Model neuron response statistics to natural images

Arvind Iyer[1] & Johannes Burge[1,2,3]

[1] Department of Psychology, University of Pennsylvania, Philadelphia PA, USA
[2] Neuroscience Graduate Group, University of Pennsylvania, Philadelphia PA, USA
[3] Bioengineering Graduate Group, University of Pennsylvania, Philadelphia PA, US

Abstract

To model the responses of neurons in the early visual system, at least three basic components are required: a receptive field, a normalization term, and a specification of encoding noise. Here, we examine how the receptive field, the normalization factor, and the encoding noise impact the model neuron responses to natural images and the signal-to-noise ratio for natural image discrimination. We show that when these components are modeled appropriately, the model neuron responses to natural stimuli are Gaussian distributed, scale-invariant, and very nearly maximize the signal-to-noise ratio for stimulus discrimination. We discuss the statistical models of natural stimuli that can account for these response statistics, and we show how some commonly used modeling practices may distort these results. Finally, we show that normalization can equalize important properties of neural response across different stimulus types. Specifically, narrowband (stimulus- and feature-specific) normalization causes model neurons to yield Gaussian-distributed responses to natural stimuli, 1/f noise stimuli, and white noise stimuli. The current work makes recommendations for best practices and it lays a foundation, grounded in the response statistics to natural stimuli, upon which principled models of more complex visual tasks can be built.

Keywords

Introduction

As interest intensifies in understanding natural signals in vision and neuroscience, it becomes increasingly important to develop a clear picture of how neural systems and their constituent components respond to real-world (photographic) images. Characterizing the statistical properties of these responses is vital for building principled models of visual processing, especially given the increasing reliance of vision and visual neuroscience on probability theory (Knill & Richards, 1996). Over the past two decades, there have been many neurophysiological (Baudot et al., 2013; Burkhardt, Fahey, & Sikora, 2006; Butts et al., 2010; Felsen, Touryan, Han, & Dan, 2005; Lesica et al., 2007; Talebi & Baker, 2012; Weliky, Fiser, Hunt, & Wagner, 2003) and computational (Brady & Field, 2000; Burge & Geisler, 2014; 2015; Clatworthy, Chirimuuta, Lauritzen, & Tolhurst, 2003; Lyu & Simoncelli, 2008; 2009b; Sebastian, Abrams, & Geisler, 2017; Tadmor & Tolhurst, 2000; Wainwright & Simoncelli, 2000) attempts to address this issue.

We report a large-scale analysis of model neuron response statistics to natural images. To model neural responses, at least three basic components are required: a receptive field, a normalization term, and a specification of encoding noise. The receptive field specifies the neuron's preferred stimulus feature, and indicates how inputs are weighted and summed (i.e. pooled) across space to determine the response (Hubel & Wiesel, 1962; 1968). The normalization term specifies how gain control is implemented to prevent the neuron's response from exceeding its dynamic range (Albrecht & Geisler, 1991; Heeger, 1992). The encoding noise specifies the uncertainty in the response to repeated presentations of the same stimulus (Tolhurst, Movshon, & Dean, 1983a).

We examine how methods for modeling the receptive field, the normalization factor, and the encoding noise impact the response statistics to natural images and the signal-to-noise ratio for stimulus discrimination. We focus our analysis on model neurons with oriented receptive fields like those in early visual cortex. We show that feature-specific normalization of each stimulus, with an easy to compute normalization factor, yields model neuron responses to natural images that are Gaussian distributed. We show that the stimulus-driven response variance is invariant to the scale of the preferred feature. We also show that responses with these statistics nearly maximize the signal-to-noise ratio for stimulus discrimination given the noisy neural response. Furthermore, we show that subtle variants of the standard response model, which are used across the vision and computational neuroscience communities, have an important impact on the results described above. To achieve scientific consensus on basic facts about natural stimulus processing early in the visual system, it is important to understand how variants of the standard response model impact response statistics.

Results

How do common receptive field modeling choices impact response statistics and natural image discrimination? To investigate, we first establish the relation between response variability and stimulus discriminability. Second, we describe a common model of neural response in early visual cortex and discuss how two different forms of normalization impact the response statistics and impact natural stimulus discriminability. Third, we discuss the statistical models of natural stimuli that can account for these results.

Stimulus discriminability from neural response

Consider a model neuron that produces a particular response distribution across tens of thousands of natural stimuli. Any early visual representation must be capable of distinguishing two arbitrary stimuli from one another. How well can a neuron distinguish to arbitrary stimuli from the ensemble of natural stimuli? To assess the signal-to-noise ratio (SNR) for stimulus discrimination given the responses of a particular neuron, we compute the discriminability of two arbitrary stimuli, randomly sampled from the natural distribution of stimuli.
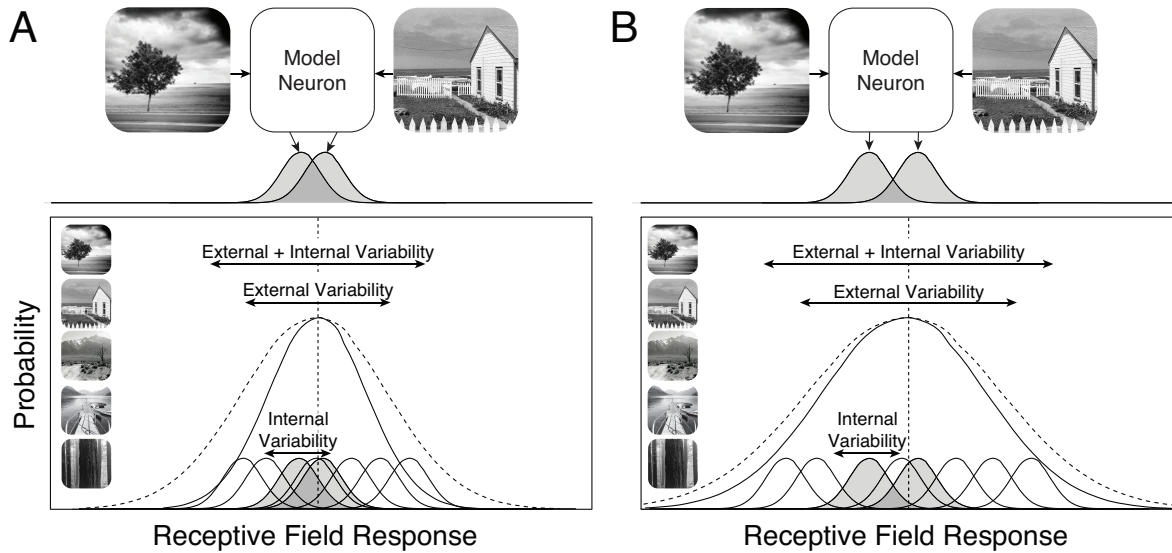
Figure 1. Stimulus-driven response variability and the signal-to-noise ratio for stimulus discrimination. **A** Model neuron yielding low stimulus-driven response variance. Two natural images elicit encoding-noise-corrupted responses (shaded bell curves) that are hard to discriminate based on the responses (top). Low stimulus-driven response variability to the natural image ensemble is associated with poor signal-to-noise for stimulus discriminability (bottom). Two randomly stimuli will be hard to discriminate on average. **B** Model neuron with high stimulus-driven response variance. The same two random stimuli are now easier to discriminate, as will randomly sampled images from the natural image ensemble.

The discriminability (i.e. d-prime) of any two stimuli based this neuron's response is given by

$$SNR_{ij} = d'_{ij} = \frac{\left| r_i - r_j \right|}{\sigma_I} \tag{1}$$

where r_i and r_j represent the expected model neuron responses to two randomly sampled stimuli and σ_I represents internal encoding noise.

The response distribution $p(r)$ to natural stimuli has a critical impact on discriminability. Under the assumption that responses to natural stimuli are Gaussian distributed $p(r) = N\left(0, \sigma_E^2\right)$ the expected discriminability across all stimuli is given by

$$E[SNR] = E[d'] = \frac{\sigma_E}{\sigma_I} \frac{2}{\sqrt{\pi}} \tag{2}$$

where σ_E is the stimulus-driven (external) response variation (e.g. external noise; see Fig. 1) and the expectation is taken across all stimulus pairs (see Supplement). (If responses are Laplace distributed, the expected discriminability is given by $E[SNR] = \frac{\sigma_E}{\sigma_I} \frac{3}{2\sqrt{2}}$; see Supplement). For an arbitrary response distribution, expected SNR can be computed using numerical methods. The fact that the stimulus-driven standard deviation is in the numerator of Eq. 2 indicates that greater stimulus-driven response variation yields better stimulus discriminability.

Increased variance is usually associated with poorer stimulus discriminability (Ernst & Banks, 2002), so Eq. 2 deserves further explanation. The source of the response variance is critical for

determining whether it helps or hurts discrimination. When response variance is primarily due to noise, stimulus discriminability is poor. When response variance is primarily stimulus-driven, stimulus discriminability is high. For example, if the sensory afferents to a model neuron's receptive field are severed, the stimulus-driven response variance will be zero, and discriminating stimuli based on its response would be impossible. Fig. 1A shows a model neuron with low stimulus-driven response variance. Fig. 1B shows a model neuron with high stimulus-driven response variance. The top row shows how encoding noise (shaded bell curves) limits discriminability for two natural stimuli. The bottom row shows the impact of stimulus-driven variability across the image ensemble. High stimulus-driven variability will tend to yield larger differences between the expected responses to two random stimuli. Thus, discriminability improves when the source of the response variability is external and stimulus-driven, and discriminability deteriorates when the source of the response variability is internal and due to noise. It should be noted that the design of the visual system is surely driven by tasks more sophisticated than stimulus discrimination but for present purposes, it is a useful task around which to organize our discussion.

Model neuron responses
Responses of neurons in early visual cortex are commonly modeled as arising from a series of processing stages that are constrained by known properties of neural response in the early system (Fig. 2A). First, a linear receptive field filters the stimulus to yield the linear response. Next, the linear response is normalized by a factor that is determined by local properties of the stimulus or by the responses of other neurons in a local pool (Albrecht & Geisler, 1991; Heeger, 1992); the normalized response is called the 'response drive'. Last, the response drive is corrupted by encoding noise. Many models of neural response also incorporate response rectification and a static output non-linearity. We consider the effect of these non-linearities in the discussion, but our primary analysis is focused on the response drive statistics.

More specifically, the response R to a particular stimulus is given by

$$R = r_{max} \overbrace{\left[\frac{\mathbf{f}^T \mathbf{c}}{N} \right]}^{\substack{\text{response} \\ \text{drive}}} + \; \varepsilon \tag{3}$$

where r_{max} is the neuron's maximum response, $\mathbf{f}$ is the receptive field, $\mathbf{c}$ is a contrast stimulus (possibly corrupted by input noise), N is the normalization factor, and $\varepsilon \sim N\left(0, \sigma_I^2\right)$ is encoding noise. The standard deviation σ_I of the internal encoding noise can be constant or it can scale with the mean response (i.e. 'Poisson-like'). The receptive field is assumed to have a vector magnitude (i.e. L2 norm) of 1.0.

The maximum response is set to a constant for all model neurons. In individual simple cells, maximum firing rate is thought to be independent of preferred spatial frequency, orientation, and other stimulus preferences. It has been observed that overall firing rate in cortex tends to decrease as spatial frequency increases. But this decrease in overall firing rate is likely to be a population effect due to a non-uniform distribution of spatial frequency preferences in cortex, to sampling bias for low spatial frequencies in neuroscience studies, or both (De Valois, Albrecht, & Thorell, 1982a; Foster, Gaska, Nagler, & Pollen, 1985; Victor, Purpura, Katz, & Mao, 1994). Thus, in the current paper, and without loss of generality, for all receptive fields we assume r_{max} equals 1.0. We focus on characterizing the statistics of the response drive $\mathbf{f}^T \mathbf{c}/N$ to natural stimuli.
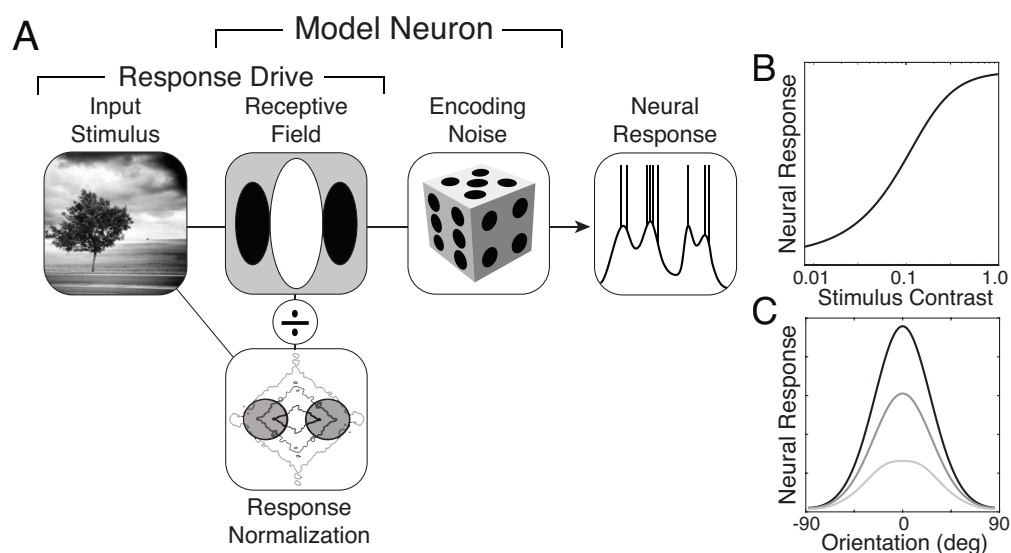
Figure 2. Model neuron. **A** Critical stages of neural response model: linear filtering, response normalization, and encoding noise. The stimulus is encoded by a linear filter, normalized by a portion of the stimulus energy, and then corrupted by encoding noise. The model neuron response provides a prediction of intracellular voltage or spike rate, depending on circumstances. **B** Contrast response function for the receptive field's preferred stimulus: a vertical Gabor. **C** Orientation tuning function for three different spatial frequencies: stimulus frequency equals the preferred spatial frequency (black), 2x the preferred spatial frequency (dark gray), and 3x the preferred spatial frequency (light gray).

The response model in equation 3 enforces the limited dynamic range of neural response in cortex and helps describe the shape of the contrast response functions of neurons in cortex (Fig. 2B). The response model also accounts for the invariance of the shapes of orientation tuning curves to Gabor or grating stimuli having different spatial frequencies (Fig. 2C).

Receptive field
A receptive field is a function that weights and sums inputs across space and time to determine a neuron's response. Responses increase when receptive field locations having positive weights are stimulated with input increments and decrease when stimulated with input decrements. The opposite happens with locations having negative weights. In early visual cortex, simple cell receptive fields are often modeled as having the shape of a Gabor—a cosine wave windowed by a Gaussian envelope (Jones & Palmer, 1987a; 1987b). Gabor receptive fields are orientation and spatial frequency selective; the selectivity is commonly quantified by the bandwidth. The orientation bandwidth specifies the range of input orientations that can elicit a response. The median orientation bandwidth in cortex is 42º (De Valois, Yund, & Hepler, 1982b). The spatial frequency bandwidth specifies the range of input spatial frequencies that can elicit a response. The distribution of simple cell bandwidths in cortex ranges between 0.8 to 1.8 octaves at half-height, with a median bandwidth of 1.2 octaves (De Valois, Albrecht, & Thorell, 1982a). We will characterize the response statistics of model neurons with vertically oriented Gabor receptive fields having the median orientation bandwidth of 42º and spatial frequency bandwidths that span the same range as simple cell receptive fields in cortex (octave bandwidths = 0.8, 1.2, and 1.8; Fig. 3AB). We examine model neuron responses having receptive fields with preferred spatial frequencies between 2 and 8 cpd.

Early models of neural response proposed that response drive is a linear function of the input stimulus (Campbell, Cleland, Cooper, & Enroth-Cugell, 1968; Hubel & Wiesel, 1962; 1968). The linear receptive field responses $R_{lin} = \mathbf{f}^T \mathbf{c}$ to natural stimuli are nicely approximated by a generalized Gaussian with tails heavier than a Laplace distribution (Fig. S1). Previous analyses of linear responses have reported similar findings (Wainwright & Simoncelli, 2000). We do not focus on the linear responses because real neurons include response normalization.
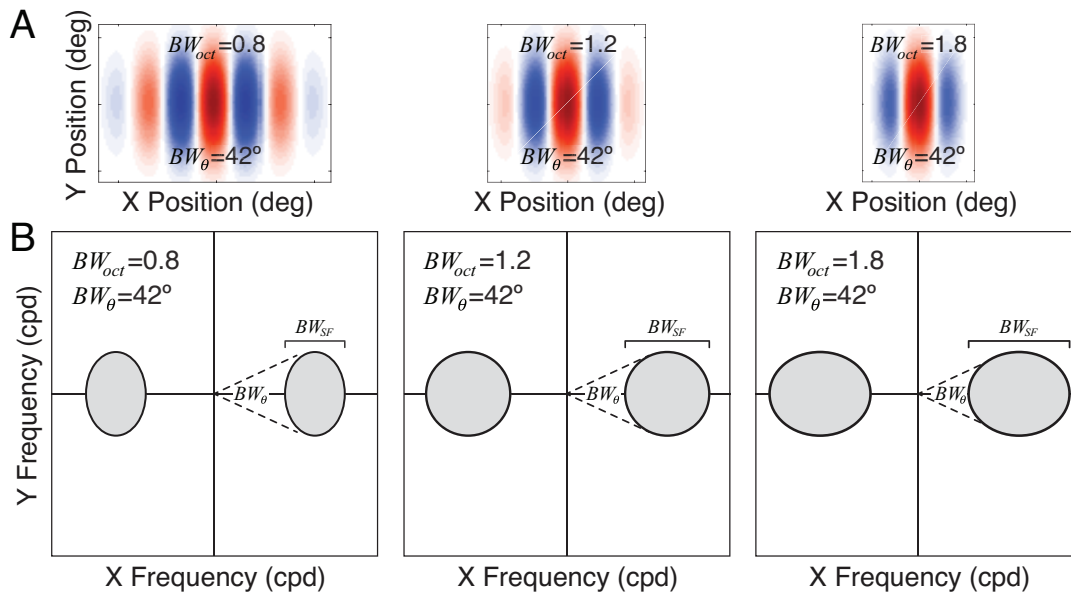
3

Figure 3. Gabor receptive fields and amplitude spectra. **A** Gabor receptive fields with octave bandwidths of 0.8, 1.2, and 1.8, and orientation bandwidths of 42°. Different octave bandwidths correspond to preferred features with different aspect ratios (see Methods). **B** Amplitude spectra of Gabor receptive fields. Orientation bandwidth BW_θ is the polar angle spanned by the amplitude spectrum at half-height. Spatial frequency bandwidth $BW_{SF} = f_{hi} - f_{lo}$ is the range of frequencies spanned by the spectrum, where f_{hi} and f_{lo} are the high and low frequencies at half-height. Octave bandwidth $BW_{oct} = \log_2\left(f_{hi}/f_{lo}\right)$ is the log-base-two ratio of the frequencies.

Response normalization

The linear model explains neural responses in some regimes, but it is insufficiently rich to capture neural response properties over a wide range of stvimulus conditions. Response normalization was originally proposed to account for the limited dynamic range of neurons in early visual cortex (Albrecht & Geisler, 1991; Heeger, 1992). Evidence for response normalization has been observed in primate retina, lateral geniculate nucleus, and early visual cortex (Albrecht & Geisler, 1991; Benardete, Kaplan, & Knight, 1992; Carandini, Heeger, & Movshon, 1997; Chander & Chichilnisky, 2001; Heeger, 1992; Mante, Bonin, & Carandini, 2008; Mante, Frazor, Bonin, Geisler, & Carandini, 2005; Nishimoto, Ishida, & Ohzawa, 2006; Shapley & Victor, 1978; Solomon, Peirce, Dhruv, & Lennie, 2004). In more recent years, normalization has been proposed to occur in higher cortical areas and be associated with computations underlying diverse behavioral phenomena (Carandini & Heeger, 2012). We examine how two types of response normalization—broadband normalization and narrowband normalization—impact the responses of model neurons to natural stimuli.

Broadband normalization is stimulus-specific but feature-independent. With broadband normalization, the model neuron responses are normalized by all the stimulus contrast in a local image region at the receptive field location, regardless of its preferred feature (Carandini et al., 1997); all orientations and spatial frequencies normalize the linear response. The broadband normalization factor is

$$N_{brd} = \|\mathbf{c}\| = \sqrt{\overbrace{\sum \mathbf{c}_i^2}^{\substack{\text{stimulus}\\\text{contrast energy}}}}$$

$$= \|A_{\mathbf{c}}\| = \sqrt{\overbrace{\sum A_{\mathbf{c}_i}^2}^{\substack{\text{stimulus}\\\text{contrast power}}}}$$

(4)

where $\mathbf{c}$ is a (possibly noisy) Weber contrast stimulus, $A_{\mathbf{c}}$ is the amplitude spectrum of the contrast stimulus, and the L2 norm operator $\|\cdot\|$ gives the square root of the sum of squares. Parseval's theorem guarantees that the total energy of the contrast stimulus equals the total power of its amplitude spectrum (Fig. 4A). Note that if the contrast stimulus is noisy (e.g. corrupted by pixel noise), the broadband normalization factor equals $N_{brd} = \sqrt{\sum \mathbf{c}_i^2 + \sigma_i^2}$ to very close approximation, where σ is the standard deviation of the input noise (Burge & Geisler, 2014).

Narrowband normalization is stimulus-specific and feature-dependent. With narrowband normalization, the model neuron responses are normalized by the stimulus contrast in the pass band of the receptive field, which means that the stimulus features that contribute most prominently to the normalization factor are those that approximately match the preferred feature (Fig. 4B). For example, if the receptive field's preferred stimulus is a vertically oriented Gabor with a carrier frequency of 4 cpd, the responses are normalized primarily by features that are near vertical and near to 4 cpd. The narrowband normalization factor is given by

$$
\begin{aligned}
N_{nrw} &= N_{brd} S \\
&= A_{\mathbf{c}}^T A_{\mathbf{f}}
\end{aligned}
$$

(5)

where $S = \dfrac{A_{\mathbf{c}}^T A_{\mathbf{f}}}{\|A_{\mathbf{c}}\|\|A_{\mathbf{f}}\|}$ is the phase-invariant similarity, the cosine similarity between the stimulus and receptive field amplitude spectra (Sebastian et al., 2017). (The amplitude spectrum of the receptive field $A_{\mathbf{f}}$ is assumed to have an L2 norm of 1.0.) Similarity is thus constrained to take a value between 0 and 1, which means that the narrowband normalization factor is always less than or equal to the broadband factor.

Broadband-normalized responses $R_{brd} \propto \mathbf{f}^T \mathbf{c}/N_{brd}$ to natural stimuli are highly non-Gaussian. The Laplace distribution provides an excellent fit to the broadband responses for all preferred spatial frequencies and octave bandwidths (Fig. 4C,D; Fig. S2). Narrowband-normalized responses $R_{nrw} \propto \mathbf{f}^T \mathbf{c}/N_{nrw}$ differ from broadband responses in two important ways. First, natural-stimulus-driven response standard deviation σ_E is approximately two and a half times higher for narrowband than broadband responses. Second, narrowband normalization yields response distributions that are very nearly Gaussian (Fig. 4C,D; Fig. S3). Related findings have been reported by other groups (Burge & Geisler, 2014; 2015; Jaini & Burge, 2017; Lyu & Simoncelli, 2008; 2009a; Sebastian et al., 2017; Wainwright & Simoncelli, 2000).

Relative to broadband normalization, narrowband normalization also improves signal-to-noise for stimulus discrimination by nearly three times, assuming constant encoding noise (Fig. 4E). The improvement in signal-to-noise is mediated both by the more Gaussian (lighter-tailed) response distributions (Fig. S4A), and by the increased stimulus-driven response variability (Eq. 2). Poisson-like encoding noise, which is more like response noise in cortex (Tolhurst, Movshon, & Dean, 1983b), yields very similar results (Fig. S4B-D).
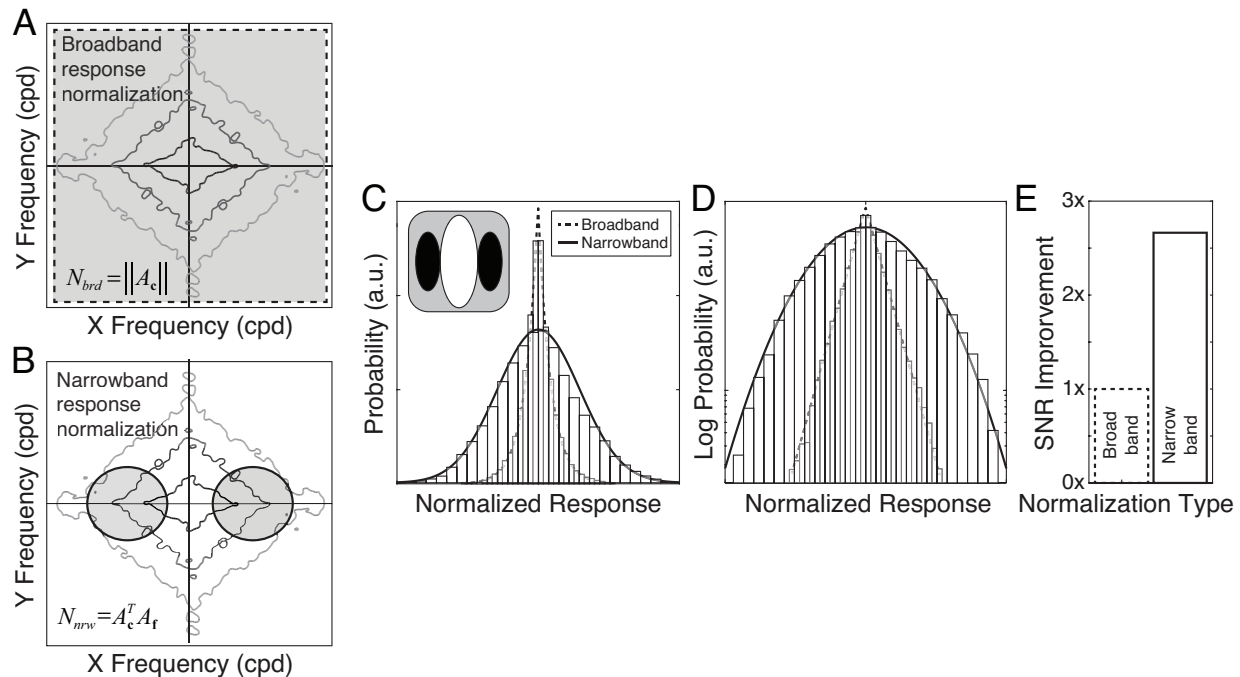
Figure 4. Broadband vs. narrowband normalization with a Gabor-shaped receptive field. **A** Broadband normalization uses all the stimulus contrast (gray area) to normalize the linear receptive field response, regardless of orientation and spatial frequency. The diamond shaped contours represent the amplitude spectrum of an individual natural image patch. **B** Narrowband normalization uses only the stimulus contrast in the pass band of the receptive field (gray area) to normalize the linear receptive field response. **C** Probability of model neuron responses. With broadband normalization, receptive field responses to natural stimuli are highly non-Gaussian, and are nicely approximated by a Laplace distribution (dashed curve). With narrowband normalization, the same receptive field yields responses to natural stimuli that are well described by a Gaussian (solid curve). **D** Same responses as in C, but with the y-axis showing log-probability over three orders of magnitude. **E** Factor improvement in signal-to-noise (SNR) for stimulus discriminability with narrowband vs. broadband normalization. Results are shown for a vertically oriented Gabor with an orientation bandwidth of 42° and an octave bandwidth of 1.2. Similar results are obtained for other receptive fields.

Why does narrowband normalization increase stimulus-driven response variance relative to broadband normalization? Because the narrowband normalization factor is always less than or equal to the broadband normalization factor (Fig. 5A; Eqs. 4,5). Therefore, across many stimuli, the response distribution will tend to have larger variance with narrowband normalization. Why does narrowband normalization result in more Gaussian responses than broadband normalization? Because narrowband normalization causes small broadband responses to be amplified significantly, and leaves large broadband responses relatively unchanged (Fig. 5B). For example, if the stimulus is a poor match to the receptive field (i.e. the broadband response approaches 0.0), it is likely that only a small proportion of the stimulus contrast is in the pass band of the receptive field. This, in turn, means that the narrowband normalization factor will be quite small compared to the broadband factor, which will increase the proportion by which the narrowband response is amplified relative to the broadband response (Eq. 5; Fig. 5B,C). On the other extreme, if the stimulus is a *perfect* match to the receptive field, the broadband response equals r_{max}, and all the stimulus contrast must be in the pass band of the receptive field. The narrowband and broadband normalization factors will thus be identical, and the narrowband response will equal the broadband response. Intermediate broadband responses are amplified by an intermediate amount (Fig. 5C). These effects mediate the differences in the shapes of the broadband and narrowband response distributions (Fig. 5D).
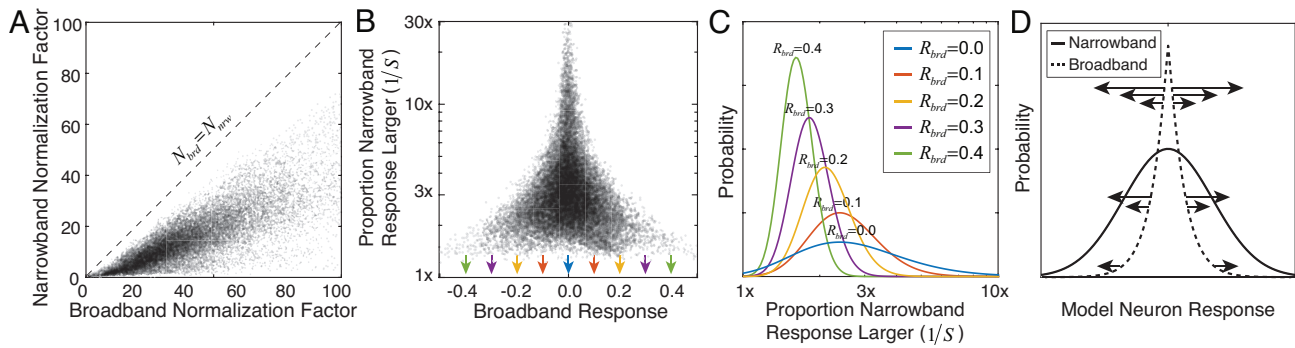
Figure 5. Broadband vs. narrowband normalization. **A** The narrowband normalization factor is smaller than the broadband normalization factor for each stimulus. **B** Proportion that the narrowband response is larger than the broadband response R_{nrw}/R_{brd} as a function of the broadband response. Large proportions occur only for small broadband responses, accounting for why narrowband normalization increases the Gaussianity of the model neuron responses. **C** Distribution of the proportional increase in the narrowband response relative to the broadband response, conditioned on different absolute values of the broadband response (colors), as fit by inverse gamma distributions (Fig. S5). The proportion is equivalent to inverse similarity. Arrows in B mark the absolute values of broadband response upon which the proportions are conditioned. **D** Schematic showing why the relationship between the narrowband and the broadband responses contributes to the increased Gaussianity of the model neuron responses. The data in A-C are for a vertically oriented cosine-phase Gabor receptive field with an orientation bandwidth of 42°, an octave bandwidth of 1.2, and a preferred frequency of 2 cpd. Results are similar for all receptive fields.

The proportional increase R_{nrw}/R_{brd} of the narrowband response vs. the broadband response depends strongly on the broadband response (Fig. 5B). Figure 5C shows conditional distributions of proportional increase for five different absolute values of the broadband responses, as fit by inverse gamma distributions (Fig. S5). When stimuli are narrowband normalized, small broadband responses are amplified more than large broadband responses.

There is an additional point worth making. The results presented in Fig. 4C,D suggest that the broadband-normalized responses can be represented as a Gaussian scale mixture of random variables. When considering natural images, the input contrast image $\mathbf{c}$ is a random variable. It follows that the broadband-normalized response R_{brd}, the narrowband-normalized response R_{nrw}, and the phase-invariant similarity S are all random variables. By combining equations 4-5, it is easy to show that these variables have the following relationships

$$R_{brd} = \frac{\mathbf{f}^T \mathbf{c}}{N_{brd}} = \frac{\mathbf{f}^T \mathbf{c}}{N_{nrw}} S = R_{nrw} S \tag{6A}$$

$$R_{brd} = R_{nrw} \sqrt{S^2} \tag{6B}$$

Equation 6B implies that the broadband-normalized responses are distributed as $R_{brd} \sim N(0, S^2)$ because the narrowband-normalized responses are approximately zero-mean Gaussian (see Fig. 4C,D). Furthermore, given that the broadband responses are approximately Laplace-distributed (see Fig. 4C,D), equation 6B also implies that the square of the phase-invariant similarity should be approximately Gamma distributed. This is because the Laplace distribution can be represented as a Gaussian scale mixture when the mixing distribution (i.e. the variance of the Gaussian) is Gamma distributed with a shape parameter of 1.0 (i.e. an Exponential distribution $S^2 \sim \Gamma(\alpha = 1, \beta) = Exp(\beta)$) (Ding & Blitzstein, 2018). Figure 6B,C shows that S^2 is indeed nicely approximated by a Gamma distribution with a shape parameter of 1.4, which is Exponential to close approximation. Thus, normalizing the broadband responses by the similarity yields Gaussian-distributed narrowband responses.
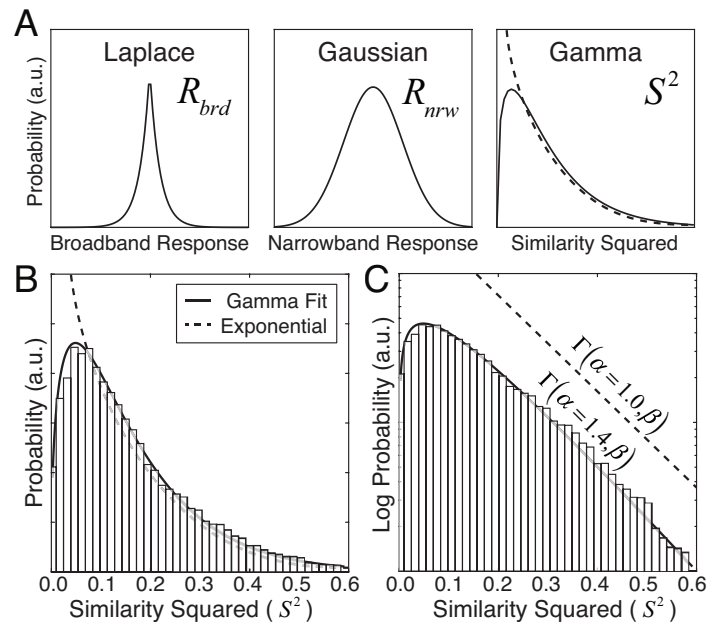
Figure 6. Broadband responses represented as a Gaussian scale mixture. **A** Laplace-distributed broadband responses can be expressed as a scale mixture of Gaussian narrowband responses with a Gamma (i.e. Exponential) distributed mixing variable. **B** Squared similarity across all natural stimulus (bars) and a Gamma distribution fit via maximum likelihood (solid curve); an Exponential distribution (a gamma distribution with a shape parameter of 1.0) is shown for reference (dashed curve). **C** Same data as A on a log-probability axis spanning two orders of magnitude with Gamma and Exponential fits. The exponential distribution is shifted vertically to reduce clutter. These results data are for a vertically oriented cosine-phase Gabor receptive field with an orientation bandwidth of 42º, an octave bandwidth of 1.2, and a preferred frequency of 2 cpd. Results are similar for all receptive fields.

Lyu & Simoncelli (2008) also modeled linear filter responses to natural images as a Gaussian scale mixture. Specifically, they modeled the linear (un-normalized) filter responses as a Gaussian scale mixture. They estimated the value of the mixing random variable, from the joint responses of a large bank of multi-scale filters. One potential advantage of the work presented here is that the narrowband normalization factor (i.e. the value of the mixing random variable) can be computed directly from the amplitude spectra of the stimulus and the receptive field (Eq. 5). Being able to compute the normalization factor directly from the stimulus may make stimulus- and feature-specific normalization easier to implement for some computational investigations.

Normalization pooling region
The receptive field specifies how inputs are weighted and pooled across space to determine the stimulus drive (i.e. $\mathbf{f}^T\mathbf{c}$) to neural response. Receptive fields are typically modeled by a matrix of positive and negative weights that determine how the value of each stimulus pixel contributes to the response (see Fig. 3). Here, we ask how the visual angle spanned by the receptive field weight matrix impacts the response statistics of model neurons to natural stimuli. The visual angle spanned by the receptive field weight matrix impacts the statistics because it determines the stimulus region from which the normalization factor is computed (Eqs. 4,5; see below).

Consider two sets of neurons employing narrowband contrast normalization having receptive fields with Gabor-shaped preferred features. In the first set, the visual angle spanned by the receptive field weight matrix becomes increasingly mismatched to its preferred feature with increases in preferred spatial frequency (Fig. 7A). In the second set, the visual angle spanned by the receptive field weight matrix is matched to the preferred feature regardless of its spatial frequency (Fig. 7B). These two sets of model neurons produce very different sets of response distributions to natural stimuli. When the weight matrix and preferred feature are matched (i.e. span the same visual angle), the normalization factor is computed from the same image region that drives the linear

receptive field response, and the response distributions have constant variance and are approximately Gaussian for all preferred spatial frequencies. When the weight matrix and preferred feature are mismatched, the normalization factor is computed from an image region larger than the preferred feature, the response variance decreases with the inverse frequency (1/f) of the preferred feature and the response distributions become less Gaussian (Fig. 7C). These results are summarized by the stimulus-driven response standard deviation and kurtosis (Fig. 7D,E). Note also that for the largest mismatch considered here, the narrowband response distribution is well approximated by a Laplace distribution. Thus, for large mismatches the benefits of narrowband response normalization are surrendered. (See Supplement for results with broadband normalization; Fig. S6.)
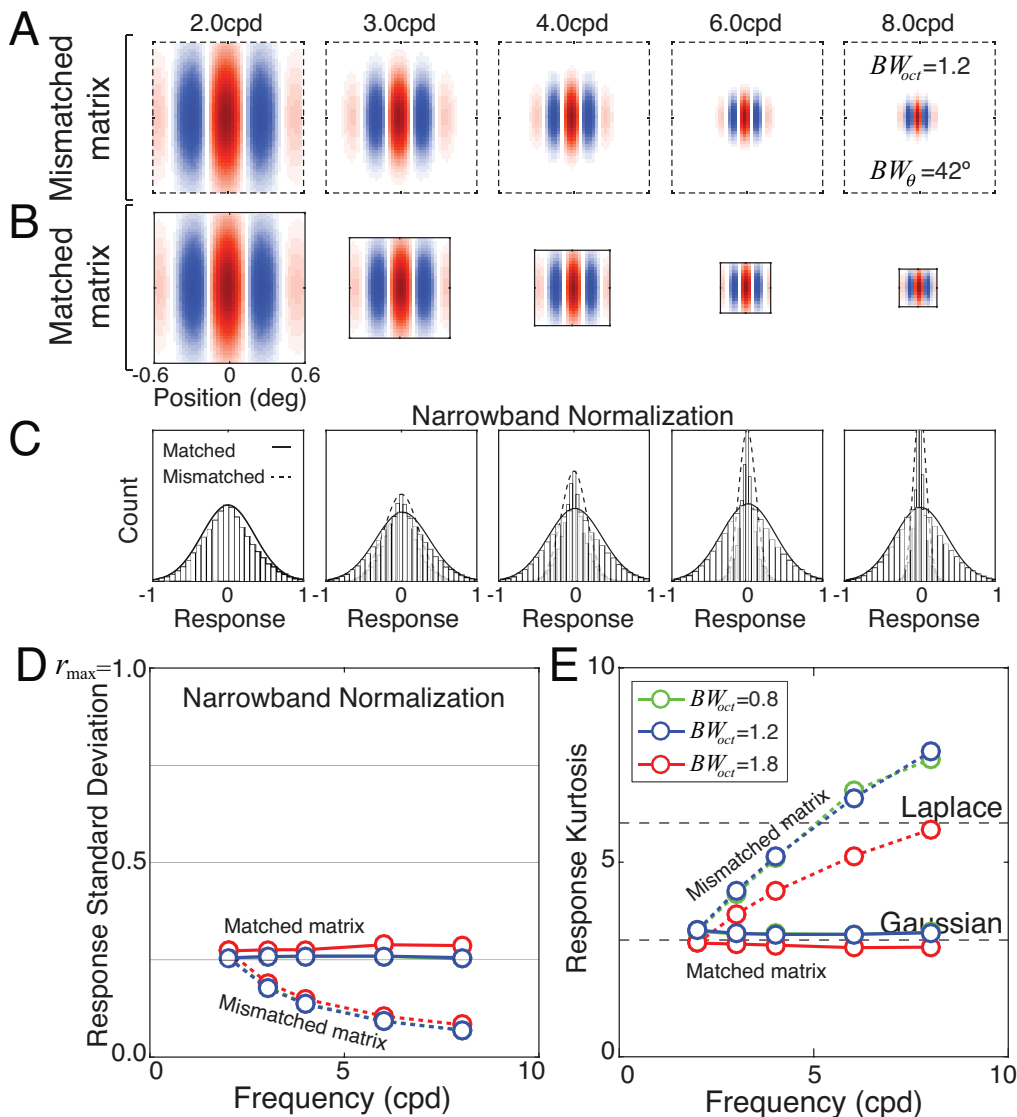

Figure 7. Narrowband response statistics with receptive field weight matrices that are **A** mismatched and **B** matched to the preferred feature: a vertically oriented Gabor with 1.2 octave bandwidth and 42° orientation bandwidth. **C** Response distributions from matched and mismatched matrices. Matched weight matrices (solid curves) yield response distributions that are invariant to the scale of the preferred feature. Mismatched weight matrices (dashed curves) yield response distributions that change shape and variance with the magnitude of the mismatch. **D** Response standard deviation as a function of preferred spatial frequency for octave bandwidth (colors) . Stimulus-driven response variance is constant with preferred frequency when the matrix is matched to the preferred feature. When the matrix is mismatched, response variance decreases with the magnitude of the mismatch. **E** Response kurtosis is the same as a Gaussian with matched weight matrices, but increases with the amount of mismatch.

The primary finding, however, is that with narrowband normalization and matched receptive fields, responses are approximately Gaussian, mean-zero, and strikingly invariant to the scale of the preferred feature. Regardless of the preferred spatial frequency, the response is equally statistically reliable. Furthermore, the results indicate that it is exceedingly rare for a neuron to be stimulated by its preferred feature in natural scenes. The stimulus-driven response standard deviation equals approximately 25% of the maximum response, which means that less than one natural stimulus in 10,000 will cause a response within 5% of the neuron's maximum response.

How should these results inform our thinking about neurophysiological processing of real-world signals? The majority of single-unit neurophysiology has focused on characterizing the stimuli to which individual neurons respond most strongly. We owe much of our knowledge about the response properties of V1 neurons to this approach. But the real-world stimuli that drive neural responses most strongly occur only very rarely. To understand the coding problems faced by nervous systems in natural viewing, it is critical to understand how neurons respond to the stimulus ensemble encountered in the real world.
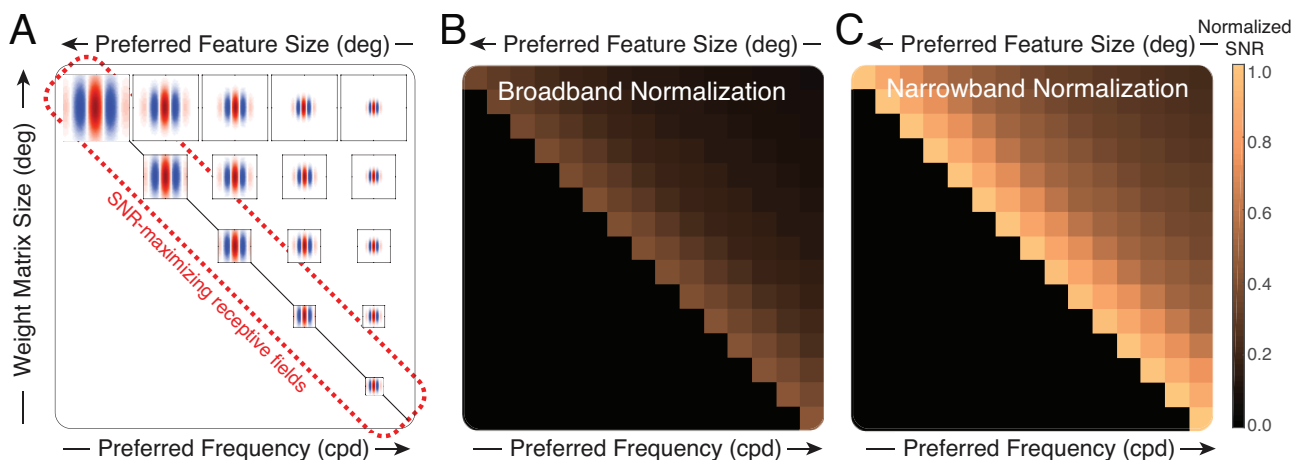

Figure 8. The impact of different receptive field modeling choices on signal-to-noise ratio (SNR). **A** Model receptive fields maximize signal-to-noise ratio (SNR) for stimulus discrimination when the visual angle subtended by the receptive field weight matrix matches the visual angle of the preferred feature (on-diagonal receptive fields). When the visual angles are mismatched, SNR is lower. **B** Normalized SNR with broadband normalization. **C** Normalized SNR with narrowband normalization as a function of preferred spatial frequency and the visual angle spanned by the weight matrix.

To summarize the impact of matching the weight matrix to the preferred feature, we plot the signal-to-noise ratio for stimulus discrimination for a range of preferred features and weight matrix sizes (Fig. 8A). There is a substantial advantage i) for narrowband over broadband normalization, and ii) for matching the visual angle of the receptive field weight matrix to the visual angle of the preferred stimulus feature (Fig. 8B,C). Thus, to maximize the signal-to-noise ratio for stimulus discriminability and to achieve scale invariant response statistics to natural stimuli, one should perform narrowband normalization with weight matrices that match the receptive field's preferred feature.

As mentioned earlier, these effects occur because of the nonlinear effects of response normalization. The visual angle spanned by the weight matrix determines the size of the image region from which the normalization factor is computed. With mismatched matrices, the normalization factor is determined from the stimulus contrast in an image region larger than the preferred feature, that will likely contain spatial frequencies lower than the preferred frequency. Natural images have 1/f amplitude spectra (D. J. Field, 1987); contrast energy at frequencies lower than the preferred frequency is likely to dominate and substantially increase the value of the normalization factor, thereby decreasing the normalized response. As the mismatch increases, the decrease in the normalized response becomes more pronounced, reducing the stimulus-driven

response variability associated with high frequency features. The normalization factor should therefore be determined from the same image region that is selected for by the preferred feature.

Another way to understand the effects is to consider two nearly identical model neurons that do not employ response normalization. Their response drives are equal to $\mathbf{f}^T\mathbf{c}$ instead of $\mathbf{f}^T\mathbf{c}/N$. The neurons prefer the same feature and differ only because one has a mismatched and the other has a matched weight matrix. The larger mismatched matrix is padded with zero-valued coefficients. Multiplying inputs with zero-valued coefficients does not change the linear response. Both neurons will thus produce identical linear responses regardless of whether the matrix is matched. The differential effects must therefore be due to the size of the image region that determines the normalization factor, relative to the size of the preferred feature of the receptive field.

Downsampling
The model neuron receptive fields considered thus far have had identical sampling resolution, so the number of pixels representing a preferred feature scales with the visual angle spanned by the receptive field (Fig. 9A,B). In the primate visual system, at a given eccentricity, simple cells with larger receptive fields pool over more relay cell inputs from the lateral geniculate nucleus (LGN) than those with smaller receptive fields (Taylor, Sedigh-Sarvestani, Vigeland, Palmer, & Contreras, 2018). Similarly, parasol ganglion cells in the retina pool over more cone receptors than midget ganglion cells (G. D. Field et al., 2010). Sometimes, however, large receptive fields pool inputs that have lower sampling resolution than their smaller counterparts. For example, large retinal ganglion cells (RGCs) in the retinal periphery pool inputs from large low-resolution cone photoreceptors, whereas small foveal RGCs of the same type pool over small high-resolution photoreceptors (Croner & Kaplan, 1995; Rossi & Roorda, 2010). The processing motif employed by the peripheral retina is roughly equivalent to downsampling, a common pre-processing method in the computer vision, image processing, and deep learning communities (Burt & Adelson, 1983). In general, downsampling reduces the number of pixels (i.e. sampling resolution) representing a particular image patch, and hence the computational requirements for processing that patch.
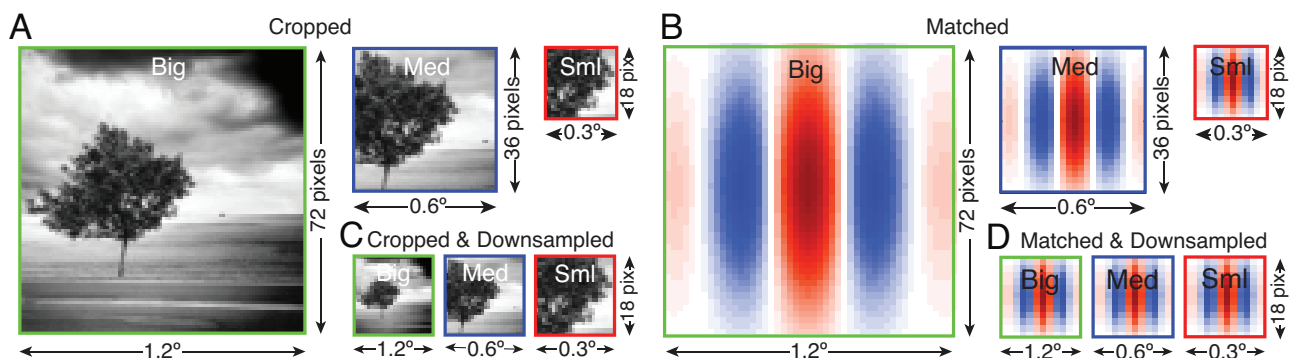

Figure 9. Downsampling cropped images for matched receptive field weight matrices. **A** Cropped images: big, medium, and small. Boxes indicate three different scales at which image data is processed. The cropped images span different visual angles but have a fixed sampling rate. Each cropped image therefore has a different number of pixels. **B** Matched receptive field weight matrices. The visual angle spanned by the weight matrix matches the visual angle spanned by the preferred feature. Each matrix also has a different number of pixels. **C** Cropped and downsampled images. Cropped and downsampled images span different visual angles, but have the same number of pixels. **D** Matched and downsampled receptive fields. All weight matrices have the same number of pixels, regardless of the spanned visual angle.
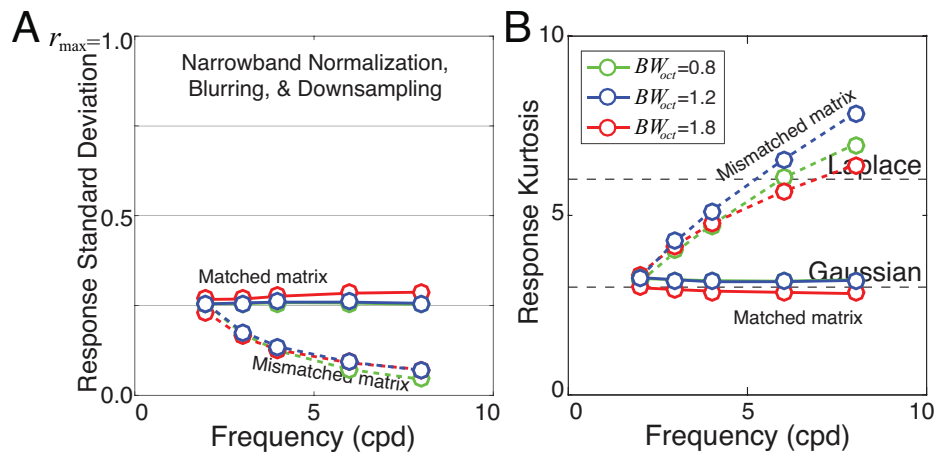
Figure 10. The impact of downsampling on response statistics. **A** Stimulus-driven response standard deviation with narrowband normalization, blurring, and downsampling. **B** Stimulus-driven response kurtosis with blurring and downsampling. The response statistics are essentially identical with and without downsampling.

We asked how downsampling the input stimuli impacts model neuron response statistics. First, we generated a new set of receptive field weight matrices where both the spanned visual angle and the sampling resolution were yoked to the preferred feature. The result was a set of receptive fields defined by weight matrices that all had an identical number of pixels. Specifically, all weight matrices had 18x18 pixels, the same number as the original matrix corresponding for the smallest preferred feature (0.3º, 8 cpd). Then, we downsampled the image patches (after appropriate blurring to prevent aliasing) to match the sampling resolution of the receptive fields (see Methods). Downsampled images and receptive fields are shown in Figure 9C,D.

With downsampling, the narrowband-normalized responses with matched matrices are constant variance Gaussian (Fig. 10A,B). The response statistics (i.e. standard deviation and kurtosis) are within one percent of the response statistics without downsampling (see Fig. 7DE). There is no disadvantage (or advantage) to downsampling in terms of the signal-to-noise for signal discrimination. Thus, at least in terms of signal-to-noise for stimulus discrimination, there is no pressure on the visual system to avoid downsampling. This result is useful for approaches that seek to learn receptive fields via non-parametric methods, where every additional pixel in a receptive field weight matrix incurs considerable computational cost (see Discussion).

Discussion
Model neurons employing narrowband response normalization yield scale-invariant Gaussian distributed responses to natural stimuli. The scale-invariant response statistics come close to maximizing the signal-to-noise ratio for stimulus discrimination with natural stimuli, but the scale-invariance depends on the normalization factor being determined from image region that matches the size of the receptive field's preferred feature. In the discussion section, we examine how these results are affected by receptive fields that are not oriented Gabors, and stimulus types that are not natural (i.e. noise stimuli). We discuss how the results reported here can explain why models fitted to neurons in cortex tend to poorly predict responses to natural stimuli, even when they beautifully predict responses to noise stimuli.

Generality of conclusions
Different subfields in vision and computational neuroscience have different methodological conventions for modeling neurons. Under many simplified circumstances, the different conventions have little or no practical impact. However, when the model neurons include the dominant features of real neurons in cortex—receptive field, response normalization, and encoding noise—the subtle differences in the modeling conventions can have a dramatic impact on their response statistics to

natural images. How, then, do modeling choices other than those considered in the body of the paper impact the response statistics?

First, we asked whether the response statistics generalize to other Gabor receptive fields. In the results section, we analyzed only the response statistics of vertically oriented even-symmetric (cosine phase) Gabor receptive fields. Do odd-symmetric (sine phase) receptive fields produce similar results? Yes. Gabor receptive fields with all phases and orientations produce equivalent results. Thus, model neurons with biologically plausible receptive fields and appropriate narrowband response normalization produce response statistics that are invariant to the preferred feature.

Next, we asked whether model neurons with other receptive field shapes produce similar results. There is a long history of using Gabor functions to describe the receptive fields of simple cells in early visual cortex (Jones & Palmer, 1987a; 1987b). However, empirical data suggests that log-Gabors may provide a better characterization of simple cell receptive fields in early visual cortex (De Valois, Albrecht, & Thorell, 1982a; Geisler & Albrecht, 1997; Hawken & Parker, 1987). It has also been argued on theoretical grounds that log-Gabor receptive fields may be better than Gabor receptive fields for encoding natural images (D. J. Field, 1987). We re-ran our analyses with log-Gabor shaped receptive fields. All results hold with log-Gabor receptive fields (Fig. S7).

Then, we asked whether the main conclusions hold for the receptive fields like those of retinal ganglion cells or relay cells in the lateral geniculate nucleus, which are radially symmetric and do not select for orientation. We repeated our analyses with center-surround Gabors and difference-of-Gaussian (DoG) preferred features embedded in matched weight matrices. Results are quite similar for oriented and un-oriented receptive fields with narrowband normalization. However, model neurons with un-oriented receptive fields also yield Gaussian response distributions with broadband normalization (Fig. S8). (Recall that oriented receptive fields with broadband normalization yield Laplace-distributed responses; Fig. 4C,D, Fig. S2) This result implies that the differences in the shape of the response distributions that distinguish broadband and narrowband normalization with oriented Gabors (c.f. Fig. 4) are due primarily to the orientation selectivity of the receptive field.

Natural vs. noise stimuli
Model neurons with narrowband response normalization and matched weight matrices yield Gaussian-distributed responses to natural stimuli. But noise stimuli are more often used in psychophysical and neurophysiological experiments because they have useful properties for methods designed to recover the stimulus features (i.e. receptive fields) that drive behavioral and neural response (Ahumada & Lovell, 1971; Schwartz, Pillow, Rust, & Simoncelli, 2006). 1/f noise has the amplitude spectrum but not the phase structure of natural images. White noise has neither the amplitude spectrum nor the phase structure of natural stimuli. It is therefore important to ask how our model neurons respond to noise stimuli.

We compared linear and normalized model neuron responses to natural, 1/f noise, and white noise stimuli. As shown above, linear responses to natural stimuli are distributed with heavier tails than Laplace distributions. Broadband-normalized responses to natural stimuli are Laplace-distributed. And narrowband-normalized responses to natural stimuli are approximately Gaussian (Figs. S1-3). In contrast, linear and broadband responses to Gaussian 1/f and white noise stimuli are Gaussian. Narrowband-normalized responses to 1/f and white noise stimuli are also approximately Gaussian (Figs. S9,S10). Thus, the differences between linear, broadband-, and narrowband-normalized responses are substantial for natural stimuli. The negligible for noise stimuli (Fig. 11).
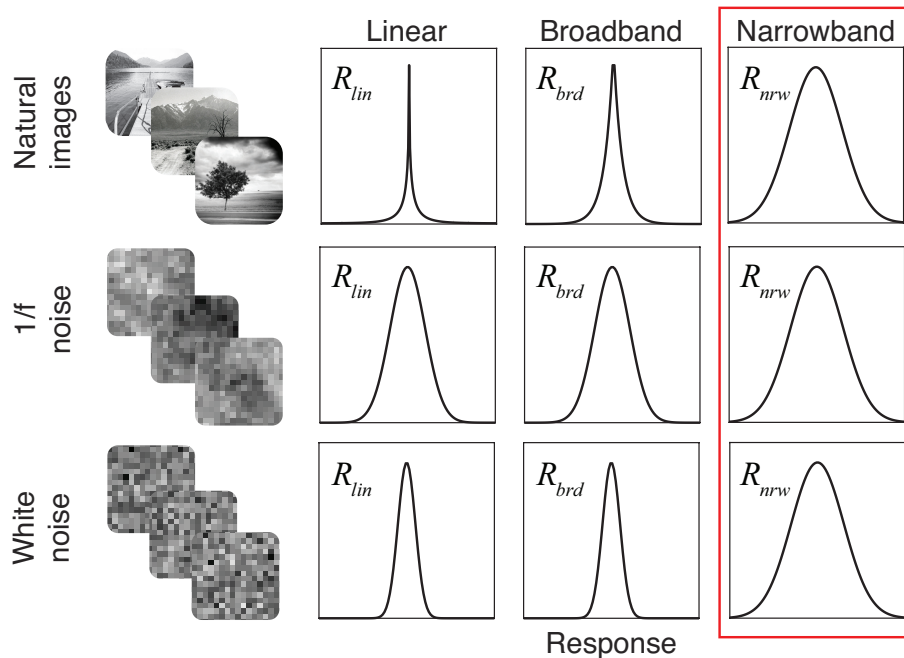
13

Figure 11. Distributions of linear, broadband-normalized, and narrowband-normalized responses to natural, 1/f noise, and white noise stimuli. Linear responses to natural stimuli are very heavy tailed (i.e. heavier than a Laplace distribution; see Fig. S1). Broadband responses to natural stimuli are Laplace-distributed (see Fig. S2). Narrowband responses to natural stimuli are Gaussian (see Figs. 4, 7, S3). Linear and broadband responses to 1/f and white noise are guaranteed to be Gaussian, assuming that the number of pixels defining each stimulus (i.e. dimensionality) is high enough. Narrowband responses to noise stimuli are all approximately Gaussian. Narrowband normalization produces responses (i.e. response drives) that are very nearly Gaussian with all three stimulus types. Narrowband normalization helps standardize the distributional form of the response statistics across stimulus types (box). Narrowband normalization should therefore improve the ability of computational models of visual information processing to generalize across stimulus types.

These findings provide a clue about why subunit models of neural response tend to generalize poorly when tested with natural stimuli. Subunit models are a popular method for performing neural systems identification. Their aim is provide a concise computational level description of the input-output relationship between stimulus and response. Subunit models are typically fit and tested using noise stimuli. But even when fitted subunit models nicely predict performance with noise stimuli, they tend to perform poorly when tested with natural stimuli (Heitman et al., 2016; Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003; Talebi & Baker, 2012).

Subunit models of neural response tacitly assume Gaussian-distributed responses (McFarland, Cui, & Butts, 2013; I. M. Park, Archer, Priebe, & Pillow, 2013; Rust, Schwartz, Movshon, & Simoncelli, 2005; Schwartz et al., 2006), even though the models do not typically include response normalization (but see (McFarland et al., 2013)). However, with Gaussian noise stimuli linear (un-normalized) response distributions are guaranteed to be Gaussian. (Broadband-normalized response distributions to noise stimuli are also guaranteed to be Gaussian, assuming that the number of pixels defining each stimulus—i.e. the dimensionality of each stimulus—is sufficiently high (Poincaré, 1912).) Thus, the absence of narrowband response normalization has little practical effect with the stimuli that subunit models are most often fit and tested. In contrast, with natural stimuli, the absence of narrowband response normalization results in response distributions that are highly non-Gaussian. The non-Gaussian response distributions violate the tacit distributional assumptions of the models. This violation contributes to the poor generalization of subunit models to natural stimuli. Incorporating narrowband normalization into these methods for neural systems identification will yield Gaussian response distributions for both natural and noise stimuli, and should therefore improve the ability of these models to generalize across these stimulus types.

Non-parametric receptive field learning: Model neuron modeling conventions

Some areas of computational neuroscience aim to learn populations of receptive fields (i.e. preferred features) that optimize a particular goal. The preferred features of these receptive fields are typically learned by iteratively updating coefficients (i.e. pixel values) in pursuit of the goal. The receptive field weight matrices typically have a fixed number of pixels and span a fixed visual angle. This convention is convenient for matrix-based programming languages (e.g. Matlab), but it often results in weight matrices that are mismatched to the preferred feature. For example, mismatched matrices are commonly reported by papers that fit subunit models to describe the stimulus-response properties of neurons early in the visual processing stream (c.f. Fig. 2a in Rust et al, 2005; (McFarland et al., 2013; I. M. Park et al., 2013; Samengo & Gollisch, 2013; Schwartz et al., 2006; Vintch, Movshon, & Simoncelli, 2015)). Analyses inspired by the efficient coding hypothesis seek receptive field populations that efficiently encode natural stimuli also commonly report mismatched weight matrices (c.f. Fig. 4a in Olshausen & Field, 1996; (Bell & Sejnowski, 1997; Lewicki, 2002; Olshausen & Field, 1997; Rehn & Sommer, 2007)). Why are mismatched weight matrices commonly reported if they carry the disadvantages detailed in the results section (c.f. Fig. 8)? The literatures mentioned above typically assume that receptive field (or subunit) responses are driven purely by linear operations; that is, they do not incorporate response normalization. (They also do not typically model encoding noise.) In the absence of normalization, matched and mismatched matrices yield identical responses (see Results).

To increase biological realism and to facilitate generalization to natural stimuli, normalization should be included in the response models for these feature-learning methods and others. When normalization is included, the normalization factor must be computed from the same stimulus region as the preferred feature. Otherwise, receptive fields with preferred features smaller than the visual angle spanned by the weight matrix will have poor signal-to-noise (Fig. 7,8). This will make it difficult to interpret the results of feature learning studies. If high frequency features are not learned, it will be unclear whether it is because of the poor signal-to-noise or because high frequency filters are fundamentally not useful. If receptive fields within a parametric family are being learned (e.g. Gabor) the parameters can be used to determine the area spanned by the feature, but this is not possible with non-parametric approaches. One way to address this issue would be learn receptive fields with localized priors, but this technique poses significant technical challenges in many contexts (M. Park & Pillow, 2011). A simpler approach would be to learn receptive fields with response normalization at multiple scales simultaneously. At each scale, receptive fields with small mismatched preferred features will yield with poor signal-to-noise, biasing the learning procedures against selecting those features relative to the scale. But across scales, large and small features will both be given a fair chance; large features would be matched to large matrices and small features would be matched to small matrices.

Neural response in early visual cortex

Across many natural images, model neuron response drive is mean zero and Gaussian distributed. The response drive is equivalent to the neural response under the assumptions that i) no response rectification occurs and ii) that the power of the static output non-linearity is 1.0. In real neurons these assumptions do not typically hold. As a consequence, the responses of real neurons in early visual to natural stimuli are not Gaussian (Felsen et al., 2005; Weliky et al., 2003). Rather, the response drive is typically rectified and the power of static output nonlinearity is close to 2.0 (Priebe & Ferster, 2008). We examined how rectification and a squaring output non-linearity changes the model neuron response statistics. Rectifying and squaring the response drive converts the Gaussian response distributions into chi-squared response distributions with one degree of freedom $\sigma_E^2 \chi_1^2$ scaled by the stimulus-driven response variance (Fig. 12). The scaled chi-squared is more similar to the response distributions observed from spiking neurons in cortex.

The scaled chi-squared distribution is equivalently described by a gamma distribution with shape and scale parameters of $1/2$ and $2\sigma_E^2$, respectively.

These considerations make several predictions. First, presenting natural stimuli to simple cells in early visual cortex should yield response rate distributions that are nicely modeled by scaled chi-squared distributions with one degree of freedom. Second, because complex cells are typically modeled as resulting from quadratic pooling of two (or more) subunit receptive field responses, complex cells should yield response rate distributions that are well modeled by scaled chi-squared distributions with two (or more) degrees of freedom. Third, the distribution of response across stimuli for different neurons should be approximately invariant to the preferred feature of each neuron. We advocate examining the distribution of response across hundreds (or thousands) of natural images and comparing their distribution to a chi-squared distribution with variable degrees of freedom.
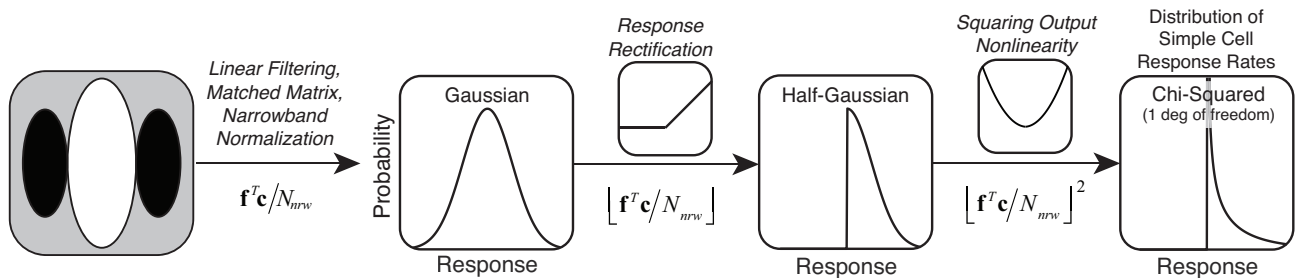

Figure 12. Relating Gaussian response drive statistics to response statistics in cortex. Gaussian responses are first rectified to create a half-Gaussian distribution. Rectified responses are then squared to create a scaled chi-squared distribution (see text). The scaled chi-squared distribution represents the predicted distribution of response rates to natural images by neurons (i.e. simple cells) in early visual cortex.

It is also important to ask how the signal-to-noise ratio for stimulus discrimination would be impacted by rectifying and squaring the response drive (i.e. the normalized linear response). If the controlling source of encoding noise occurs after rectification and squaring, signal-to-noise would be altered (Eq. 7A). On the other hand, if the controlling source of encoding noise is added at the level of the response drive, before the rectification and squaring (Eq. 7B), the signal-to-noise for stimulus discrimination will be unaffected because a monotonic transform of a noisy signal together will not alter the signal-to-noise ratio (Pelli, 1985; Rieke & Rudd, 2009).

$$r = r_{max} \left\lfloor \overbrace{\left\lfloor \frac{\mathbf{f}^T\mathbf{c}}{N} \right\rfloor}^{\text{response drive}} \right\rfloor^p + \varepsilon \tag{7A}$$

$$r = r_{max} \left\lfloor \overbrace{\left\lfloor \frac{\mathbf{f}^T\mathbf{c}}{N} \right\rfloor}^{\text{response drive}} + \varepsilon \right\rfloor^p \tag{7B}$$

A number of influential models in the neurophysiological literature propose that the controlling source of encoding noise is at the level of the membrane voltage, and should be modeled as constant, additive, and zero mean (Carandini, 2004; Mohanty, Scholl, & Priebe, 2012; Priebe & Ferster, 2008). After rectification and squaring, the noisy voltage signal predicts Poisson-like encoding noise (i.e. response variance that scales with the mean response to each stimulus) like that typically observed in cortex (Tolhurst, Movshon, & Dean, 1983b). If these models are correct, the controlling encoding noise should modeled as occurring at the level of the response drive (Eq. 7B). Lastly, these considerations make one an additional prediction. If the intracellular voltage

reflects the response drive as it has been discussed in this manuscript, the distribution of intracellular voltages across the natural stimulus ensemble should be Gaussian distributed.

Task specific analysis of natural images and scenes
Efficient coding is an influential theoretical framework for thinking about how neural systems encode natural stimuli (Attneave, 1954; Barlow, 2001). Many papers have focused on learning receptive fields that efficiently reconstruct proximal stimuli (Bell & Sejnowski, 1997; Lewicki, 2002; Olshausen & Field, 1996; 1997). However, the sensory-perceptual systems of humans and other animals must do more than efficiently encode and reconstruct sensory inputs. Sensory-perceptual systems must extract information from input stimuli that is useful for the behavioral tasks that organisms must perform to survive and reproduce. Efficient coding does not directly address this problem.

In recent years, statistical techniques have been developed that learn small populations of receptive fields that encode the stimulus features most useful for specific tasks (Burge & Jaini, 2017; Geisler, Najemnik, & Ing, 2009; Jaini & Burge, 2017). These techniques have helped to find the optimal solutions in tasks useful for estimating the three-dimensional structure of the environment—focus error estimation (Burge & Geisler, 2011; 2012), disparity estimation (Burge & Geisler, 2014), and retinal motion estimation (Burge & Geisler, 2015)—and that have been the focus of intense study for decades by the vision and neuroscience communities (Banks, Gepshtein, & Landy, 2004; Burge, Fowlkes, & Banks, 2010; Burge, Peterson, & Palmer, 2005; Cormack, Stevenson, & Schor, 1991; Flitcroft, 1990; Gekas, Meso, Masson, & Mamassian, 2017; Held, Cooper, & Banks, 2012; Iyer & Burge, 2018; Jogan & Stocker, 2015; Kotulak & Schor, 1986; Kruger, Aggarwala, Bean, & Mathews, 1997; Priebe & Lisberger, 2004; Priebe, Cassanello, & Lisberger, 2003; Rust, Mante, Simoncelli, & Movshon, 2006; Sebastian, Burge, & Geisler, 2015; Simoncelli & Heeger, 1998; Tyler & Julesz, 1978; Weiss, Simoncelli, & Adelson, 2002). In this manuscript, we examined only the response statistics of individual model neurons and considered only performance in a very simple task: discriminating one stimulus from another. The optimal solutions to these more sophisticated tasks require combining the responses from multiple receptive fields. With appropriate normalization, responses to natural stimuli from multiple receptive fields are jointly Gaussian, a fact that should simplify computations for optimally combining those receptive field responses. Thus, the results reported here should be thought of as the beginning of a more complete investigation of how visual systems process natural stimuli. Having an accurate picture of the response statistics of model neurons, and understanding how small differences in modeling conventions affect those response statistics, lay an important foundation for building principled models in the future.

Conclusion
The work presented in this manuscript suggests that computational models of neural processing should incorporate narrowband (i.e. stimulus- and feature-specific) response normalization. A simple expression for computing the narrowband normalization factor is provided, that should facilitate the inclusion of narrowband normalization in non-parametric methods for learning receptive fields of model-neuron-like units. Narrowband normalization should also improve the ability of such models to generalize from noise stimuli to natural stimuli, because narrowband normalization yields scale invariant Gaussian responses for both natural and noise stimuli.

Acknowledgments

Methods

Natural stimuli: Natural image patches were sampled from two recently published photographic databases of natural scenes (Burge & Geisler, 2011; Burge, McCann, & Geisler, 2016). Scenes were photographed on and around the University of Texas at Austin campus and contained grass, shrubs, and trees, and streets, cars, and buildings. The images were calibrated such that the intensity values were linear with luminance. The data represents 30,888 unique 1.2° image patches (72x72 pixels); 312 non-overlapping patches were randomly selected from each of 99 calibrated natural images.

Local contrast: The intensity patches are converted to Weber contrast images by luminance normalization. The contrast image is obtained by subtracting off and dividing by the local mean intensity

$$\mathbf{c}(\mathbf{x}) = \sum_{\mathbf{x} \in A} \left[\left(\frac{I(\mathbf{x}) - \bar{I}}{\bar{I}} \right) \right] \tag{8}$$

where $\mathbf{c}(\mathbf{x})$ is the local contrast image patch, $I(\mathbf{x})$ is the local intensity image patch, $\bar{I}$ is the local mean intensity, and $\mathbf{x} = \{x, y\}$ indexes spatial position in the area A spanned receptive field weight matrix. The local mean intensity is given by $\bar{I} = \sum_{\mathbf{x} \in A} \left[I(\mathbf{x}) \right]$.

Receptive fields: The receptive field of each model neuron is modeled with a weight matrix. The receptive field weight matrix is determined by the preferred feature, the visual angle spanned by the weight matrix, and by the spatial sampling rate. The preferred feature of each model neuron is modeled as a Gabor. A Gabor is a cosine wave multiplied by a Gaussian envelope

$$\mathbf{f}(\mathbf{x}) = gauss(x', y'; x_0, y_0, \theta_0, \sigma_b, \sigma_l) \cos(2\pi f_0 x' + \phi_0) \tag{9}$$

where x_0 and y_0 specify the position of the Gaussian envelope, θ_0 is the preferred orientation, σ_b is the standard deviation of the envelope in the band pass direction (orthogonal to the grating orientation), σ_l is the standard deviation of the envelope in the low pass direction (parallel to the grating orientation), f_0 is the preferred spatial frequency, ϕ_0 is the preferred phase, and $\{x', y'\}$ are transformed coordinates due to the preferred orientation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \left[\begin{pmatrix} \cos\theta_0 & \sin\theta_0 \\ -\sin\theta_0 & \cos\theta_0 \end{pmatrix} \right] \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} = \begin{bmatrix} (x - x_0)\cos\theta_0 - (y - y_0)\sin\theta_0 \\ (x - x_0)\sin\theta_0 + (y - y_0)\cos\theta_0 \end{bmatrix} \tag{10}$$

The coefficients of the receptive field weight matrix are normalized such that the L2 norm of the receptive field weight coefficients $\|\mathbf{f}(\mathbf{x})\| = \sqrt{\sum_{\mathbf{x}} \mathbf{f}(\mathbf{x})^2}$ equals 1.0. The octave bandwidth of the preferred feature is given by the log-base-two ratio of the high and low frequencies at half-height

$$BW_{oct} = \log_2\left(\frac{f_H}{f_L}\right) = \log_2\left(\frac{f_0 + \sqrt{\ln 4}\big/2\pi\sigma_b}{f_0 - \sqrt{\ln 4}\big/2\pi\sigma_b}\right) \tag{11}$$

The orientation bandwidth specifies the polar angle spanned by the Gaussian envelope at half-height and is given by

$$BW_\theta = 2\tan^{-1}\left(\frac{\sqrt{\ln 4}}{2\pi\sigma_l f_0}\right) \tag{12}$$

When the visual angles spanned by the x' and y' values are 5x the envelope standard deviations in the band pass and low pass directions, respectively, the weight matrix is matched to the preferred feature. When the spanned visual angles are greater than 5x the standard deviation in either the band pass or low pass direction, the weight matrix is mismatched to the receptive field.

We analyze the response statistics of model neurons with vertically-oriented Gabor receptive fields having 42º orientation bandwidths and 0.8, 1.2, and 1.8 octave bandwidths. Simple cells in early visual cortex have a median orientation bandwidth of 42º, a median octave bandwidth of 1.2 octaves. The distribution of cortical octave bandwidths spans approximately 0.8 to 1.8 octaves at half-height (De Valois, Albrecht, & Thorell, 1982a; De Valois, Yund, & Hepler, 1982b). We computed response statistics for mismatched receptive field weight matrices spanning 5x (e.g. 2 cpd, 72 pixels, 1.2º) to 20x (e.g. 8 cpd, 72 pixels, 1.2º) the envelope standard deviations.

The aspect ratio of the Gaussian envelope in terms of the octave and orientation bandwidths is obtained by solving Eqs. 11 and 12 for σ_b and σ_l, respectively, and then taking the ratio

$$AR = \frac{\sigma_l}{\sigma_b} = \frac{\left(2^{BW_{oct}} - 1\right)}{\left(2^{BW_{oct}} + 1\right)\tan\left(\dfrac{BW_\theta}{2}\right)} \tag{13}$$

The log-base-two aspect ratios $\log_2(AR)$ of these receptive fields are -0.5, 0.0, and 0.5, respectively, which correspond to envelopes that are wider than high by a factor of $\sqrt{2}$, circular, and higher than wide by a factor of $\sqrt{2}$, respectively.

All data in the main text is presented for rectangular image patches and receptive field weight matrices. However, there are practical disadvantages to working with matrices that are rectangular. It is more convenient to work with square patches and weight matrices. We examined how the response statistics differ between rectangular and square weight matrices. Note that nominally matched square matrices are actually slightly mismatched for octave bandwidths other than 1.2. We computed the response statistics with square image patches and weight matrices for all octave bandwidths. The differences were minor (Fig. S11).

Normalization: To obtain the normalization factor for each stimulus, we converted each contrast image into its frequency domain representation by performing a fast-fourier transform (FFT). Next, we normalized the transform such that its total power equaled the total energy of the contrast image, in accordance with Parseval's theorem. To prevent high-frequency artifacts that may be caused by the edge of the image patch, it is common to apply a cosine window before performing

the FFT. However, because of numerical issues, it is impossible to avoid occasionally exceeding the maximum response r_{max} when a window is applied. Stimulus energy near the edge of the image patch that increases the linear response, may be windowed out of the normalization factor. In this case, the normalization factor will be smaller than it should be. In some cases, it will cause the normalized response to exceed the maximum. Results were similar with and without windowing, but they were better behaved without windowing.

Encoding noise: The responses of neurons are noisy. If the same exact stimulus is presented multiple times, the neuron is likely to give a slightly different response to each presentation. We considered two types of encoding noise: constant additive noise and scaled additive noise. Both types were modeled as mean zero Gaussian noise $\varepsilon \sim N\left(0, \sigma_I^2\right)$. With constant additive noise, the encoding noise variance σ_I^2 is constant regardless of the mean response. With scaled additive noise, the encoding noise variance $\sigma_I^2 = \alpha|r| + \sigma_0^2$ scales in rough proportion to the mean, where α is the fano factor. All of the qualitative results are essentially invariant to whether constant or scaled additive noise is used. The manuscript presents results for constant additive noise.

Downsampling: Image patches were downsampled using Matlab's imresize.m function with linear interpolation. Similar results are obtained using Matlab's impyramid.m. However, with impyramid.m the downsampling factors are restricted to powers of two; we favor imresize.m because of its increased flexibility. Other downsampling methods are likely to produce very similar results.

20

References

Ahumada, A., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of ….* http://doi.org/10.1121/1.1912577

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, *7*(6), 531–546.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–193.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience*, *24*(9), 2077–2089. http://doi.org/10.1523/JNEUROSCI.3852-02.2004

Barlow, H. (2001). Redundancy reduction revisited. *Network (Bristol, England)*, *12*(3), 241–253.

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., & Frégnac, Y. (2013). Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Frontiers in Neural Circuits*, *7*, 206. http://doi.org/10.3389/fncir.2013.00206

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

Benardete, E. A., Kaplan, E., & Knight, B. W. (1992). Contrast gain control in the primate retina: P cells are not X-like, some M cells are. *Visual Neuroscience*, *8*(5), 483–486.

Brady, N., & Field, D. J. (2000). Local contrast in natural images: normalisation and coding efficiency. *Perception-London-*, *29*(9), 1041–1055. http://doi.org/10.1068/p2996

Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images. *Proceedings of the National Academy of Sciences*, *108*(40), 16849–16854. http://doi.org/10.1073/pnas.1108491108

Burge, J., & Geisler, W. S. (2012). Optimal defocus estimates from individual images for autofocusing a digital camera. Presented at the Proceedings of the IS&T/SPIE 47th Annual Meeting, Proceedings of SPIE. http://doi.org/10.1117/12.912066

Burge, J., & Geisler, W. S. (2014). Optimal disparity estimation in natural stereo images. *Journal of Vision*, *14*(2). http://doi.org/10.1167/14.2.1

Burge, J., & Geisler, W. S. (2015). Optimal speed estimation in natural image movies predicts human performance. *Nature Communications*, *6*, 7900. http://doi.org/10.1038/ncomms8900

Burge, J., & Jaini, P. (2017). Accuracy Maximization Analysis for Sensory-Perceptual Tasks: Computational Improvements, Filter Robustness, and Coding Advantages for Scaled Additive Noise. *PLoS Computational Biology*, *13*(2), e1005281. http://doi.org/10.1371/journal.pcbi.1005281

Burge, J., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *Journal of Neuroscience*, *30*(21), 7269–7280. http://doi.org/10.1523/JNEUROSCI.5551-09.2010

Burge, J., McCann, B. C., & Geisler, W. S. (2016). Estimating 3D tilt from local image cues in natural scenes. *Journal of Vision*, *16*(13), 2. http://doi.org/10.1167/16.13.2

Burge, J., Peterson, M. A., & Palmer, S. E. (2005). Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision*, *5*(6), 534–542. http://doi.org/10.1167/5.6.5

Burkhardt, D. A., Fahey, P. K., & Sikora, M. A. (2006). Natural images and contrast encoding in bipolar cells in the retina of the land- and aquatic-phase tiger salamander. *Visual Neuroscience*, *23*(1), 35–47. http://doi.org/10.1017/S0952523806231043

Burt, P. J., & Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, *COM-31*, 1–9.

Butts, D. A., Desbordes, G., Weng, C., Jin, J., Alonso, J.-M., & Stanley, G. B. (2010). The episodic nature of spike trains in the early visual pathway. *Journal of Neurophysiology*, *104*(6), 3371–3387. http://doi.org/10.1152/jn.00078.2010

Campbell, F. W., Cleland, B. G., Cooper, G. F., & Enroth-Cugell, C. (1968). The angular selectivity of visual cortical cells to moving gratings. *The Journal of Physiology*, *198*(1), 237–250.

Carandini, M. (2004). Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS Biology*, *2*(9), E264. http://doi.org/10.1371/journal.pbio.0020264

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews. Neuroscience*, *13*(1), 51–62. http://doi.org/10.1038/nrn3136

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience*, *17*(21), 8621–8644.

Chander, D., & Chichilnisky, E. J. (2001). Adaptation to temporal contrast in primate and salamander retina. *Journal of Neuroscience*, *21*(24), 9904–9916.

Clatworthy, P. L., Chirimuuta, M., Lauritzen, J. S., & Tolhurst, D. J. (2003). Coding of the contrasts in natural images by populations of neurons in primary visual cortex (V1). *Vision Research*, *43*(18), 1983–2001. http://doi.org/10.1016/S0042-6989(03)00277-3

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, *31*(12), 2195–2207.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research*, *35*(1), 7–24.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982a). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), 545–559. http://doi.org/10.1016/0042-6989(82)90113-4

De Valois, R. L., Yund, E. W., & Hepler, N. (1982b). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), 531–544.

Ding, P., & Blitzstein, J. K. (2018). On the Gaussian Mixture Representation of the Laplace Distribution. *The American Statistician*, *72*(2), 172–174. http://doi.org/10.1080/00031305.2017.1291448

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. http://doi.org/10.1038/415429a

Felsen, G., Touryan, J., Han, F., & Dan, Y. (2005). Cortical sensitivity to visual features in natural scenes. *PLoS Biol*. http://doi.org/10.1371/journal.pbio.0030342.sd002

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. a, Optics and Image Science*, *4*(12), 2379–2394.

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., et al. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, *467*(7316), 673–677. http://doi.org/10.1038/nature09424

Flitcroft, D. I. (1990). A neural and computational model for the chromatic control of accommodation. *Visual Neuroscience*, *5*(6), 547–555.

Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *The Journal of Physiology*, *365*, 331–363.

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

Geisler, W. S., Najemnik, J., & Ing, A. D. (2009). Optimal stimulus encoders for natural tasks. *Journal of Vision*, *9*(13), 17.1–16. http://doi.org/10.1167/9.13.17

Gekas, N., Meso, A. I., Masson, G. S., & Mamassian, P. (2017). A Normalization Mechanism for Estimating Visual Motion across Speeds and Scales. *Current Biology*, *27*(10), 1514–1520.e3. http://doi.org/10.1016/j.cub.2017.04.022

Hawken, M. J., & Parker, A. J. (1987). Spatial properties of neurons in the monkey striate cortex. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society (Great Britain)*, *231*(1263), 251–288.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*(2), 181–197.

Heitman, A., Brackbill, N., Greschner, M., Sher, A., Litke, A. M., & Chichilnisky, E. J. (2016). Testing pseudo-linear models of responses to natural scenes in primate retina. *bioRxiv*. http://doi.org/10.1101/045336

Held, R. T., Cooper, E. A., & Banks, M. S. (2012). Blur and disparity are complementary cues to

depth. *Current Biology*, *22*(5), 426–431. http://doi.org/10.1016/j.cub.2012.01.033

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243.

Iyer, A. V., & Burge, J. (2018). Depth variation and stereo processing tasks in natural scenes. *Journal of Vision*, *18*(6), 1–22. http://doi.org/10.1167/18.6.4

Jaini, P., & Burge, J. (2017). Linking normative models of natural tasks to descriptive models of neural response. *Journal of Vision*, *17*(12), 16. http://doi.org/10.1167/17.12.16

Jogan, M., & Stocker, A. A. (2015). Signal Integration in Human Visual Speed Perception. *Journal of Neuroscience*, *35*(25), 9381–9390. http://doi.org/10.1523/JNEUROSCI.4801-14.2015

Jones, J. P., & Palmer, L. A. (1987a). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6), 1233–1258.

Jones, J. P., & Palmer, L. A. (1987b). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6), 1187–1211.

Knill, D. C., & Richards, W. (1996). Perception as Bayesian Inference. (D. C. Knill & W. Richards, Eds.). New York: Cambridge University Press.

Kotulak, J. C., & Schor, C. M. (1986). A computational model of the error detector of human visual accommodation. *Biological Cybernetics*, *54*(3), 189–194.

Kruger, P. B., Aggarwala, K. R., Bean, S., & Mathews, S. (1997). Accommodation to stationary and moving targets. *Optometry & Vision Science*, *74*(7), 505–510.

Lesica, N. A., Jin, J., Weng, C., Yeh, C.-I., Butts, D. A., Stanley, G. B., & Alonso, J.-M. (2007). Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron*, *55*(3), 479–491. http://doi.org/10.1016/j.neuron.2007.07.013

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature Neuroscience*, *5*(4), 356–363. http://doi.org/10.1038/nn831

Lyu, S., & Simoncelli, E. P. (2008). Nonlinear Image Representation Using Divisive Normalization. (pp. 1–8). Presented at the IEEE Conf Comput Vis Pattern Recognit. http://doi.org/10.1109/CVPR.2008.4587821

Lyu, S., & Simoncelli, E. P. (2009a). Modeling multiscale subbands of photographic images with fields of Gaussian scale mixtures. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *31*(4), 693–706. http://doi.org/10.1109/TPAMI.2008.107

Lyu, S., & Simoncelli, E. P. (2009b). Nonlinear extraction of independent components of natural images using radial gaussianization. *Neural Computation*, *21*(6), 1485–1519. http://doi.org/10.1162/neco.2009.04-08-773

Mante, V., Bonin, V., & Carandini, M. (2008). Functional Mechanisms Shaping Lateral Geniculate Responses to Artificial and Natural Stimuli. *Neuron*, *58*(4), 625–638. http://doi.org/10.1016/j.neuron.2008.03.011

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., & Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience*, *8*(12), 1690–1697. http://doi.org/10.1038/nn1556

McFarland, J. M., Cui, Y., & Butts, D. A. (2013). Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLoS Computational Biology*, *9*(7), e1003143. http://doi.org/10.1371/journal.pcbi.1003143

Mohanty, D., Scholl, B., & Priebe, N. J. (2012). The accuracy of membrane potential reconstruction based on spiking receptive fields. *Journal of Neurophysiology*, *107*(8), 2143–2153. http://doi.org/10.1152/jn.01176.2011

Nishimoto, S., Ishida, T., & Ohzawa, I. (2006). Receptive field properties of neurons in the early visual cortex revealed by local spectral reverse correlation. *Journal of Neuroscience*, *26*(12), 3269–3280. http://doi.org/10.1523/JNEUROSCI.4558-05.2006

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.

http://doi.org/10.1038/381607a0

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research*, *37*(23), 3311–3325.

Park, I. M., Archer, E. W., Priebe, N., & Pillow, J. (2013). Spectral methods for neural characterization using generalized quadratic models. *Advances in Neural Information Processing Systems*, 1–9. Retrieved from http://papers.nips.cc/paper/4993-spectral-methods-for-neural-characterization-using-generalized-quadratic-models.pdf

Park, M., & Pillow, J. W. (2011). Receptive Field Inference with Localized Priors. *PLoS Computational Biology*, *7*(10), e1002219. http://doi.org/10.1371/journal.pcbi.1002219.g010

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. a, Optics and Image Science*, *2*(9), 1508–1532.

Poincaré, H. (1912). Calcul des probabilités (pp. 1–354). Paris: Gauthier-Villars.

Priebe, N. J., & Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, *57*(4), 482–497. http://doi.org/10.1016/j.neuron.2008.02.005

Priebe, N. J., & Lisberger, S. G. (2004). Estimating target speed from the population response in visual area MT. *Journal of Neuroscience*, *24*(8), 1907–1916. http://doi.org/10.1523/JNEUROSCI.4233-03.2004

Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *Journal of Neuroscience*, *23*(13), 5650–5661.

Rehn, M., & Sommer, F. T. (2007). A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. *Journal of Computational Neuroscience*, *22*(2), 135–146. http://doi.org/10.1007/s10827-006-0003-9

Rieke, F., & Rudd, M. E. (2009). The challenges natural images pose for visual adaptation. *Neuron*, *64*(5), 605–616. http://doi.org/10.1016/j.neuron.2009.11.028

Rossi, E. A., & Roorda, A. (2010). The relationship between visual resolution and cone spacing in the human fovea. *Nature Publishing Group*, *13*(2), 156–157. http://doi.org/10.1038/nn.2465

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*(11), 1421–1431. http://doi.org/10.1038/nn1786

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, *46*(6), 945–956. http://doi.org/10.1016/j.neuron.2005.05.021

Samengo, I., & Gollisch, T. (2013). Spike-triggered covariance: geometric proof, symmetry properties, and extension beyond Gaussian stimuli. *Journal of Computational Neuroscience*, *34*(1), 137–161. http://doi.org/10.1007/s10827-012-0411-y

Schwartz, O., Pillow, J. W., Rust, N. C., & Simoncelli, E. P. (2006). Spike-triggered neural characterization. *Journal of Vision*, *6*(4), 484–507. http://doi.org/10.1167/6.4.13

Sebastian, S., Abrams, J., & Geisler, W. S. (2017). Constrained sampling experiments reveal principles of detection in natural scenes. *Proceedings of the National Academy of Sciences*. http://doi.org/10.1073/pnas.1619487114

Sebastian, S., Burge, J., & Geisler, W. S. (2015). Defocus blur discrimination in natural images with natural optics. *Journal of Vision*, *15*(5), 16. http://doi.org/10.1167/15.5.16

Shapley, R. M., & Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *The Journal of Physiology*, *285*, 275–298.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761.

Smyth, D., Willmore, B., Baker, G. E., Thompson, I. D., & Tolhurst, D. J. (2003). The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *Journal of Neuroscience*, *23*(11), 4746–4759.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., & Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron*, *42*(1), 155–162.

Tadmor, Y., & Tolhurst, D. J. (2000). Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. *Vision Research*, *40*(22), 3145–3157.

Talebi, V., & Baker, C. L. (2012). Natural versus synthetic stimuli for estimating receptive field models: a comparison of predictive robustness. *Journal of Neuroscience*, *32*(5), 1560–1576. http://doi.org/10.1523/JNEUROSCI.4661-12.2012

Taylor, M. M., Sedigh-Sarvestani, M., Vigeland, L., Palmer, L. A., & Contreras, D. (2018). Inhibition in Simple Cell Receptive Fields Is Broad and OFF-Subregion Biased. *Journal of Neuroscience*, *38*(3), 595–612. http://doi.org/10.1523/JNEUROSCI.2099-17.2017

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

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

Tyler, C. W., & Julesz, B. (1978). Binocular cross-correlation in time and space. *Vision Research*, *18*(1), 101–105.

Victor, J. D., Purpura, K., Katz, E., & Mao, B. (1994). Population encoding of spatial frequency, orientation, and color in macaque V1. *Journal of Neurophysiology*, *72*(5), 2151–2166.

Vintch, B., Movshon, J. A., & Simoncelli, E. P. (2015). A Convolutional Subunit Model for Neuronal Responses in Macaque V1. *Journal of Neuroscience*, *35*(44), 14829–14841. http://doi.org/10.1523/JNEUROSCI.2815-13.2015

Wainwright, M. J., & Simoncelli, E. P. (2000). Scale Mixtures of Gaussians and the Statistics of Natural Images. In S. A. Solla, T. K. Leen, & K. R. Muller (Eds.), (Vol. 12, pp. 855–861). Presented at the Advances in Neural Information Processing Systems, unknown.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604. http://doi.org/10.1038/nn858

Weliky, M., Fiser, J., Hunt, R. H., & Wagner, D. N. (2003). Coding of natural scenes in primary visual cortex. *Neuron*, *37*(4), 703–718.

Supplement

"Model Neuron Response Statistics to Natural Images"
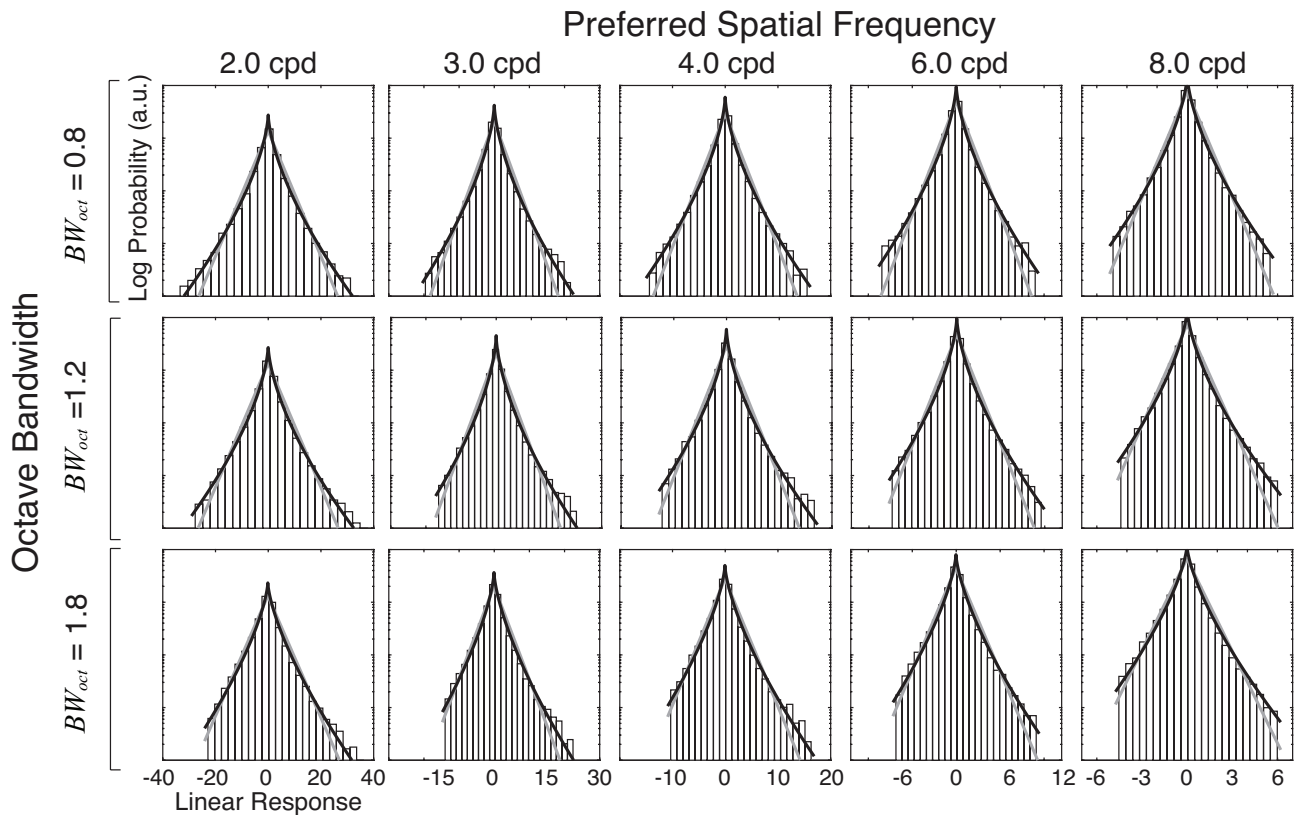
Arvind V. Iyer and Johannes Burge

Figure S1. Linear receptive field responses to natural images for all preferred spatial frequencies (columns) and octave bandwidths (rows). Note the dramatic difference in response magnitude (x-axis) as a function of preferred frequency. The responses are fit with both a Laplace and a generalized Gaussian via maximum likelihood (gray & black curves, respectively). The best-fit generalized Gaussian fit has significantly heavier tails than the best fit Laplace in all cases. The generalized Gaussian is given by $p(R) \propto \exp\left[-\left|R/\sigma\right|^p\right]$ where the Laplace distribution (p =1.0) and the Gaussian (p =2.0) are special cases. The linear responses are best fit with powers p of between 0.62 and 0.70, with a mean power of 0.65. The y-axis indicates log-probability over a four order-of-magnitude range. The Laplace fit appears as straight lines on a log-probability plot.
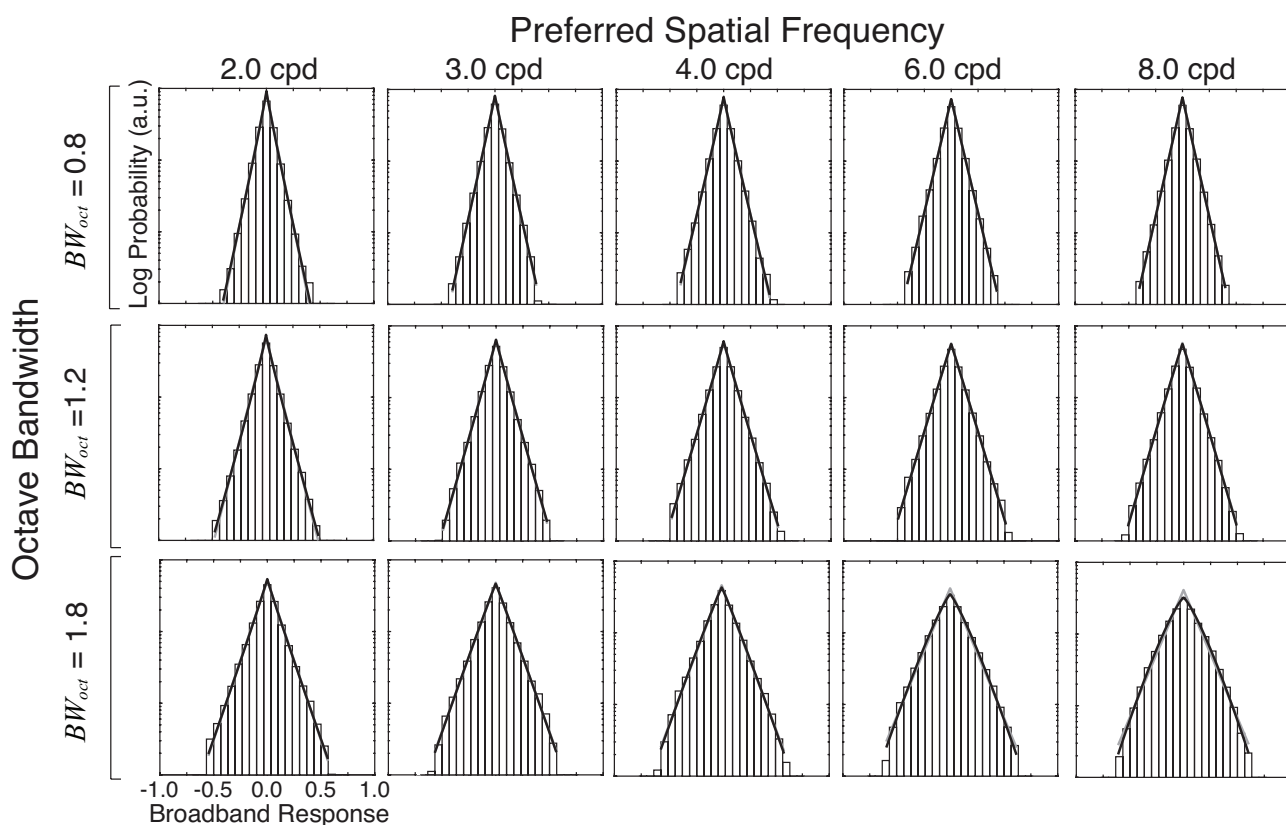
Figure S2. Broadband-normalized responses to natural images are Laplace distributed. Broadband responses with matched receptive field weight matrices for all preferred spatial frequencies (columns) and octave bandwidths (rows). The responses are fit with both a Laplace and a generalized Gaussian via maximum likelihood (gray & black curves, respectively). The generalized Gaussian is given by $p(R) \propto \exp\left[-\left|R/\sigma\right|^p\right]$ where the Laplace distribution ($p = 1.0$) and the Gaussian ($p = 2.0$) are special cases. The generalized Gaussian fit is indistinguishable from the Laplace fit in almost all cases; when the gray lines are not visible, it is because they are lying behind the black curve. The y-axis indicates log-probability over a three order-of-magnitude range. The Laplace fit appears as straight lines on a log-probability plot.
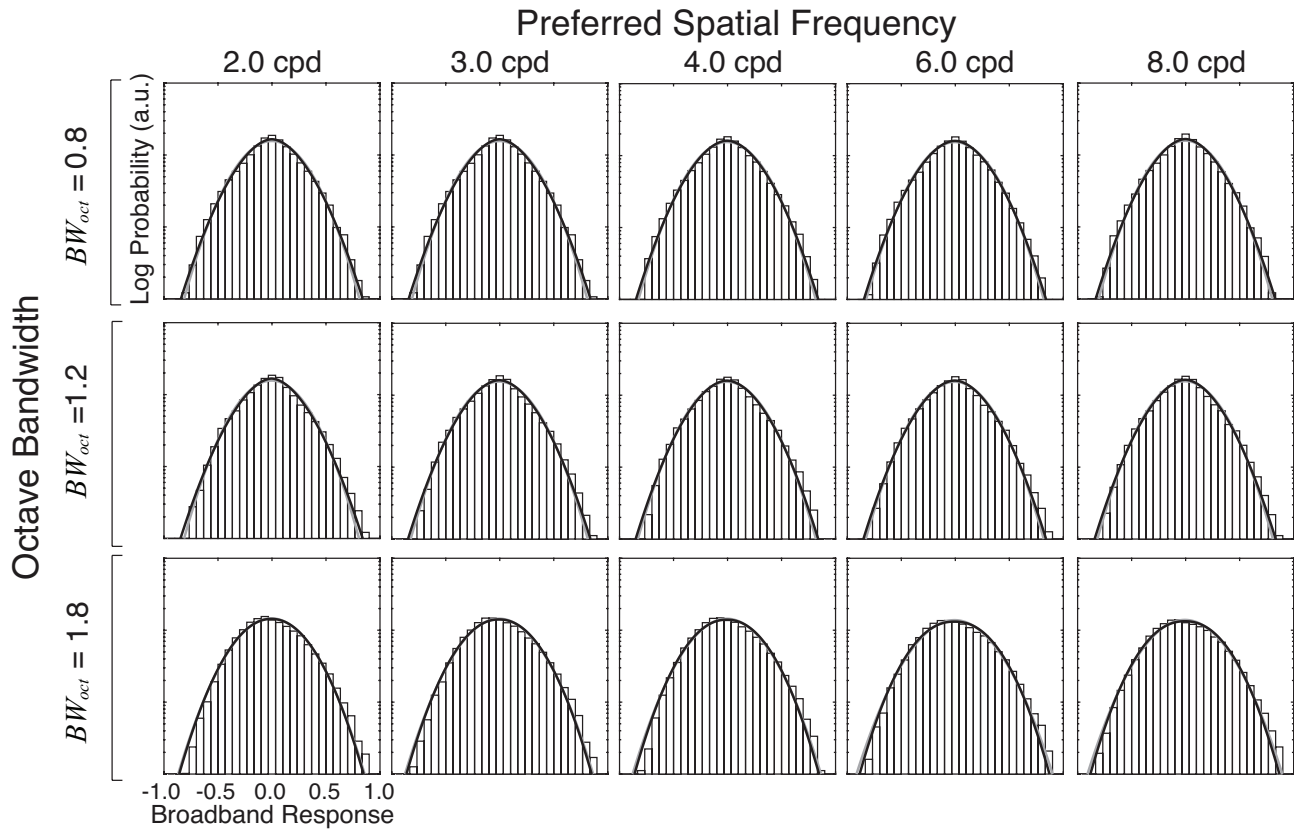
Figure S3. Narrowband-normalized responses to natural images are Gaussian-distributed. Narrowband responses with matched receptive field weight matrices for all preferred spatial frequencies (columns) and octave bandwidths (rows). The responses are fit with both a Gaussian and a generalized Gaussian via maximum likelihood (gray & black curves, respectively). The generalized Gaussian is given by $p(R) \propto \exp\left[-\left|R/\sigma\right|^p\right]$ where the Laplace distribution (p =1.0) and the Gaussian (p =2.0) are special cases. The generalized Gaussian fit is indistinguishable from the Gaussian fit in almost all cases; when the gray lines are not visible, it is because they are lying behind the black curve. The y-axis indicates log-probability over a three order-of-magnitude range. The Gaussian distribution appears as an upside-down parabola on a log-probability plot.
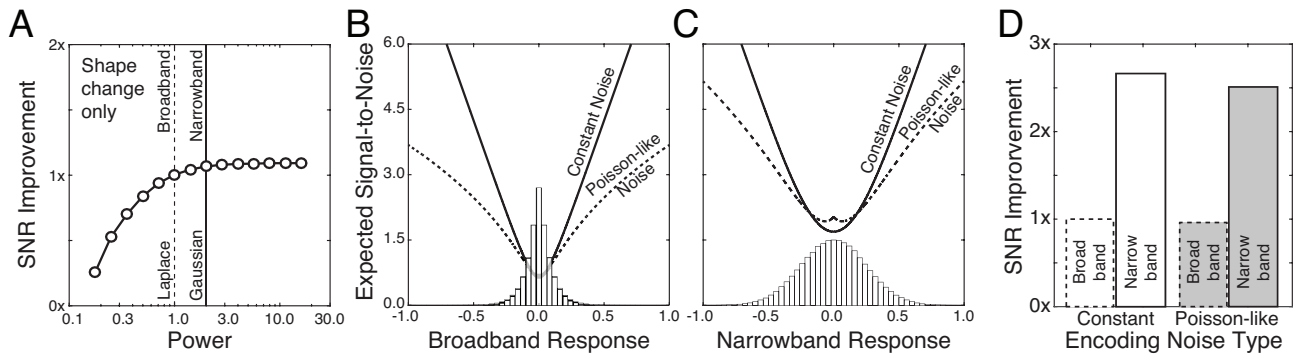
Figure S4. The impact of the shapes of the response distributions and the type of encoding noise on SNR. **A** The impact of the shape of the response distribution on signal-to-noise ratio for stimulus discriminability, assuming constant encoding noise. We computed expected SNR for a set of simulated generalized Gaussian response distributions having identical means and variances but different powers. The generalized Gaussian is given by $p(R) \propto \exp\left[-\left|R_{lin}/\sigma\right|^p\right]$. The Laplace distribution (p =1.0; kurtosis = 6.0) and the Gaussian (p =2.0; kurtosis = 3.0) are special cases. Powers larger than 2.0 approach the uniform distribution. **B** Expected SNR for broadband-normalized responses as a function of the response for constant (solid curve) and Poisson-like (dashed curve) encoding noise. The histogram shows the distribution of stimulus-driven broadband response. The constant Gaussian encoding noise was set such that the expected SNR from the broadband responses across the stimulus ensemble was equal to 1.0 (i.e. $\sigma_I \cong \sigma_E/3$). The Poisson-like is encoding noise is given by $\sigma_I = \sqrt{\alpha|R|+\sigma_0^2}$ where α is the fano factor and σ_0^2 are the fano factor and baseline variance, respectively. The values of the fano factor and the baseline variance were taken from the neurophysiological literature. Finally, the mean Poisson-like encoding noise variance across the stimulus ensemble was matched to the constant encoding noise variance. SNR for stimulus discrimination is approximately 4% lower with Poisson-like encoding noise. **C** Expected SNR for narrowband-normalized responses as a function of the response. The mean encoding noise variance was matched to the mean encoding noise for the broadband responses. SNR for stimulus discrimination is approximately 8% lower with Poisson-like encoding noise. **D** Summary of expected SNR improvement for broadband and narrowband normalization, with constant and Poisson-like encoding noise.
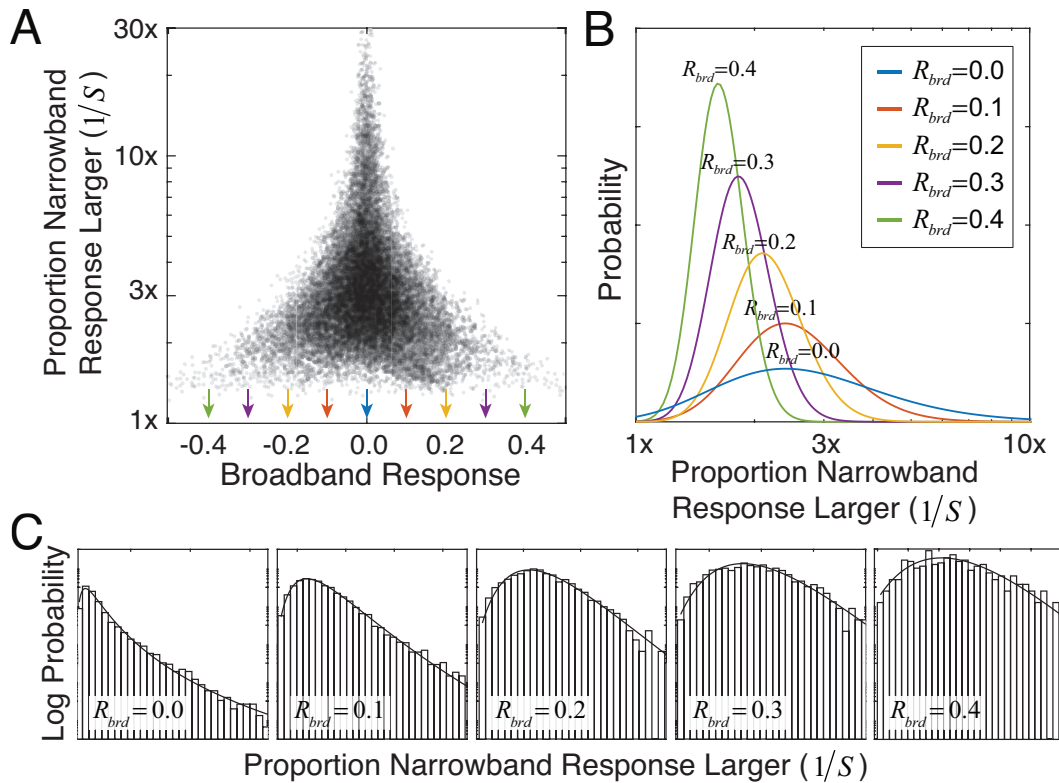
Figure S5. The impact of narrowband normalization. **A** Proportion that the narrowband response is larger than the broadband response R_{nrw}/R_{brd} as a function of the broadband response. Same data as Figure 5B in the main text. **B** Distribution of the proportional increase $p\left(R_{nrw}/R_{brd} \mid R_{brd}\right)$ in the narrowband response relative to the broadband response, conditioned on different absolute values of the broadband response (colors), as fit by inverse gamma distributions. Arrows in A mark the absolute values of broadband response upon which the proportions are conditioned. Same data as Figure 5C in the main text. **C** Maximum likelihood fits of inverse gamma distributions to the distributions of proportional increase for five absolute values of the broadband response. Small broadband responses tend to be amplified more than large broadband responses. Note that the proportional increase on the x-axis in each of the five panels ranges, respectively, from 1.2x to 33x, 1.2x to 9x, 1.2x to 4.5x, 1.2x to 3.0x, and 1.2x to 2.4x, respectively.
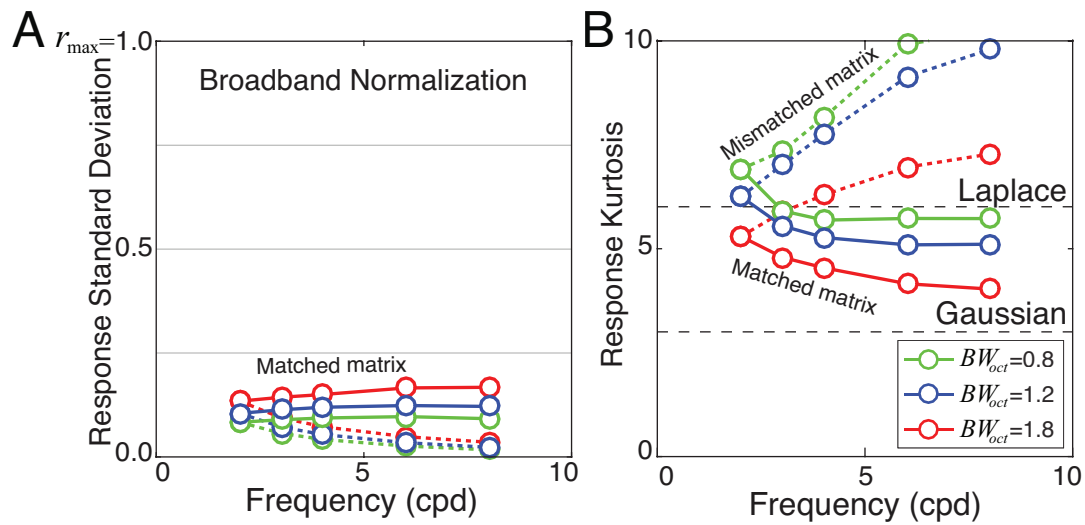
Figure S6. Response statistics with Gabor receptive fields, broadband normalization, and matched and mismatched rectangular weight matrices. **A** Stimulus-driven response standard deviation is invariant to preferred frequency when the receptive field's weight matrix is matched to the preferred feature (solid curves). Response standard deviation decreases systematically as the magnitude of the mismatch increases (dashed curves). **B** Stimulus-driven response kurtosis is consistent with a Gaussian when the weight matrix is matched to the preferred feature, but increases systematically as spatial frequency (and the magnitude of the mismatch) increases. Decreased stimulus-driven response variation (A) and increased response kurtosis (B) both decrease the signal-to-noise ratio for stimulus discrimination (Fig. S4).
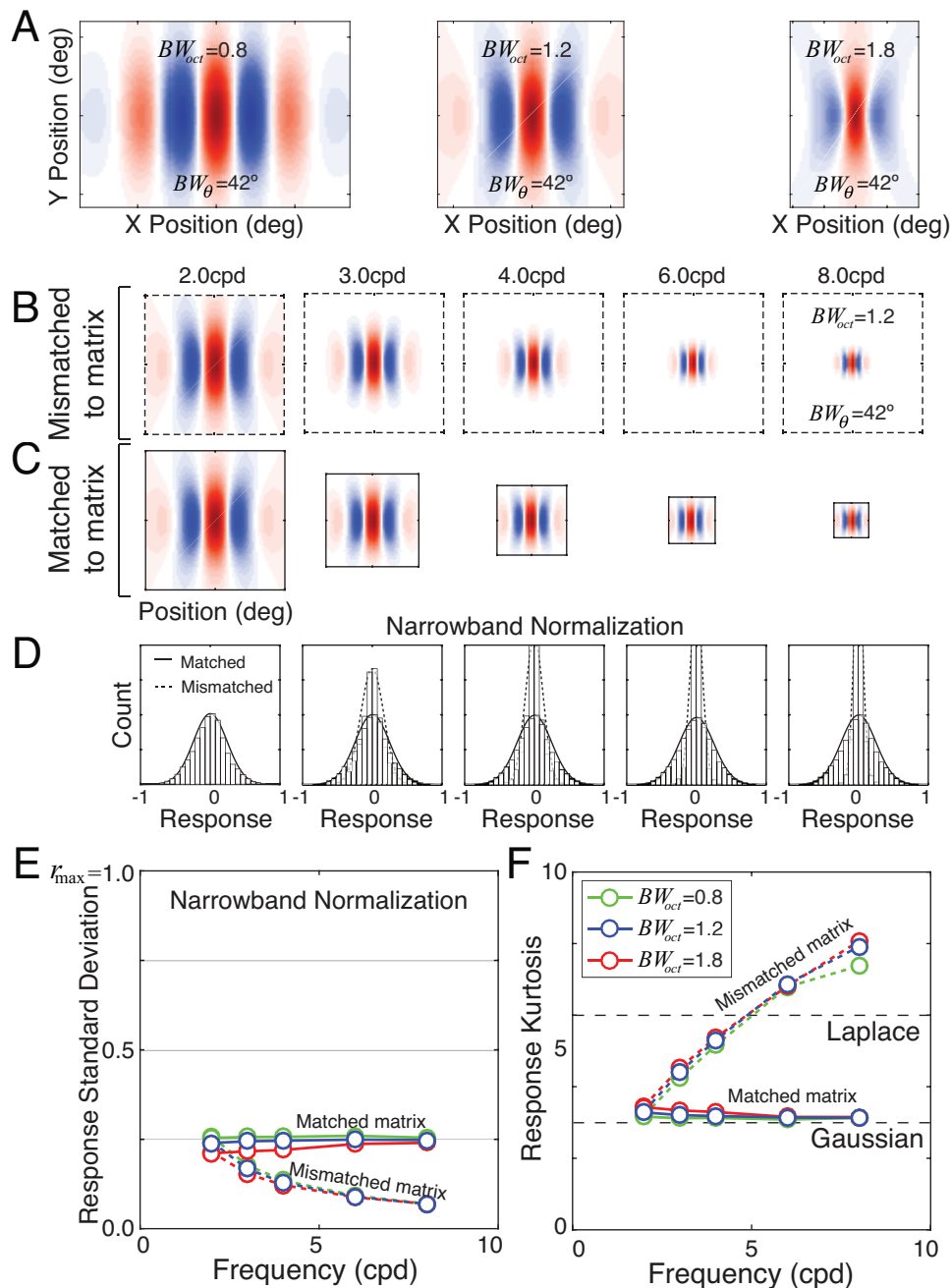
31

Figure S7. Log-Gabor receptive fields and response statistics. **A** Log-Gabor receptive fields with octave bandwidths of 0.8, 1.2, and 1.8, and orientation bandwidths of 42°. Different octave bandwidths correspond to preferred features with different aspect ratios (see Methods). The receptive fields **B** Mismatched and **C** matched receptive field weight matrices with 1.2 octave bandwidth for five preferred spatial frequencies. **D** Response distributions from matched and mismatched matrices. Matched weight matrices (solid curves) yield response distributions that are invariant to the scale of the preferred feature. Mismatched weight matrices (dashed curves) yield response distributions that change shape and variance with the magnitude of the mismatch. **E** Response standard deviation. Stimulus-driven response variance is constant with preferred frequency when the matrix is matched to the preferred feature. When mismatched, response standard deviation decreases with the magnitude of the mismatch. **F** Response kurtosis is the same as a Gaussian with matched weight matrices, but increases with the magnitude of the mismatch. Log-Gabor and Gabor response statistics are similar (see Fig. 7 in the main text).
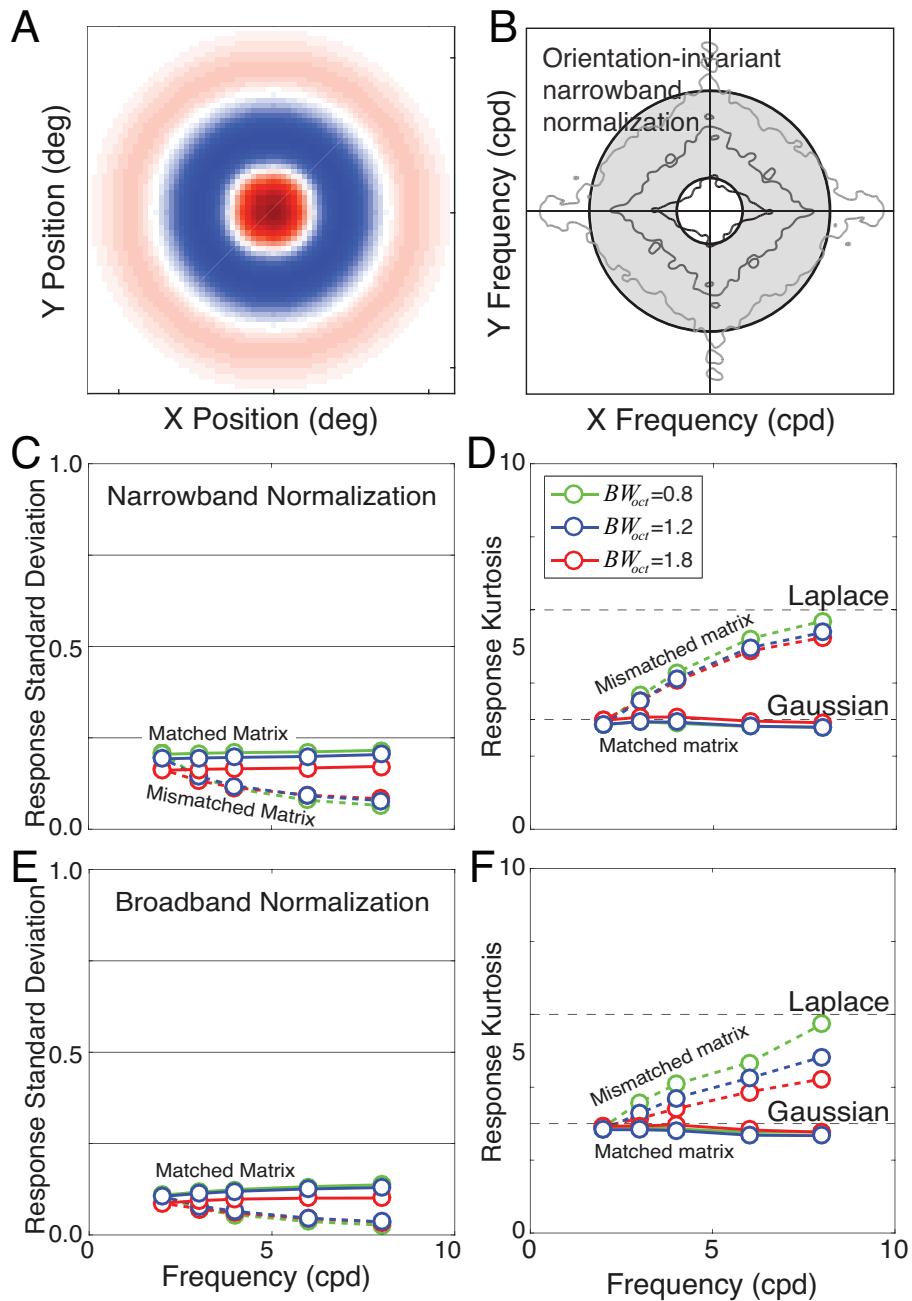
Figure S8. Response statistics for radial Gabor receptive fields with narrowband and broadband normalization. **A** Radially symmetric Gabor receptive field, similar to ganglion cell receptive fields in retina and relay cell receptive fields in lateral geniculate nucleus. **B** Schematic of pooling region in frequency space that determines the narrowband normalization factor. **C** Stimulus-driven response standard deviation with narrowband normalization for three different octave bandwidths (colors). **D** Stimulus-driven response kurtosis with narrowband normalization. **E** Response standard deviation with broadband normalization. **F** Response kurtosis with broadband normalization. Narrowband and broadband response statistics are more similar with radial Gabor receptive fields than with oriented Gabor receptive fields.
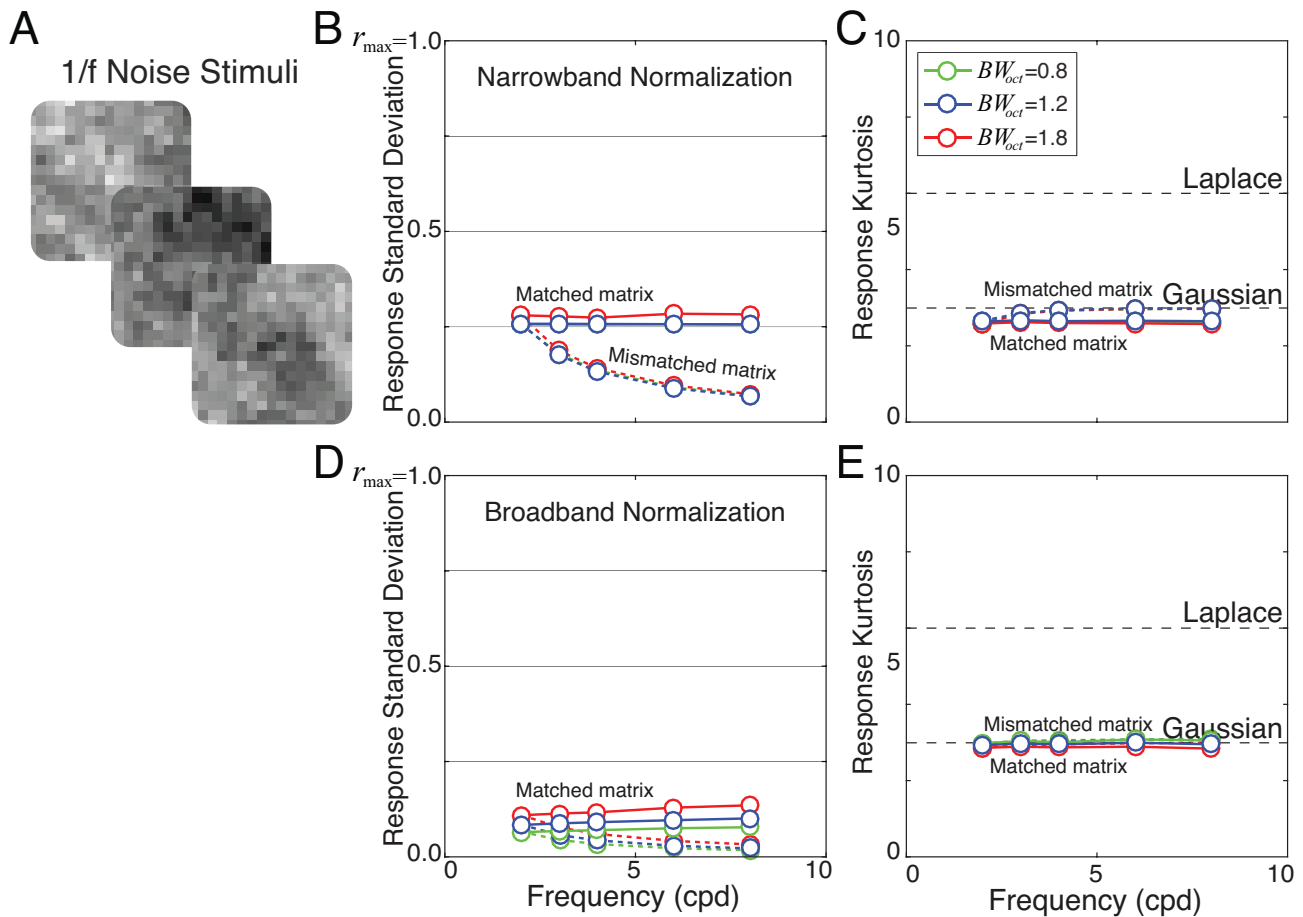
Figure S9. Response statistics to 1/f noise stimuli. Results are plotted for narrowband normalization with matched and mismatched weight matrices. **A** Example 1/f noise stimuli. 1/f noise stimuli were matched in contrast to the natural stimuli. **B** Stimulus-driven response standard deviation with narrowband normalization. **C** Stimulus-driven response kurtosis with narrowband normalization. Regardless of whether the matrices are matched or mismatched, the response kurtosis is consistent with a Gaussian. Note that it may appear surprising that with narrowband normalization, 1/f noise yields similar stimulus-driven response variance as the natural stimuli. (Broadband-normalized responses are substantially smaller with noise than with natural stimuli; see D). This is because our modeling assumes that the normalization factor is computed noiseless. This is not plausible for biological systems. Adding a small constant N_0 to the normalization term changes the response model to $R_{nrw} = \mathbf{f}^T \mathbf{c} / (N_{nrw} + N_0)$ and causes a substantial reduction in the stimulus-driven response variance to noise stimuli while leaving stimulus-driven response variance to natural stimuli relatively unaffected. The small constant has a larger effect on the response variance to noise stimuli because the (noiseless) narrowband normalization factor (i.e. similarity) tends to be much smaller for noise than for natural stimuli. Provided the constant is small enough, its addition to the response model does not appreciably change the distributional shape of the narrowband response distributions. **D** Response standard deviation with broadband normalization. The slight uptick in response standard deviation for the smaller matched matrices vanishes with downsampling. This increase is the only substantive difference we observed with downsampling. **E** Response kurtosis with broadband normalization.
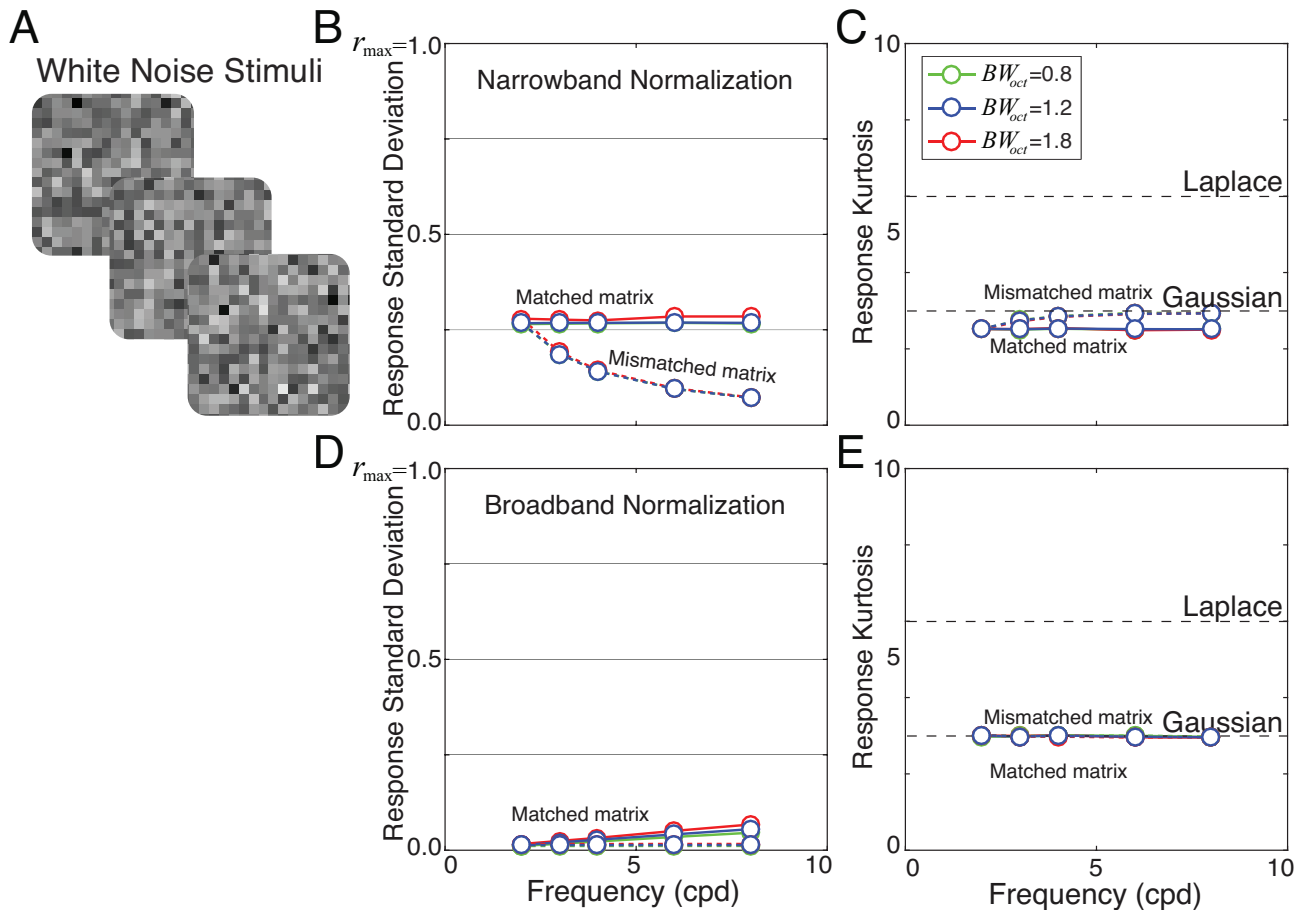
Figure S10. Response statistics to white noise stimuli. Results are plotted for narrowband normalization with matched and mismatched weight matrices. **A** Example 1/f noise stimuli. White noise stimuli were matched in contrast to the natural stimuli. **B** Stimulus-driven response standard deviation with narrowband normalization. **C** Stimulus-driven response kurtosis with narrowband normalization. Regardless of whether the matrices are matched or mismatched, the response kurtosis is consistent with a Gaussian. Note that it may appear surprising that with narrowband normalization, white noise yields similar stimulus-driven response variance as the natural stimuli. (Broadband-normalized responses are substantially smaller with noise than with natural stimuli; see D). This is because our modeling assumes that the normalization factor is computed noiseless. This is not plausible for biological systems. Adding a small constant N_0 to the normalization term changes the response model to $R_{nrw} = \mathbf{f}^T \mathbf{c} / (N_{nrw} + N_0)$ and causes a substantial reduction in the stimulus-driven response variance to noise stimuli while leaving stimulus-driven response variance to natural stimuli relatively unaffected. The small constant has a larger effect on the response variance to noise stimuli because the (noiseless) narrowband normalization factor (i.e. the similarities) tends to be much smaller for noise than for natural stimuli. Furthermore, because the phase-invariant similarity of white noise stimuli is lower than the phase-invariant similarity of 1/f noise stimuli to the receptive field, the constant drives down the stimulus-driven response variance to white noise stimuli more than to 1/f stimuli. Provided the constant is small enough, its addition to the response model does not appreciably change the distributional shape of the narrowband response distributions. **D** Response standard deviation with broadband normalization. The slight uptick in response standard deviation for the smaller matched matrices vanishes with downsampling. This increase is the only substantive difference we observed with downsampling. **E** Response kurtosis with broadband normalization.
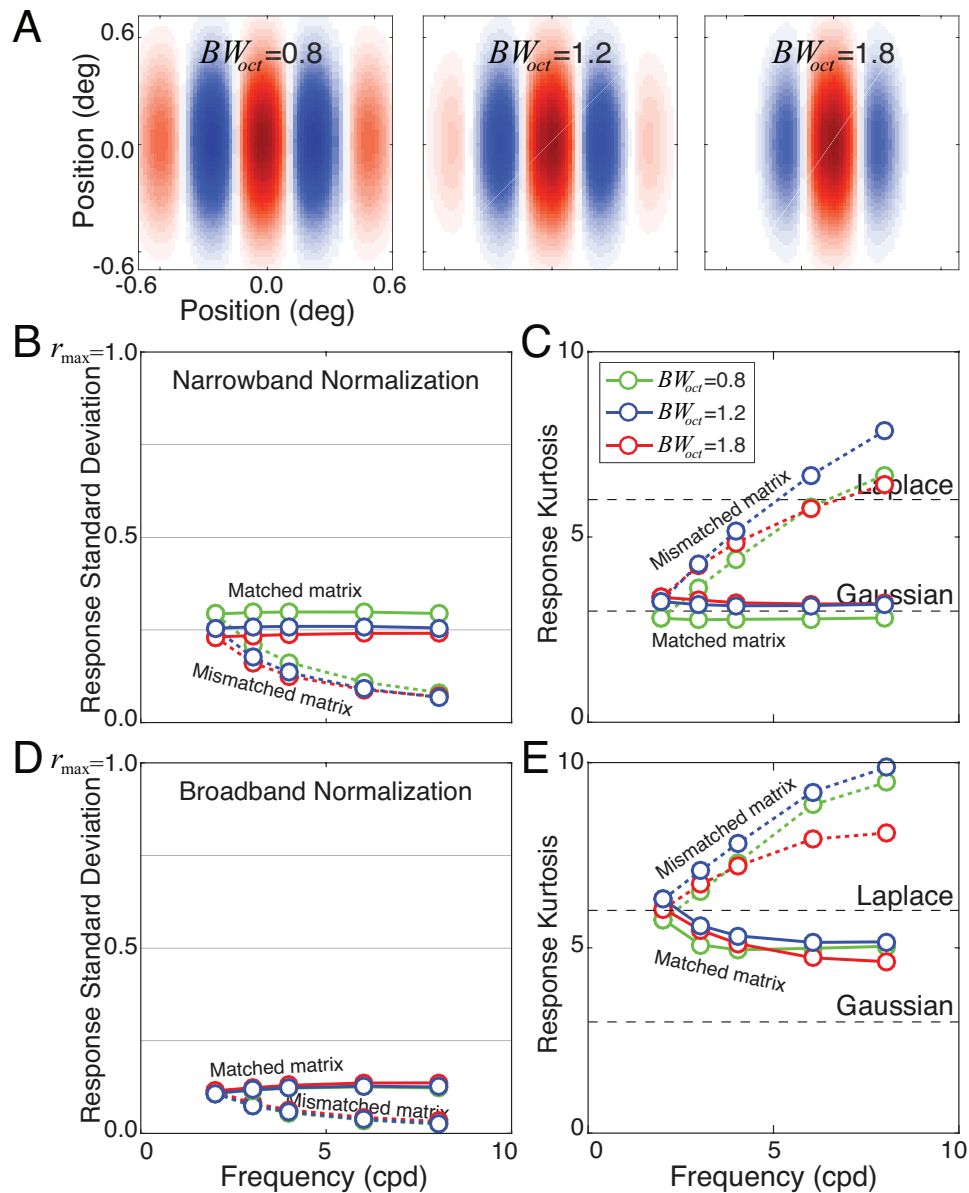
Figure S11. Response statistics with square weight matrices, which are often more convenient for modeling. Preferred features with different octave bandwidths have different aspect ratios. Thus, nominally matched square matrices are actually slightly mismatched for preferred features having octave bandwidths other than 1.2 (see Methods). **A** Preferred features in square weight matrices. **B** Stimulus-driven response standard deviation with narrowband normalization. Results for 'matched' square matrices are similar to the results with matched rectangular weight matrices presented in the main text. However, the 0.8 octave bandwidth preferred feature yields slightly higher stimulus-driven response variance and the 1.8 octave bandwidth feature yields slightly lower stimulus-driven variance than their rectangular counterparts. **C** Stimulus-driven response kurtosis with narrowband normalization. Results are very similar with square and rectangular weight matrices. **D** Response standard deviation with broadband normalization. **E** Response kurtosis with broadband normalization. Square weight matrices yield broadly similar results as rectangular weight matrices, and can probably be used for many applications interchangeably with rectangular matrices.

Supplement 1
Expected stimulus discriminability for Gaussian response distributions

Consider a neuron whose response can be modeled as a zero-mean Gaussian-distributed random variable r with standard deviation σ_E such that

$$r \sim N\left(0, \sigma_E^2\right) \tag{S1}$$

Let r_1 and r_2 be two random response samples. The response difference $u = r_1 - r_2$ is also Gaussian distributed with a variance that is twice the variance of each of the i.i.d. responses

$$u \sim N\left(0, 2\sigma_E^2\right) \tag{S2}$$

We are interested in the expected absolute difference $E\left[|u|\right] = E\left[|r_1 - r_2|\right]$ of two random responses. In general, the absolute value of a zero-mean Gaussian distributed random variable with variance σ^2 obeys a half-normal distribution with mean $\sqrt{\dfrac{2}{\pi}}\sigma$.

Given that $u = r_1 - r_2$ is a zero-mean Gaussian variable with variance $2\sigma_E^2$, we have

$$E\left[|r_1 - r_2|\right] = \sqrt{\frac{2}{\pi}}\sqrt{2}\sigma_E = \frac{2}{\sqrt{\pi}}\sigma_E \tag{S3}$$

If the neuron's response is corrupted by encoding noise of variance σ_I^2 , then the expected discriminability across stimuli for this neuron is given by

$$E\left[SNR\right] = \frac{E\left[|r_1 - r_2|\right]}{\sigma_I} = \frac{2}{\sqrt{\pi}}\frac{\sigma_E}{\sigma_I} \tag{S4}$$

37

Supplement 2
Expected stimulus discriminability for Laplace response distributions

Consider a neuron whose response can be modeled as a zero-mean Laplace-distributed random variable r with standard deviation σ_E such that

$$r \sim f(r) = \frac{1}{\sqrt{2}\sigma_E} e^{-\frac{|r|}{\sigma_E/\sqrt{2}}}, -\infty < r < \infty \tag{S5}$$

Let r_1 and r_2 be two random response samples. The response difference $u = r_1 - r_2$ is the difference of two i.i.d. responses. Let $u \sim g(u)$

$$g(u) = \int_{-\infty}^{\infty} f(x) f(x - u) dx$$

$$g(u) = \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2}\sigma_E} e^{-\frac{|x|}{\sigma_E/\sqrt{2}}} \right) \left(\frac{1}{\sqrt{2}\sigma_E} e^{-\frac{|x-u|}{\sigma_E/\sqrt{2}}} \right) dx$$

$$g(u) = \frac{1}{2\sigma_E^2} \int_{-\infty}^{\infty} e^{-\frac{|x|+|x-u|}{\sigma_E/\sqrt{2}}} dx \tag{S6}$$

The integral cannot be simply evaluated with an absolute value in the integrand. To remove the absolute value from the integrand, we split the integral depending on the values that u and x take. We note that $g(u)$ is even symmetric. Thus, solving the integral for all values of $u > 0$ will provide the solution to the integral for all values of $u < 0$. Assuming that $u > 0$, then $|x| + |x - u| = u - 2x$ when $-\infty < x < 0$, $|x| + |x - u| = u$ when $0 < x < u$, and $|x| + |x - u| = 2x - u$ when $u < x < \infty$. Evaluating for cases when $u > 0$ yields

$$g(u) = \frac{1}{2\sigma_E^2} \left(\int_{-\infty}^{0} e^{\frac{2x-u}{\sigma_E/\sqrt{2}}} dx + \int_{0}^{u} e^{-\frac{u}{\sigma_E/\sqrt{2}}} dx + \int_{u}^{\infty} e^{\frac{-2x+u}{\sigma_E/\sqrt{2}}} dx \right), \qquad u > 0$$

$$= \frac{1}{2\sigma_E^2} \left(\left[\frac{\sigma_E}{2\sqrt{2}} e^{\frac{2x-u}{\sigma_E/\sqrt{2}}} \right]_{-\infty}^{0} + \left[xe^{\frac{-u}{\sigma_E/\sqrt{2}}} \right]_{0}^{u} + \left[\frac{-\sigma_E}{2\sqrt{2}} e^{\frac{-2x+u}{\sigma_E/\sqrt{2}}} \right]_{u}^{\infty} \right), \ u > 0$$

$$= \frac{1}{2\sigma_E^2} \left(\frac{\sigma_E}{2\sqrt{2}} e^{\frac{-u}{\sigma_E/\sqrt{2}}} + ue^{\frac{-u}{\sigma_E/\sqrt{2}}} + \frac{\sigma_E}{2\sqrt{2}} e^{\frac{-u}{\sigma_E/\sqrt{2}}} \right), \qquad u > 0$$

$$= \frac{1}{2\sigma_E^2} \left(\frac{\sigma_E}{\sqrt{2}} e^{\frac{-u}{\sigma_E/\sqrt{2}}} + ue^{\frac{-u}{\sigma_E/\sqrt{2}}} \right), \qquad u > 0 \tag{S7}$$

Given that $g(u)$ is even-symmetric, we can replace u with $|u|$ in Equation S7

$$g(u) = \frac{1}{2\sigma_E^2} \left(\frac{\sigma_E}{\sqrt{2}} e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} + |u| e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} \right) \tag{S8}$$

After distributing the leading scale factor $1/\sigma_E^2$ and rearranging terms, we obtain an function where the two terms in the sum are the expressions for a Laplace distribution and a bilateral Gamma distribution

$$g(u) = \frac{1}{2}\left(\overbrace{\frac{1}{\sqrt{2}\sigma_E}e^{\frac{-|u|}{\sigma_E/\sqrt{2}}}}^{\text{Laplace distribution}} + \overbrace{\frac{1}{\sqrt{2}\sigma_E}\frac{|u|}{\sigma_E/\sqrt{2}}e^{\frac{-|u|}{\sigma_E/\sqrt{2}}}}^{\text{Bilateral Gamma Distribution}} \right) \tag{S9}$$

Recall that we are interested in the expectation of the absolute value of $|u| = |r_1 - r_2|$, not the expectation of $u = r_1 - r_2$ itself. Computing the expectation $E\big[|u|\big]$ using $g(u)$ from the definition of expectation

$$E\big[|u|\big] = \int_{-\infty}^{\infty} |u| g(u)\, du$$

$$E\big[|u|\big] = \frac{1}{2}\int_{-\infty}^{\infty} |u| \left(\frac{1}{\sqrt{2}\sigma_E}e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} + \frac{1}{\sqrt{2}\sigma_E}\frac{|u|}{\sigma_E/\sqrt{2}}e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} \right) du$$

Noting that the integrand is an even function means that twice the integral from zero to infinity equals the integral from negative infinity to infinity

$$E\big[|u|\big] = \int_{0}^{\infty} \left(\frac{|u|}{\sqrt{2}\sigma_E}e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} + \frac{1}{\sigma_E^2}|u|^2\, e^{\frac{-|u|}{\sigma_E/\sqrt{2}}} \right) du$$

On the positive real line, we can drop the absolute value symbols

$$E\big[|u|\big] = \int_{0}^{\infty} \left(\frac{u}{\sqrt{2}\sigma_E}e^{\frac{-u}{\sigma_E/\sqrt{2}}} + \frac{1}{\sigma_E^2}u^2 e^{\frac{-u}{\sigma_E/\sqrt{2}}} \right) du$$

Splitting the integral

$$E\big[|u|\big] = \frac{1}{\sqrt{2}\sigma_E}\int_{0}^{\infty} u e^{\frac{-u}{\sigma_E/\sqrt{2}}}\, du + \frac{1}{\sigma_E^2}\int_{0}^{\infty} u^2 e^{\frac{-u}{\sigma_E/\sqrt{2}}}\, du \tag{S10}$$

Each of the two definite integrals in equation S6 can be computed with the standard result

$$\int_{0}^{\infty} u^n e^{-au}\, du = \frac{n!}{a^{n+1}}, n \in W, a > 0$$

39

Plugging in

$$E\left[|u|\right] = \frac{1}{\sqrt{2}\sigma_E}\left[\frac{1}{\left(\frac{1}{\sigma_E/\sqrt{2}}\right)^2}\right] + \frac{1}{\sigma_E^2}\left[\frac{2}{\left(\frac{1}{\sigma_E/\sqrt{2}}\right)^3}\right]$$

Simplifying terms

$$E\left[|u|\right] = \frac{1}{\sqrt{2}\sigma_E}\left[\frac{\sigma_E^2}{2}\right] + \frac{1}{\sigma_E^2}\left[\frac{2\sigma_E^3}{2\sqrt{2}}\right]$$

$$= \frac{\sigma_E}{2\sqrt{2}} + \frac{\sigma_E}{\sqrt{2}}$$

$$= \frac{3}{2\sqrt{2}}\sigma_E$$

Therefore, the mean absolute difference between two i.i.d. mean-zero Laplace random variables of standard deviation σ_E is

$$E\left[|u|\right] = E\left[|r_1 - r_2|\right] = \frac{3}{2\sqrt{2}}\sigma_E \tag{S11}$$

For internal noise of standard deviation σ_I, the expected stimulus discriminability across all stimuli is given by

$$E\left[SNR\right] = \frac{E\left[|r_1 - r_2|\right]}{\sigma_I} = \frac{3}{2\sqrt{2}}\frac{\sigma_E}{\sigma_I} \tag{S12}$$