Tuning properties of noisy cells with application to orientation selectivity in rat visual cortex

Axel Etzold, Christian W. Eurich, and Helmut Schwegler

Institut für Theoretische Physik,
Universität Bremen
Postbox 330 440, D-28334 Bremen, Germany

Keywords

orientation tuning, stochastic neural responses, visual cortex

Abstract

Common measures for the tuning of cells that are used in the neuroscience literature break down even in the case of moderately noisy neurons. For this reason, a considerable proportion of recorded neuronal data remains unconsidered. One reason for the unreliability of tuning measures is that least-squares fitting of a function for the tuning curve is likely to give too much influence to outliers. We present an algorithm using a rank-weighted norm to construct a tuning curve which weighs outlying data less strongly. As a model function for the tuning curve, we take a trigonometric polynomial, whose coefficients can be determined using a linear approximation. This approach avoids the occurrence of multiple local minima in the optimization process. A test criterion is given to answer the question whether a trigonometric polynomial of lower degree can account for the data. Throughout, we apply our findings to our own experimental data recorded from a population of neurons from area 17 of the rat.

1 Introduction

Since the discovery of orientation selectivity in visual cortical neurons by Hubel and Wiesel [4], experiments to determine orientation tuning curves have been made in many animals. In experimental studies, different measures are used to assess the tuning of cells. As a result, the number of cells from a particular area classified as selective to a particular stimulus can differ greatly from one study to another. We applied some commonly used measures to our own electrophysiological recordings from area 17 of the rat and found that they pose problems when applied to the noisy responses we obtained. Here, we address two problems associated with these measures: their sensitivity to outliers in the neural responses and difficulties in finding minima in the optimization due to particular classes of tuning functions.

Sections 2, 3 and 4 treat index methods for the assessment of tuning, problems related to least-squares fits and to fits using particular classes of functions, respectively. In chapter 5, we give an algorithm to construct tuning curves avoiding the problems addressed before. Fits based on this algorithm are discussed in the last section.

2 Index methods for the assessment of orientation tuning

In experiments, it is only possible to measure the tuning of cells at a limited number of points. Some researchers use an index method to describe the orientation selectivity of cells [1]:

$$\mathrm{ORI} = \frac{\text{max resp} - \text{orth resp}}{\text{max resp}}, \tag{1}$$

where 'max resp' and 'orth resp' are responses to the optimal orientation and to the orientation orthogonal to the optimal one, respectively. Cells are classified as orientation selective if ORI exceeds a pre-defined threshold [2]. Yet, the cell's true preferred orientation might not be among the measured points and the minimal response need not lie orthogonally to the optimal response, e.g. in the case of additional directional tuning. If the firing of the neuron is noisy, values used to calculate ORI obtained from averaging over a limited number of trials still have considerable jitter. This means that often, the classification of cells may be determined because of noise present at the time of the recording rather than due to the actual stimulus. The index method does not indicate how secure the classification of a cell with respect to its sensitivity is. We suggest that this is a reason why many papers using index methods give different results in the proportion of cells from a population which they classify as sensitive to a particular stimulus.

3 Breakdown of least-squares fits

Generally, cells respond differently to repeated presentations of the same stimulus and typically, some of these answers differ greatly from the rest. Whereas an expert biologist would reject these outliers, least squares estimation gives them great importance in the fit of the model function. A single data point lying sufficiently far apart can suffice to significantly change an estimation based on the least squares method, even if the estimation is based on many data. This difficulty can be addressed by using a different norm, which we a going to present in chapter 5.

4 Problems related to the choice of a model function for the tuning curve

Tuning curves are often bell-shaped and very roughly symmetrical. Although it may be useful in a special case, the choice of a particular function as a model for a tuning curve is somehow arbitrary and needs to be done anew in every case. For example , in contrast to our data obtained from the rat, the data obtained in [7] from the cat have very little noise and the variant of a von Mises function [6] used there [7] fits the data well. If the choice of a particular function is not appropriate, the fitted tuning curve can lie far from the measured data, and special features of the tuning like asymmetry or multiple maxima might not be reflected by the model function.

Due to the structure of the modified von Mises function, an estimation of its parameters by solving normal equations is not possible. Applying algorithms that use stochastic elements (e.g. simulated annealing) for the fit instead cannot guarantee that the minimum found by the algorithm is actually a global minimum. Further, the noisier the data are, the more likely they are to induce several fitting parameter sets giving different tuning curves. We address this problem in the following chapter.

5 An algorithm for fitting tuning curves using trigonometric polynomials

In the first step, we specify a class of functions for the tuning curve whose parameters can be estimated by an approximation in a linear subspace. One such class of functions is a trigonometric polynomial.

For an experiment with c different orientations, it is given by

$$a_0 + \sum_{k=1}^{d} (a_k \cos k\theta + b_k \sin k\theta)$$

[1,5], where $d = c/2$ if c is even and $d = (c-1)/2$ if c is odd.

Let Y_{ij} be the responsee of a neuron in the i-th trial $i \in \{1,..,c\}$ upon the presentation of a stimulus with orientation $j \in \{1,..,r\}$. Y_{ij} can be written in the following form:

$$Y_{ij} = a_0 + \sum_{k=1}^{c} (a_k \cos k\theta_i + b_k \sin k\theta_i) + \varepsilon_{ij} \tag{2}$$

We write

$$\begin{aligned} \mathbf{Y} &= (Y_{11} \ \cdots \ Y_{1r} \ Y_{21} \ \cdots \ Y_{2r} \ \cdots \ \cdots \ Y_{c1} \ \cdots \ Y_{cr})^T, \\ \boldsymbol{\beta} &= (a_0 \ a_1 \ \cdots \ a_d \ b_1 \ \cdots \ b_d)^T = (a_0 \ \mathbf{a} \ \mathbf{b})^T, \\ \boldsymbol{\varepsilon} &= (\varepsilon_{11} \ \cdots \ \varepsilon_{1r} \ \varepsilon_{21} \ \cdots \ \varepsilon_{2r} \ \cdots \ \cdots \ \varepsilon_{c1} \ \cdots \ \varepsilon_{cr})^T \end{aligned}$$

and define a matrix $\mathbf{X}$ in such a way that the problem of fitting a tuning curve is equivalent to finding the parameter vector $\boldsymbol{\beta}$ in the equation $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, which minimizes $||\boldsymbol{Y} - \boldsymbol{X\beta}||$. Thus, the residual vector $\boldsymbol{\varepsilon}$ which accounts for the stochasticity of the data not explained by the parameter $\boldsymbol{\beta}$ is minimized in norm. To reduce the impact of single outliers, we construct a different norm in the following way. We rank the entities $|\varepsilon_{ij}| := |Y_{ij} - \sum_{j=0}^{2d+1} X_{ij}\beta_j|$ in a descending order. If two or more of these numbers are equal, we assign the average of their ranks to all of them. This rank is denoted by $R(|\varepsilon_{ij}|)$. Now we can use

$$\|\boldsymbol{\varepsilon}\|_{w1} = \sum_{i=1,..,c \ j=1,..r} R(|\varepsilon_{ij}|)|\varepsilon_{ij}|. \tag{3}$$

as a norm underweighing outliers. A proof that this function is a norm can be found in [8]. Based on this norm, the following algorithm for estimating $\boldsymbol{\beta}$ can be applied:

1. Set $k = 0$. Obtain an initial estimate $\hat{\boldsymbol{\beta}}^{(\mathbf{0})}$. This can be done by making a singular value decomposition of $\mathbf{X}$ [9]. From this, get an estimate of the variance $\hat{\sigma}^{(0)}$ by estimating the variance of $\varepsilon_{ij}^{(0)} := Y_{ij} - \sum_{j=0}^{2d+1} X_{ij}\hat{\boldsymbol{\beta}}_j^{(0)}$.

   In a least-squares analysis, the algorithm would end here, but note that we use a different norm which weighs distances differently. The first step yields the residuals $\boldsymbol{\varepsilon} = \hat{\boldsymbol{\varepsilon}}^{(0)}$.

2. Based on the singular value decomposition of $\mathbf{X} = \mathbf{UWV^T}$ performed in step 1, compute the matrix $\mathbf{H} = \mathbf{UW}$ which spans the column space of $\mathbf{X}$.

3. Let $\mathbf{k} \to \mathbf{k+1}$. Calculate new residuals $\hat{\boldsymbol{\varepsilon}}^{(\mathbf{k})}$ by setting

   $$\hat{\boldsymbol{\varepsilon}}^{(\mathbf{k})} = \hat{\boldsymbol{\varepsilon}}^{(k-1)} - \hat{\sigma}^{(0)} H w(R(\hat{\boldsymbol{\varepsilon}}^{(k-1)}))), \tag{4}$$

   where $w(R(\hat{\boldsymbol{\varepsilon}}^{(k-1)}))$ denotes a vector-valued function which determines the influence of outliers according to the norm chosen. In our case, it is given by $w(i) = \phi(\frac{i}{n+1})$, $\phi(u) = \sqrt{12}(u - \frac{1}{2})$ [10]. If the new residuals display a lower dispersion around the fitted model as measured by the norm $\| \cdot \|_{w1}$ the step has been successful. Otherwise a linear search can be made along a direction to find a value which minimizes the dispersion. This is the new residual in step k.

4. By replacing the norm in the normal equations used for approximation in least squares-analysis, it can be shown that the dispersion is minimized by

$$\boldsymbol{\beta} = \sigma(X^T X)^{-1} \|Y\|_{w1}. \tag{5}$$

In this equation, we replace $\mathbf{Y}$ by the residuals $\boldsymbol{\varepsilon}^{(\mathbf{k})}$ in order to obtain a correction term for $\boldsymbol{\beta}$ which is added to the last estimate of $\boldsymbol{\beta}$ to produce the new estimate for $\boldsymbol{\beta}^{(\mathbf{k})}$. Based on the new value for $\boldsymbol{\beta}^{(\mathbf{k})}$, we get a new estimate of the variance $\hat{\sigma}^{(k)}$ like in step 1.

5. If the relative drop of the dispersion from one step to the next falls below a predefined threshold, stop the algorithm and acccept the last values $\boldsymbol{\beta}^{(k)}$ and $\hat{\sigma}^{(k)}$ as final estimates $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}$. Otherwise , go to step 3.

6 Can a polynomial of lower degree account for the data ?

While the algorithm described in the last section uses a very general class of functions from which diverse tuning curves can be fitted, it is desirable to have a simple function for the modelling of a tuning curve. Especially, we expect many of the coefficients of the higher terms in the trigonometric polynomial (3) to be small. To check whether the measured data allow for a simpler model, we use the algorithm described in the previous section to fit a trigonometric polynomial (3) of lower degree than d. This gives a distribution of residuals ε_{ij}^{red} for the reduced model. If the reduced model can account for the data, these residuals will have an approximately symmetrical distribution around zero. Based on the residual distribution, we calculate the sum of the signs of the residuals in all trials. If the fitted model function accounts well for the data, the distribution of the residuals is a binomial distribution of cr residual points around zero with a probability that the sign of a residual is positive or negative which is $\frac{1}{2}$. We then accept the best approximating polynomial of the lowest degree as fitted by the algorithm presented as a fitting tuning curve, whose residual distribution cannot be identified as non-symmetrical around zero by this test.
In our recordings, we found that 14 out of 80 cells were sharply tuned ($\alpha < 1\%$), 9 further cells moderately tuned ($1\% < \alpha < 5\%$) and the rest untuned (i.e. the tuning curve was a constant for $\alpha > 5\%$).

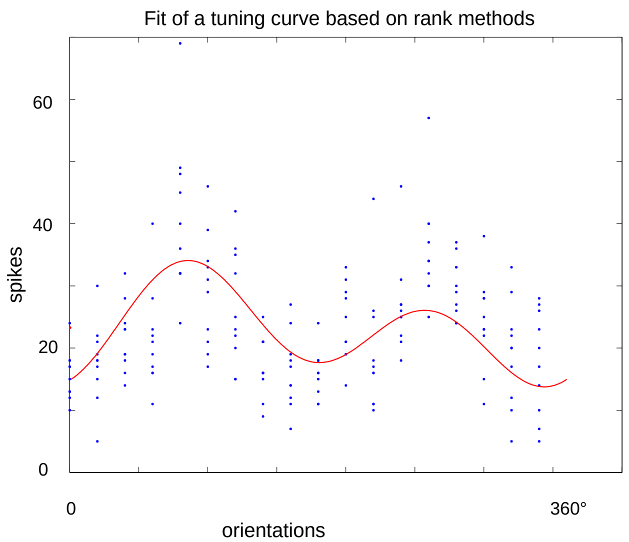

Fig. 1 Examples of recordings from a simple cell. Data were obtained from a total of 80 neurons in rat area 17 (ten animals, Brown Norway) [11]. Stimuli consisted of whole screen black and white

gratings, moving with constant velocity (5-20 deg/sec) and spatial frequency (0.08-0.6 cycl/deg). Background illumination was kept below 1 cd/m^2 and stimulus intensities ranged from 7-10 cd/m^2. Each experiment consisted of several blocks of trials in which 18 stimuli with particular moving direction were presented in a psudo-random order. In the figure shown, each dot corresponds to a value Y_{ij}. Every orientation condition was repeated 10 times. Equal responses are not shown separately. Thus, the maximum on the right in the scatterplot looks stronger on this plot than it actually is, for the lowest point in the fifth and the second lowest point in the sixth condition from the right occured three times and twice respectively, whereas double answers laid more to the median of the scattering distribution in the other conditions. The curve shows a tuning curve obtained by fitting a tuning curve of degree 2 by reducingb the degree of the fitting polynomial according to chapter 7 using a critical value of $\alpha = 10\%$. The residual distribution (not shown) still presents some trend, which could be avoided by choosing a lower critical value in the test described in section 7.

As another example, we show a fit of a cell classified *not tuned* by the using the trigonometric polynomial of the highest degree. For this neuron, this test gave a constant fit for critical levels higher than 1 percent.

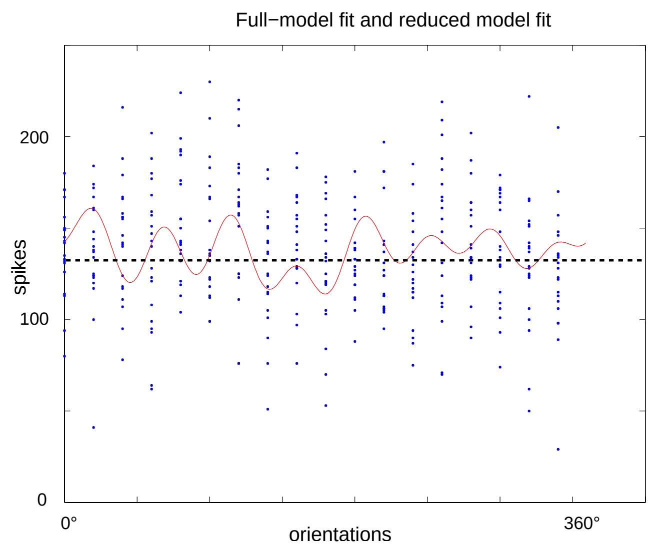

Fig.2 Examples of recordings from a simple cell. Every orientation condition was presented 20 times, equal answers are not shown separately. The continuous curve shows the fitted tuning curve using a fit of a polynomial of maximal degree. **However, none of the higher terms was significant**. We accept the dashed constant function as the tuning curve.

7 Discussion

The algorithm presented will deterministically find an approximation of the tuning curve by a trigonometric polynomial. This avoids using stochastic optimization methods. It also gives a criterion for the reduction of parameters. The approximation polynomial of maximal degree allows to have asymmetries and several maxima in the tuning curve. If the scattering of the residuals around the fitted polynomial of maximal degree does not differ significantly from a symmetric distribution, a simpler model is chosen for the tuning curve.

However, a class of polynomials approximating the tuning function more closely with respect to the given might be found. This work is currently in progress.

Acknowledgments

Data were kindly provided by Winrich Freiwald, Heiko Stemmann and Aurel Wannig from the Department of Biology , University of Bremen, Germany.

References

1. Encyclopedic dictionary of mathematics, edited by Kiyôsi Itô, MIT press , Cambridge , Massachusetts and London, 1996.
2. Girman S.V. , Sauvé Y. , Lund R.D., Receptive field properties of single neurons in rat primary visual cortex, J.Neurophysiol. 82 (1999) 301-311.
3. Henry G.H., Dreher B., Bishop P.O., Orientation specificity of cells in cat striate cortex. J. Neurophysiol. 37 (1974) 1349-1409.
4. Hubel D.H. and Wiesel T.N., Receptive fields of single neurons in the cat's primary visual cortex. J Physiol 148 (1959) 574-591.
5. Jackson D., The theory of approximation, Amer. Math. Soc. Colloq. Publ., 1930.
6. Mardia K.V. and Jupp P. E., Directional Statistics, John Wiley, Chichester, 2000.
7. Swindale N.V., Orientation tuning curves : empirical description and estimation of parameters. Biol. cybernetics (1998) 45-56.
8. Hardy, G.H. Littlewood, J.E. and Pólya, G. Inequalities, 2nd edition, Cambridge : Cambridge university press, 1952.
9. Press W.H., Teukolsky S.A., Vetterling W.T., and Flannery, B.P., Numerical recipes in C, Second edition, Cambridge university press, 1992.
10. Witting H., Nölle G., Angewandte mathematische Statistik, Teubner, Stuttgart, 1970.
11. Freiwald, W.A., Stemmann H., Wannig A., Kreiter, A.K., Hofmann U.G., Hills M.D., Kovacs G.T.A., Kewley, D.T., Bower, J.M., Eurich C.W., Etzold A. and S.D. Wilke, *Stimulus representation in rat primary visual cortex: multi-electrode recordings and estimation theory*, Neurocomputing (in press).