LETTER Communicated by Robert Shapley

Sample Skewness as a Statistical Measurement of Neuronal Tuning Sharpness

Jason M. Samonds
samondjm@cnbc.cmu.edu
Center for the Neural Basis of Cognition and Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.

Brian R. Potetz
potetz@google.com
Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, U.S.A.

Tai Sing Lee
tai@cnbc.cmu.edu
Center for the Neural Basis of Cognition and Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.

We propose using the statistical measurement of the sample skewness of the distribution of mean firing rates of a tuning curve to quantify sharpness of tuning. For some features, like binocular disparity, tuning curves are best described by relatively complex and sometimes diverse functions, making it difficult to quantify sharpness with a single function and parameter. Skewness provides a robust nonparametric measure of tuning curve sharpness that is invariant with respect to the mean and variance of the tuning curve and is straightforward to apply to a wide range of tuning, including simple orientation tuning curves and complex object tuning curves that often cannot even be described parametrically. Because skewness does not depend on a specific model or function of tuning, it is especially appealing to cases of sharpening where recurrent interactions among neurons produce sharper tuning curves that deviate in a complex manner from the feedforward function of tuning. Since tuning curves for all neurons are not typically well described by a single parametric function, this model independence additionally allows skewness to be applied to all recorded neurons, maximizing the statistical power of a set of data. We also compare skewness with other nonparametric measures of tuning curve sharpness and selectivity. Compared to these other nonparametric measures tested, skewness is best used for capturing the

Brian Potetz is now at Google, Los Angeles, CA.

Neural Computation **26, 860–906** (2014)
doi:10.1162/NECO_a_00582
© 2014 Massachusetts Institute of Technology

sharpness of multimodal tuning curves defined by narrow peaks (maximum) and broad valleys (minima). Finally, we provide a more formal definition of sharpness using a shape-based information gain measure and derive and show that skewness is correlated with this definition.

1 Introduction

Since Adrian and Zotterman (1926) discovered that the number of action potentials recorded from a muscle nerve fiber varied with the amount of force applied to the muscle, neurophysiologists have been characterizing neurons throughout the nervous system by the changes in their mean firing rates of action potentials to changes in sensory input or motor output characteristics. For many neurons, their firing rates vary for a particular parameter or feature of the sensory input or motor output and fire maximally for one particular value of that feature. That one value is expressed as the preferred value of that feature for a neuron, and a plot of mean firing rates across all possible values of that feature is known as a tuning curve.

Tuning curves throughout the brain can be represented by mathematical functions. Most tuning curves, such as orientation tuning in primary visual cortex (V1) and direction of motion tuning in area MT, are represented by a gaussian function (Henry, Bishop, Tupper, & Dreher, 1973; DeAngelis & Uka, 2003). For neurons in motor cortex, tuning curves of direction of movement are fit with a cosine function to capture the shape of the peak (Georgopoulos, Kalaska, Caminiti, & Massey, 1982). Finally, neurons tuned for horizontal binocular disparity in visual areas have disparity tuning curves that are best described by Gabor functions (Ohzawa, DeAngelis, & Freeman, 1990; Hinkle & Connor, 2001; Prince, Pointon, Cumming, & Parker, 2002; DeAngelis & Uka, 2003).

Neurophysiologists also often characterize the sharpness of these tuning curves. Sharpness is generally defined as whether a tuning curve is broad or narrow, with a narrower tuning curve being described as sharper. The motivation for characterizing sharpness is based on the presumption that a neuron with a sharper tuning curve describes the stimulus with greater specificity and precision than a neuron with a broader tuning curve. Changes in sharpness are frequently observed for tuning curves of many features, especially for the neurons involved in processing incoming sensory information. For example, studies in the visual and auditory systems find that tuning curves in general become sharper as you progress from cortical input layers to output layers (Blasdel & Fitzpatrick, 1984; Fitzpatrick, Batra, Stanford, & Kuwada, 1997). Some tuning curves for individual neurons for several different visual and auditory features sharpen over time (Ringach, Hawken, & Shapley, 1997; Suga, Zhang, & Yan, 1997; Bredfeldt & Ringach, 2002; Menz & Freeman, 2003; Samonds, Potetz, & Lee, 2009; Samonds, Potetz, Tyler, & Lee, 2013). Sharpness can also increase with increasing stimulus size (Chen, Dan, & Li, 2005; Xing, Shapley, Hawken, &

Ringach, 2005; Samonds et al., 2013). In addition, increases in attention and training can sharpen tuning curves in higher visual areas and are correlated with improved behavioral performance (Spitzer, Desimone, & Moran, 1988; Freedman, Riesenhuber, Poggio, & Miller, 2006). Finally, recent experiments have revealed that suppressing a particular class of inhibitory interneurons leads to broader orientation tuning and diminished performance in an orientation discrimination task, providing even stronger evidence that tuning sharpness is behaviorally relevant (Lee et al., 2012).

Recently we found that disparity tuning sharpened over time, sharpened with increasing stimulus size, and had decreased sharpness for binocular anticorrelation stimulation in the macaque primary visual cortex and showed that recurrent interactions among disparity-tuned neurons could account for the changes in sharpness (Samonds et al., 2013). Sharpness has traditionally been quantified with bandwidth measures of the tuning curve (Albright, 1984; Bredfeldt & Ringach, 2002). More recent studies fit the data to functions and define sharpness by one of the parameters of the function. The sigma term or variance in the gaussian function closely resembles bandwidth (Henry et al., 1973) and frequency terms in sinusoidal or Gabor functions can capture narrowing tuning curve peaks (Menz & Freeman, 2003). Unlike gaussian functions used to describe orientation tuning curves for V1 neurons, Gabor functions of disparity tuning introduce some difficulties in quantifying sharpness. Disparity tuning curves can have multiple peaks and valleys, so if we define sharper disparity tuning as conveying more information and precision about disparity, then greater sharpness is not necessarily limited to a narrowing of a peak or best described by an increase in frequency. Characteristics of increased sharpness of disparity tuning include narrowing of the peak around the preferred disparity, broadening of valleys away from the preferred disparity, and suppression of secondary peaks (see Figure 1A; Samonds et al., 2009, 2013). Parameters of Gabor functions do not adequately capture all of this behavior because the observed function of disparity is actually deviating from a Gabor with increasing sharpness.

A simpler approach has been to use nonparametric methods to quantify the sharpness and selectivity for complex tuning curves, since these measures do not require modeling tuning curves for a diverse set of neurons with a single function (Ringach et al., 1997; Rolls & Tovee, 1995; Moody, Wise, di Pellegrino, & Zipser, 1998; Lehky, Sejnowski, & Desimone, 2005; Freedman et al., 2006). We propose using the nonparametric standard statistical measurement of the sample skewness of the distribution of mean firing rates over disparity to quantify sharpness. In this letter, we demonstrate how skewness quantitatively captures the characteristics of disparity tuning sharpening described in Figure 1A and sharpness of other features that have simple and complex tuning functions. We also illustrate how increases in skewness are related to changes in parameters for Gabor functions typically used to describe disparity tuning and difference-of-gaussians functions that have been used to capture changes in sharpness. We discuss

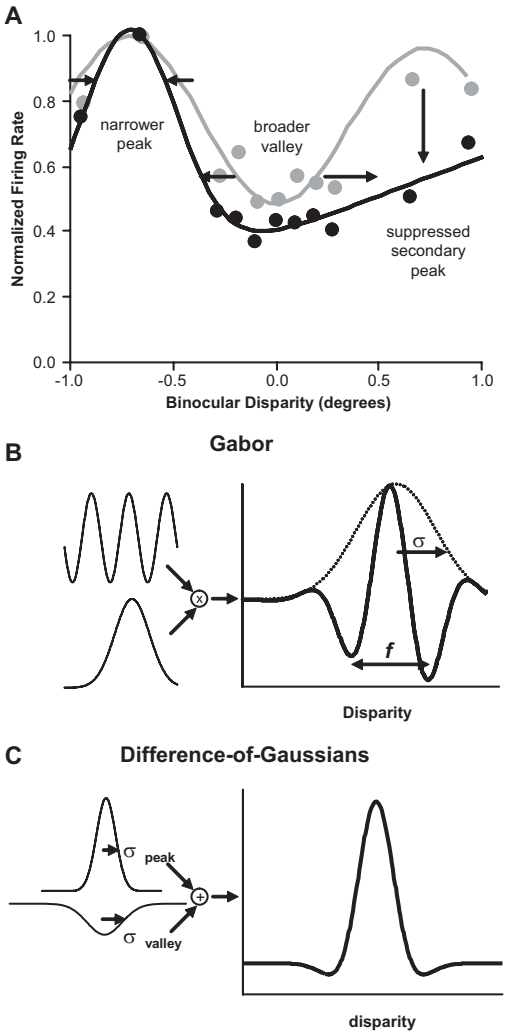


Figure 1: An example of the sharpening of disparity tuning and deviation from a Gabor function observed in V1 neurons (Samonds et al., 2009, 2013). (A) Tuning measured later in the response (black data and difference-of-gaussians fit, 450–850 ms from stimulus onset) had narrower peaks, broader valleys, and suppressed secondary peaks compared to disparity tuning measured early in the response (gray data and Gabor fit, 150–250 ms from stimulus onset). (B) A Gabor function is commonly used to describe disparity tuning curves and is a combination of a sinusoidal component with a gaussian envelope. (C) A difference-of-gaussians function is commonly used to describe tuning curves that evolve from recurrent interactions among a population of tuned neurons and combines two gaussians with opposite sign.

when and why skewness is better suited for analyzing sharpness than these parametric approaches. We also show how increases in skewness are related to other statistical measurements and when and why skewness is a better measure. Finally, we formulate a clear and quantitative definition of sharpness with a shape-based information gain measure of estimating features from a tuning curve and demonstrate that skewness is highly correlated to this measure when applied to a variety of potential tuning curves. All of these qualities of skewness make it an ideal measure for quantifying the sharpness of tuning for a wide range of complex tuning functions for neurons throughout the nervous system.

2 Materials and Methods

2.1 Neurophysiological and Model Data. The disparity tuning data for this study were robust tuning curves taken from Samonds et al. (2013), which used surgical and recording procedures that were approved by the Institutional Animal Care and Use Committee of Carnegie Mellon University and are in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. The data were collected simultaneously with data reported in two previous articles where the details about the specific methods can be found (Samonds et al., 2009; Samonds, Potetz, & Lee, 2012). The first recording procedure used for two (male and female) rhesus monkeys (*Macaca mulatta*) used two to eight tungsten-in-epoxy and tungsten-in-glass microelectrodes in a chamber overlying the operculum of V1 (Samonds et al., 2009). The second recording procedure used on the third monkey (male) used a chronically implanted 10 × 10 Utah Intracortical Array (400-*mu* m spacing) inserted to a depth of 1 mm in V1 (Samonds et al., 2012). Dynamic random dot stereograms (DRDS) with 25% density, a 12 Hz refresh rate, and disparities between corresponding dots of ±0.94, ±0.658, ±0.282, ±0.188, ±0.094, and 0 degrees (10–60 repeats of each disparity) were presented for 1 second within a 3.5 degree aperture to fixating monkeys to measure disparity tuning over time.

A disparity tuning curve of a representative example model neuron was also used from Samonds et al. (2013), where details about the model can be found. In brief, the neuronal network model had a single layer of disparity energy model neurons (Ohzawa et al., 1990) with recurrent connections among binocular disparity-tuned neurons that represented what has been inferred based on cross-correlation results (Menz & Freeman, 2003; Samonds et al., 2009). The feedforward disparity tuning curve can be shown to be approximately Gabor as a function of disparity. Within a spatial location, the weight of recurrent connections between neurons was chosen to be proportional to the Pearson correlation between the feedforward tuning curves of those two neurons, but using thresholded tuning curves to approximate the reduced and sparse response to natural scenes. All neurons within a spatial location were interconnected. Across spatial locations,

neurons were connected with positive weighting set by a spatial gaussian only if they had matching disparity and spatial frequency preferences. The input response for the model neuron is the tuning curve based on the disparity energy model (Ohzawa et al., 1990), and the steady-state response is the tuning curve measured after applying several iterations of the recurrent inputs using a standard dynamic neural field model.

The orientation tuning data for this study were collected from two (male and female) rhesus monkeys using the 10 × 10 Utah array method described above (same male monkey) and a semichronic recording chamber (Salazar, Dotson, Bressler, & Gray, 2012) implanted overlying the operculum of the V1/V2, which was approved by the Institutional Animal Care and Use Committee of Carnegie Mellon University and in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. The chamber has 32 independently movable tungsten-in-glass microelectrodes. Drifting sinusoidal gratings with 100% Michelson contrast, 6.25 Hz temporal frequency, 1.3 cycles per degree spatial frequency, and orientation increments of 22.5 degrees over 360 degrees (covering the entire range of orientation at both drift directions; 10–50 repeats of each orientation) were presented for 1 second within a 4 degree aperture to fixating monkeys to measure orientation tuning over time for V1 and V2 neurons.

The object tuning data analyzed in this study were from a previous study by Freedman et al. (2006) and were collected from two female rhesus monkeys using a recording chamber implanted over the inferior temporal cortex (ITC) using procedures approved by the MIT Committee on Animal Care guidelines and in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Single-electrode recordings were made in ITC while randomly chosen pictures and photographs were presented from the Corel Image Library. For training, monkeys had to perform a delayed match-to-category task using a large set of morphed images based on three cat and three dog prototypes (Freedman, Riesenhuber, Poggio, & Miller, 2001). Monkeys had to indicate whether two successively presented stimuli were from the same cat or dog category. For recordings, and the tuning curves analyzed in this letter, 18 cat and dog stimuli from each of six levels of cat/dog blends (100:0, 80:20, 60:40, 40:60, 20:80, and 0:100) along the three morph lines that crossed the category boundary (Freedman et al., 2006) were presented to monkeys for 0.6 seconds at seven image-plane rotations relative to the trained orientation (0, 22.5, 45, 67.5, 90, 135, and 180 degrees) while monkeys performed a simple fixation task. Tuning curves for the 18 images were computed within a 100 ms window starting from 80 ms from stimulus onset from at least 10 repeats of each image.

2.2 Parametric Methods. We fit two models to disparity tuning curves. The first model is a Gabor function and has been shown to fit well to disparity tuning curves measured from neurons recorded in V1 and other

visual areas (Prince et al., 2002; Hinkle & Connor, 2001; DeAngelis & Uka, 2003) and the function is predicted by feedforward models of V1 disparity tuning (Ohzawa et al., 1990). However, Gabor functions do not do well in explaining the dynamic behavior of disparity tuning (Samonds et al., 2013). The second model, a difference-of-gaussians (and similar difference models), does well at explaining the dynamics of orientation tuning (Somers, Nelson, & Sur, 1995; Xing et al., 2005).

2.2.1 Gabor Function. The disparity energy model is composed of multiple simple cell disparity tuning curves that are the sum of left and right eye phase- or position-shifted two-dimensional Gabor filters convolved with a stimulus. The output of these simple cells is then squared and summed to produce a position-invariant disparity tuning curve for a complex cell. The output of the feedforward disparity energy model predicts that disparity tuning for this complex cell can be represented as a Gabor function (see Figure 1B; Ohzawa et al., 1990; Prince et al., 2002):

$$x(d) = R + Ae^{\frac{-(d-d_0)^2}{2\sigma^2}} \cos\left(2\pi f(d-d_0) + \varphi\frac{\pi}{180}\right), \tag{2.1}$$

where d is disparity, R is the baseline firing rate, A is the amplitude, d_0 is the preferred disparity, σ is the standard deviation of the gaussian envelope, f is the spatial frequency of the sinusoidal component, and φ is the phase shift of the sinusoidal component. An example of a Gabor fit is illustrated in Figure 1A (gray) for a disparity tuning curve measured soon after the response onset following the start of visual stimulation.

2.2.2 Difference-of-Gaussians Function. Although a Gabor is the standard function for describing disparity tuning over a diverse population of neurons, alternative functions that are not constrained to a single frequency or bandwidth component, and are therefore ideal for capturing more dynamic behavior of tuning, such as those that occur in recurrent networks, are the difference-of-gaussians or difference–of–von Mises functions (Somers, Nelson, & Sur, 1995; Xing et al., 2005). We fit our disparity tuning curves with a simplified difference-of-gaussians function (see Figure 1C):

$$x(d) = R + A_{peak}e^{\frac{-(d-d_0)^2}{2\sigma^2_{peak}}} - A_{valley}e^{\frac{-(d-d_0)^2}{2\sigma^2_{valley}}}, \tag{2.2}$$

where d is disparity, R is the baseline firing rate, A_{peak} is the amplitude of the positive gaussian component, d_0 is the preferred disparity, σ_{peak} is the standard deviation of the positive gaussian component, A_{valley} is the amplitude of the negative gaussian component, and σ_{valley} is the standard deviation of the negative gaussian component. The positive gaussian component (peak)

can capture the narrowing of the primary peak, while the negative gaussian component (valley) can capture the broadening of the valleys as the disparity tuning deviates from a Gabor function with stronger recurrent input (see Figure 1A, black).

2.3 Statistical Measurements.

2.3.1 Sample Skewness. The statistical measurement of the sample skewness is the third standardized moment of a distribution and can be computed directly from the mean firing rates $f(d)$ for each disparity d tested:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N}\sum_{d=1}^{N}(f(d)-\overline{f})^3}{\left(\sqrt{\frac{1}{N}\sum_{d=1}^{N}(f(d)-\overline{f})^2}\right)^3}, \tag{2.3}$$

where $\bar{f}$ is the mean firing rate across all N disparities. Skewness represents the direction in which a distribution is skewed (see Figures 2A and 2B, gray) with respect to a normal distribution (see Figures 2A and 2B, black). This can describe the sharpness of a tuning curve because greater skewness means that most of the data are on one side of the mean, while smaller numbers of the data are far away on the other side of the mean. With respect to a disparity tuning curve, this means that most of the disparities, such as nonpreferred disparities, have firing rates just below the mean, while only a small number of the disparities, such as the preferred disparity, have firing rates far above the mean (see Figures 2C and 2D). In addition, skewness is invariant with respect to the mean and variance of the tuning curve, so changes in skewness cannot be attributed to changes in the baseline firing rate or the amplitude of the tuning curve (difference between the maximum and minimum) over time.

For large nonparametric sets of complex stimuli, the distributions of firing rates for neurons look similar to the example unimodal distributions presented in Figures 2A and 2D, with distributions such as exponential or gamma (Franco, Rolls, Aggelopoulos, & Jerez, 2007; Lehky, Kiani, Esteky, & Tanaka, 2011). For parametric sets of stimuli such as disparity and orientation, there are more complex bimodal distributions of firing rates. Like our example tuning curve with a unimodal distribution of firing rates in Figures 2C and 2D, skewness can capture the sharpness of these tuning curves with bimodal distributions of firing rates. Indeed, skewness captures all the features of sharpness that we observe with disparity tuning (see Figure 1A). When one or a few disparities have firing rates far above the mean firing rate across the entire disparity tuning curve, skewness has a high positive value. This happens with narrow positive peaks and broad valleys (see Figure 2E, top two rows), and skewness will increase if positive peaks become narrower or valleys become broader, or both. The peak of the distribution of firing rates occurs below or to the left of the mean firing

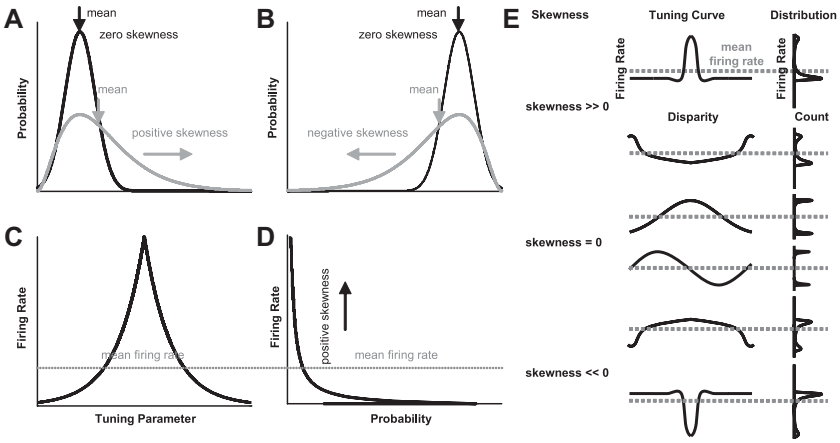


Figure 2: How tuning curve sharpness influences skewness. (A) Skewness is positive when the peak of a probability distribution is below the mean and extreme values of the distribution are skewed above the mean. (B) Skewness is negative when the peak of a probability distribution is above the mean and extreme values of the distribution are skewed below the mean. (C) Therefore, a tuning curve with a very sharp peak (preferred feature) far above the mean firing rate of the tuning curve (extreme value) and most parameters (nonpreferred features) with firing rates below the mean firing rate of the tuning curve has a high positive value of skewness. (D) The distribution of firing rates for the tuning curve is spread out toward very high firing rates compared to the mean firing rate of the tuning curve. (E) Illustrative examples to demonstrate how features in a disparity tuning curve influence the skewness computation. In the center column are tuning curves with samples of particular features; the right column shows the distribution of firing rates for the corresponding tuning curve.

rate, and the distribution is spread out for values above or to the right of the mean. Skewness also increases if the response to secondary peaks is reduced (see Figure 1A). When one or a few disparities have firing rates far below the mean firing rate across the entire disparity tuning curve, skewness has a large negative value. This happens with narrow negative peaks and broad positive peaks (see Figure 2E, bottom two rows), and skewness will decrease if negative peaks become narrower or positive peaks become broader, or both. The peak of the distribution of firing rates occurs above or to the right of the mean firing rate, and the distribution is spread out for values below or to the left of the mean. Skewness also becomes more negative if the response to secondary negative peaks is reduced. If there are equal numbers of disparities with responses equally above and below the mean (e.g., sinusoidal function), skewness will be equal to zero (see Figure 2E,

third and fourth rows). The distributions of firing rates are symmetric with respect to the mean. Overall, skewness increases for tuned excitatory neurons (positive preferred peak) and decreases for tuned inhibitory neurons (negative preferred peak) for all the changes in sharpness that we observed for disparity tuning (Samonds et al., 2009, 2013).

It is important to point out that skewness depends on the range over which the tuning curve is sampled. If the tuning curve is sampled over a wider range of stimuli that includes many baseline responses, then the widths of peaks will be narrower relative to the range of the stimuli, so measured skewness values will be stronger. Therefore, skewness measurements should be compared only when the tuning curve is sampled in the same manner (i.e., using the same set of stimuli or conditions).

2.3.2 Sample Kurtosis. The statistical measurement of the sample kurtosis is the fourth standardized moment of a distribution and can be computed directly from the mean firing rates $f(d)$ for each disparity d tested,

$$\beta_2 = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N}\sum_{d=1}^{N}(f(d)-\overline{f})^4}{\left(\sqrt{\frac{1}{N}\sum_{d=1}^{N}(f(d)-\overline{f})^2}\right)^2}, \tag{2.4}$$

where $\bar{f}$ is the mean firing rate across all N disparities. Kurtosis measures heavy tails or represents the peakedness of a distribution. It has been used widely to describe the sparseness of responses among a population of neurons (Olshausen & Field, 1996, 2004; Lewicki & Sejnowski, 2000; Willmore & Tolhurst, 2001) and to characterize selectivity of tuning curves (Lehky et al., 2005, 2011; Lehky & Sereno, 2007).

2.3.3 Circular Variance. Directional statistics are ideal for data based on variables such as orientation because very different values of the variable can have very similar representations (e.g., 1 and 359 degrees). One directional statistic that can describe the spread of data across a variable such as orientation is the circular variance (V), which can be used to quantify the bandwidth of orientation tuning curves (Ringach et al., 1997; Ringach, Shapley, & Hawken, 2002):

$$V = 1 - \frac{\left|\sum_k R_k e^{i2\theta_k}\right|}{\sum_k R_k}, \tag{2.5}$$

where R_k is the response at orientation θ_k for angles 0 degrees $\leq \theta_k <$ 180 degrees. If all responses are equal in a tuning curve (flat), the resultant (right side of equation 2.5) will be zero and the circular variance will be equal to one. Therefore, broad tuning curves will have circular variance measurements approaching one. If all responses are zero except for the preferred orientation in a tuning curve (impulse function), the resultant

will be one, and circular variance will be equal to zero. Therefore, sharp tuning curves will have circular variance measurements approaching zero.

2.3.4 Selectivity Breadth Index. There have also been statistical measurements developed explicitly for quantifying the selectivity of neuronal tuning curves. One example is the selectivity breadth index (SBI) applied to the responses of ITC neurons to a set of object images morphed along a continuum (Freedman et al., 2006). Responses were normalized to range between 0 and 1 using $R_{\min}$ and $R_{\max}$, and the median $\tilde{R}$ was subtracted from one:

$$SBI = 1 - \frac{\tilde{R} - R_{\min}}{R_{\max} - R_{\min}}. \tag{2.6}$$

SBI employs a similar strategy as skewness with respect to characterizing tuning curve sharpness. If the median firing rate in the tuning curve is close to the maximum, a lot of responses are close to the response to the preferred feature, the tuning curve is relatively broad, and SBI approaches zero. If the median firing rate in the tuning curve is close to the minimum, a lot of responses are close to the response to the least-preferred feature, the tuning curve is relatively narrow, and SBI approaches one.

2.4 Information-Theoretic Analysis. Although researchers generally agree on an intuitive definition of sharpness of a tuning curve, there is not a clear consensus about a general formal definition of tuning curve sharpness. Individual definitions tend to rely on specific parameters of the function used to model the tuning curve or are based on the specific method chosen to quantify sharpness. Here, we wanted to develop a general formal definition of sharpness of tuning curves using information-theoretic analysis of simulated tuning curves so that we could verify that skewness is an appropriate measure of sharpness. To be general, our definition had to describe sharpness regardless of the underlying function or shape of the tuning curve and had to capture information only about changes in shape.

One motivation for measuring the sharpness of a tuning curve has been to quantify the selectivity of a neuronal response toward a stimulus: a response of a neuron with a sharp tuning curve is more specific in its selectivity of excitatory stimuli than a broadly tuned neuron. Informally, we expect that neural response selectivity is related to the degree of information transmitted from a neuron by a single spike. After observing a single spike from a sharply tuned neuron, more information is acquired about the stimulus, and with greater certainty, than after observing a spike from a broadly tuned neuron. In this letter, we formalized this intuition using information theory and will show that skewness of the neural tuning curve is highly correlated with the information gain that results from a single action potential.

There has been much research and debate regarding whether sharp tuning curves are more effective at conveying information about a feature efficiently (Tolhurst, Movshon, & Dean, 1983; Bradley, Skottun, Ohzawa, Sclar, & Freeman, 1987; Scobey & Gabor, 1989; Geisler & Albrecht, 1997; Pouget, Deneve, Ducom, & Latham, 1999; Zhang & Sejnowski, 1999; Series, Latham, & Pouget, 2004; Purushothaman & Bradley, 2005; Butts & Goldman, 2006). The effectiveness of communication is commonly measured using mutual information or the related measure of Fisher information. Some studies find that for a particular stimulus, discrimination is optimal for neurons with steep tuning curves at that stimulus (Purushothaman & Bradley, 2005; Butts & Goldman, 2006). Others find that for high noise levels (or, equivalently, for small neural pools or short time windows), discrimination is optimal for neurons with sharp peaks (Butts & Goldman, 2006). And finally, certain patterns of connectivity can result in circumstances where changes in neural noise counteract any gains in information efficiency from increases in sharpness (Pouget et al., 1999; Series et al., 2004). We do not examine mutual information in this letter because its relationship with sharpness depends on all of these factors and others. No unrestricted claim can be made that sharpness is associated with higher mutual information in general. However, we can show that tuning curve skewness is associated with high information gain given a single spike.

The intuition we wish to formalize is that when the tuning curve has high skewness, a single spike carries more information. More specifically, when it is known that the neuron has fired recently, the uncertainty of the stimulus is reduced by an amount that is greater for neurons with tuning curves with high skewness. This intuition is reasonable only when properties of the neuron's tuning curve other than skewness are held constant. In particular, we assumed that the mean firing rate and the tuning curve amplitude are fixed.

We began our formalization of this intuition with the definition of the entropy of a probability distribution $p(x)$, which is a measure of its uncertainty. This is defined as

$$H[p(x)] = \mathbf{E}[-\log_2 p(x)] = -\sum_x p(x) \log_2 p(x), \tag{2.7}$$

where $H[p(x)]$ denotes the entropy of a distribution and $\mathbf{E}$ denotes expectation. Entropy is often described as the expected level of surprise under distribution p. In this case, the "surprise" of an event x is defined as $-\log_2 p(x)$, with rare events being more surprising than common ones. Surprise is defined to be logarithmic with respect to probability, matching the intuition that two independent rare events are twice as surprising as one. More formally, this definition of entropy is justified as measuring the fewest possible number of bits of information required to transmit values drawn from that distribution. Note that distributions that are nearly deterministic or concentrated at a few points are less likely to produce surprising values and therefore have lower entropy than distributions that are broader.

Recall that we wish to formalize the intuition that a single spike conveys more information when the neural tuning curve is sharp: when a sharply tuned neuron fires, it disambiguates the driving stimulus more than a spike from a broadly tuned neuron. Thus, we were interested in measuring the uncertainty in estimates of the stimulus; in particular, we wanted to measure how much that uncertainty is reduced by observing a neuron fire within some small time window of length τ milliseconds. The exact value of τ will have a negligible effect on the results as long as τ is sufficiently small to ensure that the likelihood that the neuron fires twice within the time window is very low. Let θ be some property of the stimulus (e.g., binocular disparity), and let r be the neural response to that stimulus, defined as the number of spikes within some small time window of length τ ms. Then $p(\theta|r)$ gives the probability of the stimulus having value θ given that r spikes were observed. We define the information gain of a single spike by

$$I[\theta|r > 0] = H[p(\theta)] - H[p(\theta|r > 0)], \tag{2.8}$$

where $p(\theta)$ represents our prior knowledge of the stimulus (before seeing any neural response). The event $r > 0$ represents the event that the neuron fires during the time window. Here, we will assume that $p(\theta) = 1/N$, where N is the number of possible stimuli θ.

It is important to note that many different aspects of neural response can be measured and characterized using information theory (Borst & Theunissen, 1999). Here, we chose to examine a specific aspect of neural behavior, the information gain of a single spike, which was motivated from the intuition about the significance of neural sharpness. Information theory may be used to measure either the entropy of $p(\theta|r)$ or of $p(r|\theta)$. The entropy of $p(\theta|r)$ relates to the reliability of neural decoding, or how much is known about the stimulus given the neural response. In contrast, the entropy of $p(r|\theta)$ relates to the reliability of neural encoding and the spread of neural responses given to a single stimulus. Since we wished to verify the claim that skewness is related to the information gain produced by a single spike, we chose to examine the entropy of $p(\theta|r)$.

The information gained by observing a single spike from a neuron is an important measure that influences how rapidly a neuron can signal a change in a stimulus. A postsynaptic neuron can be made to fire very shortly after receiving a small number of spikes from a small number of neurons. Neurons also convey information by silence (e.g., when a tuning curve has a primary negative peak), which can be expressed as $I[\theta|r = 0]$. Because neurons spend less time firing than not, $I[\theta|r = 0]$ is generally several orders of magnitude smaller than $I[\theta|r > 0]$. Mutual information can be expressed as the weighted sum of the information gain of a spike and the information gain of silence, weighted by the relative probabilities of spiking and not spiking, respectively. Note that mutual information is generally applied under the assumption that the response is steady and can be measured

over time windows that are long enough to approximately estimate the firing rate. However, if a change occurs in the stimulus, neurons should be able to communicate the new stimulus within a small time window, ideally before the firing rate can be estimated accurately. Because spikes are rarer and more informative than silence, neurons with higher information gain of a single spike are able to convey more information within a small time window. In other words, while mutual information tells us the total amount of information a neuron conveys on average, the information gain of a spike tells us the maximum amount of information that the neuron can convey within a small time window. In this section, our goal was to formalize the hypothesis that tuning curve skewness is associated with high information gain produced by a single spike. In other words, a neuron with high skewness is better able to disambiguate the stimulus over a small time window when at most one spike is present. Therefore, we chose a small value for time window τ.

Ideally, in order to measure the information gain of a single spike, we must measure the bivariate joint probability distribution $p(r,\theta)$. In practice, this is generally infeasible; $p(r,\theta)$ requires many parameters to learn and would require many trials of the same stimulus while recording from the same neuron to measure accurately. Instead, let us suppose that $p(r|\theta)$ follows a Poisson distribution:

$$p(r|\theta) = \frac{f(\theta)^r e^{-f(\theta)}}{r!}, \tag{2.9}$$

$$p(r > 0|\theta) = 1 - e^{-f(\theta)}, \tag{2.10}$$

where $f(\theta)$ is the neuron's tuning curve and is equal to the expected number of spikes within time window τ given stimulus θ. In order to best formalize our initial hypothesis, we would like to measure the information gain produced at some infinitesimal period of time. In practice, τ can be any small value as long as only one spike is likely to occur within that period. Using Bayes' rule, we can write:

$$p(\theta|r) = \frac{p(r|\theta)p(\theta)}{p(r)} = \frac{p(r|\theta)p(\theta)}{\sum_{\tilde{\theta}} p(r|\tilde{\theta})p(\tilde{\theta})} = \frac{p(r|\theta)}{\sum_{\tilde{\theta}} p(r|\tilde{\theta})} \tag{2.11}$$

$$p(\theta|r > 0) = \frac{1 - e^{-f(\theta)}}{\sum_{\tilde{\theta}} (1 - e^{-f(\tilde{\theta})})} = \frac{1}{Z}(1 - e^{-f(\theta)}) \tag{2.12}$$

where $Z = \sum_{\tilde{\theta}} (1 - e^{-f(\tilde{\theta})})$ and all stimuli are assumed to be equally likely, so that $p(\theta)$ is constant. We can now write the information gain of a spike as where, again, N is the number of possible stimuli θ:

$$I[\theta|r > 0] = \log_2 N + \sum_{\theta} \frac{1}{Z}(1 - e^{-f(\theta)}) \log_2 \frac{1}{Z}(1 - e^{-f(\theta)}). \tag{2.13}$$

Information gain and skewness are conceptually related in that both measures depend on the shape of a tuning curve. However, information gain is also highly sensitive to changes in either the mean firing rate or the amplitude of the tuning curve (difference between the maximum and minimum), while skewness is normalized for these characteristics. When the amplitude is low, all stimuli elicit similar neural responses, and so it is difficult to determine the stimulus θ from neural response alone and the information gain is low. Tuning curves with higher amplitudes result in much higher information gain. Similarly, when the mean firing rate is high compared to the amplitude, the differences in neural responses produced by different stimuli are low in comparison with the absolute firing rate. This causes the information gain to decrease when mean firing rates increase. Thus, a change in information gain may be due to either changes in the tuning curve shape or the tuning curve's amplitude or mean firing rate. To distinguish these two sources of change, we defined shape-based information gain $\tilde{I}[\theta|r > 0]$ as the information gain of a single spike after the tuning curve has been normalized for amplitude and mean firing rate to lie between 0 and 60 sps $\cdot\tau$ (sps: spikes per second). This gave us a measure that depends on tuning curve shape and not tuning curve amplitude or mean firing rate for direct comparison with skewness.

3 Results

We first examined how two commonly used parametric methods and our nonparametric method of skewness describe changes in disparity tuning sharpness. The methods were tested with example tuning curves with known increases in sharpness over time that converge to a steady state over several hundreds of milliseconds and are consistent with the average behavior over a population of neurons and the dynamic tuning predicted from a recurrent network model (Samonds et al., 2013). We describe which parameters captured sharpening over time and under what conditions, as well as when and why skewness captured more information about sharpness than the parametric methods did. We then illustrate how skewness can be applied to a wide range of tuning complexity by showing the results of measurements with simple orientation tuning curves with known increases in sharpness over time and complex object tuning curves with known changes in sharpness that depend on training and the orientation of the presentation of the objects. We show how skewness was a more accurate measurement than other nonparametric methods that have been used to describe sharpness of orientation and object tuning curves. Finally, using information-theoretic methods, we formally define sharpness of tuning and show how skewness was consistent with this definition.

3.1 Relationship Between Sample Skewness and Tuning Function Parameters. Before we carried out a direct and thorough comparison between

the parametric methods (see equations 2.1 and 2.2) and our statistical measurement of skewness (see equation 2.3) that quantify sharpness for model data and data from neurophysiological recordings, we looked directly at the relationship between those disparity tuning function parameters that could quantify sharpness and skewness for an illustrative example. We used the results of this example to help us interpret the results when we applied the methods to neurophysiological data.

For a Gabor function, the σ parameter describes the width of the gaussian envelope (see Figure 1B). As σ decreases, the gaussian envelope becomes narrower, and as long as secondary peaks are present in the Gabor function, they will be suppressed by the narrower envelope and skewness will increase (see Figure 3A). The f parameter describes the frequency of the sinusoidal component (see Figure 1B). In section 2.3.1, we explained that a sinusoidal signal has zero skewness because it has an equal number of data points above and below the mean at an equal distance from the mean. Regardless of the frequency, a sinusoidal component will result in zero skewness. However, when a sinusoidal component is combined with a gaussian envelope, the envelope introduces a bias that leads to nonzero skewness. The relationship between frequency and skewness depends on the relationship between frequency and the bandwidth of the gaussian envelope. When more sinusoidal cycles are observed due to a large envelope σ or high-frequency f, increases in f increase the magnitude of secondary peaks, and that leads to decreases in skewness (see Figure 3B). When fewer sinusoidal cycles are observed and the secondary peaks, are suppressed, increases in f cause the primary peak to become narrower and skewness increases (see Figure 3C). Overall, the relationship between Gabor parameters and skewness is complicated and will vary over a reasonable range of the Gabor parameter space that is typically used to describe some disparity tuning curves (Prince et al., 2002).

For a difference-of-gaussians function, the σ_{peak} parameter describes the width of the positive gaussian component and the σ_{valley} parameter describes the width of the negative gaussian component (see Figure 1C). When the positive gaussian component becomes narrower (decreasing σ_{peak}), there is a sharper peak, and skewness increases (see Figure 3D). When the negative gaussian component becomes broader (increasing σ_{valley}), there is a subtle increase in skewness (see Figure 3E). Overall, there is a consistent relationship between difference-of-gaussians parameters and skewness over a reasonable range of the parameter space that is typically used to describe some disparity tuning curves (Samonds et al., 2013).

3.2 Testing Parametric Methods and Skewness on Neurophysiological Data with Changes in Sharpness. Disparity tuning sharpness increases over time, increases with increasing stimulus size, and decreases for binocular anticorrelation stimulation in the macaque primary visual cortex; a

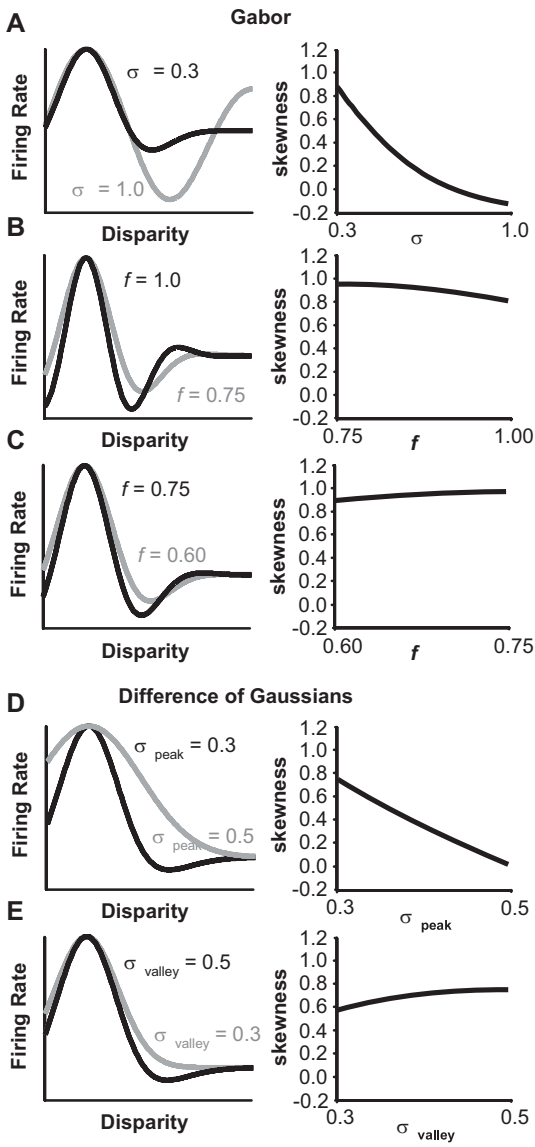


Figure 3: Relationship between sample skewness and function parameters. (A) Decreasing the width of the gaussian envelope increases skewness for Gabor functions. (B, C) Increasing the frequency of the sinusoidal component can decrease or increase skewness, respectively, for Gabor functions. (D) Decreasing the width of the positive gaussian increases skewness in a difference-of-gaussians function. (E) Increasing the width of the negative gaussian increases skewness in a difference-of-gaussians function.

recurrent model can account for those changes in sharpness (Samonds et al., 2013). Sharpening over time was the strongest result from that study, with several examples from recordings and a population average that clearly matched the behavior of the model. Therefore, to test how effectively the parameters from our functions (see equations 2.1 and 2.2) and skewness (see equation 2.3) describe sharpness of disparity tuning, we applied these methods to a model neuron with disparity tuning that is sharpened by a population of recurrently connected disparity-tuned neurons and recorded neurons with known sharpening of disparity tuning over time that matched the model's predictions (see section 2.1 and Samonds et al., 2013). The top row in Figure 4 shows the initial disparity energy model tuning curve (input) for a representative model neuron and the resulting tuning curve after several iterations of recurrent inputs with other model neurons (steady state). The following rows in Figure 4 show disparity tuning curves measured from recorded neurons in subsequently later windows of time that were consistent with model tuning curves after several iterations of recurrent interactions. All five plots show clear examples of disparity tuning that sharpens over time, making them ideal cases for testing our methods of quantifying sharpness.

Sharpening was stronger in the earliest iterations and during the earliest portion of the neuronal responses soon after the peak of the response onset. Sharpening continued over iterations or time, but at a progressively slower rate. This happened in the model because the behavior converged to a steady state where the tuning curves no longer changed with a greater number of iterations. Therefore, throughout this letter, and as we have done previously (Samonds et al., 2009, 2013), we will present all of our measurements versus log time steps (log iterations) or log time.

3.3 Using the Gabor Function to Measure Sharpness of Disparity Tuning. For the same model and four example tuning curves presented in Figure 4, we fit Gabor functions (see equation 2.1) to the data over time. The plots in Figure 5A show the data and fits for the latest time window (450–850 ms) to illustrate how well the fits describe sharpened tuning curves. The plots in Figures 5B and 5C reveal how the fit parameters f and σ evolved over time from tuning curves based on the mean firing rates computed from sliding 100 ms windows. Figure 3A showed that the bandwidth of the gaussian envelope σ can signal increased sharpness with decreases in magnitude. Indeed, we found decreases in σ over time as sharpness increased, but only for two example neurons with clear secondary peaks that were suppressed over time (compare the second and third rows of Figure 4 with the same rows in Figure 5C). Depending on which disparities were measured with respect to the tuning curve, the positive or negative peaks could have greater influence on determining the best-fit frequency parameter. For our examples, it was unclear whether the frequency increased, decreased, or did not change over time as sharpness increased

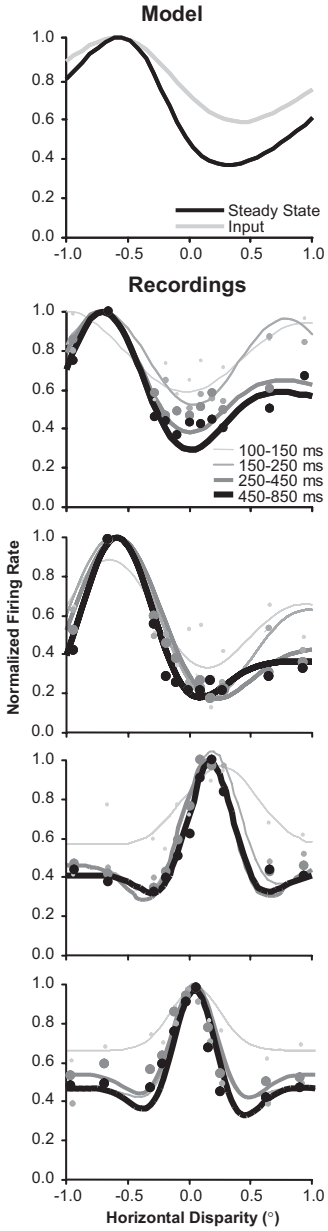


Figure 4: Examples of sharpening of disparity tuning over time for a neuron in a recurrent neuronal network model (top) and neurophysiological recordings (bottom four).

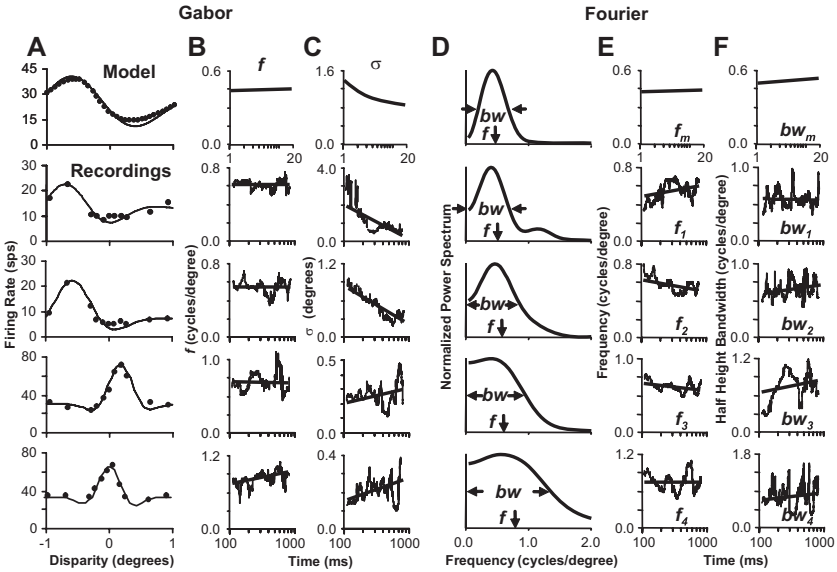


Figure 5: Sharpening was not captured by measuring the frequency of the tuning curve. (A) First, sharpening was characterized for example neurons with the frequency, f, and/or bandwidth, σ, parameters using a Gabor fit. (B) The Gabor frequency, f, versus time. (C) The Gabor bandwidth, σ, versus time. (D) Then sharpening was characterized with the centroid frequency, f, and bandwidth, bw, of the Fourier power spectrum. (E) The Fourier centroid frequency, f, versus time. (F) The Fourier half-height bandwidth, bw, versus time.

(see Figure 5B). Overall, we found that neither the bandwidth parameter nor the frequency parameter in Gabor fits was reliably and consistently able to capture the increases in sharpness of disparity tuning that we observed over time. In addition, achieving stable fits with a six-parameter Gabor model was especially problematic for short instances of time when the tuning curve was especially noisy, leading to many cases of outlier parameter estimates (e.g., frequencies more than five times the frequency estimated based on the tuning curve measured over the entire stimulation time period).

3.4 Using Fourier Analysis to Measure Sharpness of Disparity Tuning. We also computed the Fourier power spectrum for the same model and four example tuning curves presented in Figure 4 to further illustrate the difficulty of characterizing a frequency change in disparity tuning over time. The plots in Figure 5D show the Fourier power spectrum for the disparity tuning curves computed from the latest time window

(450–850 ms). The plots in Figures 5E and 7F reveal how the centroid frequency f and half-height bandwidth bw of the Fourier power spectrum evolved over time using tuning curves based on the mean firing rates computed from sliding 100 ms windows. There were similar problems with using the power spectrum of the Fourier transform of the tuning curves to characterize sharpness as there were with estimating the frequency parameter from a Gabor fit. This is because computing a Fourier transform and fitting a complex function to data are subject to the same limitations imposed by sampling and measurement uncertainty. Again, because the central peak was increasing in frequency (narrowing) and the valleys were decreasing in frequency (broadening), each influenced the Fourier transform to emphasize one or the other. In addition, the frequencies of the peak and valleys were close together, so we did not observe two distinct peaks (see Figure 5D) moving in opposite directions over time. The centroid of the power spectrum sometimes increased over time, did not change over time, and decreased over time (see Figure 5E). If the frequencies of the peak and valleys are the same initially (e.g., a Gabor function) and then the frequency of the peak increases and the frequency of the valleys decreases over time, the bandwidth of the power spectrum should increase over time as energy would be spread across the two new frequencies in the sharpened disparity tuning curve. Although bandwidth estimates of the power spectrum were very noisy even for our robust examples, there was some evidence of a consistent increase in bandwidth over time (see Figure 5F). Overall, the Fourier transform did not capture the increases in sharpness of disparity tuning that we observed over time.

3.5 Using the Difference-of-Gaussians Function to Measure Sharpness of Disparity Tuning. Again, for the same model and four example tuning curves presented in Figure 4, we also fit difference-of-gaussians functions (see equation 2.2) to the data over time. The plots in Figure 6A show the data and fits for the latest time window (450–850 ms) to illustrate how well these fits describe sharpened tuning curves. The plots in Figures 6B and 6C reveal how the fit parameters σ_{peak} and σ_{valley} evolved over time from tuning curves based on the mean firing rates computed from sliding 100 ms windows. In Figure 6B, we plotted the bandwidth of the positive gaussian component (σ_{peak}), which decreased over time as the primary positive peak narrowed. In Figure 6C, we plotted the bandwidth of the negative gaussian component (σ_{valley}), which increased over time as the valleys broadened, although this trend was less clear than what was observed for the positive gaussian component. A difference-of-gaussians function still required us to fit 6 parameters to 11 data points, which leaves it highly sensitive to the same problems that we encountered with Gabor fits with parameter initialization and outlier results. Overall, though, fits for these example disparity tuning curves that were well described by a

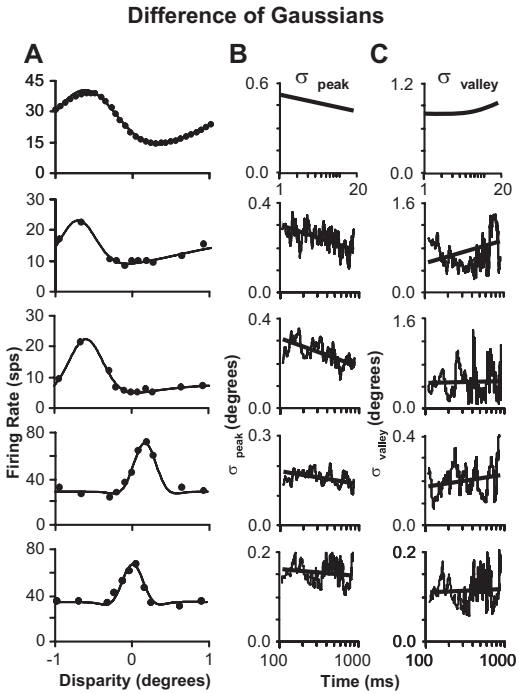


Figure 6: Sharpening was characterized for example neurons with the peak bandwidth, σ_{peak}, and/or valley bandwidth, σ_{valley}, parameters using a difference-of-gaussians fit. (A) Difference-of-gaussians fits for example disparity tuning curves. (B) The gaussian peak bandwidth, σ_{peak}, versus time. (C) The gaussian valley bandwidth, σ_{valley}, versus time.

difference-of-gaussians did capture increases in sharpness with decreases in the σ_{peak} parameter.

3.6 Using Skewness to Measure Sharpness of Disparity Tuning. Finally, for the same model and four example tuning curves presented in Figure 4, we computed the statistical measurement of skewness (see equation 2.3) on the data over time. The plots in Figure 7A show the data (left) for the latest time window (450–850 ms) and skewness (right) using tuning curves based on the mean firing rates computed from sliding 100 ms windows. For all of our example neurons with sharpened tuning over time, skewness clearly and consistently increased over time (see Figure 7A, right). Figure 7B (left) shows the normalized population average of rank-order disparity tuning over time for 184 neurons from Samonds et al. (2013) using the same time windows as in Figure 4. Disparity tuning curves were sorted before

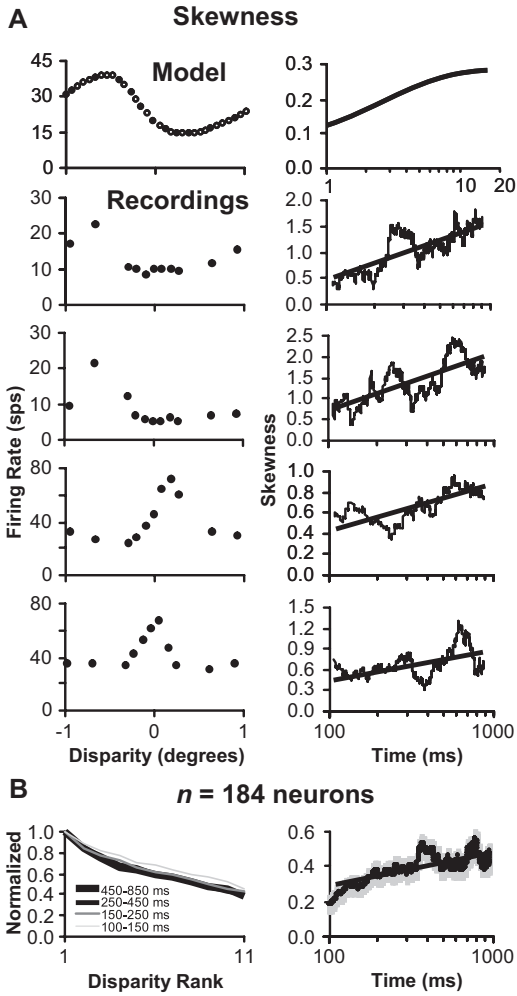


Figure 7: (A) Examples demonstrating how features in a disparity tuning curve influence the skewness computation. The plots in the left column are tuning curves, and the plots in the right column are the corresponding skewness computation (see equation 2.3) over time. (B) Normalized population average of rank-order disparity tuning curves (left) and population average of skewness (right) from Samonds et al. (2013). Error bars are standard error with respect to neurons.

averaging from the most- to least-preferred disparity (rank order tuning). The population averages of corresponding skewness measurements over time for the 184 neurons are shown in Figure 7B (right), illustrating that the population trend was consistent with the model.

Our four example neurons were tuned excitatory: they had reduced firing for most disparities and then had a higher firing for a preferred disparity. Some disparity-tuned neurons can be tuned inhibitory: they fire a lot for most disparities and then have reduced firing for a preferred disparity (Samonds et al., 2013). If these negative peaks sharpen, skewness decreases, becoming more negative (see Figure 2E). Tuned excitatory or inhibitory neurons with relatively broad peaks that sharpen can start with negative (see Figure 2E, fifth row) or positive skewness (see Figure 2E, second row) and end up with positive (see Figure 2E, first row) or negative (see Figure 2E, sixth row) skewness, respectively. When we recently applied skewness to tuning curves for a mixed population of tuned excitatory (79%) and inhibitory (21%) neurons, we analyzed tuning dynamics in two different ways (Samonds et al., 2013). First, we analyzed the subpopulations separately, where increasing skewness corresponded to greater sharpness for tuned excitatory neurons and decreasing skewness corresponded to greater sharpness for tuned inhibitory neurons. Second, we combined the subpopulations for statistical analyses by inverting the data for tuned inhibitory neurons so that increasing skewness always corresponded to greater sharpness, which is how the data were analyzed in Figure 7B.

3.7 Comparison of Parametric and Nonparametric Methods of Measuring Sharpness. Among the measures from the parametric models, the change in the difference-of-gaussians peak width σ_{peak} over time was most consistent with changes in the non-parametric statistical measurement of skewness. Both σ_{peak} and skewness captured increases in sharpness of disparity tuning over time. For all four examples from recorded neurons in Figures 6 and 7, there were decreases in σ_{peak} and increases in skewness, respectively. In addition, the two measurements were significantly correlated over time for all four examples with an average correlation of $r = -0.51 \pm 0.06$. To demonstrate that σ_{peak} and skewness capture similar changes in tuning curve shape over time, we generated a scatter plot of the changes of σ_{peak} versus the changes in skewness over time. As noted in the previous section, a six-parameter difference-of-gaussians fit is highly sensitive to parameter initialization, and not all disparity tuning curves were well described by a difference-of-gaussians function, so the fits generated instances of very noisy or extreme outlier results over time for even our most robust examples. Therefore, we simplified our analysis for the entire population of 184 neurons by dividing the responses for each neuron into two windows of time rather than making the measurements continuously over time. The early window was from 100 to 250 ms, and the late window was from 450 to 850 ms. This allowed us to carefully adjust parameter initialization for 368 disparity tuning curves so that the fits adequately captured the peak width of disparity tuning curves with the positive gaussian component. We then plotted the ratio of the difference in σ_{peak} between the

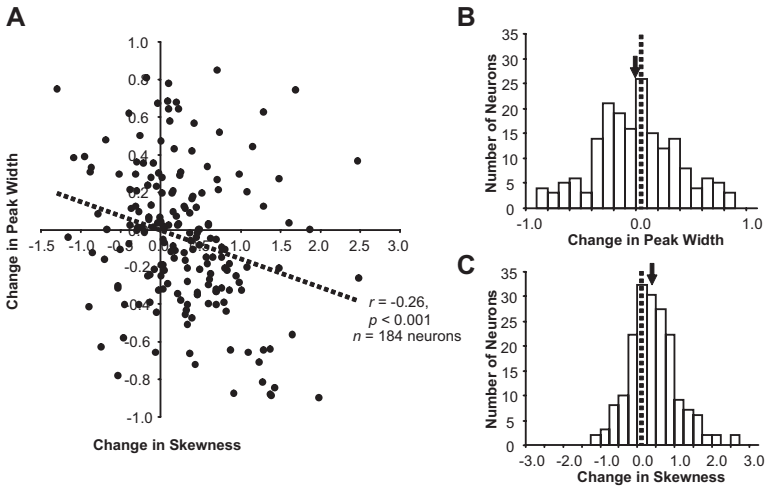


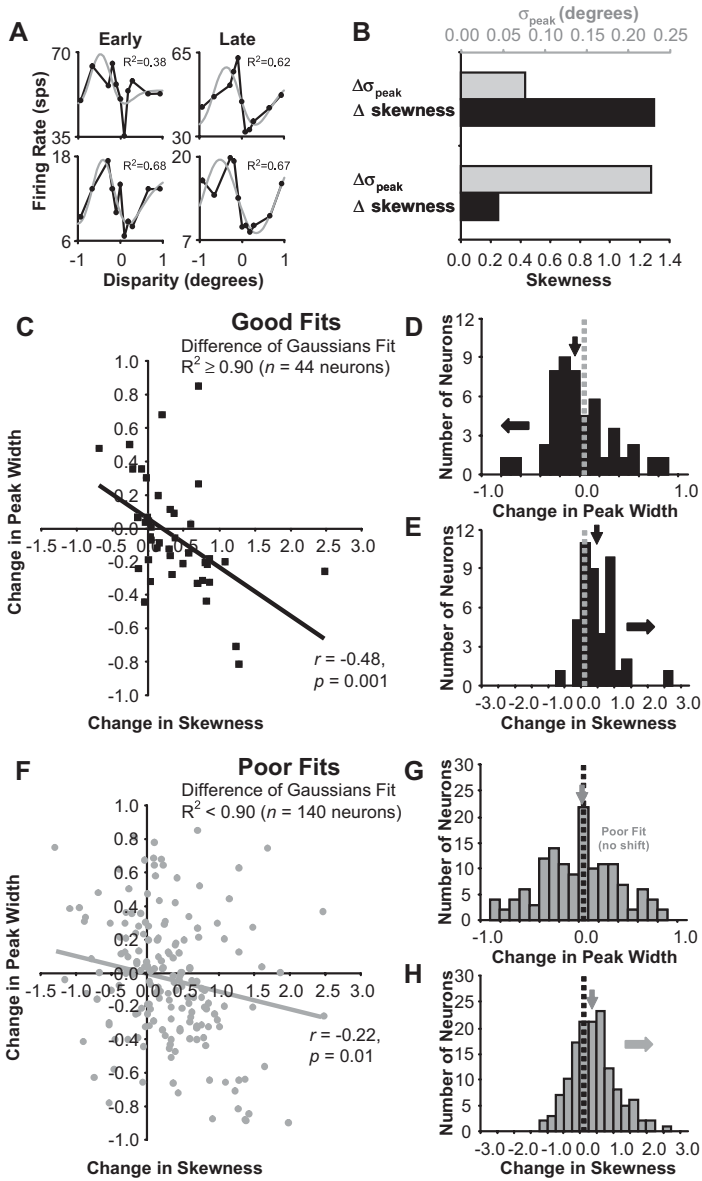
Figure 8: Comparison of peak width σ_{peak} and skewness. (A) Scatter plot of change in σ_{peak} (late – early)/(late + early) and change in skewness (late – early) from an early window of time from stimulus onset (100–250 ms) versus a late window of time from stimulus onset (450–850 ms). (B) Histogram of changes in peak width from panel A. (C) Histogram of changes in skewness from panel A.

late- and early-window tuning curves (late − early)/(late + early) versus the difference in skewness between the late- and early-window tuning curves. Figure 8A shows a significant correlation ($r = -0.26$, $p < 0.001$) between a reduction in the peak width (see Figure 8B) and an increase in skewness (see Figure 8C).

Although the results in Figure 8A reveal that σ_{peak} and skewness were correlated, there were notable differences between these two measures when we examined them more closely. One difference is that the positive shift from zero for changes in skewness is stronger than the negative shift for changes in σ_{peak}(see Figure 8C versus 8B). One reason that the parametric method produced a weaker trend of sharpening compared to skewness was that the difference-of-gaussians function does not adequately describe disparity tuning for all neurons. For our four example neurons, the difference-of-gaussians function did describe the tuning curve very well, and we were able to measure a reduction in σ_{peak} for all four examples. The examples in Figure 9A illustrate that neurons with tuning curves that are not explained well by a difference-of-gaussians function can result in a fit (gray) with a broader peak width in the later time window (see Figure 9B, gray bars) even though skewness still clearly increased in that window for these tuning curves (see Figure 9B, black bars) that are visibly sharper for the late versus the early period during stimulation (see Figure 9A, right versus left plots).

If we limited our parametric analysis to a smaller population ($n = 44$ neurons) of tuning curves where the difference-of-gaussians fit explained at least 90% of the variance in the tuning curve ($R^2 \geq 0.90$), the reduction in peak width (see Figure 9D) and its correlation with increasing skewness ($r = -0.48$, $p = 0.001$; see Figure 9C) were much clearer and the results between the two methods had a stronger relationship (compare the leftward shift in Figure 9D to the rightward shift in Figure 9E). For the 140 neurons that were not well described by a difference-of-gaussians function ($R^2 < 0.90$), there was not a clear leftward shift in change in peak width (see Figure 9G), but there was still a clear rightward shift in change in skewness (see Figure 9H). In addition, the correlation was much weaker between changes in peak width and skewness ($r = -0.22$, $p = 0.01$; see Figure 9F).

3.8 Testing Skewness on Orientation Tuning Curves. Since skewness worked well in capturing sharpness of disparity tuning, we wanted to also verify that it would capture sharpness of simpler tuning curves. Previous studies have shown that orientation tuning curves in the primary visual cortex (V1) also sharpen over time (Ringach et al., 1997, 2002). Because orientation is a circular variable, these studies used the directional statistical measurement of circular variance to characterize sharpening of orientation tuning (see section 2.3.3). A high value of circular variance means that the data are spread out evenly across all orientations and the tuning curve is broad, while a low value of circular variance means that the data are concentrated at a single orientation and the tuning curve is sharp. We applied circular variance and skewness to orientation tuning curves of neurons recorded in areas V1 and V2 to test whether skewness captured sharpening and compare the results of the two statistical measures. Circular variance (see equation 2.5) and skewness were measured over time from orientation tuning curves based on the mean firing rates of neurons computed in sliding 100 ms windows. Figures 10A to 10C show the results for four example neurons and the population average from 98 neurons. Figure 10A shows the orientation tuning curves measured in subsequently later windows of time (light-to-dark) and reveals that the tuning curves sharpened over time. Figure 10B shows that circular variance is consistent with this observation because it decreased over time, and Figure 10C shows that skewness is consistent with this observation because it increased over time. We also compared the two measurements on a neuron-by-neuron basis by plotting the slopes of circular variance versus log time versus the slopes of skewness versus log time computed from linear regression fits for each of $n = 98$ neurons and found that the two measurements were highly correlated (see Figure 10D; $r = -0.47$, $p < 0.001$). Because a logarithmic fit might not adequately describe sharpening dynamics for all neurons, we also measured the correlation between circular variance and skewness over time for each neuron. The average correlation r was significantly negative (see Figure 10E; $\mu = -0.40$, $p < 0.001$).



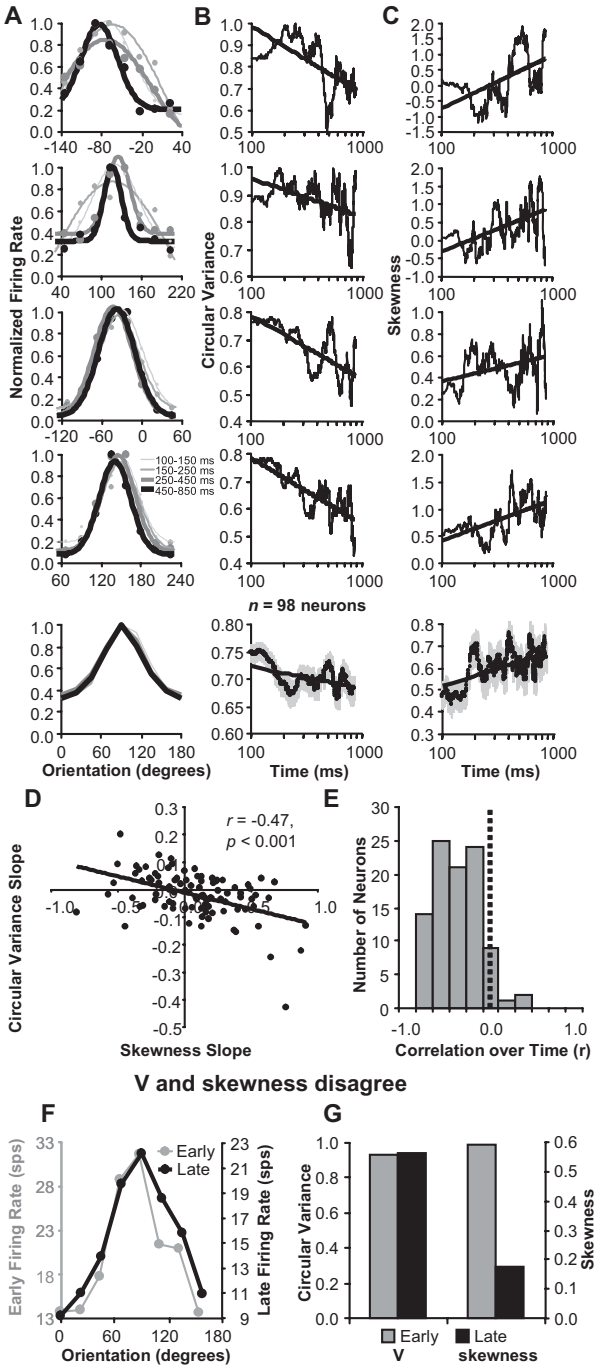
We also examined differences between circular variance and skewness by looking at when circular variance and skewness disagreed or were not strongly correlated. The example in Figure 10F actually conflicts with the population trend with an orientation tuning curve that gets broader in

shape over time (black is broader than gray). For this example, there is essentially no change in circular variance (see Figure 10G, left) but a clear decrease in skewness (see Figure 10G, right) between the early and late period. The reason the two results differ is that skewness is normalized for the mean and amplitude, while circular variance is normalized for only the amplitude. By comparing the two y-axes in Figure 10F (gray and black), it is clear that the mean firing rate decreased from the early to late time period, which is common for V1 responses over this interval (Samonds et al., 2013). Because the mean firing rate decreased, the baseline firing rate was lower so the responses were effectively spread out less across all orientations for the late tuning curve. This change led to a decrease in circular variance that canceled out the increase in circular variance resulting from the broader shape of the tuning curve. Since skewness is also normalized the for mean firing rate, it provides a more accurate measure of changes in tuning curve shape.

3.9 Testing Skewness on Object Tuning Curves. Finally, we wanted to test whether skewness would capture sharpness of tuning curves that are more complex than disparity tuning. Neurons in the inferior temporal cortex (ITC) respond selectively to complex objects (Tanaka, Siato, Fukada, & Moriya, 1991) and because it is not feasible to determine, test for, or characterize tuning for ITC neurons for all possible objects, tuning curves for ITC neurons are unlikely to be practically described by a simple function. Most ITC tuning curves depend on the set of stimuli used to measure the tuning curve and are usually presented as a matrix of mean firing rates for the complex set of objects or a sorted tuning curve ranked from the most to least preferred object (rank-order tuning).

Freedman et al. (2006) have shown that tuning curves for ITC neurons become sharper after training, but more strongly when the tuning is measured with objects presented at the same or similar orientations used during

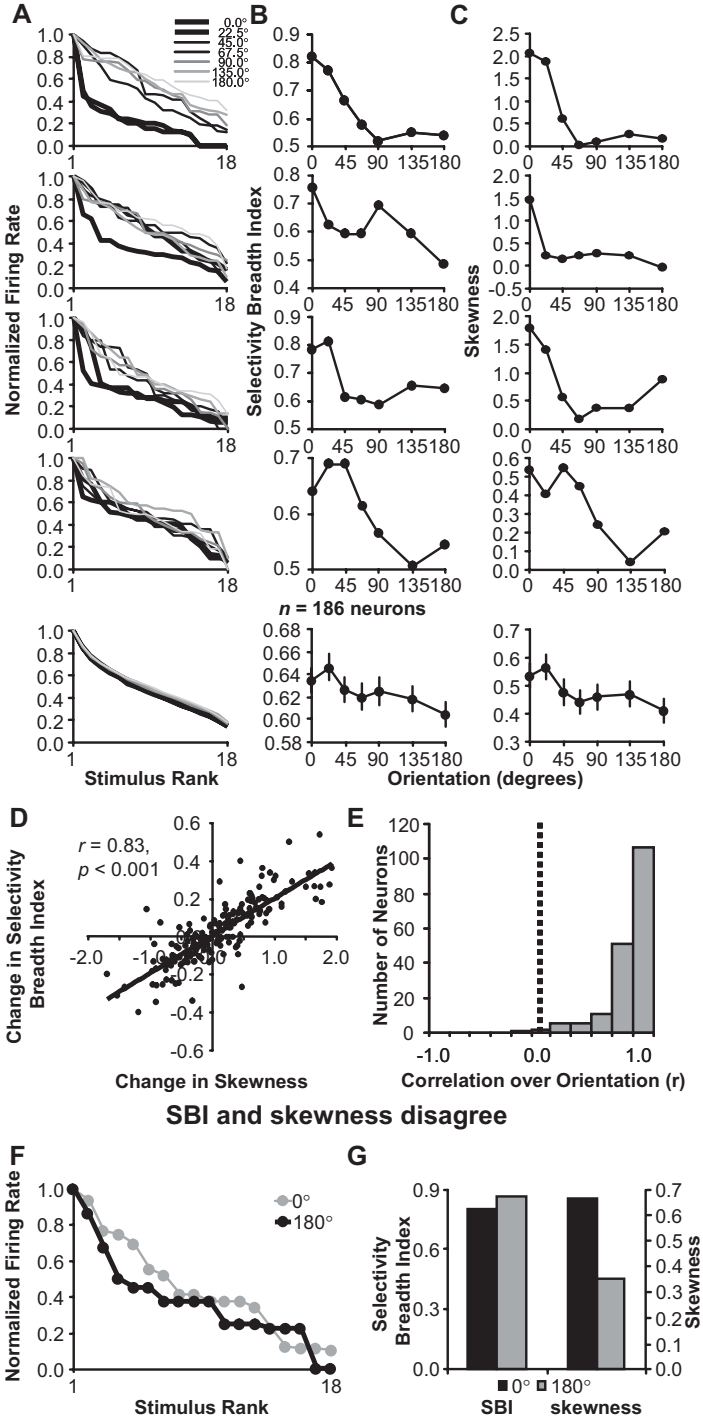
Figure 9: Comparison of peak width σ_{peak} and skewness for poor difference-of-gaussians fits ($R^2 < 0.90$). (A) Two examples of disparity tuning curves with poor difference-of-gaussians fits (gray). (B) Change in σ_{peak} (gray) and skewness (black) of late versus early tuning curves in panel A. Comparison of peak width σ_{peak} and skewness for good (C–E) and poor difference-of-gaussians fits (F–H). (C) Scatter plot of change in σ_{peak} (late − early)/(late + early) and change in skewness (late − early) from an early window of time from stimulus onset (100–250 ms) versus a late window of time from stimulus onset (450–850 ms) for disparity tuning with good difference-of-gaussians fits. (D) Histogram of changes in peak width from panel C. (E) Histogram of changes in skewness from panel C. (F) Same scatter plot as panel C, but for disparity tuning with poor difference-of-gaussians fits. (G) Histogram of changes in peak width from panel F. (H) Histogram of changes in skewness from panel F.


A
B
C
Normalized Firing Rate
Circular Variance
Skewness
100-150 ms
150-250 ms
250-450 ms
450-850 ms
n = 98 neurons
Orientation (degrees)
Time (ms)
Time (ms)
D
E
r = -0.47,
p < 0.001
Circular Variance Slope
Skewness Slope
Number of Neurons
Correlation over Time (r)
V and skewness disagree
F
G
Early
Late
Early Firing Rate (sps)
Late Firing Rate (sps)
Orientation (degrees)
Circular Variance
Skewness
Early
Late
V
skewness

training. In that study, they used an ad hoc statistical measure, the selectivity breadth index (SBI; see equation 2.6), that employs a very similar strategy to skewness (see section 2.3.4). SBI examines how close the median is to the minimum or maximum value of the tuning curve. When the median is near the maximum, the tuning curve is broad and SBI has low values. When the median is near the minimum, the tuning curve is sharp and SBI has high values. We applied SBI and skewness to object tuning curves of neurons recorded in ITC (Freedman et al., 2006) to test whether skewness captured changes in sharpness and compare the results of the two statistical measures. SBI and skewness were measured for object tuning curves computed from the mean firing rates early in the response (80–180 ms) at seven different orientations relative to the orientation of the objects the monkeys were trained on for a category discrimination task (Freedman et al., 2001, 2006). Figures 11A to 11C show the results for four example neurons and the population average from 186 neurons. Figure 11A shows the rank-order object tuning curves measured at seven orientations relative to the trained orientation (dark-to-light) and reveals that the tuning curves were sharpest at the trained orientation (dark). Figure 11B shows that SBI consistently decreased away from the trained orientation, and Figure 11C shows that skewness consistently decreased from the trained orientation. Both statistical measurements correctly captured greater sharpness for the orientation used in training observed in Figure 11A. We also compared the two measurements on a neuron-by-neuron basis by plotting the change in SBI (0 degrees versus 180 degrees from trained orientation) versus the change in skewness for each of $n = 186$ neurons and found that the two measurements were highly correlated (see Figure 11D; $r = 0.83$, $p < 0.001$). We also measured the correlation between SBI and skewness over the seven presentation orientations for each neuron. The average correlation r was significantly positive (see Figure 11E; $\mu = 0.84$, $p < 0.001$).

We also examined differences between SBI and skewness by looking at when SBI and skewness disagreed or were not strongly correlated. The

Figure 10: Examples of sharpening of orientation tuning over time for neurons from neurophysiological recordings in V1 and V2. (A) Normalized tuning curves at different windows of time. (B) Circular variance (V) versus time. (C) Skewness versus time. Population averages are presented in the bottom row of panels A–C. All orientation tuning curves were aligned with their peak at 90 degrees. Error bars are standard error with respect to neurons. (D) Comparison of slopes of circular variance (V) versus log time and slopes of skewness versus log time. (E) Histogram of correlation measured for individual neurons of circular variance and skewness over time. (F, G) Example of when V and skewness results disagree. (F) Orientation tuning was measured in a window of 100–250 ms (early) and 400–850 ms (late) from stimulus onset. (G) V and skewness measurements computed from the tuning curves in panel F.


A
B
C
0.0°
22.5°
45.0°
67.5°
90.0°
135.0°
180.0°
Normalized Firing Rate
Selectivity Breadth Index
Skewness
n = 186 neurons
Stimulus Rank
Orientation (degrees)
D
E
r = 0.83,
p < 0.001
Change in Selectivity Breadth Index
Change in Skewness
Number of Neurons
Correlation over Orientation (r)
SBI and skewness disagree
F
G
0°
180°
Normalized Firing Rate
Stimulus Rank
Selectivity Breadth Index
Skewness
SBI
skewness

example in Figure 11F is a rank-order object tuning curve that is sharper for the trained versus 180 degrees from the trained orientation (black is sharper than gray). For this example, there is essentially no change in SBI (see Figure 11G, left), but a clear decrease in skewness (see Figure 11G, right) between the trained and untrained orientations. The reason the two results differ is that SBI is computed from only three data points from the tuning curve (see equation 2.6) and skewness uses all 18 data points from the tuning curve (see equation 2.3). The three data points that SBI uses are the minimum, median, and maximum. Figure 11F shows that there is little change in those three data points between the trained and 180 degrees from trained orientation. Therefore, from the perspective of the SBI measurement, there is very little difference between the two tuning curves (see Figure 11G, left). However, the tuning curves are clearly different with one tuning curve (black) being sharper than the other (gray), which is reflected in the skewness results (see Figure 11G, right). This example illustrates how SBI can be a less robust, and therefore inaccurate, measure of overall changes in shape than skewness, since it depends on a limited sample of the tuning curve.

3.10 Comparing Skewness with Kurtosis. Another standard statistical measure that has been used previously to characterize the shape of tuning curves is kurtosis (Lehky et al., 2005, 2011; Lehky & Sereno, 2007). Both kurtosis and skewness are moments of a distribution (in addition to mean and variance) and are therefore very similar mathematically (compare equation 2.4 to 2.3). However, despite this mathematical similarity, skewness and kurtosis can capture very different properties of tuning curve shape. To illustrate the similarities and differences between skewness and kurtosis, we repeated all the skewness measurements shown in Figures 7, 10, and 11 using kurtosis instead (see Figure 12).

Figure 11: Examples of sharpening of object tuning based on presentation orientation (with respect to training orientation) for neurons from neurophysiological recordings in inferior temporal cortex (Freedman et al., 2006). (A) Normalized rank-order tuning curves. (B) Selectivity breadth index (SBI) versus presentation orientation. (C) Skewness versus presentation orientation. Population averages are presented in the bottom row of panels A–C (all data are from Freedman et al., 2006). Error bars are standard error with respect to neurons. (D) Comparison of change in SBI and change in skewness from the orientation used in training versus 180 degrees from the orientation used in training. (E) Histogram of correlation measured for individual neurons of SBI and skewness over presentation orientation. (F, G) Example of when SBI and skewness results disagree. (F) Object tuning measured at the orientation used in training versus 180 degrees from the orientation used in training. (G) SBI and skewness measurements computed from the tuning curves in panel F.

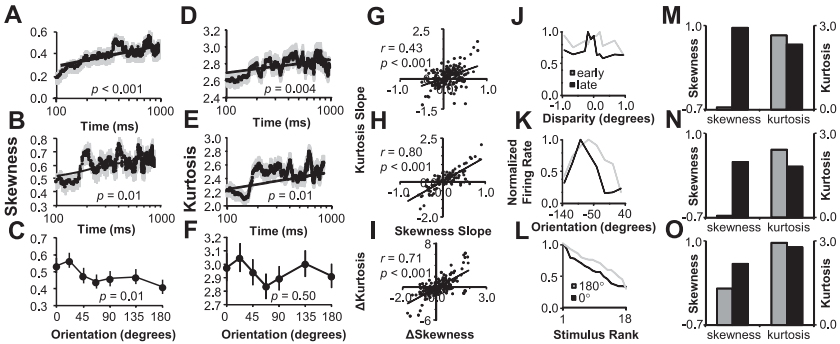


Figure 12: Comparison of skewness and kurtosis. (A) Skewness over time for disparity tuning ($n = 184$ neurons). (B) Skewness over time for orientation tuning ($n = 98$ neurons). (C) Skewness of object tuning versus presentation orientation with respect to orientation used for training ($n = 186$ neurons). (D) Kurtosis over time for disparity tuning. (E) Kurtosis over time for orientation tuning. (F) Kurtosis of object tuning versus presentation orientation with respect to orientation used for training. (G) Scatter plot of kurtosis slope versus skewness slope over log time for individual disparity tuning curves. (H) Scatter plot of kurtosis slope versus skewness slope over log time for individual orientation tuning curves. (I) Scatter plot of change in kurtosis versus change in skewness from a presentation orientation of 0 to 180 degrees with respect to orientation used for training. (J) Example early (100–250 ms) and late (400–850 ms) disparity tuning curves. (K) Example early (100–250 ms) and late (400–850 ms) orientation tuning curves. (L) Example object tuning at presentation orientations 0 and 180 degrees with respect to orientation used for training. (M) Skewness and kurtosis measurements of tuning curves in panel J. (N) Skewness and kurtosis measurements of tuning curves in panel K. (O) Skewness and kurtosis measurements of tuning curves in panel L. All error bars are standard error with respect to neurons.

For all three measurements, the average trend of kurtosis was similar (Figures 12D–F compared to Figures 12A–C) and highly correlated (see Figures 12G–I) to the average trend of skewness, as would be expected from the similar equations. The trends were statistically weaker for kurtosis compared to skewness for changes in sharpness of disparity tuning over time (significant positive slope: $p = 0.004$ versus $p < 0.001$; see Figure 12D compared to Figure 12A) and changes in sharpness of object tuning with respect to viewing orientation (significant difference between 0 and 180 degrees from training: $p = 0.50$ versus $p = 0.01$; see Figure 12F compared to Figure 12C). For changes in sharpness of orientation tuning over time, the trends were statistically equal in strength for skewness and kurtosis (see Figures 12B and 12E). Likewise, the correlation between kurtosis and skewness was stronger for changes in sharpness of orientation tuning over

time compared to changes in sharpness of disparity tuning over time or changes in sharpness of object tuning with respect to viewing orientation (see Figure 12H compared to Figures 12G and 12I).

We also examined individual cases where kurtosis and skewness strongly disagreed to gain some insight into why kurtosis produces a weaker statistical trend for some experiments. In Figure 2E, we showed that skewness increases for sharpened peaks but decreases for sharpened valleys or increases for broadened valleys. As a fourth-order moment, kurtosis increases for both sharpened peaks and sharpened valleys. As we showed in Figure 1, the peaks sharpen while the valleys broaden for disparity tuning over time. In Figure 12J, we present an example of disparity tuning that sharpens over time. In this example, there is an increase in skewness (see Figure 12M, left) and a decrease in kurtosis (see Figure 12M, right) over time. This is because the two valleys broaden while only the peak sharpens over time. In Figure 12L, we see a similar discrepancy arise between kurtosis and skewness for object tuning that sharpens for presentations from 0 to 180 degrees with respect to the trained orientation for the same reason that kurtosis and skewness disagreed for disparity tuning. In this case, because we are looking at tuning based on stimulus rank, the peak corresponds to rank 1 and the valley corresponds to rank 18. For stimuli presented at the trained orientation (see Figure 12L, black; 0 degrees), the response to different objects drops at a relatively faster rate (more vertical) near the peak (i.e., sharpening) and a relatively slower rate (more horizontal) near the valley (i.e., broadening) compared to the untrained orientation (see Figure 12L, gray; 180 degrees) resulting in an increase in skewness (see Figure 12O, left) and a decrease in kurtosis (see Figure 12O, left). Although kurtosis and skewness were highly correlated for orientation tuning (see Figure 12H) and orientation tuning has traditionally been modeled by the unimodal gaussian function (Henry et al., 1973), there were still cases where kurtosis and skewness could strongly disagree because of valleys for even orientation tuning. First, the underlying mechanisms of sharpened orientation tuning result in valleys, and these sharpened orientation tuning curves are modeled more accurately with difference-of-gaussians or difference–of–von Mises functions (Somers et al., 1995; Xing et al., 2005). These valleys can produce the same effect observed in Figures 12J and 12L. Second, if the orientation tuning curve spans more than 180 degrees (see Figure 12K, gray), the edges of this broad tuning curve effectively can be viewed as negative peaks or valleys that result in negative skewness and increased kurtosis (see Figure 12N). As the tuning curve becomes sharper over time (see Figure 12K, black), skewness increases and becomes positive (see Figure 12O, left), but kurtosis decreases (see Figure 12O, right). This discrepancy between kurtosis and skewness can be reduced by testing a wider range of orientations, but in many cases, a valley will still be present over this wider range because the response will start to increase again for orientations drifting in the opposite direction. Overall, skewness was more

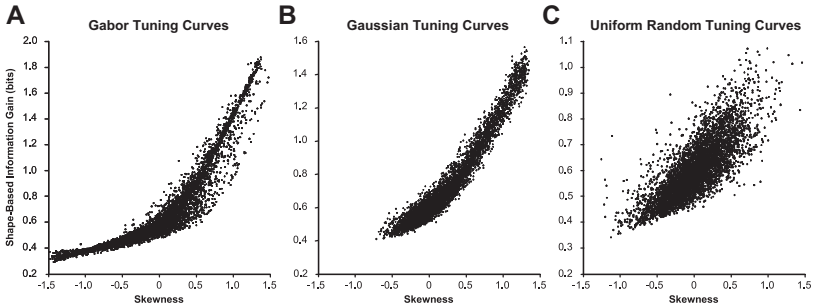


Figure 13: For various families of neural tuning curves, skewness is closely related to the information gain produced by a single spike once tuning curves are normalized by amplitude and mean firing rate.

effective at capturing the general observed changes in sharpness for these three experiments.

3.11 The Relationship Between an Information-Theoretic Definition of Sharpness and Skewness. Finally, we tested whether skewness was consistent or correlated with our information-theoretic definition of sharpness (see section 2.4). We generated a set of random tuning curves of various complexities and compared the shape-based information gain $\tilde{I}[\theta|r > 0]$ with the skewness of these tuning curves. In Figure 13, we show scatter plots of shape-based information gain versus skewness for 5000 randomly generated tuning curves for three classes of tuning curves. In all classes, there were $N = 13$ data points across the feature θ in the tuning curve, and the shape-based information gain was measured within $\tau = 10$ ns. Thirty percent uniform white noise was added to the tuning curves to simulate measurement error. Rather than normalize the heights of the tuning curves f by the minimum and maximum, which are unstable and sensitive to outliers, we normalized them by the 15th and 85th percentiles, respectively. In Figure 13A, $f(\theta)$ was set as a Gabor function with a random phase, wavelength between two and four stimulus units, and a gaussian envelope with a standard deviation selected uniformly between 1 and 4. In Figure 13B, $f(\theta)$ was set as a gaussian function with a standard deviation selected uniformly between 1 and 4 and a preferred stimulus chosen randomly from the central seven stimuli. In Figure 13C, $f(\theta)$ was selected independently for each θ from a uniform distribution from 0 to 60 sps$\cdot\tau$. In all three classes, the shape-based information gain was highly correlated with skewness.

We can also use a Taylor series argument to show that the shape-based information gain tightly correlates with skewness. First, let us define

$$\tilde{f}(\theta) = \frac{f(\theta) - \bar{f}}{\sigma}, \tag{3.1}$$

where $\bar{f}$ is the mean firing rate and σ is the standard deviation. Note that equation 3.1 is the z-score of the neural responses. Our goal is to compute the information gain of a single spike for the tuning curve $A \cdot \tilde{f}(\theta) - B$, where B is the baseline firing rate of the tuning curve or the minimum of $\tilde{f}$ and A is the amplitude of the tuning curve or the maximum of $\tilde{f}$ minus the minimum of $\tilde{f}$. Let us define

$$g(x) \equiv (1 - e^{-Ax+B}) \log_2(1 - e^{-Ax+B}) \tag{3.2}$$

so that

$$\tilde{I}[\theta|r > 0] = \sum\nolimits_{\theta} g(\tilde{f}(\theta)). \tag{3.3}$$

We can then write the Taylor series of $g(x)$ as

$$g(x) \approx \sum_{n=0}^{\infty} x^n \frac{g^{(n)}(0)}{n! \log 2}, \tag{3.4}$$

where $g^{(n)}(0)$ is the nth order derivative of g evaluated at zero. Note that $g^{(n)}(0)$ does not depend on x. Because $\tilde{f}$ is normalized, we know that $\frac{1}{N}\sum \tilde{f}(\theta) = 0$ and $\frac{1}{N}\sum \tilde{f}(\theta)^2 = 1$ and that $\frac{1}{N}\sum \tilde{f}(\theta)^3 = \textbf{skewness}\ [f]$ (see equation 2.3). Thus, the first four terms of the Taylor series expansion are linear with respect to skewness.

It can be shown that the values of $g^{(n)}(0)$ are given by

$$g(0) = (1 - e^B) \log(1 - e^B), \tag{3.5}$$

$$g^{(n)}(0) = (-1)^{n-1} A^n e^B \left(\frac{p_n(e^{-B})}{(e^{-B} - 1)^n} + \log(1 - e^B) \right), \tag{3.6}$$

where each $p_n(s)$ is a polynomial of order n that is defined recursively as

$$p_1(s) = 1, \tag{3.7}$$

$$p_{n+1}(s) = nsp(s) - s^2 p'(s) - p(s) + sp'(s) - (s-1)^{n-1}. \tag{3.8}$$

This can be verified by repeatedly differentiating g. Because $\tilde{f}$ has unit variance, B must be below -1. When $B \leq -1$, $\frac{t_n}{n! A^n}$ can be shown to quickly approach zero. Thus, $\frac{t_n}{n!}$ must approach zero faster than exponentially, or A^n. Recall that τ is generally set to some near-infinitesimal value. Thus, A should be much smaller than 1, and so $\frac{t_n}{n!}$ must converge quickly. In summary, the first three terms of the Taylor series expansion are independent of f, the fourth term is proportional to skewness[f], and subsequent terms are

comparatively small. Thus, shape-based information gain $\tilde{I}[\theta|r > 0]$ must closely correlate with skewness.

In Figure 13, the information gain flattens or continues to decrease from zero skewness to negative skewness values. These data points include tuned inhibitory neurons, which exhibit high firing rates for nonpreferred stimuli and a decreased firing rate for their preferred stimulus. We previously found that for tuned inhibitory disparity tuning, the negative primary peak around the preferred disparity sharpened and skewness decreased to progressively more negative values (Samonds et al., 2013). As noted in section 2.4, neurons can also convey information by silence when a tuning curve has a primary negative peak. Because silence occurs more often than spikes, the information gain of silence is generally much weaker than the information gain of a spike. Nonetheless, the two types of information gain can be related with respect to tuning curve sharpness. A tuning curve for a tuned excitatory neuron $f_e(\theta)$ can be expressed in terms of a tuned inhibitory neuron $f_i(\theta)$ as

$$f_e(\theta) = -A \cdot f_i(\theta) + B. \tag{3.9}$$

For any positive A and B, the skewness of $f_e(\theta)$ will be the arithmetic opposite of the skewness of $f_i(\theta)$. In addition, the shape-based information gain of silence for $f_i(\theta)$ would increase for progressively more negative skewness in a similar manner as the shape-based information gain of a spike for $f_e(\theta)$ increases for progressively more positive skewness. Therefore, decreases in skewness can still be applied to capture increases in sharpness for those special-case tuning curves that have a reduced response for a preferred feature.

4 Discussion

We introduced the method of using the statistical measurement of the sample skewness of the distribution of mean firing rates of a tuning curve to measure changes in sharpness. We applied this nonparametric method to binocular disparity tuning curves measured from example model neurons and recorded neurons in V1, orientation tuning curves measured from recordings in V1 and V2, and object tuning curves from recordings in ITC with visible changes in sharpness. We found it difficult to quantify changes in sharpness for complex tuning curves such as disparity tuning with a single parameter from functions such as a Gabor because the tuning explicitly deviated from that function as it sharpened. Skewness provided us with a simpler measurement that can increase the statistical power of a set of data compared to parametric methods by including more neurons, and skewness captures several observed sharpening behaviors. Finally, we demonstrated that changes in skewness directly correspond to changes in information about tuning curve shape.

4.1 Implications of Measuring Tuning Curve Sharpness. Sharpness of tuning can have implications on feature discrimination (Spitzer et al., 1988; Freedman et al., 2006; Lee et al., 2012) and optimizing feature representations (Tolhurst et al., 1983; Bradley et al., 1987; Scobey & Gabor, 1989; Geisler & Albrecht, 1997; Purushothaman & Bradley, 2005; Butts & Goldman, 2006). It is important to reiterate, however, that there are many different aspects of neural responses that can be measured and characterized using information theory (Borst & Theunissen, 1999), and many different properties, in addition to sharpness, can influence discrimination based on tuning curves (Pouget et al., 1999; Zhang & Sejnowski, 1999; Kang, Shapley, & Sompolinsky, 2004; Series et al., 2004; Purushothaman & Bradley, 2005; Butts & Goldman, 2006). In this letter, we illustrate how sharpness of tuning influences the information gain of a spike, if all these other neural properties are held constant between tuning curves with differing sharpness.

There is also importance to quantifying sharpness of tuning beyond behavioral changes in feature discrimination and the potential changes in information provided about the feature. The relationship between sharpness and stimuli properties (Chen et al., 2005; Xing et al., 2005; Samonds et al., 2013), the temporal dynamics of sharpness (Ringach et al., 1997; Menz and Freeman, 2003; Samonds et al., 2013), the dependence of sharpness based on cortical layer (Blasdel & Fitzpatrick, 1984; Fitzpatrick et al., 1997) or cell types (Lee et al., 2012), and changes in sharpness from attention (Spitzer et al., 1988) or training (Freedman et al., 2006) can all provide valuable information about local underlying recurrent inputs. Recurrent circuits can play a role in contour integration and interpolation (Lee & Nguyen, 2001; Li, Piech, & Gilbert, 2006), surface interpolation (Samonds, Tyler, & Lee, 2014), and development (Li, Van Hooser, Mazurek, White, & Fitzpatrick, 2008) so understanding their spatial and temporal characteristics is important for understanding many aspects of visual processing.

Olshausen and Field (1996) have proposed that a sparseness constraint for encoding sensory information could explain the development of simple cell receptive field properties for neurons in the primary visual cortex. Using a sparse coding strategy has many computational advantages in representing and storing information among a population of neurons (Barlow, 1972; Vinje & Gallant, 2000; Olshausen & Field, 2004). Sparseness of responses among a population of neurons and sharpness of tuning of individual neurons to complex or natural sets of stimuli are conceptually related quantities (Rolls & Tovee, 1995; Lehky et al., 2005, 2011; Franco, Rolls, Aggelopoulos, & Jerez, 2007) and many of the same nonparametric measures that have been used to capture sharpness and selectivity of tuning curves have also been used to measure the sparseness of responses among a population of neurons (Olshausen & Field, 1996, 2004; Bell & Sejnowski, 1997; Vinje & Gallant, 2000; Simoncelli & Olshausen, 2001; Lehky et al., 2005, 2011; Franco et al., 2007; Lehky & Sereno, 2007; Tolhurst et al., 2009). Likewise, the arguments we make about skewness being a good measure for capturing

changes in tuning curve sharpness would also apply to skewness being a good measure of population sparseness.

4.2 Difficulties of Parametric Methods in Describing Sharpness. The primary problem with fitting a Gabor function to tuning curves or computing the Fourier transform of a tuning curve to characterize sharpness is that both methods assume that an increase in sharpness corresponds to an increase in the frequency of the disparity tuning function. This assumption was not true based on our observations of disparity tuning dynamics (see Figure 5; however, see also Menz & Freeman, 2003). The primary peak in the disparity tuning function was increasing in frequency, but the valleys were decreasing in frequency (see Figure 1). Prince et al. (2002) also previously noted that their Gabor fits deviated from their data in that measured tuning curves had side flanks that were wider than the peak. A difference-of-gaussians function was better at capturing the observed dynamics of disparity tuning (see Figure 6), but it still produced statistically weaker results than skewness because the function does not describe disparity tuning well for all neurons (see Figures 8 and 9).

Overall, parametric methods can be inadequate in cases where the source of those changes (e.g., recurrent interactions) causes tuning curves to explicitly deviate from any simple or complex functions that model the tuning curve. Parametric methods will succeed only when the magnitude of changes in sharpness is larger than the differences between the observed data and fitted function to explain the tuning, as well as when sharpening can be incorporated explicitly and simply into the function that describes the tuning curve. Indeed, parametric methods such as the difference-of-gaussians fits were comparable to skewness if we limited our analysis to only those neurons that were well described by the difference-of-gaussians function (see Figures 9C to 9E). However, that criterion severely reduced our data from 184 to 44 neurons, which means that we would have needed four times as many data to observe informative trends of disparity tuning dynamics (Samonds et al., 2013). In addition, not all disparity tuning curves for neurons will necessarily be described well by one particular function.

4.3 Advantages and Disadvantages of Sample Skewness in Capturing Sharpness. Skewness captures several behaviors of a complex model in a single value because it increases for narrower peaks, broader valleys, and suppression of secondary peaks. A minimum of two parameters is necessary to capture more than one of these sharpness characteristics. This can be viewed as an advantage for skewness because you can quantify changes in sharpness with that one value produced by a relatively simple computation (see equation 2.3) that requires no fits, no parameter initialization, and no interpolation. However, this can also be viewed as a disadvantage for skewness in that the skewness value alone does not describe which specific aspect of the tuning curve shape changed that caused skewness to increase.

Therefore, parametric methods, even when applied to a subset of data where model fits are adequate, can complement skewness in providing additional information about specific changes in tuning curve shape.

Another advantage of skewness and nonparametric methods in general is that they can be applied to responses to any set of stimuli, whereas parametric methods must use a set of stimuli defined by at least one of the model parameters. This allows nonparametric methods to be applied to more complex sets of stimuli such as natural images and movies, which can be used to characterize neuronal responses under more ecological conditions (Lehky et al., 2005). This also makes nonparametric methods more suitable for characterizing tuning for more complex neurons that are still poorly understood with respect to the precise sensory parameters that modulate their responses such as IT neurons (Lehky et al., 2005, 2011). However, this flexibility of nonparametric measurements leads to a dependence of the measure on the set of stimuli. The relative number of nonpreferred stimuli compared to preferred stimuli can lead to large changes in skewness values. For a parametric method such as the bandwidth of an orientation tuning curve, the measurement will be more stable across different sets of stimuli. This makes it more difficult to compare skewness across different experiments and neurons and makes it a more suitable measurement for characterizing changes in sharpness rather than absolute levels of sharpness.

The main advantage of skewness, and most standard statistical measurements in general, is their simplicity because they characterize the shape based on variation from the mean or median of the tuning curve and do not depend on any specific model of tuning. This advantage of skewness in capturing sharpness compared to any of the parameters from model fits (see Figure 9) allowed us to examine more neurons, which allowed us to observe a variety of more subtle dynamics in disparity tuning sharpening that have provided us deeper insight into the underlying recurrent circuitry in V1 (Samonds et al., 2013). Finally, we demonstrated that skewness can capture sharpness for tuning curves from multiple areas of the brain that encode visual information with varying levels of complexity (see Figures 7, 10, and 11). Because diversity in tuning curve shapes for even one feature may not be described well by a single function, this flexibility of skewness in capturing sharpness for a variety of functions allows it to be applied to all tuning curves regardless of their shape. Finally, we demonstrated that skewness for a wide range of tuning curve functions is strongly correlated with the information gain per spike in a coding framework that benefits from sharp tuning (see Figure 13) and derived a direct relationship between skewness and information gain. This formalization provides a clear definition of tuning curve sharpness and theoretical justification for applying skewness to describe changes in tuning curve shape.

4.4 Comparison of Sample Skewness to Other Nonparametric Measures of Sharpness. First, we applied skewness to orientation tuning

curves and compared the results to another statistical measure (circular variance) that several previous authors have applied to describe the sharpening of orientation tuning (Ringach et al., 1997, 2002; Gur, Kagan, & Snodderly, 2005; Tao, Cai, McLaughlin, Shelley, & Shapley, 2006; Zhu, Xing, Shelley, & Shapley, 2010; Chavane et al., 2011; Nauhaus, Nielsen, Disney, & Callaway, 2012; Nandy, Sharpee, Reynolds, & Mitchell, 2013). Circular variance and skewness produced similar results in capturing visible sharpening of orientation tuning over time (see Figure 10). However, there are important differences between the measures:

- Circular variance is limited to circular variables like angle, while skewness does not depend on the variable. Therefore, circular variance could not be applied to disparity and object tuning curves.
- Circular variance is limited to unimodal tuning curves. For example, skewness would still conclude a tuning curve was sharper with two sharpened peaks while circular variance would not necessarily conclude that the tuning curve was sharper.
- Circular variance is normalized for amplitude but not for the mean, while skewness is normalized for both, which means skewness more accurately describes changes in tuning curve shape such as sharpness (see Figures 10F and 10G).

The statement that skewness is a more accurate description of changes in tuning curve shape such as sharpness does not imply that skewness is also a better measurement of tuning curve selectivity or discrimination compared to circular variance, as well as many other nonparametric measures of tuning curve selectivity that do not normalize for changes in mean or baseline firing rates (Rolls & Tovee, 1995; Rainer, Assad, & Miller, 1998; Moody, Wise, di Pellegrino, & Zipser, 1998). As we noted in section 2.4, the mean and amplitude of the tuning curve can strongly influence our measure of information gain. Tuning curve discrimination based on the Chernoff distance is also highly sensitive to baseline firing rate (Kang et al., 2004). This leads to a strong correlation between circular variance and Chernoff distance compared to sharpness based on tuning curve width, although the correlation between Chernoff distance and sharpness increases when making very fine discriminations. Therefore, one nonparametric measurement is not necessarily more appropriate than the other in describing selectivity; rather they can be complementary and capture changes to different properties of tuning curves that may be influenced by different underlying mechanisms and may serve different neural computational roles.

Next, we applied skewness to object tuning curves and compared the results to the nonparametric measure called the selectivity breadth index (SBI) that previous authors have applied to describe sharpness of object tuning (Freedman et al., 2006). SBI and skewness capture very similar statistical properties, so it was not surprising that they also produced similar results in capturing visible changes in sharpness such as for object tuning with respect

to viewing orientation (see Figure 11). Although the two measures are based on capturing similar behavior, they can be influenced very differently by different changes in the tuning curve. These differences become more pronounced with a greater number of data points included in the tuning curves. For example, SBI can be strongly influenced by changes or noise for only the minimum, median, and maximum value in the tuning curve. With enough data points, there would be little influence on skewness from changes or noise to just those same three data points because skewness incorporates all data points from the tuning curve into its computation. This means that skewness would generally be more robust to outlier noise and a more accurate measurement of sharpness than SBI (see Figures 11F and 11G).

We also compared skewness to kurtosis for disparity, orientation, and object tuning curves (see Figure 12), which has been used previously to characterize the selectivity of tuning curves (Lehky et al., 2005, 2011; Lehky & Sereno, 2007) and has also been used widely to describe sparseness of responses among a population of neurons (Olshausen & Field, 1996; Lewicki & Sejnowski, 2000; Willmore & Tolhurst, 2001; Olshausen & Field, 2004). Although kurtosis and skewness are very similar mathematically, they measure different properties of a distribution and will capture different properties about tuning curve shape. Skewness, kurtosis, and the other moments are related to the Taylor series coefficients of the moment-generating function of a distribution, which provides a way to alter the skewness of any distribution without affecting the kurtosis, and vice versa as long as the moments exist. Just as there are families of distributions with fixed skewness that have differing kurtosis, there are other families of distributions with fixed kurtosis that have varying skewness. The parameters of the beta distribution, for example, can be expressed in terms of kurtosis and skewness, which is often how beta distributions are fit from data, using the method-of-moments. For a simple example of how to alter skewness without altering kurtosis, consider that for any probability distribution function $p(x)$, the distribution $p(-x)$ will have equal kurtosis but opposite skewness. The examples in Figures 12M and 12N (left) nearly illustrate this example: skewness changes from negative to positive between the two conditions. Kurtosis does not remain constant between these conditions but does decrease (see Figures 12M and 12N, right), which is a change in the opposite direction as the change in skewness. Vice versa, symmetrical sharpening or broadening of $p(x)$ will alter kurtosis without altering skewness. Thus, kurtosis cannot capture all properties that are captured by skewness and skewness cannot capture all properties that are captured by kurtosis. Instead it is more accurate to say that in spite of the similarity in their mathematical forms, skewness and kurtosis capture different (and independent) properties of distributions.

In terms of neural tuning, kurtosis increases when both peaks (near-maximal firing) and valleys (near-minimal firing) of a tuning curve sharpen. In contrast, skewness increases when peaks sharpen and valleys broaden.

The results of kurtosis and skewness will be tightly coupled for unimodal tuning curves with positive peaks (single peak and no valleys) and sampling that includes enough data below the mean (e.g., stimuli with a baseline response) so that skewness is always positive. Indeed, for orientation tuning curves, which most closely followed those conditions, we did observe the most similar results and strongest correlation between kurtosis and skewness (see Figure 12H). However, because we observed disparity tuning curves that strongly exhibited sharpening peaks and broadening valleys (e.g., see Figure 1A), we found skewness to be a more appropriate measure of changes in neural tuning shape for our experiments. The increases to kurtosis from a sharpened peak were canceled out by decreases in kurtosis from broadened valleys, which lead to statistically weaker trends for a population of neurons (see Figure 12). This problem was strongest for disparity and object tuning curves, which had clear multimodal features (see Figures 12J and 12L), but even factored into measurements for orientation tuning curves that also exhibit some multimodal features (see Figure 12K). Because skewness and kurtosis capture changes to different properties of tuning curve shape, they, like some of the other nonparametric measures described in this section, can be complementary. We have described the specific tuning curve shape properties that skewness captures and demonstrated how those properties influence the information gain of a spike. Lehky et al. (2005) have described the tuning curve properties that kurtosis captures and demonstrated how those properties influence the entropy of a tuning curve.

We have compared or described only a small subset of alternative nonparametric measures that can quantify sharpness of tuning curves. We chose particular examples to compare to skewness to highlight some general differences between common measures of sharpness and selectivity. There are potentially endless numbers of statistical measures that could also adequately quantify sharpness of tuning, including other measures of skewness. Many of these nonstandard measures, though, are less robust than skewness and prone to similar errors as those that we observed with SBI (see Figures 11F and 11G). Because many nonparametric measures capture distinct properties of tuning curve shape, it is important to understand their mathematical properties and what definition of sharpness or selectivity is being expressed when applying those measures. We have made a case about the specific benefits and provided an information-theoretic definition for the standard measure of sample skewness in capturing multiple sharpness properties for a diverse set of tuning curve shapes and complexities.

5 Conclusion

We applied the statistical measurement of skewness to our binocular disparity tuning curves because we were unable to reveal properties of the underlying network interactions in V1 when attempting to fit commonly used Gabor functions over time. The tuning curves were diverging from

Gabor functions over time in a complex manner that could not be captured by a single parameter. In this letter, we provide an intuitive rationale, empirical evidence, a statistical motivation, and an information-theoretic argument for using skewness to quantify the sharpness of tuning, which can be applied to a wide range of tuning curves throughout the nervous system.

Acknowledgments

We appreciate the technical assistance provided by Karen McCracken, Ryan Poplin, Matt Smith, Ryan Kelly, Nicholas Hatsopoulos, and Charles Gray; the ITC data provided by David Freedman; and helpful feedback from Bruce Cumming on earlier versions of the letter. This work was supported by NIH F32 EY017770, NSF CISE IIS 0713206, NIH R01 EY022247, AFOSR FA9550-09-1-0678, NIH P41 EB001977, and a grant from the Pennsylvania Department of Health through the Commonwealth Universal Research Enhancement Program.

References

Adrian, E. D., & Zotterman, Y. (1926). The impulses produced by sensory nerve endings. Part 3: Impulses set up by touch and pressure. *Journal of Physiology (London)*, *61*, 465–483.

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*(6), 1106–1130.

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, *1*(4), 371–394.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, *37*(23), 3327–3338.

Blasdel, G. G., & Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. *Journal of Neuroscience*, *4*(3), 880–895.

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, *2*(11), 947–957.

Bradley, A., Skottun, B. C., Ohzawa, I., Sclar, G., & Freeman, R. D. (1987). Visual orientation and spatial frequency discrimination: A comparison of single neurons and behavior. *Journal of Neurophysiology*, *57*(3), 755–771.

Bredfeldt, C. E., & Ringach, D. L. (2002). Dynamics of spatial frequency tuning in macaque VI. *Journal of Neuroscience*, *22*, 1976–1884.

Butts, D. A., & Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLOS Biology*, *4*(4), 639–646.

Chavane, F., Sharon, D., Jancke, D., Marre, O., Frégnac, Y., & Grinvald, A. (2011). Lateral spread of orientation selectivity in V1 is controlled by intracortical cooperativity. *Frontiers in Systems Neuroscience*, *5*(4), 1–26.

Chen, G., Dan, Y., & Li, C. Y. (2005). Stimulation of non-classical receptive field enhances orientation selectivity in the cat. *Journal of Physiology*, *564*(1), 233–243.

DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, *89*(2), 1094–1111.

Fitzpatrick, D. C., Batra, R., Stanford, T. R., & Kuwada, S. (1997). A neuronal population code for sound localization. *Nature*, *388*(6645), 871–874.

Franco, L., Rolls, E. T., Aggelopoulos, N. C., & Jerez, J. M. (2007). Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex. *Biological Cybernetics*, *96*(6), 547–560.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*(5502), 312–316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2006). Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral Cortex*, *16*(11), 1631–1644.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience*, *14*(5), 897–919.

Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *Journal of Neuroscience*, *2*(11), 1527–1537.

Gur, M., Kagan, I., & Snodderly, D. M. (2005). Orientation and direction selectivity of neurons in V1 of alert monkeys: Functional relationships and laminar distributions. *Cerebral Cortex*, *15*(8), 1207–1221.

Henry, G. H., Bishop, P. O., Tupper, R. M., & Dreher, B. (1973). Orientation specificity and response variability of cells in striate cortex. *Vision Research*, *13*(9), 1771–1779.

Hinkle, D. A., & Connor, C. E. (2001). Disparity tuning in macaque area V4. *NeuroReport*, *12*(2), 365–369.

Kang, K., Shapley, R. M., & Sompolinsky, H. (2004). Information tuning of populations of neurons in primary visual cortex. *Journal of Neuroscience*, *24*(15), 3726–3735.

Lee, S. H., Kwan, A. C., Zhang, S., Phoumthipphavong, V., Flannery, J. G., Masmanidis, S. C., . . . Dan, Y. (2012). Activation of specific interneurons improves V1 feature selectivity and visual perception. *Nature*, *488*(7411), 379–383.

Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Sciences USA*, *98*(4), 1907–1911.

Lehky, S. R., Kiani, R., Esteky, H., & Tanaka, K. (2011). Statistics of visual responses in primate inferotemporal cortex to object stimuli. *Journal of Neurophysiology*, *106*(3), 1097–1117.

Lehky, S. R., Sejnowski, T. J., & Desimone, R. (2005). Selectivity and sparseness in the responses of striate complex cells. *Vision Research*, *45*(1), 57–73.

Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology*, *97*(1), 307–319.

Lewicki, M. S., & Sejnowski, T. J. (2000). Learning overcomplete representations. *Neural Computation*, *12*(2), 337–365.

Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, *50*(6), 951–962.

Li, Y., Van Hooser, S. D., Mazurek, M., White, L. E., & Fitzpatrick, D. (2008). Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature*, *456*(7224), 952–956.

Menz, M. D., & Freeman, R. D. (2003). Stereoscopic depth processing in the visual cortex: A coarse-to-fine mechanism. *Nature Neuroscience*, *6*(1), 59–65.

Moody, S. L., Wise, S. P., di Pellegrino, G., & Zipser, D. (1998). A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. *Journal of Neuroscience*, *18*(1), 399–410.

Nandy, A. S., Sharpee, T. O., Reynolds, J. H., & Mitchell, J. F. (2013). The fine structure of shape tuning in area V4. *Neuron*, *78*(6), 1102–1115.

Nauhaus, I., Nielsen, K. J., Disney, A. A., & Callaway, E .M. (2012). Orthogonal microorganization of orientation and spatial frequency in primate primary visual cortex. *Nature Neuroscience*, *15*(12), 1683–1690.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, *249*(4972), 1037–1041.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607–609.

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, *14*(4), 481–487.

Pouget, A., Deneve, S., Ducom, J.-C., & Latham, P. E. (1999). Narrow versus wide tuning curves: What's best for a population code? *Neural Computation*, *11*(1), 85–90.

Prince, S. J. D., Pointon, B. G., Cumming, B. G., & Parker, A. J. (2002). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *Journal of Neurophysiology*, *87*(1), 191–208.

Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience*, *8*(1), 99–106.

Rainer, G., Assad, W. F., & Miller, E. K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, *393*(6685), 577–579.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, *387*(6630), 281–284.

Ringach, D. L., Shapley, R., & Hawken, M. J. (2002). Orientation selectivity in macaque V1: Diversity and laminar dependence. *Journal of Neuroscience*, *22*(13), 5639–5651.

Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, *73*(2), 713–726.

Salazar, R. F., Dotson, N. M., Bressler, S. L., & Gray, C. M. (2012). Content-specific fronto-parietal synchronization during visual working memory. *Science*, *338*(6110), 1097–1100.

Samonds, J. M., Potetz, B. R., & Lee, T. S. (2009). Cooperative and competitive interactions facilitate stereo computations in macaque primary visual cortex. *Journal of Neuroscience*, *29*(50), 15780–15795.

Samonds, J. M., Potetz, B. R., & Lee, T. S. (2012). Relative luminance and binocular disparity preferences are correlated in macaque primary visual cortex, matching natural scene statistics. *Proceedings of the National Academy of Sciences USA*, *109*(16), 6313–6318.

Samonds, J. M., Potetz, B. R., Tyler, C. W., & Lee, T. S. (2013). Recurrent connectivity can account for the dynamics of disparity processing in V1. *Journal of Neuroscience*, *33*(7), 2934–2946.

Samonds, J. M., Tyler, C. W., & Lee, T. S. (2014). *Evidence of stereoscopic surface disambiguation and interpolation in the responses of V1 neurons*. Manuscript submitted for publication.

Scobey, R. P., & Gabor, A. J. (1989). Orientation discrimination sensitivity of single units in cat primary visual cortex. *Experimental Brain Research*, *77*(2), 398–406.

Series, P., Latham, P. E., & Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: Coding efficiency and the impact of correlations. *Nature Neuroscience*, *7*(10), 1129–1135.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216.

Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, *15*(8), 5448–5465.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavorial and neuronal performance. *Science*, *240*(4850), 338–340.

Suga, N., Zhang, Y., & Yan, J. (1997). Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *Journal of Neurophysiology*, *77*(4), 2098–2114.

Tanaka, K., Siato, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*(1), 170–189.

Tao, L., Cai, D., McLaughlin, D. W., Shelley, M. J., & Shapley, R. S. (2006). Orientation selectivity in visual cortex by fluctuation-controlled criticality. *Proceedings of the National Academy of Sciences USA*, *103*(34), 12911–12916.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, *23*(8), 775–785.

Tolhurst, D. J., Smyth, D., & Thompson, I. D. (2009). The sparseness of neuronal responses in ferret primary visual cortex. *Journal of Neuroscience*, *29*(8), 2355–2370.

Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287*(5456), 1273–1276.

Willmore, B., & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network: Computation in Neural Systems*, *12*(3), 255–270.

Xing, D., Shapley, R. M., Hawken, M. J., & Ringach, D. L. (2005). Effect of stimulus size on the dynamics of orientation selectivty in macaque V1. *Journal of Neurophysiology*, *94*(1), 799–812.

Zhang, K., & Sejnowski, T. J. (1999). Neural tuning: To broaden or sharpen? *Neural Computation*, *11*(1), 75–84.

Zhu, W., Xing, D., Shelley, M., & Shapley, R. (2010). Correlation between spatial frequency and orientation selectivity in V1 cortex: Implications of network model. *Vision Research*, *50*(22), 2261–2273.

Received July 18, 2013; accepted December 9, 2013.