli, which leads to the concept of 'pathways' of information flow [3, 2].

However, the question of how to derive statements on the functional significance of a class of neurons from its tuning characteristics is still controversially discussed. A classical approach to this question is that of detector cells [4]. The idea that amphibia make use of 'bug detector' cells in the retina or in the optic tectum, for example, has dominated early studies of the neurobiology of these animals [5]. A characteristic feature of such detectors is their extremely narrow tuning. In recent years, theoretical arguments against the detector concept have been put forward. It was argued that a population of neurons with large receptive fields allows a greater resolution [6]. Thus, there are different

views on the optimal neuronal tuning width. A thorough understanding of this point is crucial for the identification of the function of neuronal pathways on the basis of tuning widths. It is obvious that progress in this field could provide fundamental concepts for understanding functional aspects and coding in neural systems. Recent progress in this field has been achieved by calculating the Fisher information [8] of the neural activity pattern, which serves as a lower bound of the error for estimating the stimulus parameters from the activity [7].

Here we calculate the Fisher information of a population of spiking neurons encoding multiple stimulus features. The consideration of different tuning widths in the different dimensions enables a detailed study of the estimation errors obtained as a function of the population's tuning properties and may solve the question of optimal tuning widths. We also study the empirically encountered situation of 'hidden stimulus dimensions'. The analysis yields a quantitative account on the kind of information conveyed by the neural population.

2. Theoretical Framework

Consider a D-dimensional stimulus vector, $\mathbf{x} = (x_1, \ldots, x_D)$, which is to be encoded by a population of N spiking neurons. All feature values are measured in a fixed system of units, so that the x_i are dimensionless numbers. The mean firing rate of neuron k ($k = 1, \ldots, N$) is determined by its tuning function, $f^{(k)}$, which is assumed to be of the form

$$f^{(k)}(\mathbf{x}) = F \phi \left(\sum_{i=1}^{D} \frac{(x_i - c_i^{(k)})^2}{\sigma_i^2} \right), \tag{1}$$

where $\max_z \phi(z) = 1$, F is the maximum firing rate, and $\mathbf{c}^{(k)}$ is the center of the tuning curve. Note that $f^{(k)}(\mathbf{x})$ is not radially symmetric because there is a separate tuning width σ_i for each stimulus dimension. A good approximation for many measured tuning functions is a Gaussian, $\phi(z) = \exp(-z/2)$. The tuning functions of the population are distributed in the D-dimensional stimulus space according to some density function, $\eta(\mathbf{x})$; here, we shall restrict ourselves to a uniform distribution, $\eta \equiv \text{const}$.

The neurons are assumed to spike stochastically, i.e., measuring firing rates within a finite observation time interval τ does not always yield $\tau f^{(k)}(\mathbf{x})$ spikes. Instead, one gets an $\mathbf{x}$-dependent probability distribution $P(\mathbf{n}; \mathbf{x})$ for the spike count vector $\mathbf{n} = (n_1, \ldots, n_N)$. Here we assume that the neurons fire their action potentials independently, so that $P(\mathbf{n}; \mathbf{x}) = \prod_k P^{(k)}(n^{(k)}; \mathbf{x})$. Note that the input of individual neurons may be correlated (i.e., the tuning curves may overlap), and that only the spike generation mechanisms must be independent. Moreover, the spike count probabilities are assumed to depend only on the local value of the tuning function and on the counting time, $P^{(k)}(n^{(k)}; \mathbf{x}) = S(n^{(k)}, f^{(k)}(\mathbf{x}), \tau)$. An example for a spike generation process that satisfies this assumption is a Poisson process, where $S(n, z, \tau) = (\tau z)^n \exp(-\tau z)/n!$.

In order to assess the accuracy of the representation of the stimulus parameter x_i, we calculate the Fisher information matrix [8]. For a uniform distribution of tuning functions, all off-diagonal elements vanish, while the diagonal elements are given by

$$J_i(\mathbf{x}) := E\left[\left(\frac{\partial}{\partial x_i}\ln P(\mathbf{n};\mathbf{x})\right)^2\right] \qquad (2)$$

[7]. Here, $E[\ldots]$ denotes the expectation value over $P(\mathbf{n};\mathbf{x})$. The relevance of Fisher information for the coding quality is demonstrated by the Cramér-Rao inequality, which states that $1/J_i(\mathbf{x})$ is a lower bound for the mean square error ϵ_i^2 that is made when estimating x_i from the neural activity $\mathbf{n}$,

$$\epsilon_i^2 \geq J_i(\mathbf{x})^{-1}. \qquad (3)$$

This statement is independent of the method of estimation — it applies to all unbiased estimators of the parameter x_i. Thus, a high degree of Fisher information allows a good estimate of x_i, while low Fisher information implies poor coding accuracy.

3.   Information Content of Neural Responses

Since the neurons spike independently, the population Fisher information in the i-th dimension, $J_i(\mathbf{x})$, can be written in terms of single-neuron contributions denoted by $J_i^{(k)}(\mathbf{x})$ for neuron k. These are also given by (2), the only difference being that $P(\mathbf{n};\mathbf{x})$ must be replaced by $P^{(k)}(n^{(k)};\mathbf{x})$. For the neuronal ensemble described above, one finds that

$$J_i^{(k)}(\mathbf{x}) = \frac{1}{\sigma_i^2}A_\phi\left(\xi^{(k)2},F,\tau\right)\xi_i^{(k)2}, \qquad (4)$$

where $\xi_i^{(k)} := (x_i - c_i^{(k)})/\sigma_i$ and $\xi^{(k)2} := \sum_j \xi_j^{(k)2}$ measure the rescaled distance of $\mathbf{x}$ from the tuning curve center, and A_ϕ abbreviates the function

$$A_\phi(z,F,\tau) := 4F^2 \sum_{n=0}^{\infty} S[n,F\phi(z),\tau]\,T[n,F\phi(z),\tau]^2\phi'(z)^2, \qquad (5)$$

with $T(n,z,\tau) := \frac{\partial}{\partial z}\ln S(n,z,\tau)$.

The Fisher information $J_i(\mathbf{x})$ for the whole population is given by the sum of all single-neuron contributions (4). If the centers of the tuning curves are uniformly distributed in stimulus space with density η, it becomes independent of $\mathbf{x}$, and one obtains

$$J_i(\mathbf{x}) \equiv J_i = \frac{\eta\left(\prod_{j=1}^{D}\sigma_j\right)}{\sigma_i^2}K_\phi(F,\tau,D)D, \qquad (6)$$

where K_ϕ denotes the integral ($\xi^2 := \sum_j \xi_j^2$)

$$K_\phi(F, \tau, D) := \frac{1}{D} \int_{-\infty}^{\infty} \mathrm{d}\xi_1 \ldots \int_{-\infty}^{\infty} \mathrm{d}\xi_D \ A_\phi(\xi^2, F, \tau)\xi_1^2. \qquad (7)$$

Equations (3) and (6) yield an estimate of the single-dimension encoding accuracy obtained by the population of neurons. It is valid for the general class of probability distributions $P^{(k)}(n^{(k)}; \mathbf{x})$ described above.

The Fisher information content J_i, and thus the coding accuracy, depend on $D + 1$ main parameters of the neural population: The tuning function density η and the set of tuning widths σ_i. Equation (6) shows that increasing η, which corresponds to increasing the number of neurons N, always improves the code. The dependence on the σ_i is more complex and will be discussed in the next section.

4.   Tuning Width and Coding Accuracy

In the special case that tuning widths are equal, i.e. $\sigma_i = \sigma$, one recovers from (6) the result of Zhang and Sejnowski [7], $J_i \propto \sigma^{D-2}$, which states that small tuning widths are favorable for $D = 1$, while accuracy is independent of σ for $D = 2$, and broad tuning improves the code for $D \geq 2$.

This situation changes dramatically if we consider the more realistic case of different tuning widths for the individual stimulus dimensions, i. Suppose, for example, that only σ_i is varied, while σ_j for $i \neq j$ remain fixed. In this situation, one finds that $J_i \propto \sigma_i^{-1}$ and $J_j \propto \sigma_i$. This implies that decreasing σ_i always improves the encoding of x_i and deteriorates the encoding of x_j for $i \neq j$, no matter how large the overall number of encoded dimensions is. On the other hand, a broad tuning in dimension i can improve the coding of all dimensions $j \neq i$.

This behavior can be understood as follows. The mean error for estimating x_i from the activity of an individual neuron with sufficiently high firing rate will be in the range of the tuning width σ_i, so that this effect leads to an increase of the neuron's Fisher information $J_i^{(k)}(\mathbf{x}) \propto \sigma_i^{-2}$. On the other hand, the tuning width determines how many neurons belong to the population that is actually activated by the stimulus. Roughly speaking, for all dimensions m, large values of $J_m^{(k)}(\mathbf{x})$ are localized in a region around the center of the tuning curve (i.e., if $\xi^{(k)2} < 1$). Therefore, only $N_{\mathrm{code}} \approx \eta \prod_j \sigma_j$ neurons convey the major part of the population's Fisher information on the stimulus, and the active population size will vary $\propto \sigma_i$.

Since the population's Fisher information is the sum of the single-neuron contributions over the activated population, these two effects counteract and determine the behavior of J_i and J_j (where $j \neq i$). Obviously, J_j is only influenced by the latter effect, so that $J_j \propto \sigma_i$. In the case of J_i, the variation of the individual neurons' information outweighs that of the number of encoding neurons N_{code} and one gets $J_i \propto \sigma_i^{-1}$.

Given a fixed density of tuning curves, η, the optimal strategy for accurate encoding of a single dimension i therefore consists of a small tuning width σ_i for this dimension and of large tuning widths in all other encoded dimensions. The narrow tuning in dimension i provides each active neuron with maximal information on the stimulus, while the broad tuning in the other dimensions ensures that the sub-population that is activated by the stimulus is large enough. However, it must be remembered that the validity of (6), and thus of all results presented so far, critically depends on the assumption that the tuning curves are evenly distributed. This assumption becomes inevitably violated if the tuning is so narrow (or the tuning curve density η is so small) that 'gaps' appear between the tuning functions of individual neurons. Such gaps can be shown to lead to a divergence of the spatially averaged mean square error. Thus, there is a limit to improving the encoding quality by decreasing the tuning width.

In most experimental situations, the total number of encoded dimensions D will be unknown. Consider, for example, a neuron that responds to visual stimuli, where it is impossible to specify all D dimensions of the stimulus (color, shape, velocity, location, etc.) or even to estimate D at all. This situation is generally dealt with by probing the response only with respect to $d < D$ known dimensions and to ignore the unknown dimensions $d + 1, \ldots, D$. In this case, the Fisher information (6) can be written as

$$J_i = \frac{1}{\sigma_i^2} \left(\prod_{k=1}^{d} \sigma_k \right) X, \qquad X := \eta \left(\prod_{k=d+1}^{D} \sigma_k \right) K_\phi(F, \tau, D) D \,. \qquad (8)$$

The fact that X is unknown prohibits estimating the absolute quality of the representation via the Cramér-Rao inequality (3). However, X can be eliminated by taking the quotient of the J_i. Denoting the mean square error of an optimal estimator with $\epsilon_{i,\min}^2 := 1/J_i$, this yields

$$\frac{\epsilon_{i,\min}^2}{\sum_{j=1}^{d} \epsilon_{j,\min}^2} = \frac{\sigma_i^2}{\sum_{j=1}^{d} \sigma_j^2} \,. \qquad (9)$$

Thus, the tuning widths of the known dimensions are a measure of relative coding quality. Equation (9) can be used to interpret tuning properties measured in recording experiments. Given the tuning widths of a neural population, it yields a quantitative measure of how accurate the neurons encode each stimulus dimension. This may be used to derive a hypothesis on the function of the population within the neural network. If, for example, (9) yields that a class of neurons encodes stimulus location very accurately, while shape is represented with a considerably larger minimal error, it may be concluded that the class plays a role in the localization of the stimulus rather than in its recognition.

5.   Conclusion

Our analysis of Fisher information contained in the response of a neural population sensitive for D-dimensional stimuli yields an estimate of the encoding ac-

curacy obtained for the different stimulus features. Specifically, accuracy in one dimension can be increased by narrowing the receptive field in the respective dimension while broadening the receptive fields in the remaining dimensions. Relative accuracies are obtained if only part of the dimensions are accessible.

The question which encoding strategy a given neural population adopts remains to be answered empirically. Specifically, it may be possible to identify pathways of information flow in the nervous system [3, 2] by means of a quantitative analysis of the features encoded by a given neural population and the populations with which it interacts. An interesting question is whether part of the information contained in a population can be extracted thereby locally increasing the amount of information in the system. Such a feature would, for example, be advantageous for the purpose of object localization and motor control [9], whereby other stimulus features may be neglected.

References

[1] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts. *Proc. Inst. Radio Eng. NY*, 47:1940, 1959.

[2] G. Roth, U. Dicke, and W. Grunwald. *J. Comp. Neurol.* 404:489.

[3] M. S. Livingstone and D. H. Hubel. *Science*, 240:740, 1988. U. Dicke. *J. Comp. Neurol.* 404:473.

[4] H. B. Barlow. *Perception*, 1:371, 1972.

[5] J.-P. Ewert and W. von Seelen. *Kybernetik*, 14:167, 1974. J.-P. Ewert, H.-W. Borchers, and A. von Wietersheim. *J. Comp. Physiol.*, 126:43, 1978.

[6] G. E. Hinton, J. L. McClelland, and D. E. Rumelhart. In D. E. Rumelhart and J. L. McClelland (eds), *Parallel Distributed Processing, Vol. 1*, pp. 77–109. MIT Press, Cambridge, MA, 1986. H. S. Seung and H. Sompolinsky. *Proc. Natl. Acad. Sci. USA*, 90:10749, 1993. H. P. Snippe and J. J. Koenderink. *Biol. Cybern.*, 66:543, 1992. C. W. Eurich and H. Schwegler. *Biol. Cybern.*, 76:357, 1997.

[7] K. Zhang and T. J. Sejnowski. *Neural Comp.*, 11:75, 1999.

[8] G. Deco and D. Obradovic. *An Information-Theoretic Approach to Neural Computing.* Springer-Verlag, New York, 1996.

[9] E. Salinas and L. F. Abbott. *J. Comp. Neurosci.*, 1:89, 1994. E. Salinas and L. F. Abbott. *J. Neurosci.*, 15:6461, 1995.