

Interpretation of neuronal activity in neural networks

C.C.A.M. Gielen ^{*}, R. Glasius, A. Komoda

*RWCP [†] Novel Functions SNN [‡] Laboratory, Dept. of Medical Physics and Biophysics,
University of Nijmegen, Geert Grooteplein Noord 21, NL 6525 EZ Nijmegen, The Netherlands*

Received 15 November 1994; accepted 2 October 1995

Abstract

In this paper we will compare the performance of two techniques (the Maximum Likelihood Estimation (MLE) and the Population Vector (PV)) for estimating the interpretation of neuronal activity in a population of neurons. Although such a comparison has been made before, so far only homogeneous distributions of receptive fields have been investigated. Since the performance of both methods depends on the distribution of the receptive fields we have tested the performance for homogeneous and inhomogeneous distributions. The results demonstrate that in general the MLE method outperforms the Population Vector. However, the MLE method depends heavily on the shape of the receptive field properties of the neurons, which is not the case for the PV method. Moreover, the MLE method may give rise to artefactual results for inhomogeneous distributions of receptive fields. For the PV method the shape of the receptive field is not as important. Moreover, for the Population Vector the optimal width of the receptive field remains more or less constant when the decrease in density is small relative to the optimal width. In this case the information decreases proportionally with the density of receptive fields.

Keywords: Neural network; Population vector; Maximum likelihood estimation; Receptive field

1. Introduction

Research on neural networks is driven by efforts to understand the information processing in the nervous system and by efforts to use neural networks for various

^{*} Corresponding author. Email: stan@mbfys.kun.nl

[†] Real World Computing Partnership.

[‡] Dutch Foundation of Neural Networks.

types of applications. Therefore, a distinction has to be made between biological neural networks, which are developed and studied in order to simulate, reproduce and understand the function and operation of (a particular part of) the brain, and artificial neural networks, which are developed for a particular application. For the latter, a good performance is of primary interest and a biological plausibility is of minor importance.

One of the main advantages of artificial neural networks over other techniques for complex problem solving is that neural networks can be trained by presenting examples. No explicit algorithm for the solution of the problem is required. When training has led to a neural network with a good performance, the system behaves as a black box which is performing the required input-output mapping. For some applications it is sufficient to have such a system. However, for many applications it would be important to learn about the implicit rules that the neural network has developed during learning in order to understand and to explain the performance of the neural network. Therefore, it would be helpful if the weights in the neural network could be interpreted in terms of rules which can be extracted from the network.

Neurobiologists studying the neuronal information processing are facing a similar problem, since they have to interpret the neuronal activity, which has been recorded from single neurons, and the connectivity between neurons, which becomes evident in a correlated firing of neurons, in terms of macroscopic sensory or motor events. This is implicit in the approach by any neurobiologist, who is trying to interpret the neuronal activity such that he can attribute a function to the neurons under investigation.

Several approaches to these problems have been presented in the past. Bialek et al. [1,2] have analysed the information-theoretic optimal performance and optimal methods for reconstructing information from a neuronal spike train. In their approach a linear filter was constructed to estimate the maximal amount of information in a single action potential. Since it is generally believed that the information is encoded by the activity of an ensemble of neurons, rather than by the activity of a single neuron, several authors have developed methods to interpret the neuronal code of an ensemble of neurons. The most well-known method is the Population Vector, which is simply the summation of the optimal sensory input or motor output of neurons in a population (optimal in the sense that the neuron gives the largest response to that particular sensory stimulus or motor response) weighted by the firing rate of each neuron. Successful illustrations of these methods can be found in Gielen et al. [6] who tried to find a sensory interpretation of neuronal activity in the *nervus acusticus* of the cat, and in Georgopoulos et al. [4,5] who related the activity of motocortical cells to the direction of arm movements. The main problem with the Population Vector method is that it assumes a homogeneous sampling of the sensory or motor space. In the case of motor cortex, the preferred directions of motocortical cells appear to be distributed homogeneously in 3-D space (see Schwartz et al. [13]). However, in the acoustic nerve or in monkey visual cortex (area 17) the receptive fields are certainly not homogeneously distributed in the frequency domain or in the retinal

field of view (see e.g. [3,7]). As a consequence, the Population Vector gives an estimate for the sensory stimulus, which is biased by the inhomogeneous distribution of receptive fields. The same problem arises for other approaches, such as maximum likelihood estimators (see e.g. Seung and Sompolinsky [14]), which frequently make the implicit assumption of a homogeneous sampling.

In this paper we will compare two frequently used methods to provide an interpretation to the neuronal activity in a population of neurons. First we will start with the Maximum Likelihood Estimation and we will investigate the effect of an inhomogeneous distribution of receptive fields on the information coded by the neuronal population. Then we will investigate the conditions under which the Population Vector can give reliable estimates for the interpretation of neuronal activity. The results will be compared with other procedures which have been proposed in the past.

2. Results

In order to compare our results to those obtained by others we will follow the model proposed by Seung and Sompolinsky [14] which is qualitatively similar to the cases discussed by Salinas and Abbott [11]. It consists of a set with N neurons responding to visual stimuli with a particular orientation θ in a range from 0 to 2π . Each neuron i has its own preferred orientation θ_i which elicits the largest responses. The response of the population of the N neurons is described by a probability distribution $P(\mathbf{r}|\theta)$, where the vector \mathbf{r} is used to represent the responses r_1, r_2, \dots, r_N of the N neurons. These responses are assumed to be independent Poisson random variables, i.e. each neuron has its own independent noise source. The expected value of the response to a stimulus θ (or mean response $\langle r_i \rangle$, where $\langle \cdot \rangle$ stands for the average with respect to $P(\mathbf{r}|\theta)$) is represented by $f_i(\theta)$. In Seung and Sompolinsky the distribution $g(\theta)$ of preferred directions θ_i was homogeneous over the interval $[0, 2\pi]$. In this study we will investigate inhomogeneous distributions for $g(\theta)$ and their effect on the information content of the neural responses.

The response of a neuron depends on the preferred direction θ_i of the cell and on the stimulus θ . Like Seung and Sompolinsky, we will assume that all cells have the same response function $f_i(\theta) = f(\theta - \theta_i)$, defined by

$$f(\theta) = \begin{cases} f_{\min} + (f_{\max} - f_{\min}) \cos^m((\pi\theta)/(2a)), & |\theta| < a \\ f_{\min} & \text{otherwise} \end{cases} \quad (1)$$

The exponent m affects the shape of the response curve.

In this context we like to explain that this particular choice for a direction selective response is in no way a restriction. In a similar way direction selective responses of motocortical cells for arm movements in 3-D space (see e.g. Georgopoulos et al. [4,5]) can be modelled and the results of our analysis can be applied equally well to other types of neuronal activity (see e.g. [10,15,16]).

In order to have a measure for the quality of the interpretation of neuronal activity we have used the Fisher Information [9]. If a parameter θ parameterises a signal $S(\theta)$ and when data $\{V_j\}$ from a set Γ_j gives information about S , then the Fisher information J_j with

$$J_j = \int_{\Gamma_j} \left[\frac{\partial \log p(V_j | \theta)}{\partial \theta} \right]^2 p(V_j | \theta) dV_j$$

is a measure of the information supplied by the data V_j about the parameter θ . For a normal distribution, the inverse of J_j is equal to the variance of the probability distribution.

The Shannon Information which is defined as $\log(p(V|S)/(p(V)))$, is an absolute measure of the statistical dependence between V and S , which reduces to zero when V and S are independent. The proper Shannon measure to distinguish between S and $S + \Delta S$ is the average of the difference of information in V about S and $S + \Delta S$. This can be shown to be equivalent to the Fisher Information which can also be interpreted as a measure for the sensitivity of data V_j to small changes in θ (see below). This is the reason why we have chosen the Fisher information as a measure for the information contained in a neuronal response.

2.1. Maximum likelihood estimation

The Maximal Likelihood Estimation (MLE) method is a standard statistical approach for obtaining parameter estimates from experimental data. Under certain conditions it can be shown to be optimal in the sense that it minimizes the mean square error of the estimated parameters. The general problem is to estimate a parameter, say λ , from a set of random (usually experimentally obtained) variables, say x^i , which are a function of the unknown parameter λ . The problem is to find that value for λ , which maximizes the probability that the model has generated the observed data.

In our case the estimated parameter is the orientation θ which most likely corresponds to the stimulus θ_s , which caused the response \mathbf{r} . This is equivalent to estimating the value θ_{est} , which maximizes the probability $P(\theta_{\text{est}} | \mathbf{r}(\theta_s))$.

For a large population of neurons the variance in the estimate for θ which gives the optimal $P(\theta | \mathbf{r})$ is given by $\langle (\hat{\theta} - \theta)^2 \rangle = 1/J[\mathbf{r}](\theta)$, where $J[\mathbf{r}](\theta)$ is the Fisher information defined by

$$J[\mathbf{r}](\theta) = \left\langle \left(\frac{\partial}{\partial \theta} \log P(\theta | \mathbf{r}) \right)^2 \right\rangle$$

which under some general assumptions can be written as

$$J[\mathbf{r}](\theta) = \left\langle - \frac{\partial^2}{\partial \theta^2} \log P(\theta | \mathbf{r}) \right\rangle \quad (2)$$

(see [9]). The Fisher information is a functional of $P(\theta | \mathbf{r})$ and can be interpreted as the amount of information in \mathbf{r} about the stimulus θ .

With the definition of the response function $f_i(\theta_s)$ for the response of neuron i to stimulus θ_s and using the assumption of independence of the r_i such that $J[\mathbf{r}](\theta_s) = \sum_i J[r_i](\theta_s)$, straightforward calculation gives

$$J[\mathbf{r}](\theta_s) = N \int_0^{2\pi} \frac{d\phi}{2\pi} \frac{f'(\theta_s - \phi)^2}{f(\theta_s - \phi)} g(\phi) \quad (3)$$

where $g(\phi)$ represents the density of neurons with preferred orientation ϕ . In Appendix 1 the full expression for $J[\mathbf{r}](\theta_s)$ is given.

Fig. 1 shows the mean total information $J[\mathbf{r}](\theta)/(Nf_{\max})$ for $m = 2$ as a function of parameter a , the ‘width’ of the receptive field of the neurons, for a homogeneous distribution of neurons. Because of the homogeneous distribution of tuning curves or receptive fields the mean total information $J[\mathbf{r}](\theta)/(Nf_{\max})$ is the same for all stimulus orientations θ . This figure is the same as that shown by Seung and Sompolinsky [14]. The mean information increases for smaller a -values because the information is proportional to the square of the derivative of the receptive field. For smaller a -values, the tuning curve becomes tuned more narrowly and as a consequence, the information increases. For very large a -values the receptive field becomes very broad such that the derivative of the tuning curve approaches zero and the mean amount of information becomes zero.

When the receptive fields are not distributed homogeneously over the interval $[-\pi, \pi]$, the results become quite different. Fig. 2 shows the results for a gaussian distribution of receptive fields with a peak at orientation $\theta_s = 0$ and with a

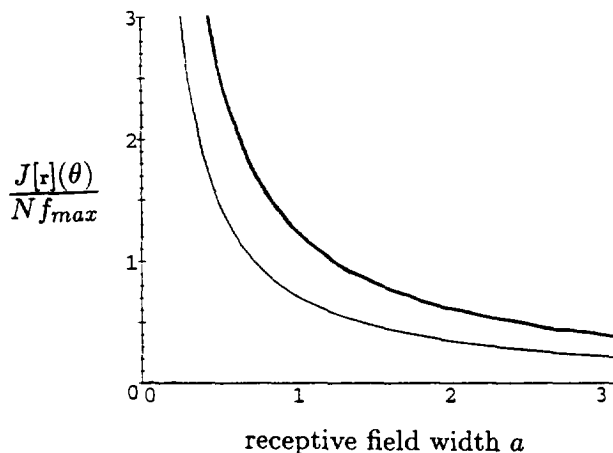


Fig. 1. Mean total information $J[\mathbf{r}](\theta)/(Nf_{\max})$ for the Maximum Likelihood Estimator with a homogeneous distribution of tuning curves as a function of parameter a for $m = 2$ with $f_{\min}/f_{\max} = 0.1$ (thin line) and $f_{\min}/f_{\max} = 0.01$ (thick line).

standard deviation σ . The distribution $g(\phi)$ of preferred directions of the tuning curves over the range $[-\pi, \pi]$ is given by

$$g(\phi) = \frac{\exp\left(\frac{-\phi^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi} \operatorname{erf}\left(\frac{\pi}{\sqrt{2}\sigma}\right)} \quad (4)$$

and is assumed to be periodic modulo 2π . Each panel in Fig. 2 shows the mean total information $J[\mathbf{r}](\theta)$ in the response of the set of N neurons for five stimulus orientations θ_s at 0.1 (A), 0.3 (B), 0.5 (C), 0.75 (D) and 1.5 (E), respectively. Because of the inhomogeneous distribution of receptive fields ($\sigma = 1.0$), the mean total information depends on the particular stimulus orientation θ_s . Notice that the absolute value of the information in Fig. 2 decreases for larger values of θ_s due to the smaller density of neurons for larger θ_s values. For large values of a the mean information approaches zero for each of the orientations θ_s . This is because the slope of the tuning curves becomes less steep for larger values of a . The mean information increases for smaller a values due to the steeper slope of the tuning curve.

Fig. 3 shows results similar to those shown in Fig. 2 for a different distribution of receptive fields ($\sigma = 0.1$). Each of the panels shows a peak in $J[\mathbf{r}](\theta)$ which shifts to larger a -values for larger values of the stimulus orientation θ_s . For very large a -values, the information decreases to zero for reasons explained before. For all θ_s -values the mean information approaches infinity for $a \downarrow 0$, just as for the homogeneous distribution. However, for larger θ_s -values the rapid increase of the mean information can only be seen for very small values of a ($a \leq \exp(-\theta_s^2/2\sigma^2)$). This is related to the fact, that the number of responding neurons decreases for larger values of θ_s due to the lower density of receptive fields. These results demonstrate that the MLE method gives most information for very small receptive fields, irrespective of the stimulus orientation θ_s and the density of receptive fields.

The second peak in the mean total information, which shifts to larger values of a for larger stimulus orientation θ_s , originates from the contribution from neurons with an optimal tuning near the orientation $\theta = 0$ and with a width a such the slope of the tuning curve is steepest near the stimulus θ_s . The latter becomes evident in the fact that the optimal value for a increases with the stimulus orientation θ_s . Obviously, this peak reflects an artefact since the neurons, which give the largest responses have their optimal tuning at $\theta_s = 0$. Moreover, this peak in the information cannot be used in practice, since the optimal width of the receptive field can only be chosen after the result (the best estimate for the stimulus orientation θ_s) is known. Moreover, in our case with a symmetric density around orientation zero, the code is ambiguous: The same neurons tuned near $\theta = 0$ will be active for both θ_s as well as for $-\theta_s$, and the information contained in the neuronal response is the same for θ_s and $-\theta_s$.

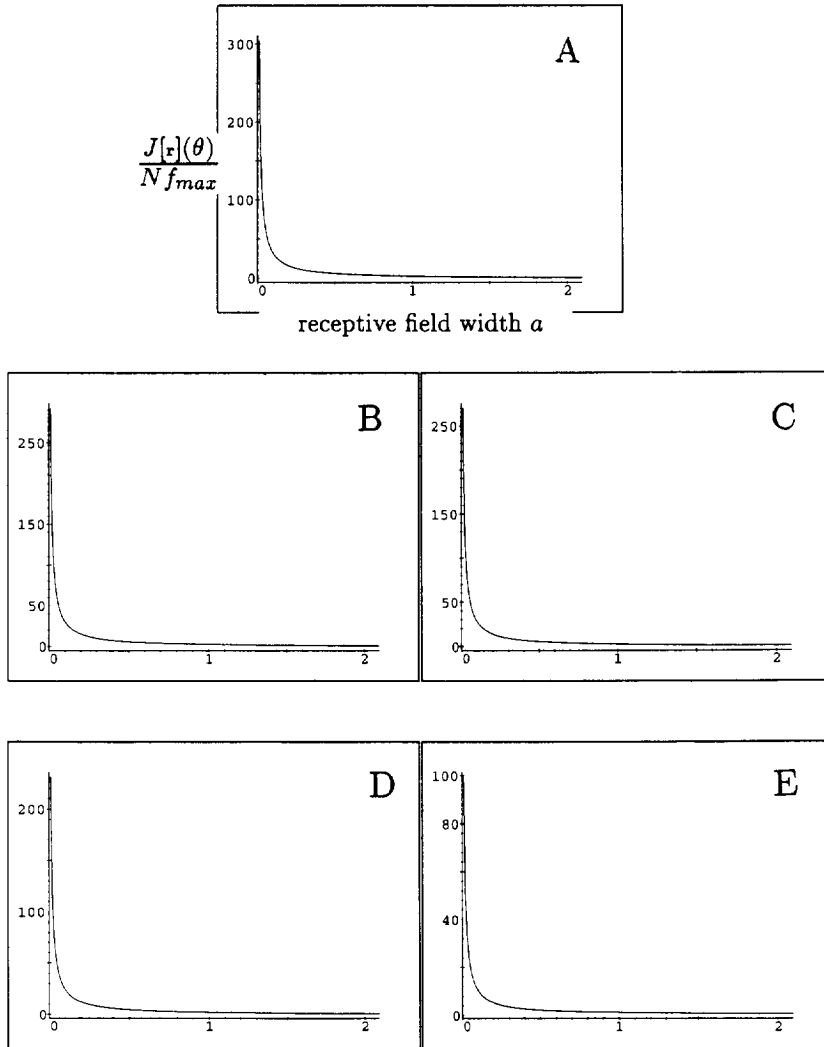


Fig. 2. Mean total information $J[\mathbf{r}](\theta)/(Nf_{\max})$ as a function of parameter a for five orientations ($\theta_s = 0.1$ (A), 0.3 (B), 0.5 (C), 0.75 (D) and 1.5 (E) radians). The distribution of tuning curves is given by

$$g(\phi) = \frac{\exp\left(-\frac{\phi^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2} \operatorname{erf}\left(\frac{\pi}{\sqrt{2\sigma^2}}\right)}$$

with $\sigma = 1.0$. The ratio $f_{\min}/f_{\max} = 0.01$; $m = 2$.

2.2. Population Vector

A biologically plausible alternative to MLE is the Population Vector. Following Seung and Sompolinsky a more convenient representation of the stimulus θ is the

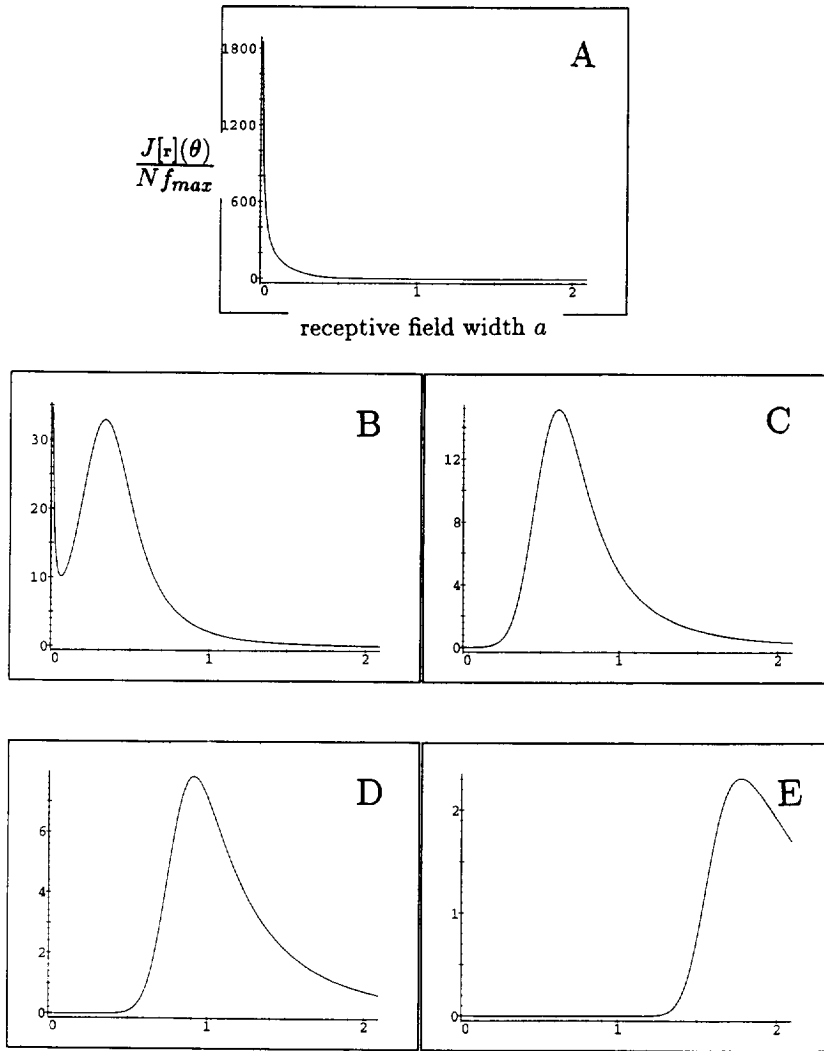


Fig. 3. Mean total information $J[r](\theta)/(Nf_{max})$ as a function of parameter a for five orientations ($\theta_s = 0.1$ (A), 0.3 (B), 0.5 (C), 0.75 (D) and 1.5 (E) radians). The distribution of tuning curves is given by

$$g(\phi) = \frac{\exp\left(-\frac{\phi^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2} \operatorname{erf}\left(\frac{\pi}{\sqrt{2\sigma^2}}\right)}$$

with $\sigma = 0.1$. The ratio $f_{min}/f_{max} = 0.01$; $m = 2$.

complex representation. When using the complex equivalent $z = e^{i\theta}$, the population vector \hat{z} is given by

$$\hat{z} = \frac{1}{N|\hat{f}_1|} \sum_{k=1}^N r_k e^{i\theta_k} \tag{5}$$

where \hat{f}_1 is the first Fourier component of the tuning curve f , with the n -th Fourier component defined by $\hat{f}_n = 1/(2\pi) \int_0^{2\pi} e^{in\theta} f(\theta) d\theta$. Since the tuning curve $f(\theta)$ is symmetric, its first Fourier component \hat{f}_1 is real and positive. As shown by Seung and Sompolinsky, the mean information $J[\hat{z}]$ for a homogeneous distribution is given by

$$J[\hat{z}] = N \frac{2\hat{f}_1^2}{\hat{f}_0 - \hat{f}_2} \quad (6)$$

(see also Appendix 2 where we derive a measure for the variance in the estimate of the Population Vector, which is also equal to or larger than the inverse of the mean total information contained in the responses.) This equation says that the information is proportional to the square of the ratio of the population-averaged signal (\hat{f}_1) to the population-average noise ($\sqrt{\hat{f}_0 - \hat{f}_2}$). As pointed out by Seung and Sompolinsky, this means that the information contained in the population vector does not critically depend on the shape of the tuning curve, contrary to the information contained in the maximum likelihood estimator.

Fig. 4 shows the mean information

$$\frac{J[\hat{z}](\theta)}{Nf_{max}}$$

as a function of the width a of the tuning curve for two ratio's of f_{min}/f_{max} . The mean information approaches zero for large a values. This is because for large a values all neurons will respond to the stimulus θ_s with approximately the same response amplitude. Therefore, all neurons contribute approximately equally which,

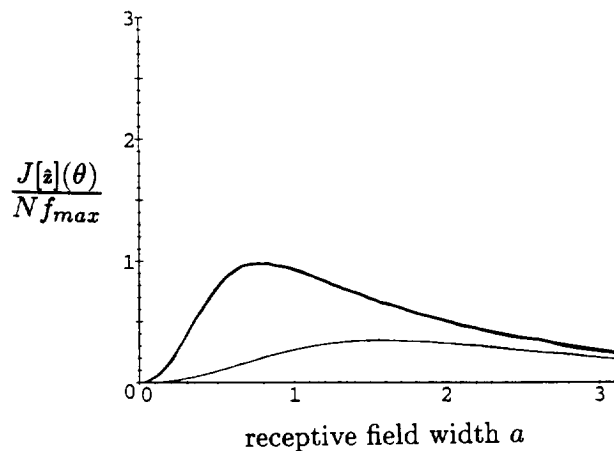


Fig. 4. Mean total information $J[\hat{z}](\theta)/(Nf_{max})$ for the population vector model with a homogeneous distribution of tuning curves as a function of parameter a for $m = 2$ with $f_{min}/f_{max} = 0.1$ (thin line) and $f_{min}/f_{max} = 0.01$ (thick line).

for a homogeneous distribution, gives rise to a cancellation of the contribution of the neurons to the Population Vector. As a consequence, the Population Vector becomes small or zero and does not contain information any more. For very small

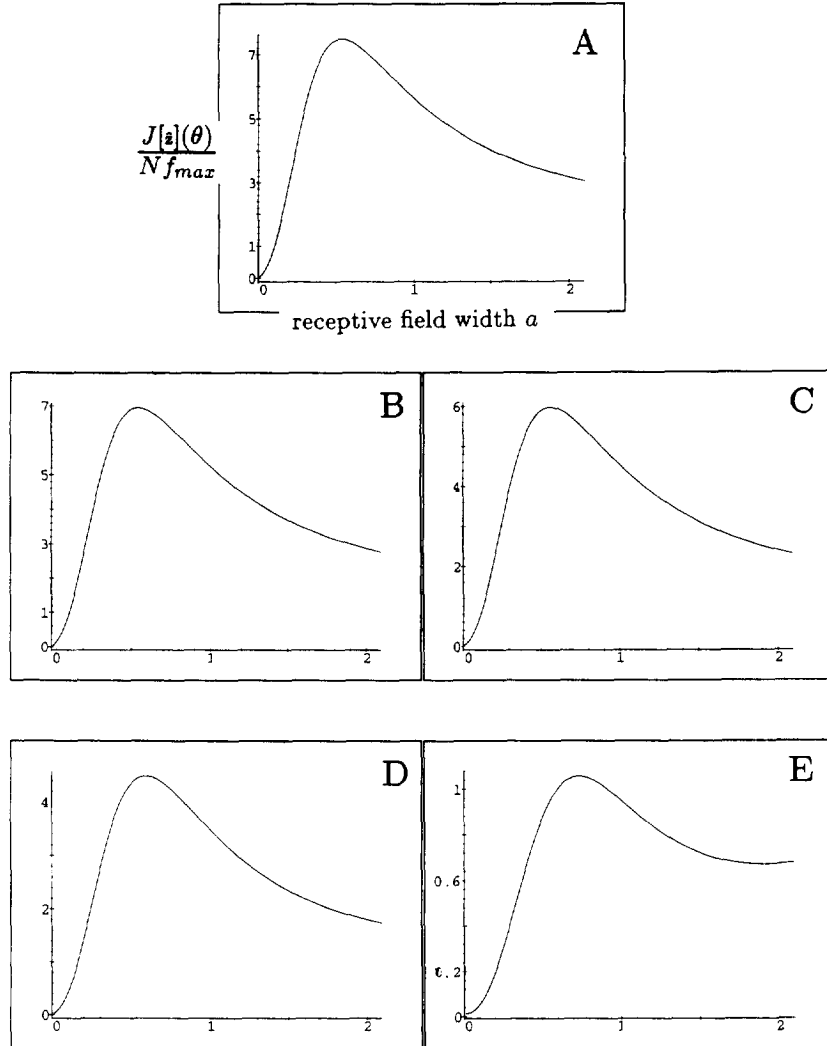


Fig. 5. Mean total information $J[\hat{z}](\theta)/(Nf_{max})$ as a function of parameter a for five orientations ($\theta_s = 0.1$ (A), 0.3 (B), 0.5 (C), 0.75 (D) and 1.5 (E) radians). The distribution of tuning curves is given by

$$g(\phi) = \frac{\exp\left(-\frac{\phi^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2} \operatorname{erf}\left(\frac{\pi}{\sqrt{2\sigma^2}}\right)}$$

with $\sigma = 1.0$. The ratio $f_{min}/f_{max} = 0.01$; $m = 2$.

a values, the neurons have a narrowly tuned response function. As a consequence only very few neurons will respond to a stimulus θ_s and the information will decrease due to the poor signal-to-noise ratio for smaller a values.

The results for an inhomogeneous distribution are shown in Figs. 5 and 6. In Fig. 5 the width of the Gaussian distribution of tuning curves ($\sigma = 1.0$) is relatively broad with respect to the stimulus interval $[-\pi, \pi]$. Therefore, this distribution is not too much different from a homogeneous distribution for small θ_s -values. The peak value of the mean information decreases for larger stimulus orientations due to the lower density at orientations farther away from $\theta_s = 0$. Otherwise the traces for the information have almost the same shape, with the optimal value for a near 0.7 for all θ_s .

For a distribution deviating more from a homogeneous one ($\sigma = 0.1$; see Fig. 6) the peak value for the information decreases rapidly with stimulus orientation θ_s . For $\theta_s = 0.1$ (i.e. near the highest density) the information is large and the optimal

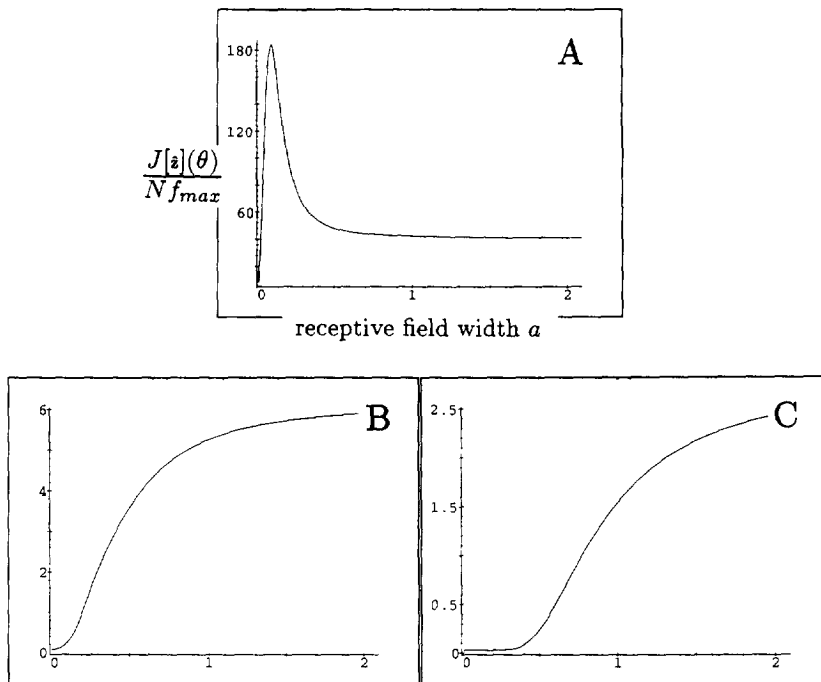


Fig. 6. Mean total information $J[\hat{z}](\theta)/(Nf_{max})$ as a function of parameter a for three orientations ($\theta_s = 0.1$ (A), 0.3 (B), and 0.5 (C) radians). The distribution of tuning curves is given by

$$g(\phi) = \frac{\exp\left(-\frac{\phi^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2} \operatorname{erf}\left(\frac{\pi}{\sqrt{2\sigma^2}}\right)}$$

with $\sigma = 0.1$. The ratio $f_{min}/f_{max} = 0.01$; $m = 2$.

value for a is near 0.1. However, for $\theta_s = 0.3$, where the density of tuning curves has decreased by more than 95%, larger values for a give rise to a larger amount of information since more neurons are responding giving rise to a better signal-to-noise ratio. The increase of information with a in Fig. 6(b) and 6(c) changes to a decrease for large values of a (outside figure) for reasons explained before. The value for a where the information reaches its maximum depends on the particular value of θ_s , since it depends on a compromise between a reduction in noise due to a larger number of responding neurons for larger values of a and a decreasing value of the Population Vector due to cancellation of contributions from neurons with opposite preferred directions.

3. Discussion

In agreement with a previous study by Seung and Sompolinsky we have found, that the MLE gives a better performance than the Population Vector for small tuning curves (i.e. small a values). Although MLE always outperforms the Population Vector, the performance of the two methods approaches each other for large values of a . The fact that MLE outperforms the Population Vector estimate can be explained by the fact that the information for the Population Vector is based on the average response, whereas for MLE the mean information is the average of the information contained in each of the responses. Therefore, the latter must always be larger than the information contained in the Population Vector. However, the Population Vector has the advantage above the Maximum Likelihood Estimation that the former method is rather insensitive to the precise shape of the tuning curve. This is in agreement with results of Sanger [12] who demonstrated in a very different way, that the fact that the Population Vector produces good estimates does not imply that the tuning curves have a cosine function shape. Actually they can have any shape as long as the higher order Fourier components of the tuning curves for the neurons are uncorrelated. In this study we extend these results for inhomogeneous distributions of tuning curves, by showing that the information for the MLE is mainly determined by the steepness of the slope of the tuning curves. For low densities, this gives rise to artefactual results in that large receptive fields are favoured such that the main contribution comes from neurons with a optimal tuning near the highest density.

For the Population Vector the optimal value for a does not change very much as long as the change in the distribution of tuning curves is small relative to the width of the tuning curve a . This is expressed by the relation

$$\left| \frac{\partial g(x)}{\partial x} \cdot 2a \right| \leq 1.$$

For the gaussian distribution with $\sigma = 1.0$ the gradient

$$\left| \frac{\partial g(x)}{\partial x} \right|$$

is always smaller than 1.1, which indeed is smaller than twice the optimal value for a . Only the amount of information decreases proportionally with the density of tuning curves.

In addition to the Maximum Likelihood Estimator and the Population Vector several other procedures have been proposed to estimate probabilities of different stimuli given the neuronal response. Salinas and Abbott [11] proposed an ‘optimal linear estimator’ (OLE), which is a linear estimator which minimizes the squared difference between the estimated and actual stimulus, averaged over all trials and over all possible stimuli. This method is related to the Population Vector. It has a correction factor to correct for the correlations in firing rate between all pairs of neurons. In their paper, Salinas and Abbott demonstrate that the OLE method corrects for any artefacts due to nonuniform distributions of neuronal tuning curves. Also, the OLE method requires fewer neuronal responses than the population vector to reach the same accuracy for the predicted stimulus. This makes the OLE method another useful procedure to interpret neuronal activity. A quantitative comparison between the OLE method and the procedures outlined in this manuscript cannot be made yet with the available material.

In another approach Kjaer et al. [8] studied the encoding of spatial pattern information by complex cells in the primary visual cortex of awake monkeys. A simple feedforward network with one hidden layer appeared to give the best performance in terms of the error between estimated and actual stimulus components. A comparison of the accuracy of this method with the other techniques has not yet been done. Although this approach may be useful for artificial neural networks, it has no relevance for biological neural networks because training the feedforward network requires supervised learning.

Both in the study by Seung and Sompolinsky and in this study we find that the information for the MLE approaches infinity for small values of a whereas it becomes zero for the Population Vector. This raises the question, whether the two measures for information are really the same. The explanation is, that the information for MLE (see Eq. (3)) does not depend on the stochastic response parameter r , but on the average value of the response $f(\theta_s)$. Since the number of responding neurons increases proportionally with a and since the information decreases faster than $1/a$, the increasing amount of information related to the steeper slope for small a values dominates over the decreasing amount of information by the smaller number of responding neurons for smaller values of a . For the population vector, however, the information is contained in the average response. Therefore, small values are not favourable since too few neurons are responding, whereas large values for a are not favourable either since for large values of a neurons with opposite preferred directions become active such that the Population Vector decreases for large values of a . Obviously, the optimal value for a is somewhere in between.

It is important to notice, that the Fisher information J only provides a lower bound for the variance of the estimate, which is approached only for a large number of neurons. It is not known yet, whether the convergence of the variance

towards $1/J$ for a total number of neurons N is equally fast for both methods. Further simulation studies have to be done to clarify this point.

The lessons of this study for the neurophysiologist are twofold. First, the Maximum Likelihood Estimator may perform better than the Population Vector. However, this is true only for very small receptive field sizes, which requires many neurons in order to code the complete sensory domain. The latter is very expensive in terms of neural ‘hardware’. Secondly, the results of this study show that for normal receptive fields the performance of the MLE method is comparable to that of the PV method. The advantage of the Population Vector is that it is less sensitive to the receptive field size. Especially for an adaptive system, this may be an advantage. For the Population Vector the size and shape of the receptive field are not important as long as the width is small relative to the scale in which the distribution density varies.

Appendix 1. Maximum likelihood estimation

Given the model described in the RESULTS section, the information in the response of neuron i to a stimulus with orientation ϕ according to the Fisher Information is given by

$$J[r_i](\phi) = \begin{cases} (f_{\max} - f_{\min})^2 \frac{m^2 \cos^{2m-2}\left(\frac{\pi}{2a}\phi\right) \sin^2\left(\frac{\pi}{2a}\phi\right) \frac{\pi^2}{4a^2}}{f_{\min} + (f_{\max} - f_{\min}) \cos^m\left(\frac{\pi}{2a}\phi\right)}, & |\phi| < a \\ 0 & \text{otherwise} \end{cases}$$

For a homogeneous distribution of tuning curves the total mean information $J[r]$ for the whole population is given by

$$J[r] = N \int_{-\pi}^{\pi} J[r_i](\phi) \frac{d\phi}{2\pi} = \frac{N}{2\pi} \int_{-a}^a (f_{\max} - f_{\min})^2 \frac{m^2 \cos^{2m-2}\left(\frac{\pi}{2a}\phi\right) \sin^2\left(\frac{\pi}{2a}\phi\right) \frac{\pi^2}{4a^2}}{f_{\min} + (f_{\max} - f_{\min}) \cos^m\left(\frac{\pi}{2a}\phi\right)} d\phi \quad (7)$$

For a normalized inhomogeneous distribution $g(\phi)$, the mean total information, when a stimulus θ_s is presented, is given by

$$J[r](\theta_s) = N \int_{-\pi}^{\pi} J[r_i](\theta_s - \phi) g(\phi) d\phi \quad (8)$$

Appendix 2. Population vector

For a homogeneous distribution of tuning curves ($\theta_i = 2\pi i/N$) the estimate according to the population vector is given by

$$\langle \hat{z} \rangle = \frac{e^{i\theta_s}}{N |\hat{f}_1|} \int_0^{2\pi} f(\phi) e^{i\phi} \frac{d\phi}{2\pi} = (\hat{f}_1 / |\hat{f}_1|) e^{i\theta_s} \quad (9)$$

with $\hat{f}_1 = \frac{1}{2\pi} \int_0^{2\pi} e^{i\theta} f(\theta) d\theta$. The covariance matrix

$$C = \begin{pmatrix} \langle (x - \bar{x})^2 \rangle & \langle (x - \bar{x})(y - \bar{y}) \rangle \\ \langle (x - \bar{x})(y - \bar{y}) \rangle & \langle (y - \bar{y})^2 \rangle \end{pmatrix}$$

is given by

$$\frac{1}{2N\hat{f}_1^2} \begin{pmatrix} \hat{f}_0 + \hat{f}_2 \cos(2\theta_s) & \hat{f}_2 \sin(2\theta_s) \\ \hat{f}_2 \sin(2\theta_s) & \hat{f}_0 - \hat{f}_2 \cos(2\theta_s) \end{pmatrix}$$

From

$$\sigma_\theta^2 \approx \left(\frac{\partial \theta(x, y)}{\partial x} \right)^2 \sigma_x^2 + 2 \left(\frac{\partial \theta(x, y)}{\partial x} \right) \left(\frac{\partial \theta(x, y)}{\partial y} \right) \sigma_{xy}^2 + \left(\frac{\partial \theta(x, y)}{\partial y} \right)^2 \sigma_y^2 \quad (10)$$

we find

$$J[\mathbf{r}](\theta_s) = \frac{1}{\sigma_\theta^2} \approx \frac{2N\hat{f}_1^2}{\hat{f}_0 - \hat{f}_2} \quad (11)$$

For an inhomogeneous distribution of tuning curves the population vector for a stimulus θ_s is given by the weighted sum

$$\begin{aligned} \langle \bar{z}(\theta_s) \rangle &= \left\langle \frac{1}{N} \frac{1}{\bar{n}} \sum_{k=1}^N r_k(\theta_s) e^{i\theta_k} \right\rangle \\ &= \frac{f(\theta_s - \theta_k) = f(\theta_k - \theta_s)}{N} \frac{1}{\bar{n}} \sum_{k=1}^N f(\theta_k - \theta_s) e^{i\theta_k} \\ &\stackrel{N \rightarrow \infty}{=} \frac{1}{\bar{n}} \int_{\theta_1}^{\theta_N} f(\theta - \theta_s) e^{i\theta} g(\theta) d\theta \\ &= \frac{e^{i\theta_s}}{\bar{n}} \int_{-\pi - \theta_s}^{\pi - \theta_s} f(\phi) e^{i\phi} g(\phi + \theta_s) d\phi \end{aligned} \quad (12)$$

where \bar{n} is a normalization factor.

Using the relations $x = \text{Re } \hat{z}$ and $y = \text{Im } \hat{z}$ the variance in x is given by

$$\begin{aligned} & \langle (x - \bar{x})^2 \rangle \\ &= \left\langle \left(\frac{\sum_{k=1}^N r_k (R \cos(\theta_k) + I \sin(\theta_k))}{N |\hat{f}_1|^2} - \frac{\sum_{k=1}^N f_k (R \cos(\theta_k) + I \sin(\theta_k))}{N |\hat{f}_1|^2} \right)^2 \right\rangle \\ &= \frac{1}{N^2 |\hat{f}_1|^4} \sum_{k=1}^N f_k (R^2 \cos^2(\theta_k) + I^2 \sin^2(\theta_k) + 2RI \cos(\theta_k) \sin(\theta_k)) \end{aligned}$$

using the shorthand notation

$$R \equiv \text{Re}(\hat{f}_1)$$

$$I \equiv \text{Im}(\hat{f}_1)$$

$$f_k \equiv f(\theta_s - \theta_k)$$

Similarly the variation in y is given by

$$\begin{aligned} & \langle (y - \bar{y})^2 \rangle \\ &= \frac{1}{N^2 |\hat{f}_1|^4} \sum_{k=1}^N f_k (R^2 \sin^2(\theta_k) + I^2 \cos^2(\theta_k) - 2RI \cos(\theta_k) \sin(\theta_k)) \end{aligned}$$

and the variation in xy is given by

$$\begin{aligned} & \langle (x - \bar{x})(y - \bar{y}) \rangle \\ &= \frac{1}{N^2 |\hat{f}_1|^4} \sum_{k=1}^N f_k ((R^2 - I^2) \cos(\theta_k) \sin(\theta_k) - RI(\cos^2(\theta_k) - \sin^2(\theta_k))) \end{aligned}$$

Using Eq. (10) for the variance in θ we find

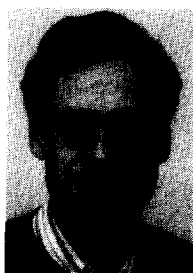
$$\begin{aligned} & \sigma_\theta^2 \\ & \approx \frac{1}{N |\hat{f}_1|^4} \left(\frac{I^2 - R^2}{2} \text{Re}(\hat{f}_2) - IR \text{Im}(\hat{f}_2) + \frac{R^2 + I^2}{2} \hat{f}_0 \right) \end{aligned} \quad (13)$$

For a homogeneous distribution the imaginary part I is zero, in which case Eq. (13) reduces to the equation for the homogeneous distribution (see Eq. (11)).

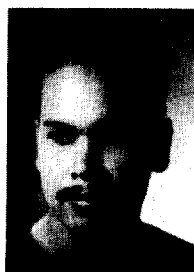
References

- [1] W. Bialek, Theoretical physics meets experimental neurobiology, in: Jen E (ed) *Lectures in Complex Systems, SFI Studies in the Science of Complexity, Vol. 2*. (Addison-Wesley, Redwood City CA, 1989) pp. 413–595.

- [2] W. Bialek, F. Rieke, R.R. de Ruyter van Steveninck and D. Warland, Reading a neuron code *Science* 252 (1991) 1854–1857.
- [3] P.M. Daniel and D. Whitteridge, The representation of the visual field on the cerebral cortex in monkeys, *J. Physiol. (Lond.)* 159 (1961) 203–221.
- [4] A.P. Georgopoulos, A. Schwartz and R.E. Kettner, Neuronal population coding of movement direction, *Science* 233 (1986) 1416–1419.
- [5] A.P. Georgopoulos, R.E. Kettner and A. Schwartz, Primate motor cortex and free arm movements to visual targets in three-dimensional space: II. Coding of the direction of movement by a neuronal population, *Neurosci.* 8 (1988) 2928–2937.
- [6] C.C.A.M. Gielen, G.H.F.M. Hesselmann and P.I.M. Johannesma, Sensory interpretation of neural activity patterns, *Math. Biosciences* 88 (1988) 15–35.
- [7] N.Y.-S. Kiang, *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA, 1965).
- [8] T.W. Kjaer, J.A. Hertz and B.J. Richmond, Decoding cortical neuronal signals: network models, information estimation and spatial tuning, *J. Computational Neurosci.* 1 (1994) 109–139.
- [9] S. Kullback, *Information Theory and Statistics* (Wiley & Sons, New York, N.Y., 1959)
- [10] S.R. Lehky and T.J. Sejnowski, Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity, *J. Neurosci.* 10 (1990) 2281–2299.
- [11] E. Salinas and L.F. Abbott, Vector reconstruction from firing rates, *J. Comp. Neurosci.* 1 (1994) 89–107.
- [12] T.D. Sanger, Theoretical considerations for the analysis of population coding in motor cortex, *Neural Computation* 6 (1993) 29–37.
- [13] A.B. Schwartz, R.E. Kettner and A.P. Georgopoulos, Primate motor cortex and free arm movements to visual targets in three-dimensional space: I. Relations between single cell discharge and direction of movement, *J. Neuroscience* 8 (1988) 2913–2927.
- [14] H.S. Seung and H. Sompolinsky, Simple models for reading neuronal population codes, *Proc. Natl. Acad. Sci USA* 90 (1993) 10749–10753.
- [15] F. Theunissen and J.P. Miller Representation of sensory information in the cricket cercal sensory system: II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons, *J. Neurophysiol.* 66 (1991) 1690–1703.
- [16] J.A.M. van Gisbergen, A.J. van Opstal and A.M.M. Tax, Collicular ensemble coding of saccades based on vector summation, *Neuroscience* 21 (1987) 541–555.



Stan Gielen received a B.S. degree in Physics, Mathematics and Chemistry in 1973 at the University of Nijmegen. A Masters Degree in Experimental Physics and a Ph.D. were obtained at the University of Nijmegen in 1976 and 1980, respectively. In the period from 1980 to 1988 he worked at the University of Utrecht (The Netherlands) and at Northwestern University, Evanston, USA. Since 1988 he is a full professor in biophysics. His main research activities are in the coordination of human limb movements, in neurobiology of neural networks, learning in neural networks, relation between AI and neural networks, and applications of neural networks. He is the head of a research group of about 25 people in the Dept. of Medical Physics and Biophysics. In addition he is director of the Dutch Foundation for Neural Networks, which coordinates research activities on Neural Networks in The Netherlands by initiating industry funded research activities in collaboration with several dutch universities.



Roy Glasius received a Masters Degree in Physics at the University of Nijmegen in 1989. Since then he is involved as a Ph.D. student in research on Neural Networks for trajectory planning, obstacle avoidance and on the interpretation of neuronal activity.

Andrzej Komoda received a M.Sc. and Ph.D in theoretical physics of random magnetic systems at the Wroclaw University (Poland) in 1979 and 1984 respectively. In 1987 the Academic Authority of the Katholieke Universiteit Leuven (Belgium) awarded him a postdoctoral scholarship where he worked on random and self-avoiding walks. In October 1988 he started to work as a research physicist at the 'Limburgs Universitair Centrum' in Diepenbeek (Belgium) on associative memory and pattern recognition using models of neural networks. Since October 1991 he is working in the Department of Medical Physics and Biophysics/Foundation of Neural Networks in Nijmegen. His main interest are in the theory of spin glass systems and in the theory of information processing in biological and artificial neural networks.