

You might find this additional information useful...

Supplemental material for this article can be found at:

<http://jn.physiology.org/cgi/content/full/00379.2009/DC1>

This article cites 44 articles, 14 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/103/1/591#BIBL>

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/103/1/591>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of January 11, 2010 .

Hierarchical Bayesian Modeling and Markov Chain Monte Carlo Sampling for Tuning-Curve Analysis

Beau Cronin,^{1,*} Ian H. Stevenson,^{2,*} Mriganka Sur,¹ and Konrad P. Körding²

¹Department of Brain and Cognitive Sciences, Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, Massachusetts; and ²Department of Physiology, Rehabilitation Institute of Chicago, Northwestern University, Chicago, Illinois

Submitted 29 April 2009; accepted in final form 31 October 2009

Cronin B, Stevenson IH, Sur M, Körding KP. Hierarchical Bayesian modeling and Markov chain Monte Carlo sampling for tuning-curve analysis. *J Neurophysiol* 103: 591–602, 2010. First published November 4, 2009; doi:10.1152/jn.00379.2009. A central theme of systems neuroscience is to characterize the tuning of neural responses to sensory stimuli or the production of movement. Statistically, we often want to estimate the parameters of the tuning curve, such as preferred direction, as well as the associated degree of uncertainty, characterized by error bars. Here we present a new sampling-based, Bayesian method that allows the estimation of tuning-curve parameters, the estimation of error bars, and hypothesis testing. This method also provides a useful way of visualizing which tuning curves are compatible with the recorded data. We demonstrate the utility of this approach using recordings of orientation and direction tuning in primary visual cortex, direction of motion tuning in primary motor cortex, and simulated data.

INTRODUCTION

A primary goal of many neuroscience experiments is to understand the relationship between the firing properties of a neuron and a single external variable. For sensory neurons, this external variable is typically a property of the stimulus, such as the orientation of a bar or a grating (cf. Hubel and Wiesel 1959, 1962). For motor neurons, this external variable typically refers to a property of the executed movement, for example the direction of hand movement (e.g., Georgopoulos et al. 1986). The relationship between the external variable and the firing rate, the tuning curve (see Fig. 1A), is one of the most frequently used tools to describe the response properties of neurons.

In these experiments, researchers often wish to answer questions such as the following. (1) Is the recorded neuron selective for the stimulus property in question? For example, a number of studies have asked what percentage of neurons in visual cortex are tuned to orientation (e.g., Maldonado et al. 1997; Ohki et al. 2005). (2) If the neuron is selective, how can its selectivity be described quantitatively? For example, many studies describe how well a neuron is tuned to the orientation of a stimulus (Carandini and Ferster 2000; Ringach et al. 2002). (3) How much uncertainty remains in this quantification after all of the evidence from the data has been taken into account? Often studies aim at giving error bounds on estimated neuronal properties. For example, it has long been a debate if neurons have sharper orientation tuning late in response than earlier in the response (Gillespie et al. 2001; Ringach et al.

1997a). Answering such questions demands reliable margins of certainty (Schummers et al. 2007). (4) Given two or more hypotheses concerning the functional or qualitative form of the neuron's selectivity, for which of these hypotheses does the data provide the most evidence (Amirikian and Georgopoulos 2000; Swindale 1998)? For example, it has been asked if orientation-selective cells in monkey striate cortex are tuned in the shape of a so-called Mexican hat (Ringach et al. 1997a). (5) Do the neuron's selectivity properties change after an experimental manipulation, either qualitatively or quantitatively? For example, several studies have asked whether adaptation, pharmacological intervention, or attention affect the tuning of neural responses to visual stimuli (Dragoi et al. 2001; McAdams and Maunsell 1999; Nelson et al. 1994; Sharma et al. 2003; Somers et al. 2001), while others have posed similar questions in the motor domain (Gandolfo et al. 2000; Li et al. 2001).

In addition to well controlled experiments, answering the questions in the preceding text requires statistical tools. Generally, there are two classes of methods for statistical analysis of tuning curves: parametric methods, which assume an explicit tuning function with several parameters, and nonparametric methods, which allow for essentially complete freedom in the form the tuning curve can take. Here we focus on parametric methods, such as circular Gaussian models for the orientation selectivity of simple cells in primary visual cortex (Hubel and Wiesel 1959, 1962) or cosine tuning models for direction of movement in primary motor cortex. These parametric models are particularly powerful when we know the form of tuning curves from past research—until the form is known nonparametric analysis such as spike-triggered averaging may be more appropriate. From a statistical point of view, we wish to estimate the parameter values given the data, estimate confidence intervals for the parameters, and select the best model from a set of possible models.

There are a number of well-established techniques in the statistical literature for solving these problems such as maximum likelihood estimation (MLE) (e.g., Swindale 1998), reverse-correlation (Simoncelli et al. 2004), bootstrapping (Stark and Abeles 2005), model comparison methods, and many others (Kass et al. 2005). These approaches have provided principled answers to scientific questions across a wide range of species, brain areas, and experimental manipulations. In this paper, we introduce a set of methods based on Bayesian statistics, which complement and extend these previous approaches for tuning-curve analysis.

Bayesian models start by formalizing how the variables we are interested in depend on one another and how uncertainty

* B. Cronin and I. H. Stevenson contributed equally to this work.

Address for reprint requests and other correspondence: I. Stevenson, 345 E Superior St Attn: ONT-931, Chicago, IL 60611 (E-mail: i-stevenson@northwestern.edu).

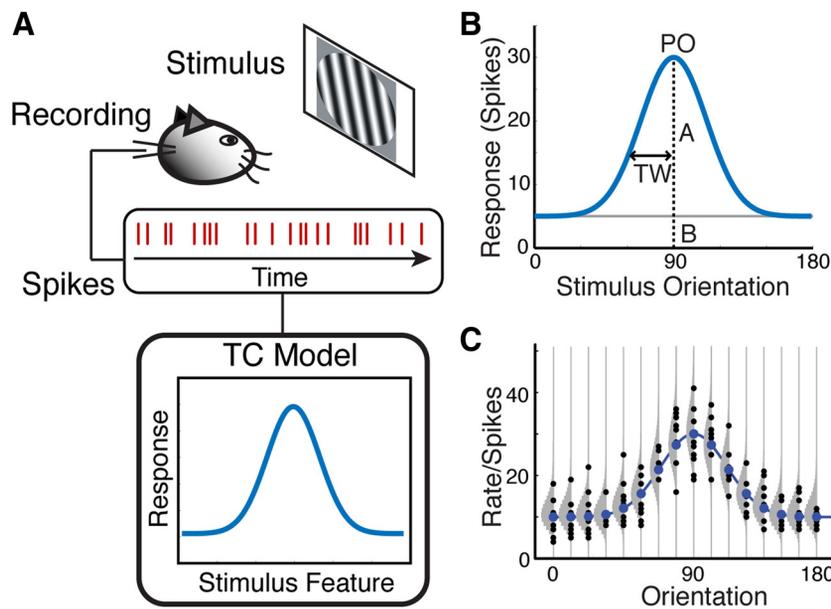


FIG. 1. Introduction. *A*: a schematic illustrating the setup of the sensory neuroscience experiments for which the analyses discussed here are applicable. An animal, either anesthetized or awake, is presented with stimuli that vary along a single feature dimension. Neuronal responses to these stimuli are recorded and then described using tuning-curve functions. *B*: an example tuning-curve model, the circular Gaussian function. The function can be described by the values of 4 parameters: the baseline (*B*), amplitude (*A*), preferred orientation (*PO*), and tuning width (*TW*). *C*: a schematic of the generative process assumed by the hierarchical statistical model. A neuron exhibits tuning described by a particular tuning-curve function. The firing rate in response to a stimulus of a particular value is determined by the value of this tuning-curve function when evaluated at this location. The response for a particular trial is then generated according to the assumed probability model, here a Poisson process.

arises. In neuroscience, this implies specifying the assumptions we have about how neural responses are generated from stimuli or movements. We first describe the framework for making these assumptions (a Bayesian hierarchical modeling framework). Second, we present a Markov chain Monte Carlo (MCMC) sampling process for performing the statistical inference procedures necessary for estimation of parameters and their credibility intervals, as well as, hypothesis testing. The central difference between the Bayesian approach presented here and more commonly used methods is that we attempt to describe the entire distribution over tuning curves (the posterior probability distribution) rather than a single tuning-curve estimate (such as the maximum likelihood estimate). Each point in this posterior probability distribution describes how compatible a given tuning curve is with the measured data. Samples from this distribution allow for straightforward estimation of parameters and error bars, as well as, hypothesis testing and model selection.

For many analyses, Bayesian methods yield the same results one would obtain when using more traditional ones. However, there are two cases where Bayesian methods can be particularly helpful: when there is limited data and when there are many parameters. Both cases are typical of problems in neuroscience where recording times are limited and neurons exhibit complicated tuning properties. Moreover, in both cases, estimating the parameter of tuning curves is difficult. Without considering distributions over parameters or using priors most models may over-fit the data. In the descriptions that follow, we focus on illustrating the utility of Bayesian methods for limited data and assume noninformative (flat) priors over the parameters. When there are many parameters, informative priors are often useful, but the application of these priors may require specific knowledge about the problem. However, even without using informative priors, many problems may benefit from Bayesian approaches, given the fact that electrophysiological experiments obtain datasets of limited size.

The statistical methods described here have been successfully applied in a range of scientific fields (Gelman et al. 2004; Liu 2002; MacKay 2003; Metropolis et al. 1953; Robert and

Casella 2004). In electrophysiology, MCMC sampling has been used for fitting peristimulus time histograms (Dimatteo et al. 2001), fitting nonparametric tuning curves (Kaufman et al. 2005), spike sorting (Nguyen et al. 2003; Wood et al. 2006), and neural decoding. However, these methods, and Bayesian methods in general, have not yet been adopted by neurophysiologists on a large scale. One major barrier to their adoption has been the difficulty in implementing MCMC sampling routines that are both appropriate and computationally efficient. To alleviate this obstacle, we make available a software toolbox that will allow researchers to easily and routinely apply all of the methods discussed here in the analysis of their own data.

METHODS

The spiking patterns of neurons will depend both on the properties of a stimulus and on the tuning properties of the neuron. In a typical experiment, we observe signals from neurons (e.g., spikes), and we want to ask which tuning properties are exhibited by the recorded neuron. In Bayesian terms, we want to ask which tuning properties are most consistent with the measured data given assumptions about their functional form (e.g., Gaussian or sigmoid). These Bayesian models assume an explicit generative process that defines how the observed firing patterns are (statistically) related to both the stimuli and the underlying tuning properties of the recorded neurons. We then use numerical methods to optimally estimate the values of the tuning properties as well as the uncertainty which remains (i.e., the error bars).

Hierarchical model

The key feature of our analysis method is that it describes a (hypothesized) probabilistic relationship between the parameters of a chosen tuning-curve (TC) function, any external variables (e.g., stimuli), and the neuronal responses collected in an experiment (see Fig. 1, *B* and *C*).

Given experimental data, such as spike counts, knowledge of the stimuli presented over the course of the experiment, and an assumption about the functional form of the TC, we want to know how likely each set of tuning-curve parameters is. Example tuning-curve func-

tions, including the circular Gaussian and the sigmoid function, are described in the supplemental materials.¹

In statistical terms, this means that we need to estimate the parameters of the tuning function p_1 through p_K (such as tuning width and preferred orientation) from the observed neural responses x and stimuli S (both vectors of length N , the number of experimental trials). We also assume that we may have some prior knowledge $[\Pr(p_j|\phi_j)]$ about the properties of these parameters, as characterized by a hyperparameter ϕ .

We can use Bayes' rule to estimate the posterior distribution of the TC parameter variables

$$\Pr(p_1 \dots p_K | x, S, \phi_1 \dots \phi_K) \propto \Pr(x | p_1 \dots p_K, S, \phi_1 \dots \phi_K) \prod_{j=1}^K \Pr(p_j | \phi_j) = \prod_{i=1}^N \Pr(x_i | R_{TC}(S_i; p_1 \dots p_K)) \prod_{j=1}^K \Pr(p_j | \phi_j) \tag{1}$$

R_{TC} is the value which results from evaluating the TC function for the particular stimulus value S_i (i.e., the stimulus presented on trial i), given parameter values p_1 through p_K . The term on the left-hand side of the equation is the joint posterior distribution of the parameter values, whereas the two terms on the right-hand side are the data likelihood and the prior distributions of the parameters.

This model can be adapted to suit a particular experimental context in two main ways. First the TC function should be chosen to reflect as accurately as possible the hypothesized selectivity properties of the neuron in question. For example, for neurons in primary visual cortex, a sigmoid function might be chosen when the stimulus contrast is being manipulated (e.g., Sclar et al. 1989), while a circular Gaussian (CG) function might be chosen when stimulus orientation is varied (e.g., McAdams and Maunsell 1999). Second, an appropriate probability model should be chosen to describe the relationship between the TC model output and the observed data. For example, a Poisson probability model might be appropriate when the data consists of spike counts, whereas a normal (i.e., Gaussian) model might be preferred when the data consist of continuous values such as fluorescence readings from confocal or two-photon microscopy (see the supplemental materials for more details).

MCMC sampling

Because the TC functions commonly used to model neural responses can impose complex nonlinear relationships between the parameters of interest, it is usually impossible to compute the joint posterior in Eq. 1 analytically. Sampling methods such as MCMC offer a powerful and flexible set of tools by which to overcome these limitations (Liu 2002; Metropolis et al. 1953; Robert and Casella 2004). These methods can be used to compute an approximation of the joint posterior parameter distribution.

Sampling methods operate by performing a search through parameter space. In this regard, they are similar to optimization techniques such as least-squares fitting methods. Unlike standard optimization algorithms, which only move "uphill" from less optimal solutions to better ones, sampling methods move both up- and downhill to explore the parameter space. The resulting random walk through parameter space is biased such that the amount of time spent in any region of the space is proportional to the probability density of the parameter values corresponding to that region. As a result, MCMC sampling returns an approximation of the joint probability distribution of the model parameters, given the data, any prior information, and assumptions about the form of the model (see Fig. 2A).

Sampling for each parameter p_i proceeds by starting with the current sampled value, $p_i^{(t)}$, and proposing a new value p_i^* . While this proposal can in principle be generated in a number of ways, we will

use normally distributed proposals the mean of which is the current sample value, i.e., $p_i^* \sim N(p_i^{(t)}, \sigma_0)$. Next, the probability in Eq. 1 is computed for the two different values of p_i , yielding $q_i^* = \Pr(p_1^{(t)} \dots p_i^* \dots p_K^{(t)} | x, S, \phi_1 \dots \phi_K)$ and $q_i^{(t)} = \Pr(p_1^{(t)} \dots p_i^{(t)} \dots p_K^{(t)} | x, S, \phi_1 \dots \phi_K)$. If q_i^* is greater than $q_i^{(t)}$, then the proposal is accepted, and $p_i^{(t+1)} \leftarrow p_i^*$. If q_i^* is less than $q_i^{(t)}$ (i.e., the proposal represents a "downhill" move), then it might still be accepted with a probability equal to the ratio of the two posterior probabilities, $q_i^*/q_i^{(t)}$; otherwise, it is rejected. If the proposal is rejected, then the previous sample value is maintained, i.e., $p_i^{(t+1)} \leftarrow p_i^{(t)}$.

It can be shown that sampling according to this procedure will result in parameter samples the values of which approximate the joint posterior distribution of the parameters, given the data (Metropolis et al. 1953). The samples can therefore be used as a proxy for the complete joint parameter distribution, and questions about this true distribution can be answered by referring to the samples. Figure 2B shows the results of sampling using a circular Gaussian tuning-curve function on simulated data, for which ground truth is known.

It should be noted that optimization techniques, such as those that use the mean squared error (MSE), also make (implicit) assumptions about the joint probability distribution of the model parameters. In particular, such methods assume that the parameters are normally distributed, an assumption which may not be appropriate in many cases. Figure 2B also shows the fit obtained with such a method.

In the discussion that follows, we assume that we have applied an MCMC procedure like the one just described to obtain M samples for each of the K tuning-curve model parameters: $\{\hat{p}_1^{(1)} \dots \hat{p}_1^{(M)}\} \dots \{\hat{p}_K^{(1)} \dots \hat{p}_K^{(M)}\}$.

Bayesian model selection

Once we use Bayesian sampling methods to estimate tuning functions, we can straightforwardly use these methods to also ask if one model is better at explaining the data than another model. To compare different models in this way, it is necessary to compare how well each of them can explain the data. In a Bayesian setting, this implies integrating over the specific parameter values (in Eq. 1). Bayesian model selection relies on the ratio of the probabilities of each model holding true, given the experimentally observed data (MacKay 2003)

$$\frac{\Pr(M_2|x)}{\Pr(M_1|x)} = \frac{\Pr(M_2)\Pr(x|M_2)}{\Pr(M_1)\Pr(x|M_1)} \tag{2}$$

Bayes Factor

Here, M_1 and M_2 are two models to be compared, and x is the observed data, as before. The first term on the right-hand side is the prior preference for one model over the other. In all cases, we assume that the two models being compared are, a priori, equally probably; under this assumption, this term is equal to 1. The second term, called the Bayes factor, is the ratio of the likelihood of the observed data, given each model (i.e., with all model parameters integrated out). This Bayes factor measures how much support the data provides for one model relative to the other. For example, if this quantity is 10, then it would be reasonable to say that the data provides evidence that model 2 is 10 times as likely to explain the data as model 1.

In general, it can be very difficult to actually perform the necessary integration to obtain $\Pr(x|M)$. For models of the complexity described here, however, very reasonable approximations of this integral can be obtained directly from the MCMC samples. One reasonable approximation of this quantity is simply the harmonic mean (Kass and Raftery 1995) of the sampled probabilities

$$\Pr(x|M) = \left(\frac{1}{m} \sum_{i=1}^m \Pr(x|p^{(i)}, M)^{-1} \right)^{-1} \tag{3}$$

This quantity is easy to compute from the output of the MCMC sampling routine, and this procedure provides good estimates of how preferable one model is relative to another (especially for low-

¹ The online version of this article contains supplemental data.

dimensional models). For higher dimensional models, other methods can be more stable (Chib and Jeliazkov 2001).

Simulations

Figs. 4 and 7 are based on the model of orientation/direction tuned simple cells used throughout the text. We use a circular Gaussian tuning curve ($b = 1$, $a = 4$, $\mu = 90^\circ$, $\sigma = 20^\circ$) (Fig. 4A and Fig. 7, rows 2 and 3) or a direction-selective (asymmetric) circular Gaussian tuning curve ($b = 1$, $a_1 = 4$, $a_2 = 2$, $\mu = 90^\circ$, $\sigma = 20^\circ$; Figs. 4B and 7, bottom). Figure 7 (top) uses an untuned neuron with $b = 1$. All models use Poisson noise, and stimulus directions were selected uniformly between 0 and 180° or 0 and 360° . One trial corresponds to a single random sample from a Poisson distribution with the mean given by the tuning-curve model at the randomly selected stimulus orientation. After 10,000 burn-in samples, 20,000 samples from the posterior were drawn. To reduce correlations between the samples only 400 samples (every 50th) were used for later estimation and model selection. The likelihood ratios in Fig. 7 were computed using cross-validated data from the same underlying tuning functions.

Bootstrap and nonlinear optimization

The bootstrap method is a general and powerful approach to statistical inference that relies on the basic operation of creating surrogate data sets by resampling from the observed data (Efron and Tibshirani 1986). Given a set of (presumed independent) observations $x = \{x_1 \dots x_n\}$, a single bootstrap sample $x^{(b)}$ is created by sampling, with replacement, from the original collection of data until the sample contains the same number of data points as the original set. For example, $x^{(1)} = \{x_1, x_1, x_3, x_4, x_4\}$ is a valid bootstrap sample when $n = 5$.

The bootstrap can be used to compute the uncertainty (i.e., error bars) about any summary of the observed data, $T(x)$. The general bootstrap procedure is as follows: 1) construct B bootstrap samples from the original observations, as described in the preceding text. 2) For each sample b , compute $T(x^{(b)})$; 3) construct confidence intervals from the sample summaries.

Our complete method for assessing the confidence intervals on the tuning-curve parameters combines nonlinear optimization for curve fitting (to compute the parameter estimates for each bootstrap sample), and the bootstrap for assessing the degree of uncertainty about those estimates. 1) Draw B bootstrap samples from the observed data. 2) For each bootstrap sample, use constrained nonlinear optimization with a squared-error objective function to compute the tuning-curve parameter values that best fit the data. 3) For each tuning-curve parameter of interest, use the estimate provided by the optimization routine on each bootstrap sample to construct confidence intervals.

This procedure results in a set of parameter samples that can be compared directly to those samples obtained from the Metropolis procedure.

RESULTS

To estimate tuning curves, several assumptions need to be made. We need to assume a dimension along which firing rates may vary (e.g., direction of movement). We need to assume a functional form of the tuning curve (e.g., cosine). Last, we need to assume how the tuning curve affects neural firing (e.g., assuming Poisson noise). To estimate the parameters of a tuning curve, we then use MCMC sampling techniques in the context of a hierarchical generative model that describes our prior knowledge about tuning curves; details are provided in METHODS and the supplemental methods section.

To give a more concrete example, a common generative model for data recorded from simple cells in mammalian primary visual cortex is a circular Gaussian tuning curve,

characterized by the baseline value, amplitude, preferred orientation, and tuning width. If spike count data are collected, then assuming that measurements vary with Poisson noise might be appropriate (see Fig. 1). Together the tuning-curve function and the noise model describe how likely the observed neural signals are given the parameters of the tuning curve (preferred direction, baseline firing rate, etc). Applying Bayes' rule, this likelihood can then be combined with prior assumptions to calculate the posterior probability distribution over the parameters. We can thus calculate how likely each combination of parameters is given observations from a neuron.

Ideally we would take every possible combination of parameters and calculate how likely the measured spikes are given the assumed combination. However, in many cases this is simply impossible: if we want to consider 100 values for each parameter and have five parameters we would have 10^{10} combinations of parameter values. MCMC is a well-defined way of having many of the advantages of considering all possible values without requiring the associated run times. Our algorithm generates samples from the posterior using the Metropolis method, a particular example of an MCMC technique. This method produces parameter samples by performing a random walk through parameter space, proposing new parameter values as random perturbations of previously accepted values. By accepting and rejecting these samples on the basis of their joint-posterior probability, the algorithm effectively builds a picture of this distribution (Fig. 2A). By generating more samples, the distribution can be approximated with arbitrary precision. At the conclusion of the sampling procedure, the collection of parameter samples obtained by the walk can be used to estimate the most probable parameter values, as well as the residual uncertainty—the error bars (Fig. 2B).

In the following sections, we demonstrate how these general methods can be applied by experimentalists to answer a range of scientific questions that frequently arise in the analysis of tuning-curve data. We use data obtained from orientation-tuned neurons in the visual system (V1), from direction-selective neurons in the motor system (M1) as well as several simulations to provide specific examples. For illustration purposes, these examples use circular Gaussian tuning-curve functions or cosine tuning functions and a Poisson noise model with noninformative priors on all parameters. These forms of tuning curves are frequently used for the study of primary visual cortex and motor cortex, respectively. The assumption that neurons spike according to Poisson statistics with a changing intensity function is the basis of many recent studies of neural systems. The general approach of using hierarchical generative models with MCMC for inference applies to a wide range of tuning-curve and neural models and can easily be extended to higher dimensional problems (cf. Schummers et al. 2007). Additional tuning-curve functions and neural models are discussed in the supplemental materials and the supplied toolbox supports the use of each of these models.

Estimating parameter values and credibility intervals

In many experiments, one of the first challenges confronting an experimentalist is to quantify the tuning properties of each recorded neuron. It is important to compute not only the most probable parameters but also the range of plausible parameters (the error bars or credibility intervals). We can pose this

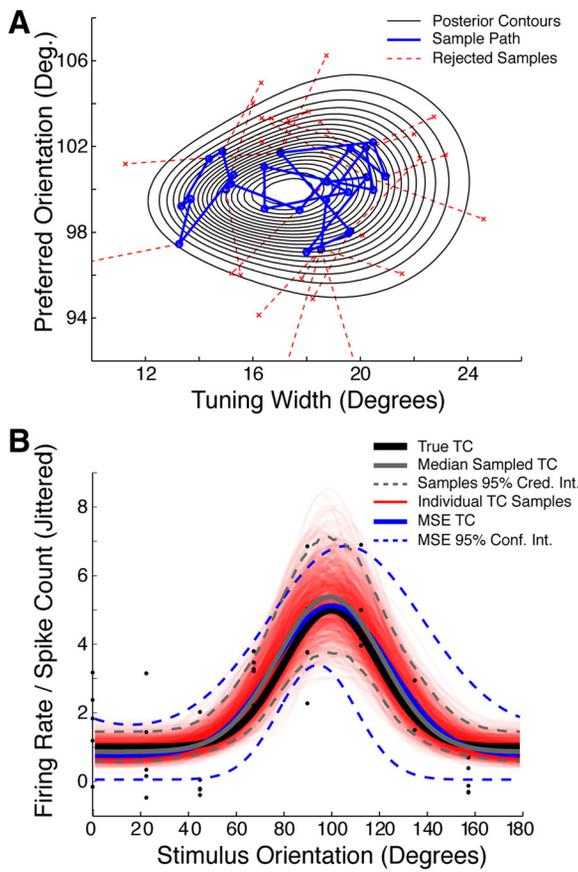


FIG. 2. Markov chain Monte Carlo (MCMC) sampling. *A*: metropolis sampling proceeds by making a biased random walk on the surface defined by the joint posterior probability density of the model parameters. Shown here are 2 dimensions of this surface for a circular Gaussian tuning-curve function; the preferred orientation and tuning width parameters are shown, while the amplitude and baseline parameters are omitted for clarity. The gray lines are the contours of uniform density. The blue path shows the trajectory of the samples with rejected proposals shown in red. Note that rejected samples always represent “downhill” moves, whereas accepted samples can be either up- or downhill. If enough samples are collected, the number of samples per region will converge to match the height of the density contours in the region; in other words, the samples will approximate the distribution. *B*: each sample corresponds to a different circular Gaussian tuning curve. Shown here are sampled tuning curves (light red lines), which are obtained by applying the generative model to simulated spike count data (black dots) generated from a known tuning curve (solid black line). Both the sampling method (solid gray line) and MSE-based nonlinear optimization (solid blue line) perform well in recovering ground truth in this case. However, the 95% error bounds generated from the samples (dashed gray lines) are narrower than those obtained from optimization, indicating that the hierarchical model and sampling approach is able to make more efficient use of the data. In addition, plotting the individual samples provides a nuanced understanding of the range of possible tuning curves which could have produced the observed data.

question more precisely by asking, for each parameter of the chosen tuning-curve function, what range of parameter values is consistent with the set of observed neuronal responses? For example, what is the plausible range of preferred orientations for an orientation selective neuron in V1?

Questions of this type can be answered directly from a set of MCMC samples because they approximate the joint posterior distribution over the parameter values. Parameter values that are sampled more often correspond to those that are more likely to account for the data. To obtain, say, a 95% credibility interval for parameter p_i , the samples $\{\hat{p}_i^{(1)} \dots \hat{p}_i^{(M)}\}$ are first

sorted from lowest to highest to obtain the rank order $\langle \hat{p}_i^{[1]} \dots \hat{p}_i^{[M]} \rangle$. Next, the lowest and highest 2.5% of the samples are discarded. For example, if 1,000 samples were collected, then the lowest and highest 25 would be discarded. The remaining extreme values define the desired error margin (see Fig. 3, *A* and *B*). In the 1,000-sample example, these would be the 26th and 975th sample from the ranked list. Importantly, this method does not need to make the assumption that the posterior is Gaussian. Depending on the model, the distribution may even be asymmetric or multimodal. As such, this method can deal with many of the potentially non-Gaussian distributions occurring in neuroscience.

In addition to credibility intervals, several other statistical estimates can be calculated from the samples. We can find a maximum a posteriori (MAP) estimate as well as the median or mean values of each parameter (Fig. 2*B*). When the prior over parameters is noninformative, the MAP estimate is equivalent to a MLE. However, in many cases, the mean or median posterior is a more robust estimate of the parameters. These estimators are optimal in that the posterior mean minimizes the squared error between the estimated parameter and its true value while the posterior median minimizes linear loss (Berger 1985). Like the maximum likelihood estimate, these estimators are asymptotically unbiased and efficient in most cases. The difference between MLE and Bayes estimators is generally apparent only for small sample sizes and disappears as more data are collected. In simulations of orientation-tuned neurons, for instance, the MLE and Bayes tuning-curve estimates are the same after ~ 40 trials (Fig. 4). For small sample sizes, using Bayes estimators, which take the distribution over parameters into account, can improve estimation.

Visualization of potential tuning curves

The presented approach also allows for simple visualization of the set of potential tuning curves that are compatible with the observed data. This visualization can be a useful tool for understanding the quantitative power that is provided by the data from a single cell (as used in Figs. 3 and 5). Because the posterior distribution over potential tuning curves may not be Gaussian and the parameters may not always be linearly independent, error bars on each parameter provide only part of the picture. For instance, the distribution of potential tuning curves may be skewed or multimodal (Fig. 5, *B* and *E*). This is one advantage of representing an entire distribution over the tuning curves described by a certain model.

Assaying quantitative changes in response properties

Quantifying tuning properties using the methods in the preceding text is often merely the first step in answering scientific questions about physiological data. Frequently, we wish to know whether a particular manipulation has resulted in a significant change in tuning properties. For example, it might be asked whether visual neurons change their orientation selectivity properties in response to attentional influences, sensory pattern adaptation, or perceptual learning (Dragoi et al. 2000, 2001; McAdams and Maunsell 1999) or if the preferred directions of neurons in motor cortex change in a force field (Rokni et al. 2007). In these kinds of experiments, neuronal responses would be measured in both the control and test

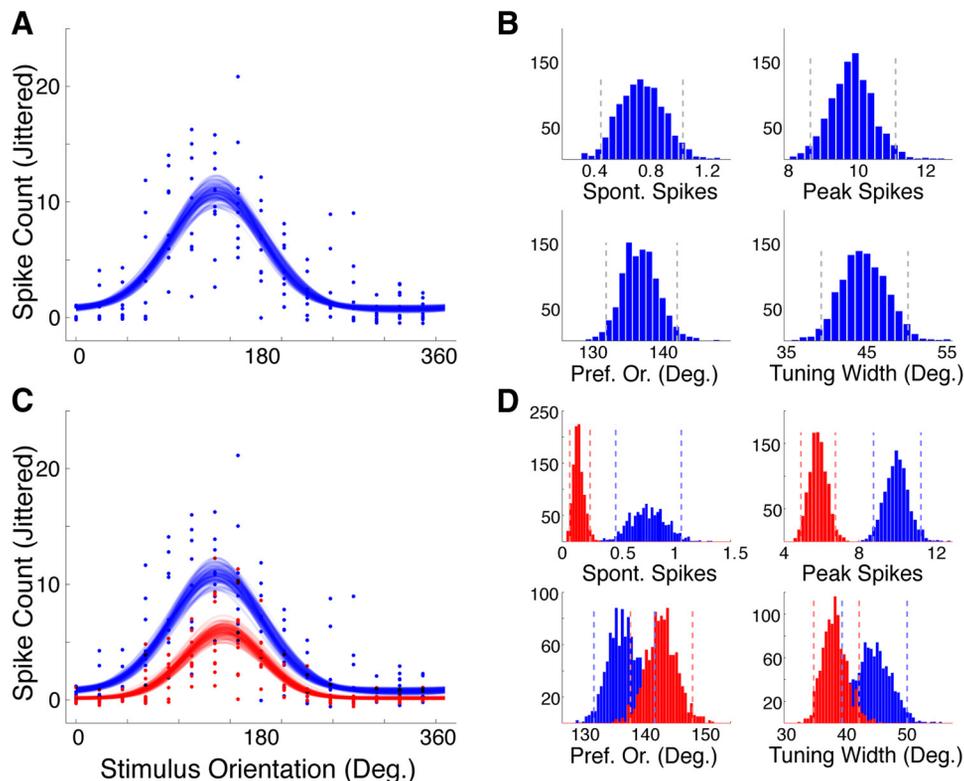


FIG. 3. Parameter estimation. *A*: spike counts (dots) and sampled tuning curves (lines) from a neuron in anesthetized mouse visual cortex, recorded in response to gratings of different orientations. On each trial, a different orientation was chosen (x axis), and the number of spikes emitted in response was recorded (y axis). These data were analyzed using a circular Gaussian TC function and a Poisson noise model. *B*: parameter sample histograms and credibility intervals for the sampled tuning-curve parameters. Each panel displays the samples for a different parameter: the spontaneous spike count, the number of spikes emitted in response to the preferred orientation, the identity of the preferred orientation, and the tuning width of the response. Dashed gray lines indicate the 95% credibility intervals for each of these parameter values. *C*: spike counts and sampled tuning curves of the same cell, recorded under both control (blue) and adapted (red) conditions. The tuning curves suggest that the cell has undergone a significant change in its response properties. *D*: histograms of each of the sampled tuning-curve parameter values and their corresponding credibility intervals, similar to *B*, but here comparing the values obtained in the control and adapted conditions. Both the baseline firing and amplitude at preferred orientation have changed significantly between the 2 conditions as indicated by the fact that the credibility intervals are disjoint. The preferred orientation and tuning width parameters are not significantly different between the 2 conditions because the intervals are overlapping. For all example data shown here, extracellular recordings were performed in mouse visual cortex, using methods similar to those described in Wang et al. (2006).

conditions, and MCMC samples would be generated separately for these two sets of data. To determine whether the data supports a conclusion of significant changes in tuning, credibility intervals are computed for each of the two conditions as in the preceding text. To ascertain whether a significant shift has occurred, it is sufficient to observe whether these two intervals overlap. If there is no overlap (i.e., the intervals are disjoint), then the data support the conclusion that the corresponding response property has changed between the two

conditions (see Fig. 3, *C* and *D*). If the two intervals do overlap, then this conclusion is not supported. This negative result is consistent with two possibilities. First, the underlying response property might really be unaffected by the experimental manipulation, which is a true negative result. Second, the tuning property has indeed changed but not enough data has been collected to reduce uncertainty to the point where this change can be reliably detected. In neither of the cases will this method wrongly report a change.

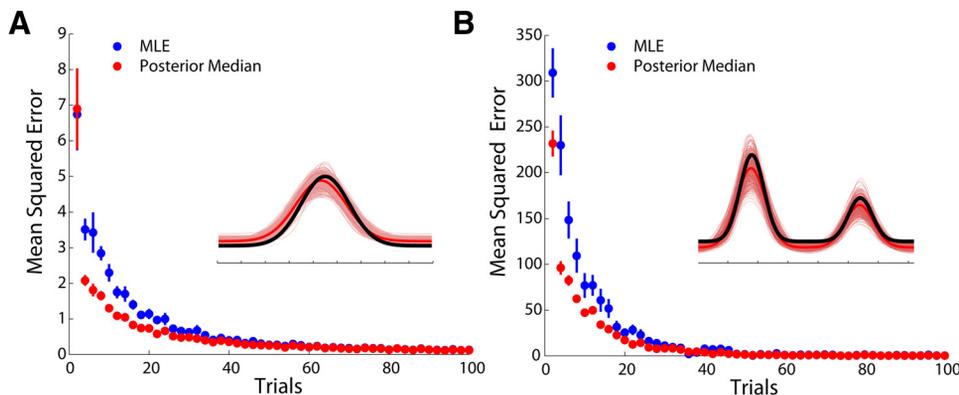


FIG. 4. Maximum likelihood and posterior median estimators. Especially for small sample sizes, the posterior median (or mean) is often more accurate than the maximum likelihood estimate. *A* and *B*: the mean squared error of the estimated tuning curve for these 2 estimators as a function of the number of trials (simulated data) for a tuned neuron and a direction-selective neuron (see insets). We assume a Poisson noise model. Solid black lines, the true tuning curve; dark red lines, the posterior median estimate after 50 trials, and light red lines indicate samples from the posterior. Error bars denote SE over 100 simulations (see METHODS for more details).

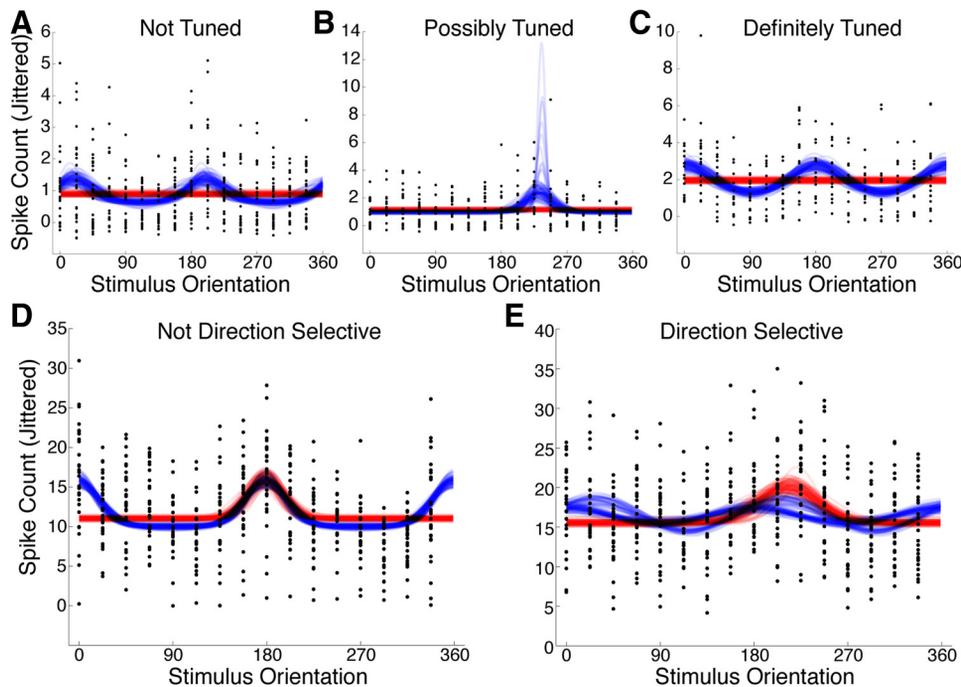


FIG. 5. Model selection for V1 neurons. *A*: spike counts (dots) and sampled tuning curves (lines) for an example cell that is not significantly tuned, as indicated by the fact that the evidence provided by the data for a tuned model (blue lines) is less than that for an untuned model (red lines). These data were analyzed using a circular Gaussian TC function, and a Poisson noise model. *B*: as in *A* except that the data provide marginally more support for the tuned model. *C*: as in *A* except that the data provide overwhelming support for the tuned model over the untuned model. *D*: as in *A–C* except that 2 different tuning models are being compared. Blue lines correspond to samples from a circular Gaussian TC model with periodicity of 180° , and red lines indicate a circular Gaussian with periodicity of 360° . For the cell shown here, the data provide more evidence for the 180° model, indicating that the cell is not significantly tuned for the direction of motion of the stimulus. *E*: as in *D* except that the data for this neuron provide more evidence for the 360° periodicity model, indicating that the cell is selective for the direction of motion of the stimulus.

Assessing selectivity to the stimulus feature using model selection

In many cases, it is of primary scientific interest to determine whether a neuron is actually selective for the stimulus property that is being varied in the experiment. For example, many but not all neurons in striate cortex are selective for particular stimulus orientations (Maldonado et al. 1997), and $\sim 20\%$ of neurons in primary motor cortex are not cosine tuned (e.g., Georgopoulos et al. 1986). Before drawing conclusions about the quantitative details of each neuron's orientation-dependent response, it is necessary to assess whether the cell is selective in the first place.

Bayesian model selection (BMS) can be used to determine how much evidence is provided by the data for one response model over another (Kass and Raftery 1995; MacKay 2003). To determine whether a cell is indeed selective for a stimulus property, BMS can be employed to compare models with two different tuning-curve functions: one in which the response varies with the stimulus value, and another in which the response is assumed to be insensitive to the particular stimulus (see Fig. 5, *A–C*).

BMS is similar to traditional hypothesis testing methods such as the likelihood ratio test in that we compare the probability of observing the data under each model. However, unlike the likelihood ratio test, BMS uses Bayes factors—the ratio of probability assigned to the data by each model, integrating over all possible parameter values. In general, it is difficult to compute Bayes factors because this integration can be nonlinear and high dimensional. For models with relatively small numbers of parameters, however, such as those used here, approximate Bayes factors can be computed directly from the MCMC samples (see Eq. 2). Furthermore, the quality of these estimates can be improved by increasing the number of samples computed. We have found that excellent results can be obtained for the models discussed here with very reasonable computational resources (<1 min/cell on a desktop machine).

A key property of BMS is that it appropriately penalizes models with more degrees of freedom. This “regularization” ensures that models with extra parameters are not favored merely because they are more expressive, which is a well-known complication in model comparison procedures (Buckland et al. 1997; MacKay 2003; Pitt and Myung 2002). This penalty is in some ways similar to the Akaike information criterion (AIC) or Bayesian information criterion (BIC). However, unlike these two methods, the Bayes factor applies to any priors over parameters or model types.

Comparing different selectivity models

A number of scientific questions can be posed in terms of BMS. For example, some neurons in primary visual cortex of the cat are direction selective, meaning that their responses to a grating of a particular orientation depend on the direction of motion of the grating. Other cells are not direction selective, and their responses are not modulated by the direction of motion of the grating. BMS can be used to distinguish between these two types of cells. In this case, two different tuning-curve functions would be used, each of which corresponds to one of the two hypotheses. The first is a circular Gaussian function with a periodicity of 180° ; this function represents a nondirection-selective response because its values are the same for stimuli of opposite direction. The second TC function is a circular Gaussian with a periodicity of 360° , which represents responses for only one direction of motion (see Fig. 5, *D* and *E*). In the case of neurons in motor cortex, one might compare a cosine tuning function and a constant tuning function (Fig. 6). When BMS is applied in this setting, using the methods described above, its output indicates the strength of evidence provided by the data that a particular cell is tuned for direction of motion.

In real data, the “true” tuning curve is generally unknown. However, we can illustrate some of the features of BMS using simulated data. To this end, we performed four simulations where

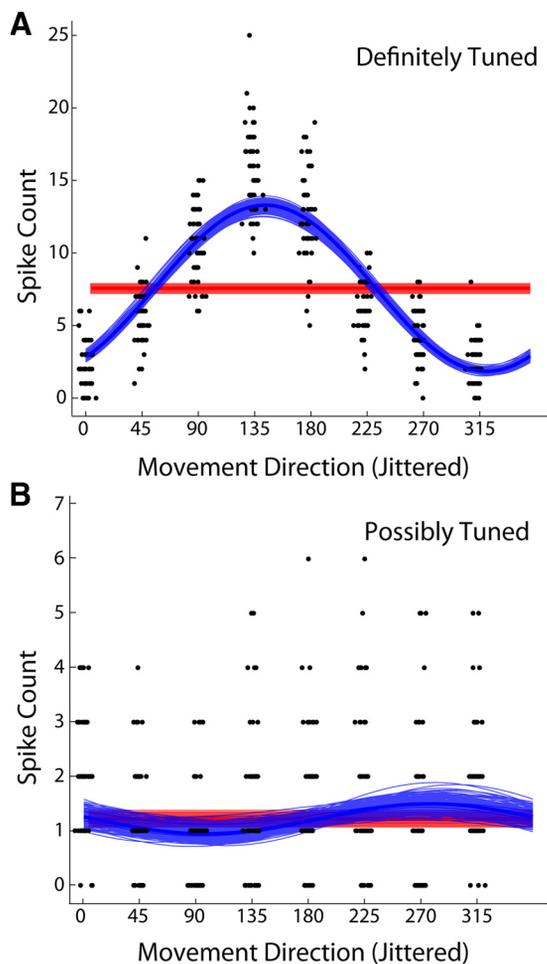


FIG. 6. Model selection for M1 neurons. Spike counts (dots) as a function of movement direction for 2 cells recorded while a monkey performed a center-out reaching task (414 trials). *A*: an example cell that is significantly tuned (Bayes' factor >100); *B*: an example cell which is only possibly tuned (Bayes' factor >10). The 2 models being compared are a cosine tuning-curve function (blue lines) and a constant tuning function (red lines) both with a Poisson noise model.

the true tuning-curve and noise model are known (Fig. 7). We simulated spikes from an untuned neuron (Fig. 7, top), an orientation-tuned neuron (Fig. 7, 2nd row), a nondirection-selective neuron (Fig. 7, 3rd row), and a direction-selective neuron (Fig. 7, bottom). For each of these sets of simulated data, we then generated MCMC samples from two models: untuned and tuned models for the first two cases (Fig. 7A) and nondirection-selective and direction-selective models for the second two cases (B). A Bayes factor >1 (or a log-ratio >0) suggests that model 1 is preferred, while a Bayes factor <1 (log-ratio <0) suggests that model 2 is preferred. Similar to cross-validated likelihood ratio tests, the Bayes factor selects the correct underlying tuning-curve model after a relatively small number of trials (Fig. 7, right).

Comparison with the bootstrap

One established alternative to the methods presented here is bootstrapping. Bootstrap methods are a very general set of techniques by which to estimate the residual uncertainty of arbitrary statistical summaries, including credibility intervals

for model parameters (Efron and Tibshirani 1997). These techniques have reached fairly wide use in neuroscience and solve some of the same problems that the techniques presented here address, particularly the estimation of error-bars (Brown et al. 2004; Kass et al. 2005; Paninski et al. 2003; Ringach et al. 2003; Schwartz et al. 2006; Stark and Abeles 2005; Ventura et al. 2005). Briefly, the bootstrap involves generating a series of surrogate data sets ("bootstrap samples") by sampling with replacement from the observed data; estimates of the uncertainty—error bars—are then computed by computing separate estimates from each of these samples. As such, bootstrap methods are a natural point of comparison for the hierarchical modeling and sampling methods described here. We have performed a series of comparisons between the two methods, using simulated orientation data for which the "true" tuning curve was known. A complete description of the bootstrap procedure used for these comparisons is provided in METHODS.

Figure 8 shows the results of several comparisons between the two methods. While the bootstrap performs similarly to posterior sampling when a large amount of data are available (Fig. 8A), the advantage of sampling becomes clear when less data are available. In this situation, the fully Bayesian approach is better able to recover the underlying tuning curve. This effect has been noted previously (Kass et al. 2005) and can generally be avoided by collecting more data, but it points to a more fundamental difference in the way these methods test tuning-curve models.

In addition to robustness to small sample sizes, the Bayesian method has several advantages over the bootstrap. Bootstrapping generally requires a nonlinear optimization, which can be subject to local minima in the error function. Manual intervention or more complicated methods, such as double bootstrap, may be necessary to correct for biases and dependencies in bootstrap estimates (Davison and Hinkley 1997). The Bayesian approach, on the other hand, does not include any optimization. Assuming appropriate mixing, the MCMC samples approximate the full posterior distribution over the parameters. Moreover, these techniques can easily be adapted to a wide range of tuning-curve functions, noise models, and priors.

DISCUSSION

Here we have shown how methods based on MCMC can help answer frequently occurring questions in the realm of electrophysiology. We have shown that the same method can be used to estimate properties of neural tuning for several kinds of neurons, obtain bounds on their values, examine if they change from one condition to the next and ask which model best explains the data. With the provided toolbox these methods are easy to implement and have the potential to significantly improve practical data analysis.

Comparison with other methods

The primary difference between the Bayesian sampling approach and more traditional approaches such as reverse-correlation, maximum likelihood estimation, and bootstrapping is that we attempt to consider all the tuning curves that would be compatible with the data. This analysis can reproduce the results from MLE by searching for the parameter values that are associated with the highest sample density and a noninfor-

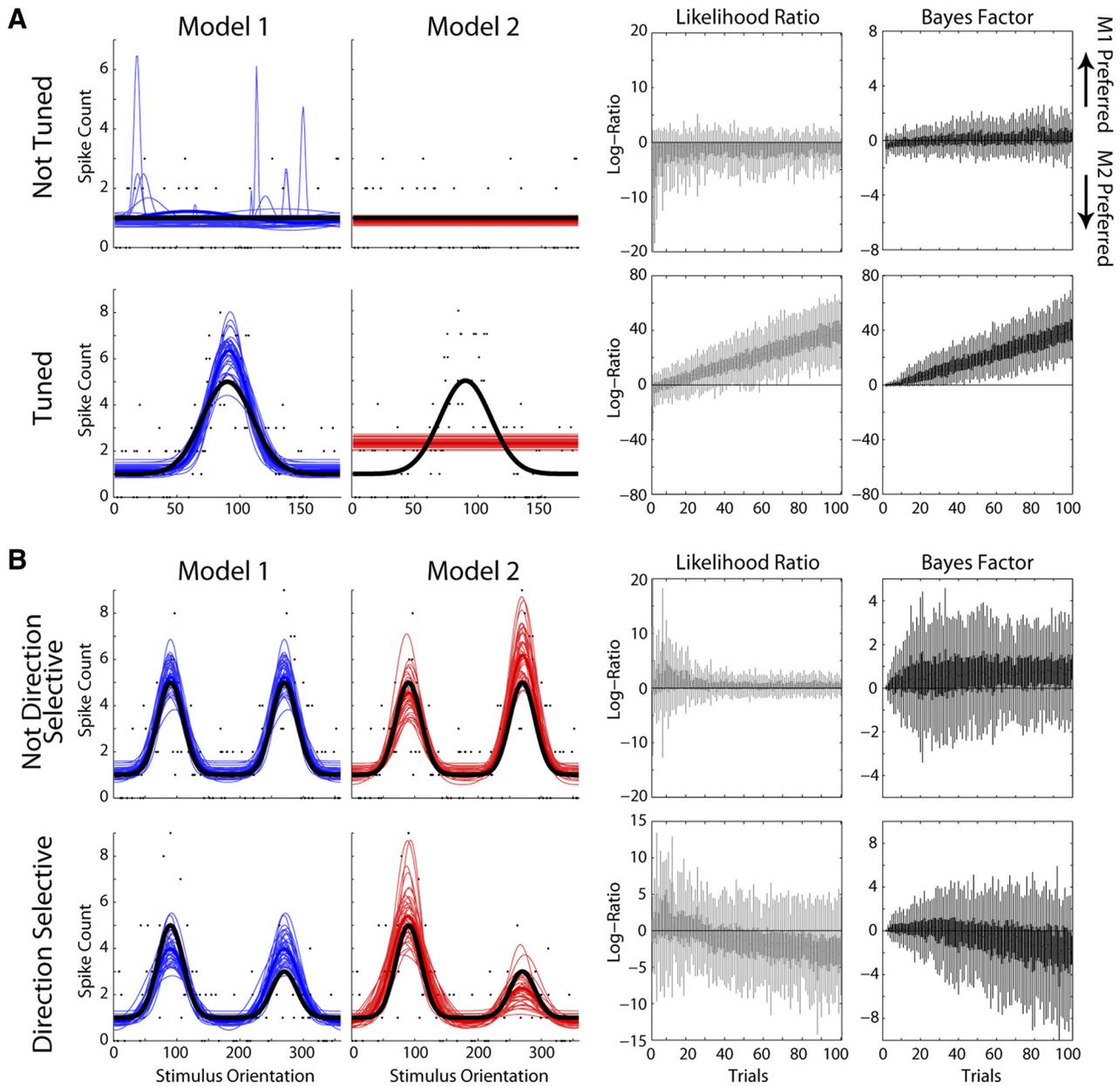


FIG. 7. Model selection with simulated data. *A* and *B*: 2 scenarios that one might encounter in analyzing data from primary visual cortex. In *A*, we want to determine whether a neuron is tuned (model 1) or untuned (model 2). Blue and red lines show samples from the posterior distribution of models 1 and 2 given 100 trials of simulated spike data (dots). The thick black lines show the true tuning curves. Using the posterior samples we can then compute the likelihood estimates and the evidence for each model. Both the Bayes factor and cross-validated likelihood ratio prefer the untuned model (model 2) when the simulated neuron is untuned (*top row, left*) and prefer the tuned model (model 1) when the simulated neuron is tuned (*2nd row, left*). In *B*, we want to determine whether a neuron is not direction selective (model 1) or is direction selective (model 2). Again, the cross-validated likelihood ratio and the Bayes factor both prefer the correct model after 100 trials. Boxes on the likelihood ratio and Bayes factor denote lower and upper quartiles (over 100 simulations). Whiskers denote the most extreme values within 1.5 times the interquartile range. Outliers have been removed for clarity (see METHODS for more details).

mativ prior. It obtains the same results as bootstrapping when there is enough data. However, for small sample sizes, Bayes estimators, such as the posterior median, and Bayesian credibility intervals are often more robust than these alternatives.

In hypothesis testing, Bayesian sampling methods converge to the results from likelihood ratio tests as well as AIC/BIC, under certain assumptions. For nonlinear or hierarchical models, the Bayes factor offers a more principled approach to model comparison. As experimental methods advance and stimulus-response models become more sophisticated, accurate model selection will become more and more important. By

looking at the distribution over potential tuning curves and not just the most likely tuning curve, physiologists may be able to gain insight into scientific questions as well as the statistical properties of tuning-curve models.

Another key feature of the sampling approach is its flexibility. Most analyses are based on an implicit assumption of Gaussian noise (Amirikian and Georgopoulos 2000; Swindale 1998), an assumption that is inappropriate for some common forms of neuronal data (e.g., spike counts). The hierarchical model described here can incorporate a range of probability models, such as the Poisson and negative binomial distribu-

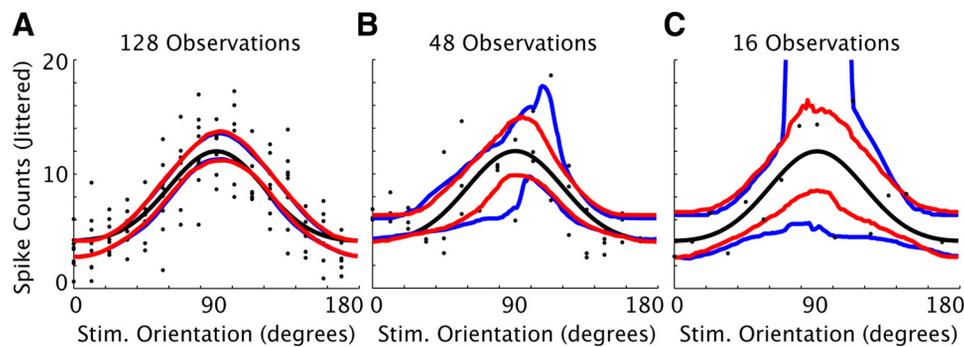


FIG. 8. Comparison to the bootstrap. Each panel shows the results of applying the Bayesian methods and the bootstrap to a different simulated data set. In each case, the simulated data were generated from the true underlying tuning curve (black line); the simulated data points are shown with black dots. The 95% error margins are shown for both the Bayesian methods (red) and the bootstrap (blue). *A*: 128 observations total; *B*: 48 observations total; and *C*: 16 observations total. The 2 methods are very similar in performance when large amounts of data are available, but the Bayesian method is better able to recover the true tuning curve when less data are collected.

tions in addition to Gaussian. This modeling flexibility also means that the framework described here can be directly applied to data other than spike counts, including fluorescence measurements from confocal and two-photon microscopy. There are a wide range of easily applicable tuning functions, including sigmoids, bilinear, von Mises functions, rectified cosines, and so on (see the supplemental methods for a full description of the tuning functions and likelihood models).

Although the methods we present here are flexible in the choice of tuning functions and noise models, they are generally much more constrained than nonparametric methods. In addition to MCMC for nonparametric tuning-curve estimation (Dimatteo et al. 2001; Kaufman et al. 2005), a large body of work exists on reverse-correlation approaches. These approaches aim to directly map stimulus properties to neural responses (Dayan and Abbott 2001; Eggermont et al. 1983; Marmarelis and Marmarelis 1978; Ringach et al. 1997b; Simoncelli et al. 2004). A wide range of toolboxes support the analysis of neural data using such nonsampling based methods (for useful software packages, see: <http://chronux.org/> <http://pillowlab.cps.utexas.edu/code.html>, <http://find.bccn.uni-freiburg.de/>). More recently, techniques have been developed that constrain reverse-correlation estimates using specific noise models and priors (Paninski 2004; Sahani and Linden 2003; Smith and Lewicki 2006). In many ways, these two strategies, nonparametric estimation of response functions with priors and fully Bayesian parametric estimation of response functions, represent a continuum of modeling frameworks that extends from minimally constrained modeling to very constrained modeling. This continuum allows experimentalists to decide how to model their data depending how much is known about the neural system or experimental manipulation in question. For instance, in a system where the exact parametric form of a tuning curve is unknown, experimentalists may want to use nonparametric or semi-parametric models rather than assuming a poorly fitting fully parameterized tuning function.

Inference assumptions and prior knowledge

All statistical inference methods, including nonparametric methods and bootstrapping, require assumptions. For example, to compute bootstrap estimates that were comparable to the

results of the Bayesian methods, it was necessary to perform nonlinear least-squares optimization to estimate the parameter values for each bootstrap sample (Swindale 1998). All optimization methods make use of an error function; the squared-error function that is used by this method is equivalent to an assumption of Gaussian noise in the data. In short, there is no way to estimate parameter values, or their residual uncertainty, or to compare different models with respect to their ability to explain the data without making assumptions. Bayesian methods generally make all such assumptions explicit.

One specific kind of inferential assumption that bears special attention is the prior distribution placed on the tuning-curve parameters. In most cases (including all of those used in the examples), a noninformative prior such as a uniform distribution over an appropriate range is sufficient. In this case, the posterior is equivalent to the likelihood, and one can easily find the maximum likelihood sample or compute a likelihood ratio between two models. However, in some cases, such as when there is very limited data or tuning functions are complicated, it may be necessary to include a more informative prior on one or more of the parameters. In these cases, it is incumbent on the experimenter to justify the use of a prior, such as by appealing to past measurements, literature, or to biophysical properties of the neurons under study. While such prior information may be controversial, we note that Bayesian methods provide the opportunity—though not the requirement—to introduce such information in a coherent and statistically principled way.

Extensions to more complex models

The general approach of using MCMC sampling to draw from the posterior distribution of model parameters can be applied to models that are more complex than those described here. These methods have been used previously to infer the parameter distribution of a model that described the orientation tuning dynamics of neurons in cat primary visual cortex in response to oriented reverse-correlation stimuli (Schummers et al. 2007). Nonparametric approaches to tuning-curve fitting using MCMC and splines have also been developed (Kaufman et al. 2005). These analyses use the same principles as those described here, illustrating that these techniques are applicable to a wide range of neurophysiological data analysis problems. We note, however, that more complex models (i.e., models containing more parameters than those described here) do

require more care in the selection of inference methods. In particular, special attention must be paid to the proposal distribution that is used to generate candidate parameter samples; additional detail can be found in Schummers et al. (2007). This study demonstrates that, once an appropriate proposal distribution has been identified, sampling methods scale quite well, and can be successfully applied to models with ≤ 100 parameters or more.

Many if not most statistical estimation techniques are currently viewed as approximations to optimal Bayesian inference. The approach we used here has the benefit that even for small amounts of data, in the limit of sufficient computer power (large number of samples), it will provably converge to the correct answers. These tools should thus be useful for the general analysis of tuning curves, which is a central statistical objective in neuroscience.

ACKNOWLEDGMENTS

The authors thank C. Runyan, J. Schummers, and R. Mao for supplying the data used in the V1 examples and A. Cherian and L. Miller for the data used in the M1 examples.

GRANTS

B. Cronin and M. Sur were supported by National Eye Institute Grant EY-07023.

REFERENCES

- Amirikian B, Georgopoulos AP. Directional tuning profiles of motor cortical cells. *Neurosci Res* 36: 73–79, 2000.
- Berger JO. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer, 1985.
- Brown EN, Kass RE, Mitra PP. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat Neurosci* 7: 456–461, 2004.
- Buckland ST, Burnham KP, Augustin NH. Model selection: an integral part of inference. *Biometrics* 53: 603–618, 1997.
- Carandini M, Ferster D. Membrane potential and firing rate in cat primary visual cortex. *J Neurosci* 20: 470–484, 2000.
- Chib S, Jeliazkov I. Marginal likelihood from the Metropolis-Hastings output. *J Am Stat Assoc* 96: 270–281, 2001.
- Davison AC, Hinkley DV. *Bootstrap Methods and Their Application*. Cambridge, UK: Cambridge Univ. Press, 1997.
- Dayan P, Abbott LF. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press, 2001.
- Dimatteo I, Genovese CR, Kass RE. Bayesian curve-fitting with free-knot splines. *Biometrika* 88: 1055–1071, 2001.
- Dragoi V, Rivadulla C, Sur M. Foci of orientation plasticity in visual cortex. *Nature* 411: 80–86, 2001.
- Dragoi V, Sharma J, Sur M. Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron* 28: 287–298, 2000.
- Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sci* 1: 54–75, 1986.
- Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York: Chapman and Hall, 1997.
- Eggermont JJ, Johannesma PM, Aertsen AM. Reverse-correlation methods in auditory research. *Q Rev Biophys* 16: 341, 1983.
- Gandolfo F, Li C, Benda BJ, Schioppa CP, Bizzi E. Cortical correlates of learning in monkeys adapting to a new dynamical environment. *Proc Natl Acad Sci USA* 97: 2259–2263, 2000.
- Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Boca Raton, FL: Chapman and Hall, 2004.
- Georgopoulos AP, Schwartz AB, Kettner RE. Neuronal population coding of movement direction. *Science* 233: 1416–1419, 1986.
- Gillespie DC, Lampl I, Anderson JS, Ferster D. Dynamics of the orientation-tuned membrane potential response in cat primary visual cortex. *Nat Neurosci* 4: 1014–1019, 2001.
- Hubel DH, Wiesel TN. Receptive fields of single neurons in the cat's striate cortex. *J Physiol* 148: 574–591, 1959.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160: 106–154, 1962.
- Kass R, Raftery A. Bayes factors. *J Am Stat Assoc* 90: 773–795, 1995.
- Kass RE, Ventura V, Brown EN. Statistical issues in the analysis of neuronal data. *J Neurophysiol* 94: 8–25, 2005.
- Kaufman CG, Ventura V, Kass RE. Spline-based non-parametric regression for periodic functions and its application to directional tuning of neurons. *Stat Med* 24: 2255–2265, 2005.
- Li CS, Padoa-Schioppa C, Bizzi E. Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. *Neuron* 30: 593–607, 2001.
- Liu J. *Monte Carlo Strategies in Scientific Computing*. New York: Springer, 2002.
- MacKay DJC. *Information Theory, Inference, and Learning Algorithms*. Cambridge, MA: Cambridge, 2003.
- Maldonado PE, Godecke I, Gray CM, Bonhoeffer T. Orientation selectivity in pinwheel centers in cat striate cortex. *Science* 276: 1551–1555, 1997.
- Marmarelis PZ, Marmarelis VZ. *Analysis of Physiological Systems: The White-Noise Approach*. New York: Plenum, 1978.
- McAdams CJ, Maunsell JH. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19: 431–441, 1999.
- Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* 21: 1087–1092, 1953.
- Nelson S, Toth L, Sheth B, Sur M. Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science* 265: 774–777, 1994.
- Nguyen DP, Frank LM, Brown EN. An application of reversible-jump Markov chain Monte Carlo to spike classification of multi-unit extracellular recordings. *Network Comput Neural Syst* 14: 61–82, 2003.
- Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433: 597–603, 2005.
- Paninski L. Maximum likelihood estimation of cascade point-process neural encoding models. *Network Comput Neural Syst* 15: 243–262, 2004.
- Paninski L, Fellows MR, Hatsopoulos NG, Donoghue JP. Spatiotemporal tuning of motor cortical neurons for hand position and velocity. *J Neurophysiol* 91: 515–532, 2003.
- Pitt MA, Myung IJ. When a good fit can be bad. *Trends Cogn Sci* 6: 421–425, 2002.
- Ringach DL, Hawken MJ, Shapley R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature* 387: 281–284, 1997a.
- Ringach DL, Hawken MJ, Shapley R. Dynamics of orientation tuning in macaque V1: the role of global and tuned suppression. *J Neurophysiol* 90: 342–352, 2003.
- Ringach DL, Sapiro G, Shapley R. A subspace reverse-correlation technique for the study of visual neurons. *Vision Res* 37: 2455–2464, 1997b.
- Ringach DL, Shapley RM, Hawken MJ. Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22: 5639–5651, 2002.
- Robert CP, Casella G. *Monte Carlo Statistical Methods* (2nd ed.). New York: Springer, 2004.
- Rokni U, Richardson AG, Bizzi E, Seung HS. Motor learning with unstable neural representations. *Neuron* 54: 653–666, 2007.
- Sahani M, Linden JF. Evidence optimization techniques for estimating stimulus-response functions. *Adv Neural Inform Process Syst* 15: 317–324, 2003.
- Schummers J, Cronin B, Wimmer K, Stimberg M, Martin R, Obermayer K, Koerding K, Sur M. Dynamics of orientation tuning in cat V1 neurons depend on location within layers and orientation maps. *Front Neurosci* 1: 145–159, 2007.
- Schwartz O, Pillow JW, Rust NC, Simoncelli EP. Spike-triggered neural characterization. *J Vision* 6: 484–507, 2006.
- Scelar G, Lennie P, DePriest DD. Contrast adaptation in striate cortex of macaque. *Vision Res* 29: 747–755, 1989.
- Sharma J, Dragoi V, Tenenbaum JB, Miller EK, Sur M. V1 neurons signal acquisition of an internal representation of stimulus location. *Science* 300: 1758–1763, 2003.
- Simoncelli EP, Paninski L, Pillow J, Schwartz O. Characterization of neural responses with stochastic stimuli. In: *The New Cognitive Neurosciences* (3rd ed.), edited by Gazzaniga M. Cambridge, MA: MIT Press 2004, p. 327–338.
- Smith EC, Lewicki MS. Efficient auditory coding. *Nature* 439: 978–982, 2006.

- Somers DC, Dragoi V, Sur M.** Orientation selectivity and its modulation by local and long-range connections in visual cortex. In: *Cat Primary Visual Cortex*, edited by Payne B, Peters A. New York: Academic, 2001, p. 471–520.
- Stark E, Abeles M.** Applying resampling methods to neurophysiological data. *J Neurosci Methods* 145: 133–144, 2005.
- Swindale NV.** Orientation tuning curves: empirical description and estimation of parameters. *Biol Cybern* 78: 45–56, 1998.
- Ventura V, Cai C, Kass RE.** Statistical assessment of time-varying dependency between two neurons. *J Neurophysiol* 94: 2940–2947, 2005.
- Wang KH, Majewska A, Schummers J, Farley B, Hu C, Sur M, Tonegawa S.** In vivo two-photon imaging reveals a role of arc in enhancing orientation specificity in visual cortex. *Cell* 126: 389–402, 2006.
- Wood F, Goldwater S, Black MJ.** A non-parametric Bayesian approach to spike sorting. *IEEE Eng Med Biol Syst* 1: 1165, 2006.