



OPEN ACCESS

EDITED BY

Kenichi Ohki,
The University of Tokyo, Japan

REVIEWED BY

Jonathan D. Victor,
Cornell University, United States
Tatsuo K. Sato,
Kagoshima University, Japan
Ryan Rowekamp,
Salk Institute, United States

*CORRESPONDENCE

Stephen D. Van Hooser
✉ vanhoosr@brandeis.edu

RECEIVED 09 December 2024

ACCEPTED 23 June 2025

PUBLISHED 21 July 2025

CITATION

Wu Z and Van Hooser SD (2025) Bayesian estimation of orientation and direction tuning captures parameter uncertainty.
Front. Neural Circuits 19:1542332.
doi: 10.3389/fncir.2025.1542332

COPYRIGHT

© 2025 Wu and Van Hooser. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Bayesian estimation of orientation and direction tuning captures parameter uncertainty

Zongting Wu¹ and Stephen D. Van Hooser^{2*}

¹Department of Biochemistry, Brandeis University, Waltham, MA, United States, ²Department of Biology, Brandeis University, Waltham, MA, United States

This study explores the efficacy of Bayesian estimation in modeling the orientation and direction selectivity of neurons in the primary visual cortex (V1). Unlike traditional methods such as least squares, Bayesian estimation adeptly handles the probabilistic nature of neuronal responses, offering robust analysis even with limited data and weak selectivity. Through the analysis of both simulated and experimental data, we demonstrate that Bayesian estimation not only accurately fits the neuronal tuning curves but also effectively captures parameter certainty or uncertainty of both strongly and weakly selective neurons. Our results affirm the complex interdependencies among response parameters and highlight the variability in neuronal behavior under varied stimulus conditions. Our findings provide guidance as to how many response samples are necessary for Bayesian parameter estimation to achieve reliable fitting, making it particularly suitable for studies with constraints on data availability.

KEYWORDS

tuning curves, motion selectivity, motion processing, spatial processing, orientation tuning

1 Introduction

Neurons in the visual system of all examined mammals, as well as neurons in other systems such as head-direction cells and place cells in the hippocampus, often exhibit tuning for the orientation or direction of a stimulus (Hubel and Wiesel, 1959, 1962; Wilson et al., 2018; Hubel and Wiesel, 1968; Henry et al., 1974; Ohki et al., 2005) or of the movement of an animal in a coordinate system (Taube et al., 1990; McNaughton et al., 1983). When selectivity is strong, there are many methods for quantifying the orientation and direction angle preference, the degree of tuning relative to an orthogonal angle, or the sensitivity of the tuning as the stimulus changes from the preferred angle (Carandini and Ferster, 2000; Ringach et al., 2002; Swindale, 1998; Mazurek et al., 2014).

Today, many optophysiology and multi-channel recording studies with massive electrode channel counts often characterize the response properties of dozens or hundreds of neurons simultaneously (Stringer et al., 2019; de Vries et al., 2019; Siegle et al., 2021), some of which may have high selectivity and some of which may not. Some methods of quantifying the degree of orientation and direction tuning are robust regardless of whether tuning is strong or weak. However, reliably estimating tuning parameters such as angle preference and tuning width is not possible with least squares methods when tuning is weak (Mazurek et al., 2014). In a prior methods paper, our group recommended simply not trying to quantify parameters other than the degree or amount (or magnitude) of tuning if a statistical test (such as Hotelling's T2 test or an ANOVA across stimuli) did not show evidence of some significant tuning (Mazurek et al., 2014).

Nevertheless, there are times when one may want to quantify properties such as preference angle or tuning width even for cells that exhibit weak tuning. For example, one may want to make a statement about how tuning has changed when following a cell over the course of development, before and after an experimental manipulation, or as another parameter (say, stimulus spatial frequency, or temporal frequency) is varied (Li et al., 2008; Moore et al., 2005). If a cell is initially poorly tuned, such that we have no certainty about its angle preference, but later exhibits strong tuning for upward motion, we may want to contrast this situation with another cell that exhibits strong tuning for downward motion but later exhibits strong tuning for upward motion. If we throw up our hands when tuning is weak, we lose the ability to make quantitative statements about these situations.

Bayesian estimation methods allow quantification of a full joint distribution of the parameters of a tuning model, so that any uncertainty or certainty of a given parameter is estimated (Bishop, 2006). Full Bayesian estimation requires much more computation than least squares fitting or maximum likelihood estimation (Cronin et al., 2010), but advances in computing power and GPU methods now allow these methods to be applied routinely and at speeds that are similar to least squares approaches. We describe a Bayesian estimation approach for orientation and direction tuning curves that can be widely applied to other sensory tuning curves. We provide enough detail that an analyst who is unfamiliar with Bayesian estimation can apply the method to a new problem. We provide comparisons to the bootstrap technique (Efron and Tibshirani, 1993; Press and Flannery, 1992), which also allows one to describe the certainty or uncertainty of tuning parameters.

2 Materials and methods

Neurons or other systems can exhibit responses that vary with stimulus direction and/or orientation. These responses can be spikes or voltage deflections, modulation of activity at a stimulus frequency, or responses obtained from an indicator such as a calcium indicator. There are two major classes of methods for analyzing orientation and direction tuning curves: vector methods (Batschelet, 1981; Swindale, 1998; Mazurek et al., 2014) and fit methods (Carandini and Ferster, 2000; Ringach et al., 2002; Swindale, 1998; Mazurek et al., 2014). Vector methods, such as the calculation of circular variance (Ringach et al., 2002) are excellent for quantifying the degree or strength of orientation or direction tuning for weakly and strongly tuned cells (Mazurek et al., 2014), but in this paper we will focus on an equation model of orientation and direction tuning based on a wrapped Gaussian (Swindale, 1998; Batschelet, 1981; Carandini and Ferster, 2000), see Swindale (1998) for other equation choices:

$$R(\theta; C, R_p, R_n, \theta_{pref}, \sigma) = C + R_p * e^{-\frac{(\text{angdiff}(\theta - \theta_{pref}))^2}{2\sigma^2}} + R_n * e^{-\frac{(\text{angdiff}(\theta - \theta_{null}))^2}{2\sigma^2}}. \quad (1)$$

Five parameters are included in the equation, which are C , R_p , R_n , θ_{pref} and σ . The offset, C , is the portion of the response that is constant and does not vary with orientation or direction. R_p is

the response above offset for the preferred direction and R_n is the response above offset for the null direction, that is, the direction that is opposite of the preferred direction. θ_{pref} is the preferred direction. In direction space, the range of θ is $[0, 360^\circ)$, and $\theta_{null} = \theta_{pref} + 180^\circ$. The function $\text{angdiff}(\theta) = \min(\theta, \theta - 360, \theta + 360)$ computes the absolute angular difference around the circle. We also sometimes refer to the response in the preferred direction and the null direction, respectively, as

$$R_{pref} = R(\theta_{pref}) \quad (2)$$

and

$$R_{null} = R(\theta_{null}). \quad (3)$$

Various different groups use different definitions of θ , with some using Cartesian coordinates where 0° refers to vertical bars moving to the right, and increasing θ implies a counter-clockwise change in angle, and other groups such as own using compass coordinates, where 0° refers to horizontal bars moving upwards and increasing θ implies a clockwise change in angle.

2.1 Bayesian formulation

Bayes' theorem is stated mathematically as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (4)$$

We constructed a Bayesian model of Equation 1. We employed a change-of-variable by letting $R_n = \alpha R_p$ with $0 \leq \alpha \leq 1$, so that we can always be sure that $R_n \leq R_p$ as we compute the likelihood of the model. Our new equation is as follows:

$$P(C, R_p, \alpha, \theta_{pref}, \sigma | D) = \frac{P(D|C, R_p, \alpha, \theta_{pref}, \sigma)P(C, R_p, \alpha, \theta_{pref}, \sigma)}{P(D)}. \quad (5)$$

In words, the equation says that the probability of the parameters $(C, R_p, \alpha, \theta_{pref}, \sigma)$ given the data D we observed is equal to the probability of seeing the data given the parameters, multiplied by the prior probability of the parameters and divided by the prior probability of the data (Bishop, 2006).

The left side of Equation 5 describes what we are seeking: how likely is it that each parameter (say, C), takes any of a set of values (say, $\{C_1, C_2, C_3, \dots\}$). In order to calculate this probability, we need to be able to calculate the probability of the data given the parameters (called the *likelihood function*) and the prior probability of the parameters.

In our case, the data D is the set of empirical mean responses across trials to the particular angles we examined. We indicate each of i angles tested by θ_i and the mean response to that angle as r_{θ_i} . We need to calculate the probability, according to our model, that we observed r_{θ_i} given the parameters. Equation 1 tells us that the expected response $\hat{r}_{\theta_i}$ for parameters $(C, R_p, \alpha, \theta_{pref}, \sigma)$ is

$$\hat{r}_{\theta_i} = R(\theta_i; C, R_p, \alpha, \theta_{pref}, \sigma). \quad (6)$$

Next, we need to take this expected response and calculate the likelihood of the actual response $r_{\theta i}$ in units of probability. Many studies of the variability of neural firing rates have found that the variance (and standard deviation) of spike responses is linearly related to the mean response rate on a log-log plot (Softky and Koch, 1993; Shadlen and Newsome, 1998). We therefore assumed that the model responses would be drawn from a normal distribution with mean $\hat{r}_{\theta i}$ and a standard deviation σ_{noise} that was a function of the mean response. To estimate $\sigma_{noise}(m)$ for our data, we started by examining the log of the standard deviation of each response $\sigma_{noise,i}$ of each cell as a function of the log of the mean response m_i :

$$\log(\sigma_{noise,i}) = a + b * \log(m_i).$$

If one were to obtain the best-fit values of a and b then one could express the expected variability as a function of the mean response as

$$\sigma_{noise}(m) = 10^a * m^b.$$

We modified this classic function (Softky and Koch, 1993; Shadlen and Newsome, 1998) slightly so that the noise is not identically 0 at 0 mean rate, which leads to strange parameter estimations, and to achieve better fits of the noise at the lowest mean rates, arriving at a final noise equation with three fitted constants C , K , and S :

$$\sigma_{noise}(m) = C_n + K * m^S. \quad (7)$$

For each set of recordings, we pooled our observations of all $\sigma_{noise,i}$ and m_i over all responses and all cells and derived a single set of C_n , K , and S that was applied to all cells. For example, when dealing with calcium imaging data from ferret visual cortex (Li et al., 2008), we used all cells and all responses to calculate the constants we used for the noise model for those cells. We did not derive these constants for each cell individually (as the limited data for a single cell would produce a poor fit), or apply C_n , K and S across situations (for example, between recordings of spiking neurons and recordings using calcium indicators).

Analysts are encouraged to examine the fit of the noise on a log-log plot for each new application. While our data were well fit, and previous studies have found this function describes the relationship between mean spike rates and variability for data in cortex (Softky and Koch, 1993; Shadlen and Newsome, 1998), factors such as calcium indicator saturation (Nauhaus et al., 2012) or different firing properties in different circuits might necessitate using a modified noise function.

In our experiments, we usually record responses over several trials T rather than recording a single trial, and report the mean response averaged over trials $r_{\theta i}$. Because we average over trials, we expect the response we measure to be less noisy than the results of a single trial. According to the central limit theorem, the variability decreases by a factor of $1/\sqrt{(T)}$. Therefore, the probability that a response at a single stimulus angle was generated by a particular set of model parameters is

$$P(r_{\theta i}|C, R_p, \alpha, \theta_{pref}, \sigma) = N(r_{\theta i}; \hat{r}_{\theta i}, \sigma_{noise}(\hat{r}_{\theta i})/\sqrt{(T)}) \quad (8)$$

where $N(x; \mu, \sigma_{normal})$ is the probability density function of the normal distribution with mean μ and standard deviation σ_{normal} . We assume that the response measurements are independent, so that the probability of seeing the set of responses $\{r_{\theta 1}, r_{\theta 2}, \dots, r_{\theta n}\}$ for each cell is

$$P(D|C, R_p, \alpha, \theta_{pref}, \sigma) = \prod_{i=1}^{i=n} N(r_{\theta i}; \hat{r}_{\theta i}, \sigma_{noise}(\hat{r}_{\theta i})/\sqrt{(T)}). \quad (9)$$

Note that the expected response in the noise equation is based on the model response. In this report, we did not make any assumptions about the *a priori* likelihood of the different values of the parameters $(C, R_p, \alpha, \theta_{pref}, \sigma)$. We estimated our distributions for discrete sets of values of each parameter (for example, C could take $\{C_1, C_2, C_3, \dots, C_{CM}\}$, R_p could take $\{R_{p1}, R_{p2}, R_{p3}, \dots, R_{pRPM}\}$, etc), and we assumed each value was equally likely so that

$$P(C, R_p, \alpha, \theta_{pref}, \sigma) = \frac{1}{(CM)(RPM)(\alpha_M)(\theta_M)(\sigma_M)}, \quad (10)$$

where CM , RPM , α_M , θ_M , and σ_M are the number of values of each parameter used to estimate the distribution. Perhaps in the future, by examining many populations, the expected prior distributions could be estimated differently from a uniform distribution, but we did not want to impose any biases at this time.

Finally, we did not explicitly compute $P(D)$, but we assumed that our calculation of $P(C, R_p, \alpha, \theta_{pref}, \sigma|D)$ would estimate the real probability distribution, and reported the estimated distribution normalized to 1. We end up with a matrix of likelihoods $L(c, r, \alpha_j, t, s)$ with one entry for each combination of parameters examined.

2.2 Practical computation and storage of Bayesian results

We calculated our Bayesian distributions using ranges of parameters that varied slightly depending upon the units of the data. In our case, we used mean spike rates and fractional change in calcium indicators from baseline. The ranges we used are shown in Table 1.

To perform Bayesian estimation for a single cell, one must perform calculations for all combinations of the 5 parameter ranges, and Table 1 indicates that billions of calculations are required for each cell. This creates burdens for computation and for storage of the results. Fortunately, modern GPUs are well-suited to performing these calculations quickly. We wrote GPU code using Matlab (MathWorks, Natick, MA) that performs these calculations in less than a minute per cell on an ordinary laptop with a GPU, so the computation is not problematic for today's computing environments.

Storing the results requires careful decisions, however. Each probability for every parameter combination is a floating point number; if stored on disk with double precision, then the results for a single cell would require on the order of 10 GB of storage. For

TABLE 1 Parameter space sampled for Bayesian parameter estimation, for spiking neurons (left column) and calcium imaging (right column).

Parameter	Spiking neurons	Calcium imaging
$\min(C)$	0.1	$-MX$
$\max(C)$	10	$+MX$
CM	60	60
$\min(R_{pref})$	0.1	0.001
$\max(R_{pref})$	20	$+3 * MX$
$R_{pref}M$	60	60
$\min(\alpha)$	0	0
$\max(\alpha)$	1	1
αM	15	21
$\min(\theta_{pref})$	0	0
$\max(\theta_{pref})$	359	359
$\theta_{pref}M$	72	72
$\min(\sigma)$	1	1
$\max(\sigma)$	60	60
σM	60	60
	Number of calculations	Number of calculations
	$2.3 * 10^8$	$3.3 * 10^8$

Parameters were varied from a minimum value to a maximum value using the number of steps indicated for each parameter with an M . For example, CM is the number of parameters evaluated for parameter C . For calcium imaging, MX indicates the empirical maximum mean response for each cell. The total number of calculations necessary for each cell is shown at the bottom. The offset C is allowed to vary between $-MX$ and MX because cells may exhibit stimulus responses that are increased or decreased relative to the nearest blank/control stimulus.

this reason, we did not store all of the results but instead stored a much briefer summary of key statistics.

First, we stored the marginal likelihood of each parameter individually by summing over all other parameters and normalizing the distribution to 1. For example, the marginal likelihood $\hat{L}c(c)$ of parameter C is computed as

$$Lc(c) = \sum_{r=1}^{r=RPM} \sum_{\alpha_j=1}^{\alpha_j=\alpha_M} \sum_{t=1}^{t=\theta_M} \sum_{s=1}^{s=\sigma_M} L(c, r, \alpha_j, t, s)$$

$$\hat{L}c(c) = \frac{Lc(c)}{\sum_{c=1}^{c=CM} Lc(c)}. \quad (11)$$

While we lose information about the joint distributions of these parameters, the summary provides considerable savings, needing only $(CM) + (RPM) + (\alpha_M) + (\theta_M) + (\sigma_M)$ entries for storage instead of $(CM)(RPM)(\alpha_M)(\theta_M)(\sigma_M)$ entries (the sum instead of the product), which with our parameters means a difference between kilobytes and gigabytes per cell. The marginal distributions say a lot about the response properties of the cell, and, if we really need the joint distribution, we can recompute it.

Second, we performed estimates of the likelihoods of index values of interest while we had the full matrix of parameter estimations in memory, and stored histograms of these values.

OI (Orientation Index) and DI (Direction Index) are two indexes of neuronal direction/orientation selectivity. Their magnitudes are used to characterize the sensitivity of neurons to stimuli motion in different directions/orientations in space. In this paper, we define these values as follows:

$$OI = (R_{pref} + R_{null} - (R_{orth+} + R_{orth-})) / (R_{pref} + R_{null}) \quad (12)$$

and

$$DI = (R_{pref} - R_{null}) / R_{pref}. \quad (13)$$

For summary statistics, we computed histograms of the likelihood that OI and DI took values between 0 and 1 in bins of size 0.05. By using the full likelihood matrix to calculate these histograms but only storing the histogram results, we obtained a good picture of our certainty or uncertainty about the value of each index that uses the full joint information, while storing only 20 values (1/0.05) per index.

Third, we also stored the most likely value of the full joint parameters also known as the maximum likelihood estimate (MLE). These parameters are a single set of values of $C, R_p, \alpha, \theta_{pref}, \sigma$ that provided the largest joint likelihood, but by themselves do not account for our uncertainty or certainty of their values.

2.3 von Mises test

To test the framework's ability to characterize the activity of tuning curves that are generated by similar but not exact matches for the model, we generated some simulated orientation tuning curves using the von Mises function (Batschelet, 1981; Swindale, 1998):

$$R_{v-m}(\theta) = A \exp(k[\cos(2(\theta - \varphi)) - 1]). \quad (14)$$

2.4 Relative error in parameter estimation

In order to quantify the performance of the Bayesian parameter estimation, we devised a relative error measurement. Given the *true* value x_t of a given parameter, we defined the expected squared error (ESE) of the Bayesian parameter estimation as:

$$ESE = \int_a^b (\hat{L}c(x) * (x - x_t)^2 dx, \quad (15)$$

and the expected squared error for a uniform distribution as

$$ESE_u = \int_a^b (x - x_t)^2 dx. \quad (16)$$

Finally, we defined the relative error of our approach compared to an uninformed (uniform) estimate as ESE/ESE_u .

2.5 Source code and data sources

Code in MATLAB that performs the Bayesian parameter estimation in a manner callable from other projects is provided at <https://github.com/VH-Lab/vhlab-bayesoridir-matlab>. The repository also includes all code that was used to generate the figures in the paper.

Some data for the figure on noise were obtained from the Allen Brain Observatory (de Vries et al., 2020) using the Brain-Observatory-Toolbox in Matlab (<https://github.com/MATLABCommunity-Toolboxes-at-INCF/Brain-Observatory-Toolbox>).

3 Results

To illustrate the approach and to explore how the Bayesian estimation tracks a tuning curve's parameters under a ground truth situation, we generated simulated data for analysis. In Figure 1, we show two example tuning curves: one is well-tuned for orientation and direction (Figure 1A) while the other is poorly tuned (Figure 1B). The output of Bayesian estimation is not a single value of each parameter but instead is a probability distribution over the parameters. The marginal likelihood for θ_{pref} and R_{pref} is plotted for each cell, along with histograms of the estimated orientation index (OI) and direction index (DI). For example, in Figure 1C, we do not obtain a single “best fit” or “most likely” value for θ_{pref} but a likelihood distribution, where the estimated probability that θ_{pref} takes any value in the range. The entire shape of the curve, not just the peak, is important to the interpretation. When the tuning is strong, there is great certainty about the parameters and the values of the index values. When the tuning is weak, there is much more uncertainty. In particular, in the weakly-tuned case, there is great uncertainty about whether the orientation and direction index values are near 0 or might be moderate. Least-squares methods would give us a single value of all of these parameters, and our uncertainty about the orientation and direction index values would not be described.

One may ask, why bother to study the tuning properties of weakly-tuned cells? To demonstrate such a case, we imagine measuring the direction tuning of a single cell while another visual stimulus parameter (temporal frequency) is changed. For low temporal frequency, the cell exhibits high direction selectivity; for a moderate temporal frequency, the cell exhibits orientation tuning but not direction tuning; for a further temporal frequency, the cell is weakly responsive.

To evaluate the predictive accuracy and sensitivity of the Bayesian estimation model, we simulated response curves for two different types of cells and performed parameter estimation. All error bars in the text are represented using standard errors of the mean (SEM).

3.1 Bayesian estimation for simulated strong and weak spatial selectivity

As shown in Figure 1, multiple graphs illustrate the response characteristics of simulated neurons. The left column, labeled

as “Well-tuned,” features graphs that exhibit neuronal responses with higher peak firing rates and marked preferences for specific directions and orientations of stimuli. These responses are quantified in terms of frequency (Hz) and marginal likelihood across different orientations and directions, denoted by directional and orientation indexes. The right column, labeled as “Poorly-tuned,” displays neurons with significantly lower response amplitudes, less distinct tuning curves, indicating a reduced selectivity for stimulus features. As shown in Figures 1A, B, the black curves represent the ideal response curve of simulated data without any noise. To approximate real data more closely, we used simulated sampling points with an added 50% Gaussian noise for fitting. Each data point is based on five samples, covering a total of 36 angles. These sampling points are represented as blue dots in the figure. The noisy data, alongside results from a linear noise model, serve as inputs for Bayesian estimation, which are then interpolated based on given parameter ranges to produce the fits displayed as red curves. The parameter settings for the well-tuned simulated cells are ($C = 1$, $R_{pref} = 10$, $R_{null} = 5$, $\theta_{pref} = 90^\circ$, and $\sigma = 30^\circ$), while those for the poorly-tuned cells are ($C = 1$, $R_{pref} = 1$, $R_{null} = 0$, $\theta_{pref} = 90^\circ$, and $\sigma = 30^\circ$).

The Bayesian MLE fit for both types of simulated cells closely align with the ideal curves, particularly for the well-tuned cell in terms of the preferred angle and tuning width values. Figures 1C–F describe the marginal probability distributions for θ_{pref} and R_p . The probability peaks accurately estimate the “true values” of θ_{pref} and R_{pref} for the well-tuned cell. The certainty or uncertainty of the parameters is indicated by the probability distribution peaks and widths. Due to the poorly-tuned cell exhibiting weak selectivity for direction, the estimated probability for θ_{pref} is dispersed across a broader range of angles compared to the well-tuned cell. The likely values of R_{pref} are considerably smaller for the cell with poorer tuning, showing its reduced response.

Figures 1G–J illustrates the probability distributions for the selectivity indexes DI and OI . It can be observed that the well-tuned cell displays distinct, singular probability peaks for DI and OI , accurately estimating the “real” values ($OI = 0.86$, $DI = 0.45$). In contrast, for the poorly-tuned cell, the likelihood estimates of both are broader, although a close estimation to the “real” value ($OI = 0.33$, $DI = 0.5$) is also observed in the probability distribution ($OI = 0.3$, $DI = 0.38$).

3.2 Storing marginal likelihoods as a summary

Storing the full joint likelihood distribution $P(C, R_p, \alpha, \theta_{pref}, \sigma | D)$ requires gigabytes of storage. For many purposes, viewing or storing only the marginal likelihood of each parameter, as we have plotted in Figure 1 is sufficient. Of course, there is some information in the full joint likelihood function that might be needed for certain applications.

We illustrate 2-dimensional views of the joint likelihood for pairs of parameters (marginalized over other parameter) in Figure 2. The relationships between some parameters are quite independent. For example, the joint likelihood of θ_{pref} and either R_p (Figure 2A) or C (Figure 2B) show little interaction among the

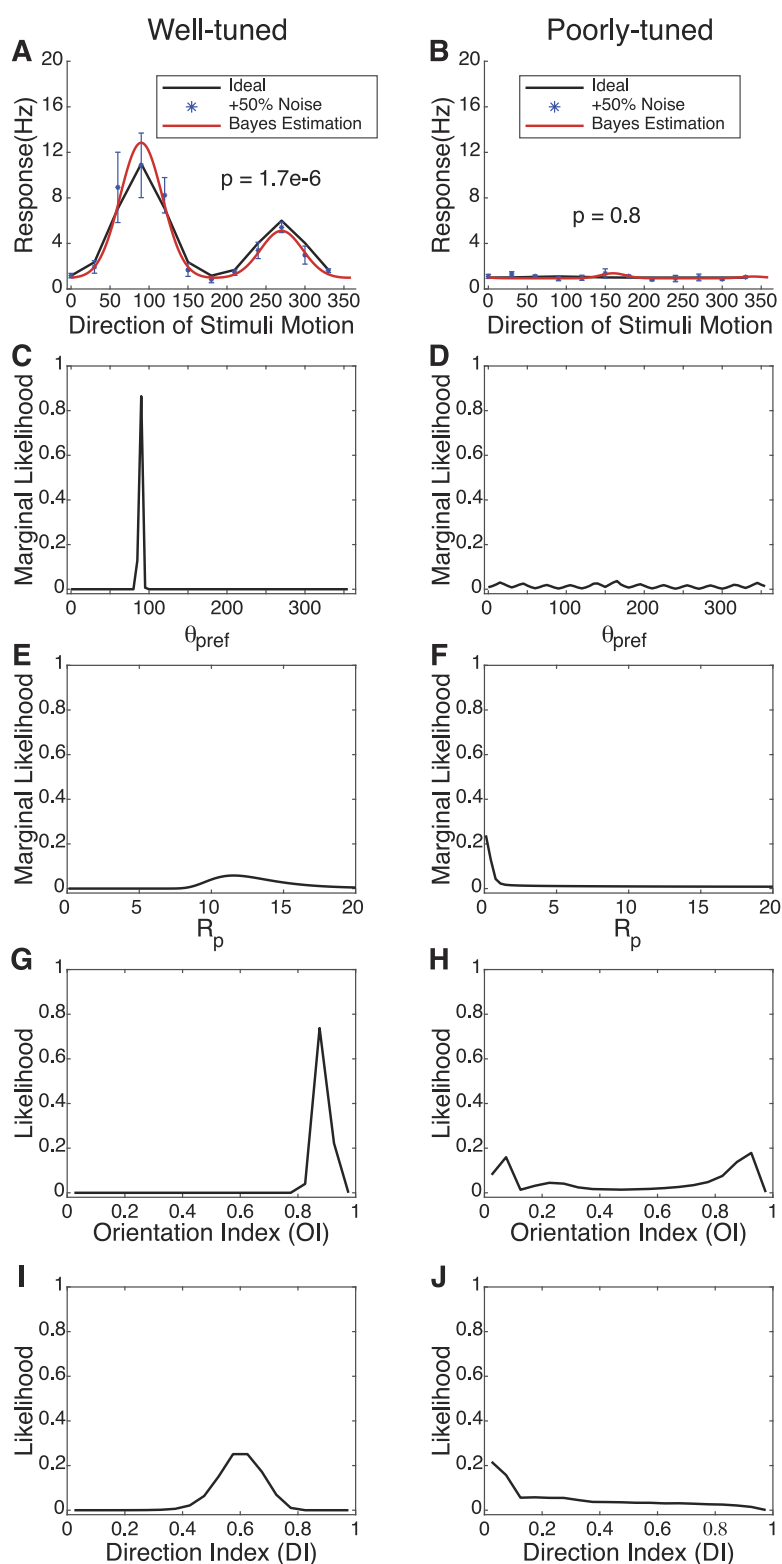


FIGURE 1

Bayesian estimation for two types of simulated cells with different levels of tuning. (A, B) Simulated response curve, data collection points with +50% Gaussian noise, and Bayesian estimation results. The thick black line indicates the “true” response of the neuron with no noise added, and the blue datapoints indicate simulated responses with noise added. The red line reflects the most likely parameters from the Bayesian parameter estimation. P values indicate the result of an ANOVA test across stimuli; the well-tuned cell is highly significant while the poorly-tuned cell is not. (C–F) The marginal likelihoods of θ_{pref} and R_{pref} in the model. (G–J) The probability distribution of OI and DI for spatial selectivity under Bayesian estimation.

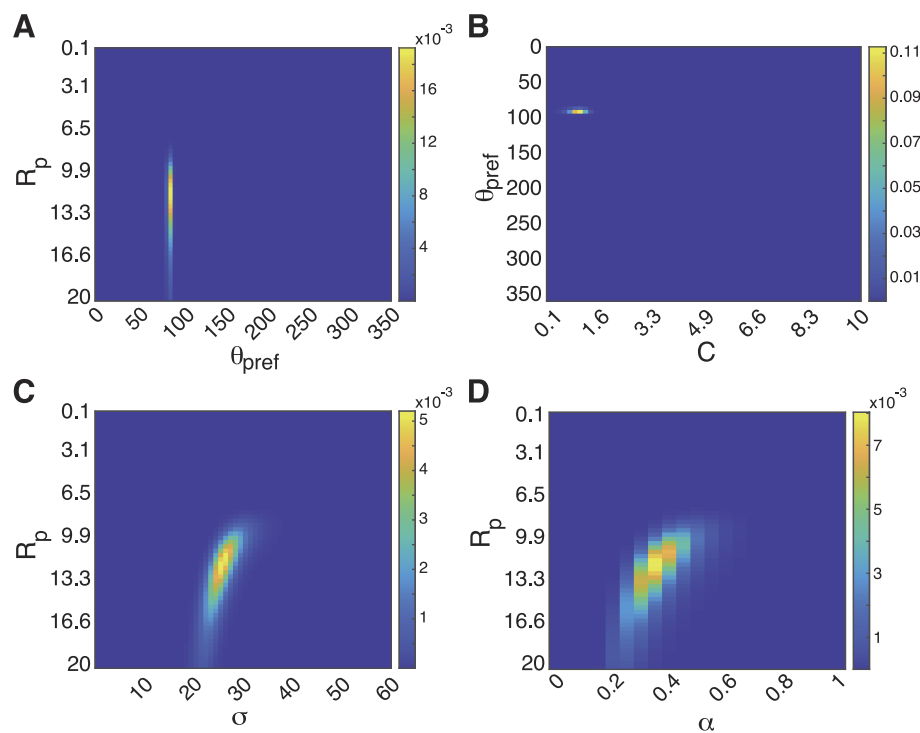


FIGURE 2

Views of joint likelihood across parameters for the strongly tuned cell in Figure 1. (A) Joint likelihood density plot of θ_{pref} and R_p , marginalized over all other parameters. The lack of a slant indicates that these likelihoods are independent and that the individual marginal likelihoods are a good summary of the whole. Scale bar indicates probability in density plot. (B) Joint likelihood of θ_{pref} and C . (C) Joint likelihood of R_p and α . Slant indicates some joint dependence of these parameters. The information in the slant is lost when summarizing only the marginal likelihoods of each parameter, and might need to be stored for some applications. (D) Joint likelihood of R_p and σ .

two parameters. For some parameters, such as between R_p and α (Figure 2C) or σ (Figure 2D), there is evidence of some correlation. Some scientists might be interested in this fine structure for some applications, and might need to store more information or to recompute it as needed.

3.3 Bayesian estimation for simulated temporal frequency

Many studies of orientation or direction selectivity in the visual cortex employ drifting sinusoidal gratings (with a particular spatial frequency and temporal frequency) as stimuli (e.g., Movshon et al., 1978). Some tuning properties, such as orientation tuning, are not very sensitive to spatial or temporal frequency, but others, such as direction selectivity, can change with stimulus parameters (Moore et al., 2005). For example, Moore et al. (2005) found that many cells in ferret visual cortex altered their direction preference angle or the amount of direction selectivity as temporal frequency was changed. Direction selectivity was often higher at low temporal frequencies than at higher temporal frequencies. If one wanted to study, say, how direction preference changed with temporal frequency, one would be out of luck with least squares approaches if selectivity became low. Here we will illustrate that the Bayesian approach allows one to study these properties quantitatively.

In Figure 3, we simulated changes in direction selectivity of a single neuron under three different temporal frequencies: low, medium, and high. In our simulations, we hypothesized that the direction selectivity is strongest at low temporal frequency ($C = 1$, $R_{pref} = 7$, $R_{null} = 1$, $\theta_{pref} = 45^\circ$, and $\sigma = 30^\circ$). The maximum response amplitude is slightly smaller at the medium temporal frequency ($C = 1$, $R_{pref} = 5$, $R_{null} = 5$, $\theta_{pref} = 45^\circ$, and $\sigma = 30^\circ$), and the direction selectivity is very low. At the highest temporal frequency ($C = 1$, $R_{pref} = 0$, $R_{null} = 0$, $\theta_{pref} = 45^\circ$, and $\sigma = 30^\circ$), responses are very weak and there is an absence of direction selectivity. As shown in Figures 3A–C, red represents the simulated sampled data points, the blue curve represents the simulated ideal curve for low temporal frequency, and yellow for medium temporal frequency, and green for high temporal frequency. To more closely replicate the actual fitting process, the number of sampling angles was reduced from 32 in Figure 1 to 16. Figures 4D–F shows the changes in the probability distributions of θ_{pref} , R_{pref} , and DI across the three different temporal frequencies, with the colors corresponding to the three temporal frequencies mentioned above. In Figure 3D, Bayesian estimation accurately identifies the preferred angle at low temporal frequency. At the medium temporal frequency, since the response rates at R_{pref} and R_{null} are close, the model indicates two probability peaks at θ_{pref} and θ_{null} . For the high temporal frequency, the probability distribution lacks distinct peaks. This is further supported in Figures 3E, F, where the probability peak of R_{pref}

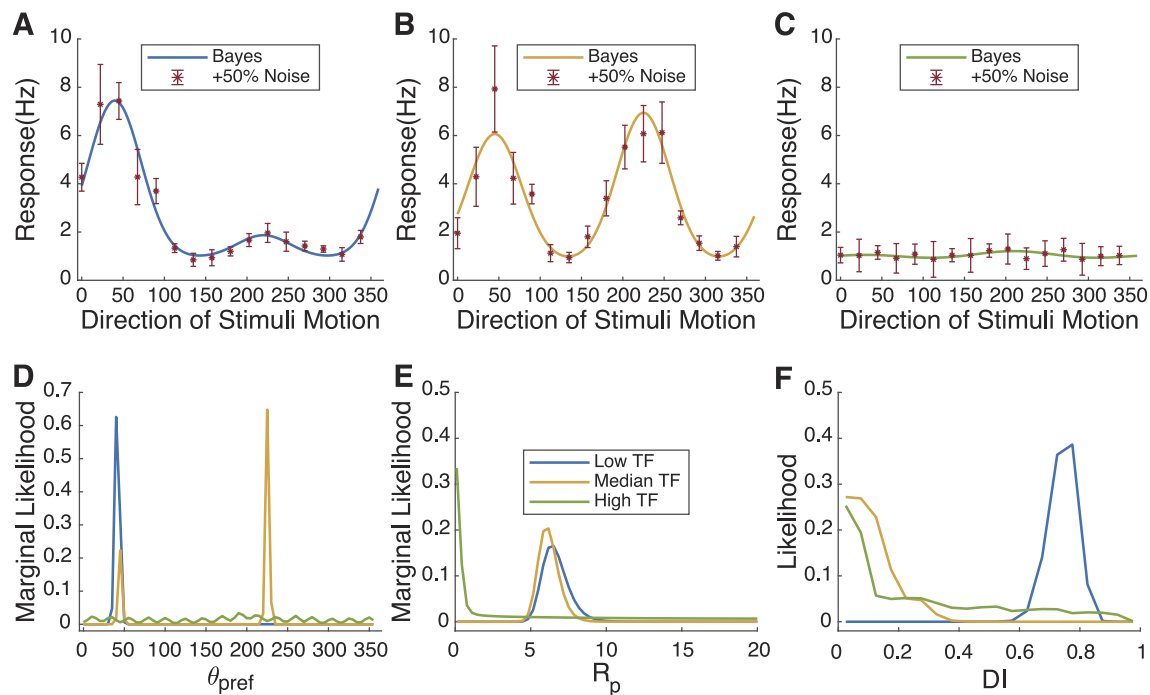


FIGURE 3
Neuronal responses and Bayesian estimation to stimuli at different temporal frequencies in simulated cell. (A–C) Simulated “real” response curves and sampling points with 50% added Gaussian noise in different temporal frequencies. (D–F) The probability distribution of θ_{pref} , R_p , and DI fitted for different temporal frequencies. The changes in the likelihood of θ_{pref} and DI track the increased uncertainty of θ_{pref} with increasing temporal frequency and the decreasing DI with temporal frequency.

and DI at high temporal frequency predominantly cluster near or at zero, indicating minimal direction selectivity. For the medium temporal frequency, the obvious peak around 6 for R_{pref} and the peak near zero for DI suggest that the ‘real’ response curve has significant orientation selectivity with similar values of R_{pref} and R_{null} . The parameters and direction index for all three temporal frequencies successfully approximate the ‘true values’.

3.4 Bayesian estimation fitting for simulated data with variable orientation index and direction index values

To better test the universality and accuracy of Bayesian estimation, we simulated two groups of neurons, each consisting of five cells with gradually increasing OI or direction index DI values, ranging from 0 to 1.

In [Figure 4](#) we simulated cells with increasing orientation selectivity index values; in [Figure 5](#), we simulated cells with increasing direction selectivity index values. In [Figure 4F](#), the likelihood of the underlying OI value tracks the true value, and θ_{pref} ([Figure 5I](#)) becomes more certain as the underlying orientation index increases. Similarly, in [Figure 4F](#), the likelihood of the underlying direction index tracks the true value, and again θ_{pref} is more certain as the direction selectivity increases.

The relative error between the distribution found by the Bayesian estimation procedure and the true value of the parameters that were used in the simulation is shown in [Figures 4K–O](#), [5K–O](#).

As expected, parameters R_p and C can be well estimated regardless of the orientation index, but the tuning width σ , angle preference θ_{pref} , and α are very uncertain when the orientation index is low. Similarly, when orientation selectivity is high and direction selectivity varies, most parameters can be well estimated except that the preferred angle θ_{pref} becomes highly uncertain when direction selectivity is very low.

Another way to quantify Bayesian parameter estimation is to ask how often the true value of the parameter distribution sits within the central $x\%$ of posterior distribution. One would expect that this would be true approximately $x\%$ of the time; in other words, that 50% of the time, the true value would lie within the central 50% of the posterior. In [Supplementary Figures 1, 2](#), we show how often the true parameter values in these simulations fall within the central 25th, 50th, and 75th percentiles of the posterior distributions. When the parameter can be well estimated (high orientation or direction selectivity), this is true, and performance falls slightly when the parameter cannot be well estimated.

3.5 Impact of sampling angle quantity on Bayesian estimation fitting results

Experimental time is extremely valuable, and it is important to know how many angle steps might be necessary to estimate the parameters to a desired precision. Each color in the [Figure 6](#) corresponds to a different number of sampling angles, with the legend indicating the number of angles used for each Bayesian

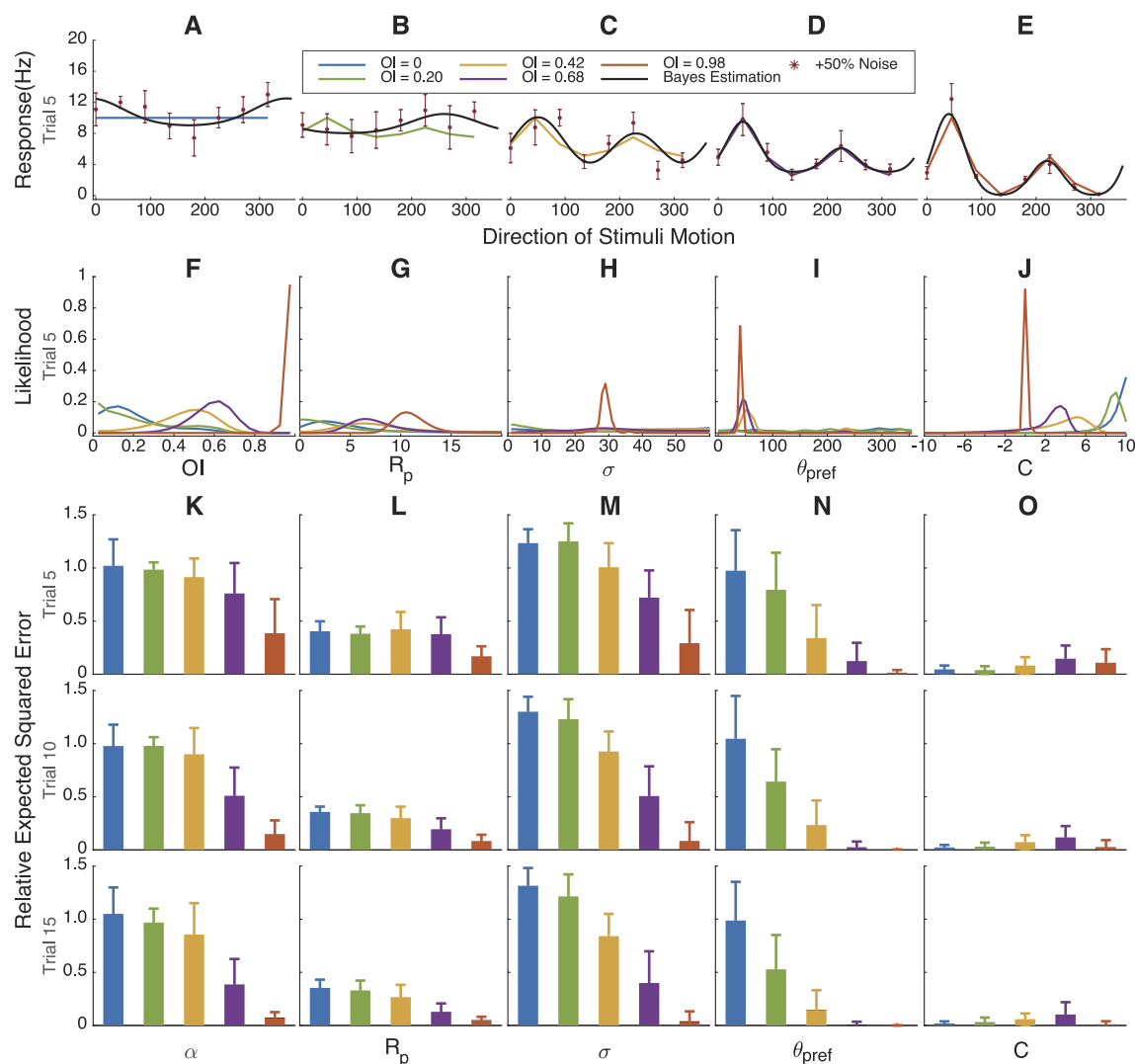


FIGURE 4

Bayesian estimation results for simulated cells with gradually increasing orientation index values. (A–E) Simulated response curve, sampling points with +50% Gaussian noise and fitting results. (F–J) The marginal likelihoods distribution of OI , R_p , σ , θ_{pref} and C . As direction selectivity progresses from high to low to non-existence, θ_{pref} , R_p , and DI track accordingly. (K–O). Relative error in parameter estimation compared to assuming a uniform distribution for each parameter. When orientation selectivity is high (red), we have certainty about the values, but when orientation selectivity is low, our certainty comes close to a uniform distribution (relative error of 1). 500 simulations of 5, 10, or 15 trials.

estimation fitting. The parameters of the simulated tuning curve are set as ($C = 2.5$, $R_{pref} = 7.5$, $R_{null} = 3.75$, $\theta_{pref} = 45^\circ$, and $\sigma = 30^\circ$).

From the likelihood distribution of OI in Figure 6F, it can be observed that four sampling angles produce a poor parameter estimation. There is also a significant likelihood of obtaining a fit with low orientation selectivity due to the accidental collection of angles where the response frequency in the “real” tuning curve is lower. Starting from eight sampling angles, a likelihood peak can be seen around 0.6, although it is somewhat broad. Similarly, in Figures 6G, J, higher sampling angles concentrate the probability peaks at specific values, and the fitting curves align well with the “real” tuning curve. This indicates that eight or more sampling angles meet a minimum requirement for Bayesian estimation.

However, under this parameter set, the probability distribution for σ remains broad, with an indistinct peak in the range from 20 to 30.

3.6 Bayesian estimation fitting results for neuronal response curves in the ferret primary visual cortex

Up to now, we have only examined simulated data for illustration. Next, we analyzed response tuning curves of primary visual cortex neurons from juvenile ferrets using calcium imaging from a prior study (Li et al., 2008). Five example neurons are

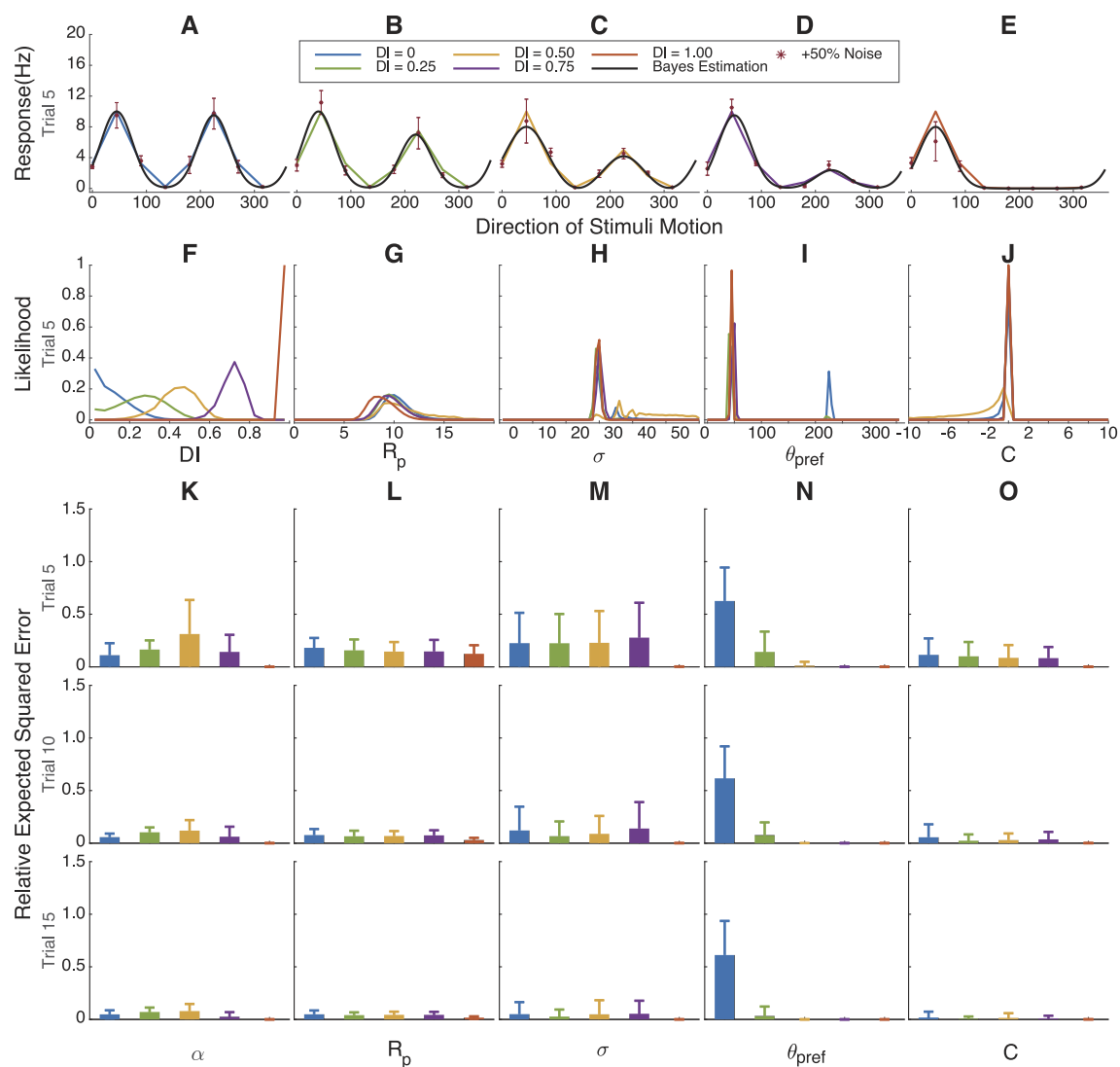


FIGURE 5

Bayesian estimation results for simulated cells with gradually increasing direction index values. (A–E) Simulated response curve, sampling points with +50% Gaussian noise and fitting results. (F–J) The marginal likelihoods distribution of DI , R_p , σ , θ_{pref} and C . (K–O) Relative error in parameter estimation compared to assuming a uniform distribution for each parameter. Because orientation selectivity is high in all cases, we have certainty about the values except for θ_{pref} . When direction selectivity approaches 0, then we have uncertainty as to the true angle preference θ_{pref} .

shown in Figures 7A–E. In Figure 7G, the likelihood distribution of R_{pref} lies between 0 to 0.2, with the green cell and purple cell showing smaller R_{pref} values. Nevertheless, all cells exhibit orientation selectivity with high confidence (Figures 7F, H–J), and have narrow ranges of preferred directions, with some cells showing a possibility of preferring the opposite direction (Figure 7I).

3.7 Bayesian parameter estimation compared to the bootstrap

Quantifying uncertainty in model parameters when the model is not a member of a well-known class, such as a mean or a generalized linear model, has been challenging. One common method is the bootstrap (Efron and Tibshirani, 1993; Press and

Flannery, 1992). Under this procedure, one generates hundreds or thousands of simulated surrogate datasets for each cell by drawing from the original data with replacement. For example, to simulate a surrogate dataset for a cell that was examined with N stimulus repetitions of S stimuli, one would, for each stimulus s , draw N response values from the N actual response values with replacement, so that some trials are drawn more than once and others drawn not at all. By performing a model fit, such as a least-squares fit, of the hundreds or thousands of surrogate datasets, one can build up a probability distribution of parameter values that reflect the certainty or uncertainty of that parameter.

At first glance, using Bayesian parameter estimation or using the bootstrap for least-squares fits might seem equally appropriate, and that the chief difference might be whether one wanted to

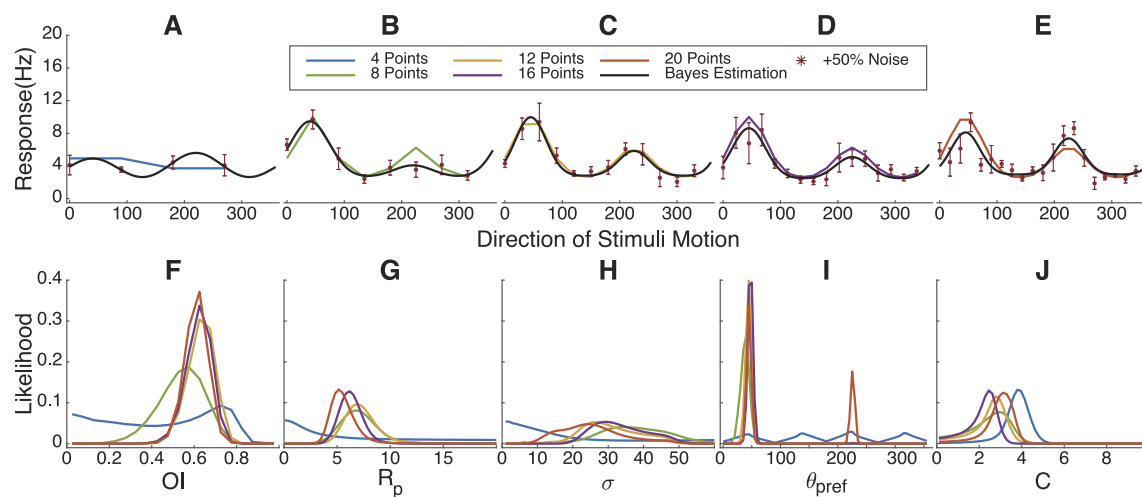


FIGURE 6

Impact of varying sampling angle quantities on Bayesian estimation results. (A–E) Simulated response curve, sampling points with +50% Gaussian noise and fitting results. (F–J) The marginal likelihoods distribution of OI , R_{pref} , σ , θ_{pref} and C . In order to estimate the tuning parameters well, one should sample at least 8 directions with visual stimulation.

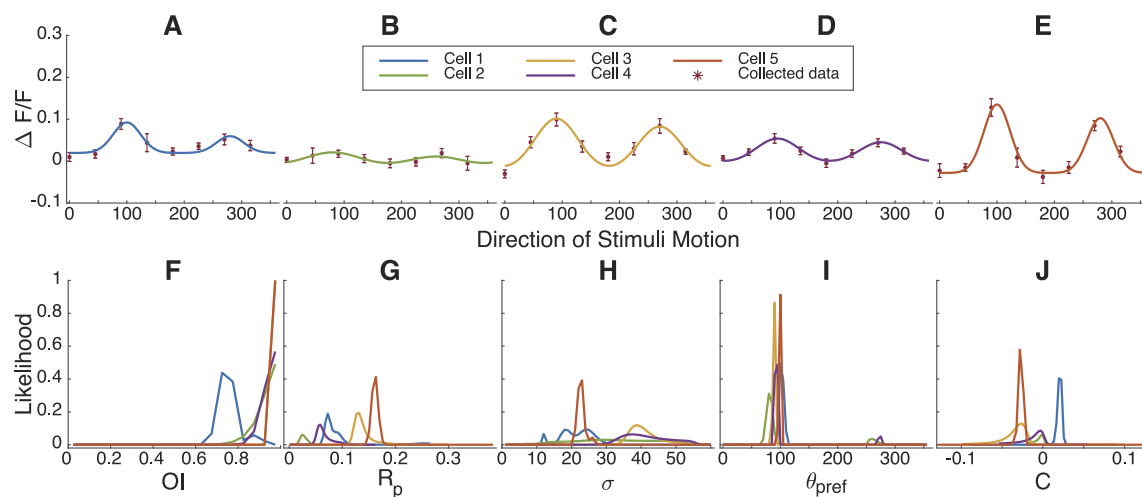


FIGURE 7

Bayesian parameter estimation of orientation and direction tuning for example cells recorded in ferret visual cortex. (A–E) The response of 5 example neurons from the juvenile ferret visual cortex to stimuli moving in different directions and Bayesian parameter estimation results. (F–J) The marginal likelihoods distribution of OI , R_{pref} , σ , θ_{pref} and C .

perform billions of forward calculations (where the calculations are not iterative and amenable to GPU methods) or thousands of least-squares fits (which are iterative and are less amenable to GPU methods). But we will argue that Bayesian parameter estimation is preferred because of the practical problems with least-squares fitting, including local minima and a frequent need to place constraints on parameters to aid with convergence (Mazurek et al., 2014).

We reanalyzed data from Li et al. (2008). In that work, the authors examined direction tuning in neurons of young ferrets before and after 3–6 hours of exposure to moving stimuli. In the naive state, cells generally exhibited relatively weak tuning for direction, whereas, after motion exposure, cells exhibited stronger

tuning. One question raised in that study was whether or not individual cells were likely to have had an eventual direction preference that (a) emerged from an unselective state, (b) was solidified from an initial bias toward that direction preference, or (c) was the result of a reversal of an initial preferred direction preference.

Figures 8A–F displays the Bayesian likelihood distributions of θ_{pref} for three sample cells before motion exposure and after motion exposure, illustrating changes in tuning properties. In Figures 8A, B, the preferred direction angle shifts 180 degrees, indicating a major and significant reorganization for that cell. Figures 8C, D shows an enhanced response strength at the original preferred angle, suggesting reinforcement of the cell's existing

tuning properties. Finally, [Figures 8E, F](#) reveals an emergence of a single direction preference, where previously the cell exhibited similar response frequencies in two opposite directions. These findings highlight the effectiveness of Bayesian estimation in capturing nuanced changes in neuronal tuning, similar to bootstrap analysis ([Figure 8D](#)).

The advantages of Bayesian parameter estimation become more apparent if we compare bootstrap and Bayesian parameter estimation across the population of 262 cells. Parameter R_p shows a monotonic relationship, indicating that both procedures do a similar job of identifying that parameter, and θ_{pref} is the same or, in a few cells that are presumably not very direction-selective, 180° apart. But, in order to converge to a meaningful fit, parameter σ was set at a floor of 22.5° because the angle step in this experiment was 45°. In many bootstrap least-squared fits ([Figure 8J](#)), this floor value was the best-fit value. However, in the Bayesian approach, we find that the MLE parameter value spans a broader range. Further, the bootstrap approach often returned the constrained minimal value for R_n , which was 0, in the bootstrap approach, whereas the MLE for the Bayesian parameter estimation spanned a wider range ([Figure 8K](#)).

Therefore, the Bayesian parameter estimation approach, which dutifully examines the likelihood of different combinations of parameters over many ranges, is immune from the local minima and constraint problems that can plague least-squares fitting. In Bayesian fitting, one has the burden of choosing an appropriate discretization of the parameter space and providing computational resources to carry out the computation, but with modern GPU methods these are not burdensome for relatively simple models.

3.8 Test of correctness on a different orientation model

Up to now, we have demonstrated our approach on tuning curves generated from the same function as the response model we are fitting. One may ask if it also does a good job on data that might be generated by similar but not identical processes. To examine the Bayesian parameter estimation performance under these conditions, we generated tuning curves from a von Mises process ([Equation 14](#)) while varying the preferred response from 0 to 10 spikes/sec and using a constant amount of noise of 2 spikes/sec across the tuning curve. Results are shown in [Figure 9](#). When A is high, θ_{pref} and R_p are well estimated by the approach. As R_p drops to 0, the estimate of θ_{pref} deteriorates as expected, indicating our uncertainty in determining the true θ_{pref} from the data.

3.9 Noise model

The Bayesian approach requires that one specify the amount of trial-to-trial noise that is expected in the model. As described in the methods, we started from the prior observations that in cortical neurons the variance or standard deviation of the response across trials is proportional to the mean of the response across trials ([Softky and Koch \(1993\)](#) and [Shadlen and Newsome \(1998\)](#) on a

log-log plot. We extended this slightly to include a constant term: $\sigma_{noise}(m) = C_n + K * m^S$.

We found that this function characterized the relationship between mean response and noise for a wide range of recordings, including 2-photon calcium imaging using Oregon Green BAPTA-1 AM from the previous section ([Li et al., 2008](#)), shown in [Figure 10A](#) and extracellular records from ferret (Van Hooser, unpublished observations) ([Figure 10B](#)). To further explore how this noise model might suite different applications, we examined several cells from the Allen Brain Observatory's 2-photon dataset ([de Vries et al., 2020](#)), where somatostatin-positive (SST) interneurons ([Figure 10C](#)), vasoactive intestinal peptide-positive (VIP) interneurons ([Figure 10E](#)), parvalbumin-positive (PV) interneurons ([Figure 10D](#)), and pan-excitatory cells (Ex) ([Figure 10F](#)) were studied in different gCamp6 mouse lines. All cell types were well described by the equation ([Figures 8C–F](#)), but, interestingly, VIP interneurons exhibited noise that was largely independent of the mean response rate such that the K and S parameters were unnecessary. Nevertheless, the strategy of fitting [Equation 7](#) to obtain noise parameters would succeed for all of these cell types.

4 Discussion

We derived a Bayesian estimation approach for orientation and direction tuning for neurons or other processes. Least-squares fitting and vector approaches do a good job of identifying model parameters when there is some substantial selectivity, but do not readily produce a measure of certainty or uncertainty when tuning is poor ([Mazurek et al., 2014](#)). Bayesian parameter estimation explicitly calculates the likelihood of different parameter values. Until recently, the billions of calculations needed to estimate these parameter distributions were an impediment to the broad use of Bayesian methods, but the new availability of fast GPUs that can perform trillions of floating-point calculations per second makes this feasible even for large populations of neurons.

4.1 Modifications of least-squares models

In order to apply Bayesian methods to tuning curves, it may be necessary to modify the original least-squares tuning model, as we did here. The traditional form of the double Gaussian equation ([Mazurek et al., 2014](#)) uses independent parameters for the preferred and null responses, and, for least-squares fitting, one can just take the larger of the two to be the preferred response parameter and the smaller to be the null parameter. However, in the Bayesian approach, it is necessary to specify the preferred angle, and we modified the null parameter to be an α times the preferred.

4.2 Selecting what to save

Another critical choice in full Bayes parameter estimation is choosing what information to save. Saving the results of billions of computations to disk is not feasible, but one can save smaller descriptions. Here, we saved the most likely parameters and the

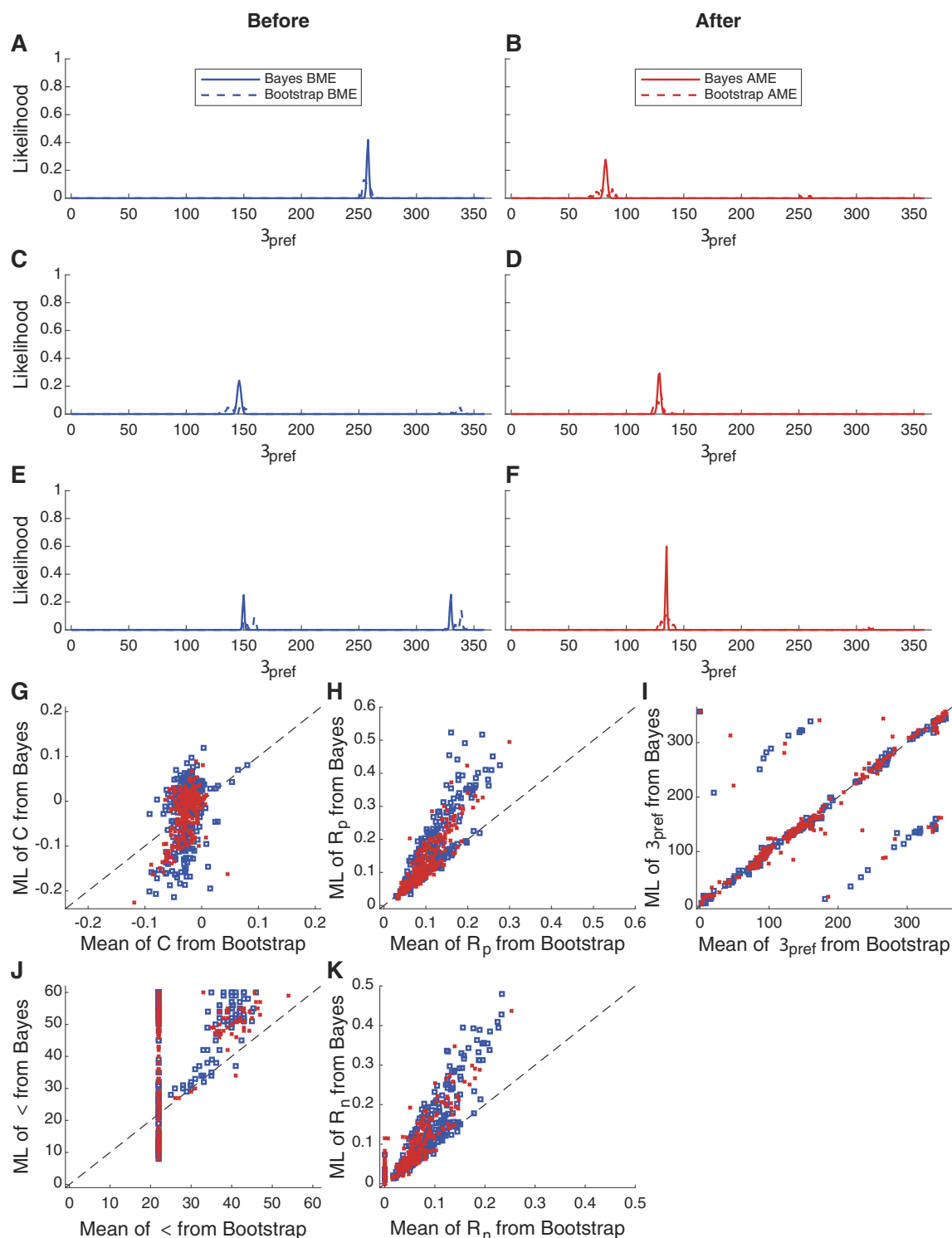


FIGURE 8

Bayesian estimation and bootstrap fitting results before and after motion exposure. (A–F) Probability distribution fitting results for the parameter θ_{pref} in three representative cells before and after motion exposure. Results from measurements made before motion exposure are shown in blue while results from measurements made after motion exposure are shown in red. (G–K) Fitting results of individual parameters across 262 cells using different models. Each panel shows the relationship between the Bootstrap mean (x-axis) and the Bayesian maximum likelihood (ML) estimate (y-axis) for various model parameters, including OL , R_{pref} , σ , θ_{pref} and C . In general, the results are proportional and the results for θ_{pref} are either identical or are occasionally shifted by 180° . However, results for parameter σ are greatly impacted by the floor of the fitting procedure employed with the bootstrap (see Mazurek et al., 2014 for the necessity of this constraint), whereas the Bayesian parameter estimation has no such issue.

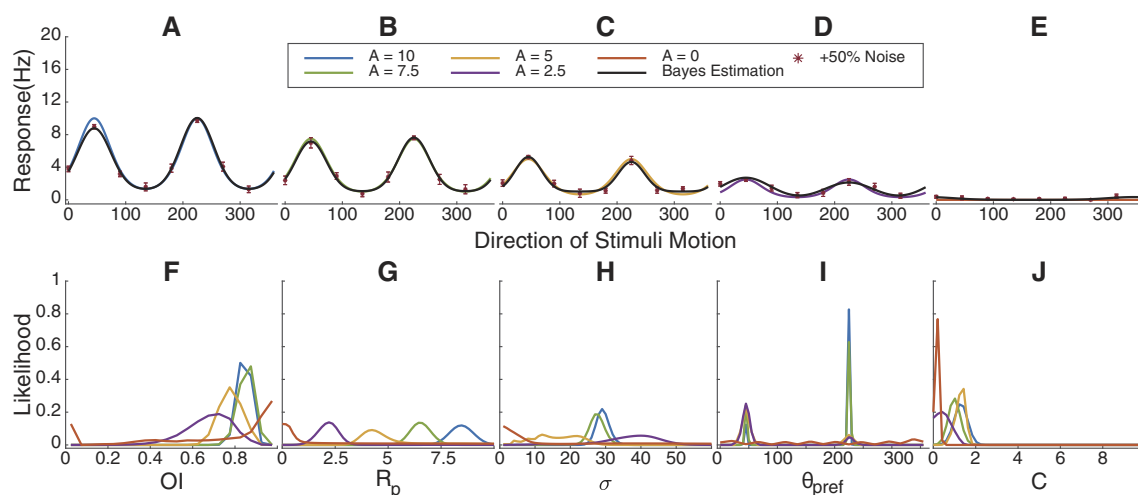


FIGURE 9

Bayesian parameter estimation on von Mises tuning curves. (A–E) Simulated tuning curves from a von Mises function with maximum firing rate (A) varied from 10 to 0 in steps of 2.5 sp/s. Data points are tuning curves disturbed by 50% noise and black lines are most-likely values of parameters determined by Bayesian parameter estimation. (F) Orientation index (OI) estimations show high values until $A = 0$. (G) Likelihoods for R_p in the Bayesian estimation closely match values of A . (H) σ . (I) θ_{pref} matches the orientation modulo 180° . (J) Estimates for C .

marginal probabilities of each parameter, but also saved a histogram of several index parameters that were calculated on the full joint distribution. These included the orientation and direction selectivity index values and the circular variance in both orientation and direction space. If one needed the full joint distribution for any reason, it can be recomputed quickly for a single tuning curve, while the marginal probability and index histograms can be saved in a database for aggregate analysis across cells.

4.3 Uncertainty and an alternative to the bootstrap

When least-squares methods dominated computation, a common technique to understand the uncertainty of each parameter was the bootstrap (Press and Flannery, 1992). The analyst would create hundreds or thousands of surrogate datasets by drawing individual observations from the originals with replacement, and perform the least-squares fitting on the surrogate datasets. In this way, the analyst would have a distribution of hundreds or thousands of potential values of the fit parameters, and could analyze the uncertainty. Bayesian parameter estimation performs a very similar function, in that one obtains a probability distribution for all parameters, and one can assess uncertainty of all parameters.

This raises the question as to which method might be preferred, Bayesian parameter estimation or least-squares fitting with bootstrap? The downside to Bayesian parameter estimation is a need to specify the values over which to perform the estimation. The resolution of the answer will be exactly that specified. With least-squares fitting, the best-fitting parameters are found to high resolution (often to many significant digits). However, least-squares fits often require constraints in order to make sure the results have physical meaning. In the case of orientation/direction fitting,

we usually place a constraint on the tuning width: it can be no less than 1/2 of the angle steps used for visual stimulation (e.g., if the visual stimulus was stepped in 30° steps, then σ must be at least 15°). The results of Figure 8 indicate that this is exactly what happened in our prior analysis, and the Bayesian parameter estimation approach would be preferred. Further, if there were some stereotypy in the fitting procedure that favors some values over others (one might imagine that certain θ_{pref} values might emerge for noisy input more often than others if the code were written in a certain way), then this stereotypy will skew the results of the bootstrap approach when the input signal is weak or noisy. Therefore, if the parameters can be specified at a suitable resolution for the problem, we recommend the Bayesian parameter estimation approach instead of the bootstrap.

4.4 When it's not worth the effort

The approach we describe here allows one to characterize a variety of properties of orientation and direction tuning, including peak response, preferred angle, and the background activity. If one only wants to assess the degree of orientation or direction tuning, then vector methods such as the circular variance (Mazurek et al., 2014; Ringach et al., 2002) are completely suitable [see simulations in Mazurek et al. (2014)].

4.5 Comparison to Bayesian decoding

Bayesian parameter estimation and Bayesian decoding are essentially the same approach that use the same mathematical equations but with slightly different goals. In Bayesian parameter estimation, one is applying Bayesian methods to learn the parameters of a model. In Bayesian decoding, one is often using a

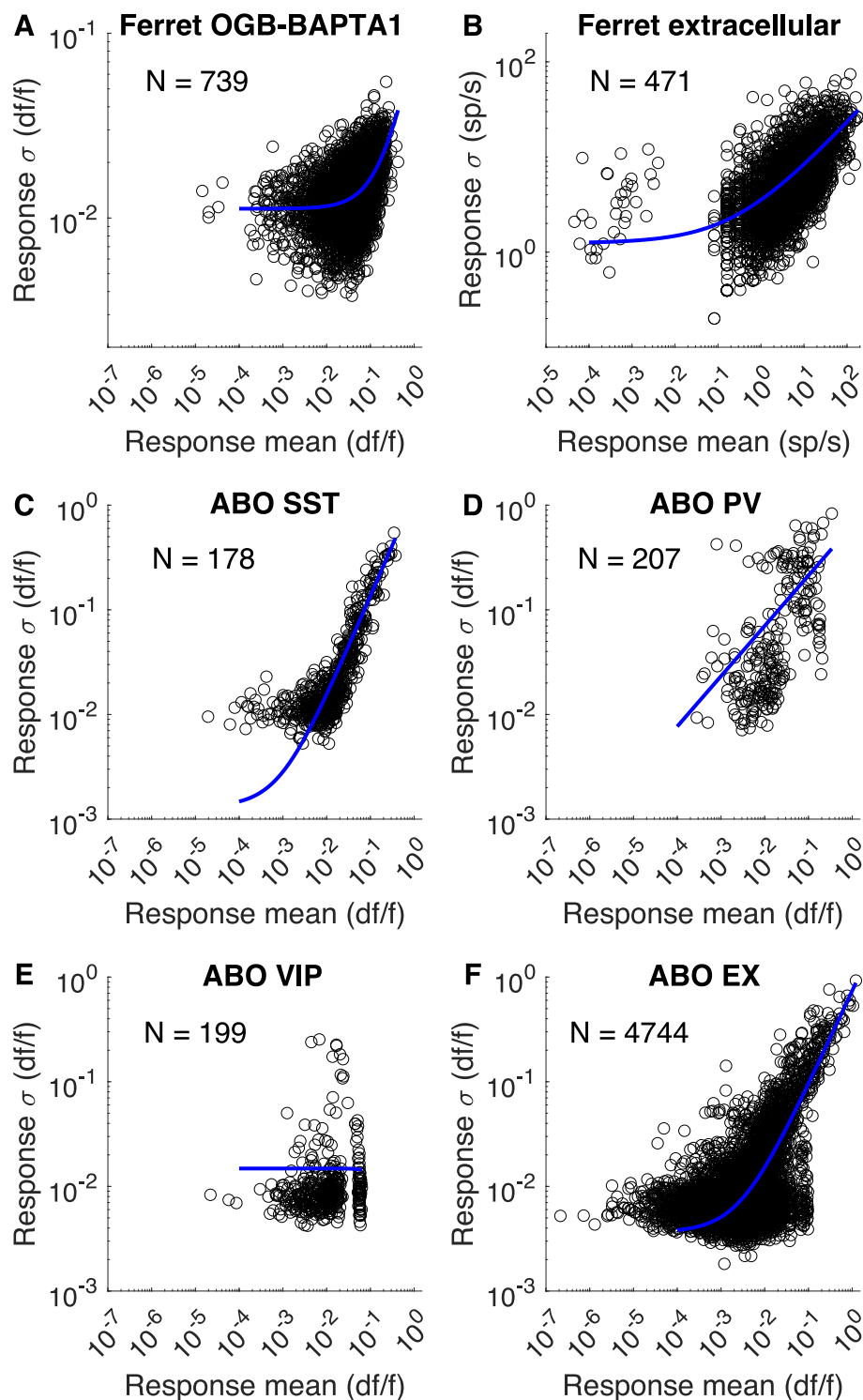


FIGURE 10

Noise models for different recordings. (A) Trial-by-trial response means and standard deviations for 739 ferret neurons labeled with Oregon Green BAPTA-1 AM (Li et al., 2008). 8 responses from each neuron are plotted (different stimulus directions). The best-fit model was $C_n = 0.011$, $K = 0.0715$, $S = 1.14$. (B) Unpublished recordings from Van Hooser lab showing mean and standard deviations of responses to visual stimuli for 471 neurons from ferret visual cortex. $C_n = 1.24$, $K = 2.31$, $S = 0.492$ (C–F) 2-photon calcium imaging responses from different mouse lines expressing gCamp6f and reported in the Allen Brain Observatory (de Vries et al., 2020). Cell numbers as indicated. For SST: $C_n = 0.00131$, $K = 1.26$, $S = 0.972$. For PV: $C_n = 1.9e-8$, $K = 0.643$, $S = 0.481$. For VIP: $C_n = 0.0148$, $K = -5.61$, $S = 3.38$. For EX: $C_n = 0.00364$, $K = 0.738$, $S = 0.901$.

model that has already been learned in order to make an inference about some feature of a system, such as the position of an animal given activity in its hippocampal place cells (Brown et al., 1998; Zhang et al., 1998; Karlsson and Frank, 2009; Davidson et al., 2009; Shin et al., 2019). Just as in Bayesian parameter estimation, one can report the “most likely” outcome (for example, the most likely position of an animal) or compute an entire distribution (for example, a distribution of likelihoods that the animal is at any given position in an environment).

4.6 Summary

We recommend the Bayesian parameter estimation approach for sensory tuning curves or other response model when it is important to understand the parameter values of the response model and the certainty/uncertainty in these parameters. We recommend this approach above least-squares for finding the most-likely mean values of the parameters, and we recommend this approach above the bootstrap for understanding parameter uncertainty in most cases. We imagine that many studies that seek to characterize the tuning of large numbers of neurons will shortly transition away from least-squares approaches to Bayesian parameter estimation. This article has described a recipe for making the switch.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author. Code and data are available at <https://github.com/VH-Lab/vhlab-bayesoridir-matlab>

Author contributions

ZW: Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review &

editing. SV: Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncir.2025.1542332/full#supplementary-material>

References

- Batschelet, E. (1981). *Circular Statistics in Biology*. New York, NY: Academic Press.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Cham: Springer.
- Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C., and Wilson, M. A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* 18, 7411–7425. doi: 10.1523/JNEUROSCI.18-18-07411.1998
- Carandini, M., and Ferster, D. (2000). Membrane potential and firing rate in cat primary visual cortex. *J. Neurosci.* 20, 470–484. doi: 10.1523/JNEUROSCI.20-01-00470.2000
- Cronin, B., Stevenson, I., Sur, M., and Körding, K. (2010). Hierarchical Bayesian modeling and Markov chain Monte Carlo sampling for tuning-curve analysis. *J. Neurophysiol.* 103, 591–602. doi: 10.1152/jn.00379.2009
- Davidson, T. J., Kloosterman, F., and Wilson, M. A. (2009). Hippocampal replay of extended experience. *Neuron* 63, 497–507. doi: 10.1016/j.neuron.2009.07.027
- de Vries, S. E. J., Lecoq, J. A., Buice, M. A., Groblewski, P. A., Ocker, G. K., Oliver, M., et al. (2019). A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat. Neurosci.* 22, 1941–1950.
- de Vries, S. E. J., Lecoq, J. A., Buice, M. A., Groblewski, P. A., Ocker, G. K., Oliver, M., et al. (2020). A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat. Neurosci.* 23, 138–151. doi: 10.1038/s41593-019-0550-9
- Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Number 57 in *Monographs on Statistics and Applied Probability*. Boca Raton, FL, United States: Chapman and Hall/CRC. doi: 10.1007/978-1-4899-4541-9
- Henry, G. H., Dreher, B., and Bishop, P. O. (1974). Orientation specificity of cells in cat striate cortex. *J. Neurophysiol.* 37, 1394–1409. doi: 10.1152/jn.1974.37.6.1394
- Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148, 574–591. doi: 10.1113/jphysiol.1959.sp006308
- Hubel, D. H., and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154. doi: 10.1113/jphysiol.1962.sp006837
- Hubel, D. H., and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195, 215–243. doi: 10.1113/jphysiol.1968.sp008455

- Karlsson, M. P., and Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* 12, 913–918. doi: 10.1038/nn.2344
- Li, Y., Van Hooser, S. D., Mazurek, M., White, L. E., and Fitzpatrick, D. (2008). Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature* 456, 952–956. doi: 10.1038/nature07417
- Mazurek, M., Kager, M., and Van Hooser, S. D. (2014). Robust quantification of orientation selectivity and direction selectivity. *Front. Neural Circ.* 8. doi: 10.3389/fncir.2014.00092
- McNaughton, B. L., Barnes, C. A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. *Exp. Brain Res.* 52, 41–49. doi: 10.1007/BF00237147
- Moore, 4th, B. D., Alitto, H. J., and Usrey, W. M. (2005). Orientation tuning, but not direction selectivity, is invariant to temporal frequency in primary visual cortex. *J. Neurophysiol.* 94, 1336–1345. doi: 10.1152/jn.01224.2004
- Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J. Physiol.* 283, 101–120. doi: 10.1113/jphysiol.1978.sp012490
- Nauhaus, I., Nielsen, K. J., and Callaway, E. M. (2012). Nonlinearity of two-photon Ca^{2+} imaging yields distorted measurements of tuning for v1 neuronal populations. *J. Neurophysiol.* 107, 923–936. doi: 10.1152/jn.00725.2011
- Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., and Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433, 597–603. doi: 10.1038/nature03274
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C*. Cambridge: Cambridge University Press.
- Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque v1: diversity and laminar dependence. *J. Neurosci.* 22, 5639–5651. doi: 10.1523/JNEUROSCI.22-13-05639.2002
- Shadlen, M. N., and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896. doi: 10.1523/JNEUROSCI.18-10-03870.1998
- Shin, J. D., Tang, W., and Jadhav, S. P. (2019). Dynamics of awake hippocampal-prefrontal replay for spatial learning and memory-guided decision making. *Neuron* 104, 1110–1125.e7. doi: 10.1016/j.neuron.2019.09.012
- Siegle, J. H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., et al. (2021). Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* 590, 224–229. doi: 10.1038/s41586-020-03171-x
- Softky, W., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J. Neurosci.* 13, 334–350. doi: 10.1523/JNEUROSCI.13-01-00334.1993
- Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature* 571, 361–365. doi: 10.1038/s41586-019-1346-5
- Swindale, N. V. (1998). Orientation tuning curves: empirical description and estimation of parameters. *Biol. Cybern.* 78, 45–56. doi: 10.1007/s004220050411
- Taube, J., Muller, R., and Ranck, J. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. description and quantitative analysis. *J. Neurosci.* 10, 420–435. doi: 10.1523/JNEUROSCI.10-02-00420.1990
- Wilson, D. E., Scholl, B., and Fitzpatrick, D. (2018). Differential tuning of excitation and inhibition shapes direction selectivity in ferret visual cortex. *Nature* 560, 97–101. doi: 10.1038/s41586-018-0354-1
- Zhang, K., Ginzburg, I., McNaughton, B. L., and Sejnowski, T. J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* 79, 1017–1044. doi: 10.1152/jn.1998.79.2.1017