

The Divisive-Normalization Model of V1 Neurons: A Comprehensive Comparison of Physiological Data and Model Predictions

Tadamasa Sawada¹ and Alexander A. Petrov²

¹School of Psychology, Higher School of Economics

²Department of Psychology, Ohio State University

October 14, 2016

Abstract

The physiological responses of simple and complex cells in the primary visual cortex (V1) have been studied extensively and modeled at different levels. At the functional level, the divisive normalization model (DNM, Heeger, 1992) has accounted for a wide range of single-cell recordings in terms of a combination of linear filtering, nonlinear rectification, and divisive normalization. We propose standardizing the formulation of the DNM and implementing it in software that takes static grayscale images as inputs and produces firing-rate responses as outputs. We also review a comprehensive suite of 30 empirical phenomena and report a series of simulation experiments that qualitatively replicate dozens of key experiments with a standard parameter set consistent with physiological measurements. This systematic approach identifies novel falsifiable predictions of the DNM. We show how the model simultaneously satisfies the conflicting desiderata of flexibility and falsifiability. Our key idea is that, while adjustable parameters are needed to accommodate the diversity *across* neurons, they must be fixed for a given *individual* neuron. This requirement introduces falsifiable constraints when this single neuron is probed with multiple stimuli. We also present mathematical analyses and simulation experiments that explicate some of these constraints.

1 Introduction

The primary visual cortex (V1) is the most studied cortical area. Beginning with the seminal studies of Hubel and Wiesel (1959, 1962), V1 neurons have been studied extensively in physiology for half a century (see Albrecht, Geisler, & Crane, 2003; Andoni, Tan, & Priebe, 2013; Angelucci & Shushruth, 2013; Ferster & Miller, 2000; Fitzpatrick, 2000; Hubel & Wiesel, 1977; Lamme, 2003; Lennie & Movshon, 2005, for reviews). A wide range of models have been proposed to account for various properties of V1 neurons at various levels of analysis (see

Albrecht, Geisler, Frazor, & Crane, 2002; Carandini et al., 2005; Carandini, Heeger, & Movshon, 1999; Graham, 1992, 2011; Grossberg, 1988; Hubel & Wiesel, 1977; Sompolinsky & Shapley, 1997, for reviews).

These models of the V1 neurons can be classified into three types: namely, *descriptive*, *functional*, and *structural* models (Albrecht et al., 2003, p. 759; see also Herz, Gollisch, Machens, & Jaeger, 2006). A *descriptive model* applies statistical regression techniques to summarize and regularize a set of empirical measurements with a mathematical equation. These measurements can be characterized by a small number of parameters of this equation. For example, the tuning bandwidth of a V1 neuron can be estimated by fitting a Gaussian function to its empirical tuning-curve. A *functional model* aims to characterize a variety of response properties within the context provided by a visual information-processing algorithm. For example, a simple cell can be modeled approximately as a linear filter followed by rectification (e.g., Movshon, Thompson, & Tolhurst, 1978c). Ideally, a *functional* model can take the stimulus image as input and calculate the response. This makes it possible to perform simulation experiments with a functional model by presenting it with a set of stimuli and examining the predicted responses. Typically, *functional* models of V1 neurons involve various combinations of linear and nonlinear operations defined via algebraic equations. The goal of a *functional* model is to characterize the neuron's response to a given stimulus. All models discussed in the present article are of this type. Note that a *functional* model can be regarded as an intermediate step toward a *structural* model (Marr, 1982, ch. 1.2). A *structural model* aims to characterize some aspect of the biophysical and/or biochemical processing mechanisms in the early visual system. Typically, it is formulated either as an algorithm representing a neural circuit or as a system of differential equations (see Ben-Yishai, Bar-Or, & Sompolinsky, 1995; Brosch & Neumann, 2014; Chance & Abbott, 2000; Ellias & Grossberg, 1975; Grossberg, 1988; Izhikevich, 2007; Kough & Poggio, 2008; Schwabe, Obermayer, Angelucci, & Bressloff, 2006; Shapley & Xing, 2013; Somers, Nelson, & Sur, 1995; Somers et al., 1998; Zhaoping, 2011, for reviews and examples). *Structural* models are beyond our scope here.

The present review is based on a functional model of the static (steady-state) properties of simple and complex cells in V1. The temporal dynamics of the neuronal response is beyond our present scope. (See, e.g., Albrecht et al., 2002; Heeger, 1993; Albrecht et al., 2003, for reviews on temporal dynamics. We will revisit this issue in Section 4, General Discussion.) Most experiments discussed below used prolonged (steady-state) stimuli such as drifting and/or flickering gratings with relatively long durations (e.g., a few seconds). Under these conditions, the transient response triggered by the stimulus onset can be bracketed out of the analysis to a good approximation, and the most important dependent variable is the steady-state firing rate of the neuron as measured by a post-stimulus time histogram. The models discussed below are functional models that take a single grayscale image as input and produce a single number for each neuron simulated that we take to be “the response” to a given stimulus.

Within this circumscribed scope, there is “a fairly well agreed on standard model of V1 response properties, usually involving a combination of linear filtering, half-wave rectification and squaring, and response normalization” (Carandini et al., 2005, p. 10590). Whereas it is still unknown how well this *divisive normalization model* (DNM) can account for the full complexity of the V1 population code for time-varying naturalistic stimuli (Olshausen & Field, 2005; but see Rust & Movshon, 2005), it is consistent with much of the available data to a good approximation. The DNM was developed over a number of years during which it combined experimental (e.g., R. L. De Valois, Albrecht, & Thorell, 1982; Hubel & Wiesel, 1959; Movshon et al., 1978c) and theoretical (e.g., Grossberg, 1973) contributions, as well as interdisciplinary explorations of the correspondence between the physiological data and the mathematical formalisms (e.g., Albrecht & Geisler, 1991; Heeger, 1992b; Carandini & Heeger, 1994). Versions of this model have been applied to a broad spectrum of data ranging from single-cell recordings (see, e.g., Albrecht et al., 2003; Heeger, 1992b; Carandini & Heeger, 2012, for reviews), multi-electrode population recordings (e.g., Busse, Wade, & Carandini, 2009; Goris, Wichmann, & Henning, 2009), and psychophysical experiments (e.g., Itti, Koch, & Braun, 2000; Olzak & Thomas, 1999, 2003; To, Lovell, Troscianko, & Tolhurst, 2010; Jesús & Laparra, 2010).

Unfortunately, there is no standard formulation of the DNM. Different authors and publications use different mathematical expressions and idiosyncratic parameterizations. There is a clear family resemblance across these model variants—the verbal description quoted above summarizes the core DNM ideas—but the main results are scattered across dozens of journal articles spanning decades of experimental and theoretical development. A typical article reports neurophysiological data collected with a certain restricted class of stimuli, and then reports the fit of a DNM variant customized for the particular application. This practice makes it difficult to compare the results across studies.

Here, we propose a standard formulation of the DNM and test its validity with respect to a comprehensive suite of empirical phenomena (listed in Table 1 below). This article is organized around a series of figures with side-by-side comparisons between data patterns illustrating various properties of real V1 neurons and the corresponding patterns simulated with the DNM.

In addition to consolidating a large and scattered literature, our systematic approach sets the stage for identifying novel falsifiable predictions of the DNM. An adequate functional model should satisfy the conflicting desiderata of flexibility and falsifiability at the same time (Popper, 1959; Roberts & Pashler, 2000). On the one hand, the model needs to be flexible enough to accommodate the variety of V1 neurons with some adjustable parameters. The data reviewed below were recorded from neurons from different species (e.g., cats, new/old world monkeys, rabbits, rodents, and ferrets) under different conditions (e.g., anesthesia vs. alertness) and different experimental protocols. Furthermore, there is substantial variability within a sample

of neurons recorded from a single animal under constant conditions. Clearly, adjustable parameters are needed to accommodate this diversity. Note also that it is easy to make the model more flexible by adding more parameters. On the other hand, if the model becomes so flexible as to be able to fit any response pattern, it would become devoid of all empirical content (Roberts & Pashler, 2000). Whereas it might still be useful as a descriptive formalism for a succinct characterization of properties such as tuning bandwidths, such a model would not constrain our theories about the functional organization of the visual system. To have empirical content, the model must be restrictive enough to rule out at least some possible data patterns.

While the model's parameters may have to be adjusted flexibly to accommodate the diversity *across* neurons, they must be fixed for a given *individual* neuron. This introduces falsifiable constraints when this single neuron is probed with a judiciously-chosen suite of stimuli. This article presents mathematical analyses and simulation experiments that explicate some of these constraints. These novel theoretical results emerge naturally from our side-by-side examination of dozens of phenomena within the framework of a single model. This type of analysis focuses on the *qualitative* patterns that can be produced by a model under a given parameterization (Pitt, Kim, Navarro, & Myung, 2006). It contrasts with the typical approach in the experimental literature where, with some notable exceptions (e.g., Tadmor & Tolhurst, 1989), one or more models were compared in terms of their *quantitative* fits to physiological data pertaining to a single phenomenon.

Our review is of potential interest to several groups of readers. First, readers interested in the neurophysiology of simple and complex cells in the primary visual cortex will find a systematic series of figures with representative data from classic experiments, as well as their interpretation under the DNM. Single-cell data published over a 50-year span were digitized from select figures in the original reports and are replotted below. Second, readers interested in functional modeling of the early visual system will find systematic exposition and motivation of the DNM, as well as its empirical grounding. Third, expert modelers of the early visual system will find mathematical derivations and simulation experiments that identify novel falsifiable predictions of the DNM. Last but not least, modelers who need an off-the-shelf front-end to a larger model (e.g, Petrov, Doshier, & Lu, 2005, 2006; Jacobs, 2009) will find a general-purpose parameterization of the DNM and a standard parameter set (Table 2 below) that is consistent with almost all phenomena listed in Table 1. The model was implemented as a software program for Matlab (The MathWorks, 2015). This software takes a static grayscale image as input and produces a matrix of firing-rate responses for a population of DNM neurons centered on a single retinal location and tuned for a range of orientations and spatial frequencies.

The rest of the article is organized as follows: Section 2 presents the DNM, its proposed parameterization, and computational implementation. Section 3 reviews over two dozen empirical phenomena and interprets them through the lens of the DNM. It also reports mathematical analyses, simulation results, and some novel

predictions. Finally, there is a general discussion (Section 4) followed by mathematical appendices.

2 Models

The essential components of the divisive normalization model (DNM) are the linear filters, the static nonlinearities, and divisive normalization. These components will be described below. But, before this is done, we must acknowledge an important pre-processing step, namely, *light adaptation* (or *luminance gain control*). Adaptation is primarily accomplished in the retina (Shapley & Enroth-Cugell, 1984). It matches the limited dynamic range of the neurons to the locally-prevalent luminance. The DNM does not model this light adaptation explicitly. It simply assumes it has been incorporated into the encoding of the input images. This assumption is justified in situations when the stimuli are embedded in a large uniform gray background, and when the visual system has adapted to the baseline luminance level L_b . The input to the model is a matrix $I(x, y)$ of local contrast around this fixed baseline:

$$I(x, y) = [L(x, y) - L_b]/L_b, \quad (1)$$

where $L(x, y)$ is the luminance at coordinates x and y . In this notation, a sinusoidal grating with *contrast* c modulates between $I_{min} = -c$ and $I_{max} = +c$, and has zero mean. The maximal possible contrast ($c_{max} = 1$) of a grating is attained when the lowest intensity is zero and the highest intensity is twice the mean.

2.1 Linear model

The majority of V1 neurons respond selectively to a variety of stimulus features including position, size, orientation, and spatial frequency (e.g., R. L. De Valois, Albrecht, & Thorell, 1982; Hubel & Wiesel, 1959, 1968; Pollen & Ronner, 1982; Watkins & Berkley, 1974). A typical V1 neuron responds to stimulation within a circumscribed region called the (*classical*) *receptive field* (RF). Different neurons have RFs centered on different positions, and V1, as a whole, forms a *topographic map* (e.g., Schwartz, Tootell, Silverman, Switkes, & De Valois, 1985). This population code is beyond our scope. We are modeling a representative individual neuron. Note that, in this article, the origin of the xy coordinate system is placed conventionally at the center of the neuron's receptive field.

Hubel and Wiesel (1959, 1962) introduced the influential distinction between *simple* and *complex cells* in V1. According to their original definition (Hubel & Wiesel, 1962), simple cells have four characteristic properties: (1) distinct excitatory and inhibitory sub-regions within the RF, (2) spatial summation within a given subregion, (3) mutual antagonism between sub-regions, and (4) the responses to novel stimuli can be predicted to a good approximation on the basis of the spatial arrangement of the sub-regions. These four properties would

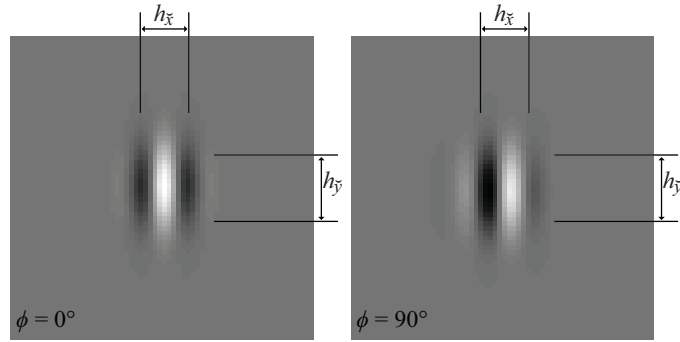


Figure 1: Examples of Gabor patches with phase $\Phi = 0^\circ$ and $\Phi = 90^\circ$, vertical orientation ($\Theta = 0^\circ$), and spatial frequency $F = 2$ cpd (Equation 2). The parameters h_x and h_y control the full width at half height (FWHH) of the Gaussian envelope along the directions that are perpendicular and parallel to the grating, respectively.

be expected from a linear spatio-temporal filter and they would motivate the application of Linear-Systems Theory (e.g., Lathi, 2005) to the study of spatial vision (e.g., R. L. De Valois & De Valois, 1988; Graham, 1989; Maffei & Fiorentini, 1973; Shapley & Lennie, 1985). Quantitative tests have identified a subpopulation of V1 neurons that exhibit these linear properties to a good approximation (e.g., Andrews & Pollen, 1979; Movshon et al., 1978c; Pollen & Ronner, 1982; see Albrecht et al., 2003, for review). This Linear-Systems approach is reinforced by a rich body of psychophysical data (e.g., Campbell & Robson, 1968; Cornsweet, 1970) and theory that supports the existence of *channels* selective for orientation and spatial frequency (R. L. De Valois & De Valois, 1988; Graham, 1989). But, two caveats should be kept in mind here (Albrecht et al., 2003). First, the match between the measured single-cell responses and the linear-systems predictions is always approximate, never exact, because they are systematic nonlinearities (discussed below). Second, the simple- vs. complex-cell distinction probably denotes the endpoints of a continuum rather than a sharp dichotomy (Mechler & Ringach, 2002). These caveats notwithstanding, this distinction has proven its theoretical utility and is widely used in V1 models.

The spatial layout of simple-cell receptive fields has been mapped out using local stimulus probes (e.g., Hubel & Wiesel, 1959, 1962, 1968; Movshon, Thompson, & Tolhurst, 1978b, see Section 3.3.6 for further references). The RF of a typical simple cell consists of alternating excitatory and inhibitory sub-regions (Hubel & Wiesel, 1962; Maffei & Fiorentini, 1973; R. L. De Valois, Albrecht, & Thorell, 1982; Sengpiel, Sen, & Blakemore, 1997). This 2D spatial pattern can be approximated well by the *Gabor function* in Equation 2 (Daugman, 1980, 1985; Field & Tolhurst, 1986; J. P. Jones & Palmer, 1987a, 1987b; Kulikowski, Marčelja, & Bishop, 1982; Marčelja, 1980; Ringach, 2002). Mathematically, a Gabor function $G_{XYF\Theta\Phi}(x, y)$ is a product of a Gaussian envelope

and a sinusoidal grating:

$$\begin{cases} G_{XYF\Theta\Phi}(x, y) = \exp \left[-4 \ln 2 \left(\frac{\check{x}^2}{h_{\check{x}}^2} + \frac{\check{y}^2}{h_{\check{y}}^2} \right) \right] \cos (F\check{x} - \Phi) \\ \check{x} = (y - Y) \sin \Theta + (x - X) \cos \Theta \\ \check{y} = (y - Y) \cos \Theta - (x - X) \sin \Theta, \end{cases} \quad (2)$$

where x and y (degrees of visual angle, deg) are positions across the image, and X and Y define the center of the RF. The grating has spatial¹ frequency F (cycles/deg, cpd), phase Φ , and orientation Θ . The parameters $h_{\check{x}}$ and $h_{\check{y}}$ control the full width at half height (FWHH) of the Gaussian envelope along the orthogonal and parallel directions as shown in Figure 1. See Appendix A for more details.

The alternating arrangement of excitatory and inhibitory RF regions makes the simple cell selectively responsive (or *tuned*) to the orientation and frequency of the stimuli. It is very informative to probe the neuron with a battery of sinusoidal gratings covering a range of orientations and frequencies. Such probing in the *frequency (Fourier) domain* complements the local probing in the space domain for simple cells, and is necessary when studying complex cells because they do not respond to local probes. There is abundant evidence that most excitatory neurons in V1 (as well as many other visual cortical areas) are tuned for orientation and spatial frequency (e.g., Albrecht, De Valois, & Thorell, 1980; Watkins & Berkley, 1974; Pollen & Ronner, 1982, 1983, see Section 3.3.3 for further references). Figure 2 illustrates typical empirical² *tuning curves* with respect to orientation and spatial frequency. For example, the tuning curve in Figure 2A has a peak at 0° and FWHH of approximately 45° , which measure this neuron’s *preferred orientation* and the *orientation bandwidth*, respectively.

With a perfectly linear filter (Lathi, 2005), the response profile in the space domain completely determines the tuning in the frequency domain (via the *Fourier transform*) and vice versa (via the *inverse Fourier transform*). Specifically, consider a linear filter whose *weighting function* is a Gabor patch in the space domain (Equation 2). The tuning function of this filter is a bivariate Gaussian in the frequency domain (Graham, 1989, p. 85). Moreover, for a given frequency F , the orientation bandwidth h_θ (in degrees) is inversely proportional to the size $h_{\check{y}}$ of the patch along the direction parallel to the grating:

$$h_{\check{y}} h_\theta F = \frac{720}{\pi^2} \ln 2 = \text{const}, \quad (3)$$

where the constant $720 \ln(2)/\pi^2$ comes from the conversion from degrees to radians and from FWHH to standard

¹A time-varying stimulus such as a drifting grating also has temporal frequency. As the temporal properties are beyond our scope, however, the term “frequency” refers to spatial frequency throughout this article unless explicitly stated otherwise.

²Published physiological data were captured from their respective figures using PlotDigitizer (<http://plotdigitizer.sourceforge.net>) and replotted here in a unified format.

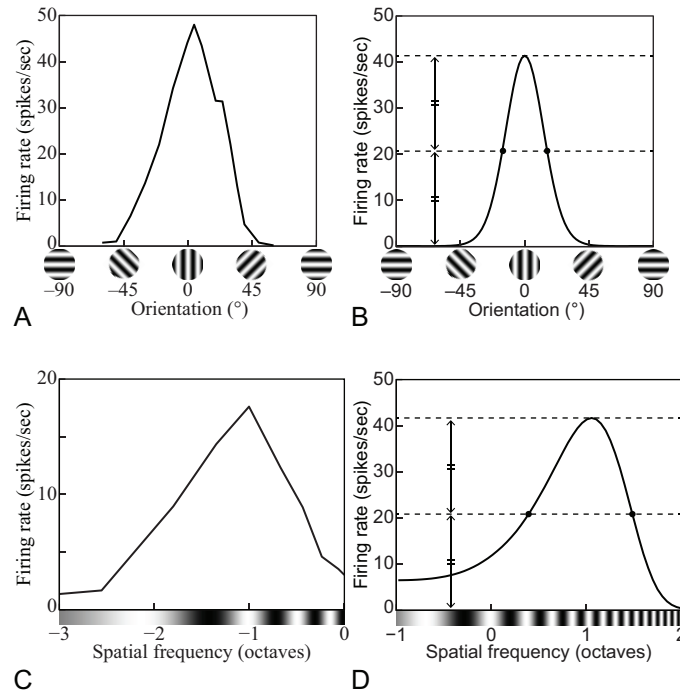


Figure 2: (A) A representative orientation tuning curve—in this case from a V1 complex cell of an anesthetized cat. Replotted from Rose and Blakemore (1974, Figure 1). (B) Orientation tuning curve of the DNM neuron with default parameters, probed with gratings with 100% contrast, 2.88° size, and neuron’s preferred frequency. (C) A representative spatial-frequency tuning curve—in this case from a V1 simple cell of an anesthetized cat. Replotted from Li and Li (1994, Figure 7C). (D) Spatial-frequency tuning curve of the DNM neuron with default parameters, probed with gratings with 100% contrast, 2.88° size, and neuron’s preferred orientation. Dotted lines depict the full- and half-heights of the curves, and the distances between the dots depict the full-widths at half-height. (DNM = divisive normalization model, defined later in the text. See Phenomena 13 and 14 in Table 1 below.)

deviation. See Appendix A for details. The reason for this inverse relationship is intuitively clear from Figure 1: To estimate the stimulus orientation with high precision (low h_θ), it is necessary to have a large baseline for measurement along the length of the bars. An analogous inverse relationship exists between the frequency bandwidth h_f and the perpendicular size h_x . Intuitively, to estimate the stimulus frequency with high precision (low h_f), it is necessary to be able to “count” many cycles within the width of the filter. It is more convenient to express the frequency bandwidth h_f in *octaves* on a \log_2 scale instead of linear units. The exact relationship (derived in Appendix A) is

$$h_x = \frac{(2^{h_f} + 1) 2 \ln 2}{(2^{h_f} - 1) \pi F}. \quad (4)$$

Real neurons are never perfectly linear but the tuning properties of simple and complex cells are in qualitative agreement with the predictions of the linear model. “The two-dimensional tuning curves are mostly moderately elongated along a radial axis [in the Fourier plane], and extreme or amorphous shapes (e.g., sausages, amoebas) are rare” (Lennie & Movshon, 2005, p. 2020). Typically, the preferred orientation does not depend much on the frequency of the test grating, and vice versa (Webster & De Valois, 1985). This means that the joint orientation-by-frequency tuning curve can be modeled as the product of two orthogonal dimension-specific curves (cf. Fig. 2), which is in agreement with the linear prediction.

The output E_S of a linear filter to a stimulus image I is a scalar quantity equal to the dot product³ of the image with the weighting function $G_{XYF\Theta\Phi}$ of the filter:

$$E_{S:XYF\Theta\Phi}(I) = \int I(x, y) G_{XYF\Theta\Phi}(x, y) dx dy. \quad (5)$$

The *weighting function* (*WF*) determines the properties of the filter. All models discussed in this article use Gabor WFs (Eq. 2). The center XY of the Gabor patch determines the center of the receptive field (RF) of the filter (in image space), the frequency F of the patch determines the preferred frequency of the filter (in Fourier space), and analogously for the orientation Θ and phase Φ . Of all images with a given energy (i.e., fixed variance of the intensity distribution), the output is maximized by the stimulus that exactly matches the WF (Cauchy-Schwarz inequality, n.d.). In this sense, the *preferred stimulus* of the filter in Equation 5 is the Gabor patch in Equation 2. The absolute value of the output can be interpreted as the similarity between the stimulus and the preferred template, and the sign indicates whether the two are *in phase* or *out of phase*. (An image is in phase with a WF when the bright spots on the image line up with the excitatory regions of the WF and the dark spots line up with the inhibitory regions; it is out of phase if the alignment is the other way around.)

³ I and $G_{XYF\Theta\Phi}$ are treated as (long) vectors in the computation of the dot product.

The linear model of a simple cell consists of a linear filtering stage (Eq. 5) followed by *halfway rectification* (Eq. 6). The linear stage is motivated by the extracellular recordings surveyed above, as well as by intracellular recordings (e.g., Jagadeesh, Wheat, & Ferster, 1998) suggesting that the fluctuations in membrane potential of simple cells around the resting potential can be modeled quantitatively in terms of linear summation of synaptic potentials. For our purposes, the output of the linear stage is referred to as the *stimulus drive* E_S to the simple cell. The stimulus drive can be positive or negative, but the firing rate of a real neuron is always non-negative. This is modeled by a rectifying nonlinearity

$$R_{LN}(I) = M_{LN} [E_{S:XYF\Theta\Phi}(I)], \quad (6)$$

where R_{LN} is the response of the cell (in spikes/sec, sps) and M_{LN} is a parameter that, under certain calibration assumptions, defines the maximum firing rate to a preferred grating. The *halfway rectification* operator $[E] = \max(0, E)$ passes positive values unchanged and converts negative values to zero.

The weighting function of a linear neuron coincides with its receptive field. Consequently, the terms *WF* and *RF* are sometimes used interchangeably in the literature. We keep them distinct because the corresponding referents are distinct for nonlinear models. *WF* is a theoretical term that is defined only with respect to a model with a linear filtering stage (Eq. 5). By contrast, the *RF* is defined operationally for any (real or simulated) neuron as the region of the visual field where stimulus presentation induces changes in firing rate. The *RF* is smaller than the *WF* in models involving nonlinearity and/or suppression as discussed below.

Complex cells differ from simple cells because of the absence of distinct excitatory and inhibitory subregions in their *RFs* (Hubel & Wiesel, 1962; Watkins & Berkley, 1974). Complex cells are relatively invariant to the phase of the stimuli (Maffei & Fiorentini, 1973; Movshon et al., 1978c; R. L. De Valois, Albrecht, & Thorell, 1982; Sengpiel et al., 1997; Ibbotson, Price, & Crowder, 2005). For example, they respond indiscriminately to light and dark bars, as long as the bar stands out from the gray background. They are sensitive to the stimulus orientation and spatial frequency, however, and their tuning curves are very similar to those of simple cells.

Complex cells are usually modeled in terms of several linear filters whose outputs are nonlinearly transformed and then combined (see Bair, 2005; Martinez & Alonso, 2003; Mechler & Ringach, 2002; Spitzer & Hochstein, 1988, for reviews). “A key feature of these models is that the underlying linear filters—not the later nonlinearities—determine the set of stimuli to which the neuron will respond” (Lennie & Movshon, 2005, p. 2023). The most influential model of this class is the *energy model* (Adelson & Bergen, 1985; Pollen & Ronner, 1983; Spitzer & Hochstein, 1985; Watson & Ahumada, 1985):

$$E_{C:XYF\Theta}(I) = \sqrt{E_{S:XYF\Theta,0^\circ}(I)^2 + E_{S:XYF\Theta,90^\circ}(I)^2}, \quad (7)$$

where $E_{S:XYF\Theta,0^\circ}$ and $E_{S:XYF\Theta,90^\circ}$ are linear filters (Eq. 5) whose WFs have identical parameters $XYF\Theta$ but orthogonal phases. Equation 7 produces phase-invariant output E_C via the trigonometric identity $\sin^2\Phi + \cos^2\Phi = 1$. The firing rate $R_{LN}(I)$ of the complex cell can be modeled by substituting E_C for E_S in Equation 6 (see, e.g., Emerson, Korenberg, & Citron, 1992; Heeger, 1992b; Lehky, Sejnowski, & Desimone, 2005; Szulborski & Palmer, 1990, for similar formulations). It is tempting to interpret Equation 7 as a formalization of the hierarchical feedforward simple-to-complex arrangement proposed by Hubel and Wiesel (1962). The physiological evidence, however, is more consistent with a non-hierarchical interpretation in terms of a continuum from strongly phase-sensitive (simple) cells to nearly phase-invariant (complex) cells (Bair, 2005; R. L. De Valois & De Valois, 1988; Martinez & Alonso, 2003; Mechler & Ringach, 2002). For our present purposes, it suffices to treat Equation 7 as a mathematically convenient functional description of the *stimulus drive* E_C to complex cells.

As a notational convenience, it is often useful to bundle the long subscripts into a *tuning preference vector*

$$\mathbf{P}_i = \langle X_i Y_i F_i \Theta_i \Phi_i \rangle, \quad (8)$$

where the phase Φ_i can take a special non-numerical value to indicate a phase-invariant (complex) cell. In this notation, an image I induces stimulus drive $E_{\mathbf{P}_i}(I)$ to a neuron with index i and preferences \mathbf{P}_i :

$$E_{\mathbf{P}_i}(I) = \begin{cases} E_{S:X_i Y_i F_i \Theta_i \Phi_i}(I) & \text{for a simple cell,} \\ E_{C:X_i Y_i F_i \Theta_i}(I) & \text{for a complex cell.} \end{cases} \quad (9)$$

Note that $E_{\mathbf{P}_i}$ is a linear operator for the simple cell. That is, for a fixed i , $E_{\mathbf{P}_i}(I' + I'') = E_{\mathbf{P}_i}(I') + E_{\mathbf{P}_i}(I'')$ for any pair of images I' and I'' , and $E_{\mathbf{P}_i}(cI) = cE_{\mathbf{P}_i}(I)$ for any number $c \geq 0$. The latter equality also holds for the complex cell.

The linear model $R_{LN}(I) = M_{LN}[E_{\mathbf{P}_i}(I)]$ provides a quantitative account of the selectivity of V1 neurons' responses to bars, edges, and gratings. Moreover, it provides "a credible account of the responses to a variety of more complicated targets, including checkerboards (K. K. De Valois, De Valois, & Yund, 1979), random dot textures and Glass patterns (Smith, Bair, & Movshon, 2004)" (Rust & Movshon, 2005, p. 1647). It dominated the field until the mid-1980s and set the stage for subsequent research that uncovered systematic departures from linearity, to which we now turn.

2.2 Hyperbolic-ratio model

The *contrast response function* (CF⁴) describes how a neuron’s response depends on the contrast of the stimulus. Consider a family of stimuli $\{cI_T\}_{c \in [0,1]}$ based on the same template image I_T but varying in contrast c . We assume throughout this section that the template I_T is normalized so that its contrast is 1.0 (cf. Eq. 1). The CFs of neurons have been measured for various templates, cortical areas, species, and conditions (e.g., Albrecht & Geisler, 1991; Albrecht et al., 2002; Albrecht & Hamilton, 1982; Carandini, Heeger, & Movshon, 1997; Dean, 1981; Derrington & Lennie, 1984; Geisler & Albrecht, 1997; C.-Y. Li & Creutzfeldt, 1984; Sclar, Maunsell, & Lennie, 1990; see Albrecht et al., 2003; Carandini et al., 1999; Graham, 2011; Heeger, 1992a; Lennie & Movshon, 2005, for reviews). Further more, Albrecht, Geisler, Frazor, and Crane (2002) pointed out that “Although there is a great deal of heterogeneity from cell to cell, it is possible to provide a description of the basic properties of the contrast response function that applies to the overwhelming majority of neurons: As the contrast increases from zero, the response increases in an accelerating fashion, remains dynamic over some limited range of contrasts, and then saturates” (p. 888). Figure 3 illustrates the characteristic sigmoidal shape characteristic of a typical CF.

The *hyperbolic-ratio model* has been widely used as a descriptive model to fit these data (e.g., Albrecht & Hamilton, 1982, see also⁵ Naka & Rushton, 1966; Legge & Foley, 1980; Graham, 2011). The response R_{HB} of this model to a stimulus cI_T with contrast c and template I_T is

$$R_{HB}(cI_T) = M_{HB}(I_T) \frac{c^{n_{HB}}}{\alpha_{HB}^{n_{HB}} + c^{n_{HB}}}, \quad (10)$$

where the *semi-saturation contrast* parameter α_{HB} expresses the contrast of the image that produces one-half of the saturation level $M_{HB}(I_T)$. Note that the latter depends on the template I_T . When contrast is plotted on a log axis, the exponent n_{HB} controls the slope of the CF and α_{HB} controls its location (Figure 4, see Appendix B for properties of the CF plotted on a linear contrast axis). Note also that the response to a stimulus with zero contrast (i.e., a uniform gray field) is assumed to be zero.

The hyperbolic-ratio model (Eq. 10) involves two nonlinear operations: exponentiation and division. The former accounts for the accelerating shape of the CF at low contrasts—that is, for the fact that the CF slope gets steeper and steeper as the contrast increases from zero. Using the logarithmic scale of the contrast c , the maximal slope is $n_{HB}/4$ at $c = \alpha_{HB}$. The divisive operation of the hyperbolic-ratio model saturates the CF at high contrasts. The CF of the hyperbolic-ratio model with these properties can represent shapes of CFs of real

⁴The abbreviation CRF should be to be avoided because it tends to be misread as a “classical receptive field.”

⁵The function $R(c) = c^n / (\alpha^n + c^n)$ is sometimes called the *Naka-Rushton* function. In biochemistry, it is called the *Michaelis-Menten* function and is the centerpiece of a standard model of enzyme kinetics.

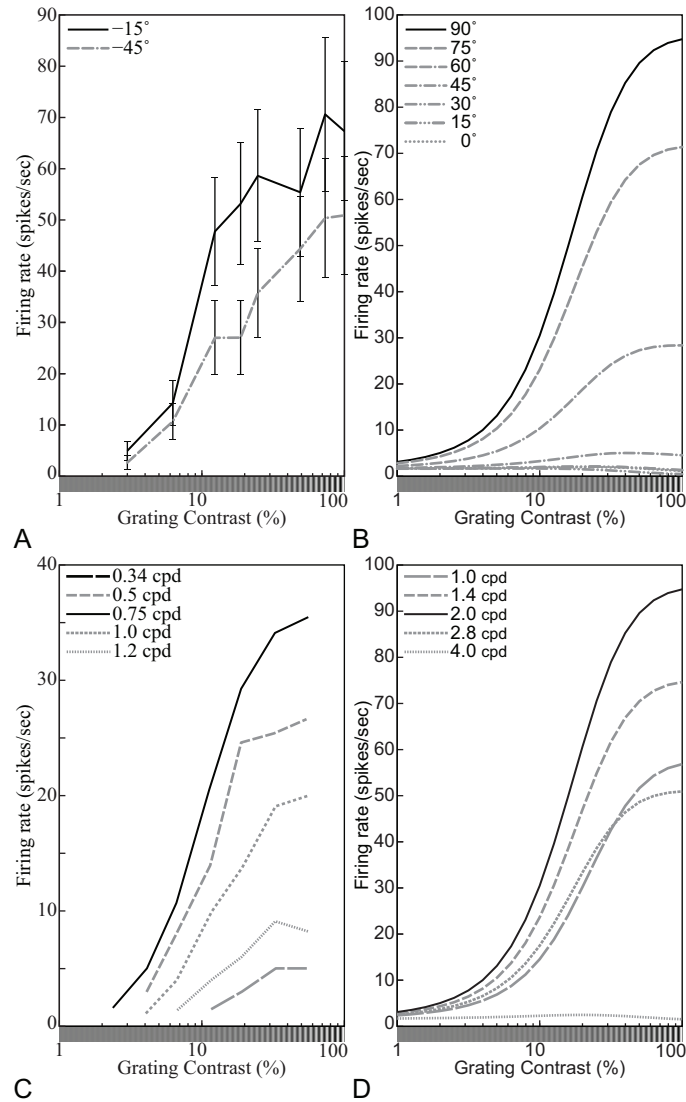


Figure 3: Representative contrast response functions (CFs). (A) Responses of a simple cell to drifting sinusoidal gratings spanning a range of contrasts at two orientations (see legend). Replotted from Carandini, Heeger, and Movshon (1997, Figure 4B, anesthetized macaque, error bars = \pm SE). (B) CFs of the DNM neuron with default parameters, probed with gratings with the neuron's preferred frequency (1.0 octave) and orientations shown in the legend. The size of the stimuli was equal to the measured RF diameter (0.81°). (C) Responses of a V1 neuron to drifting sinusoidal gratings with the neuron's preferred orientation and spatial frequencies shown in the legend. Replotted from Albrecht and Hamilton (1982, Figure 7A). (D) CFs of the DNM neuron with default parameters, probed with gratings with the neuron's preferred orientation (0°) and spatial frequencies shown in the legend. The size of the stimuli was equal to the measured RF diameter (0.81°). (DNM = divisive normalization model, defined later in the text. See Phenomena 7, 10, and 11 in Table 1 below.)

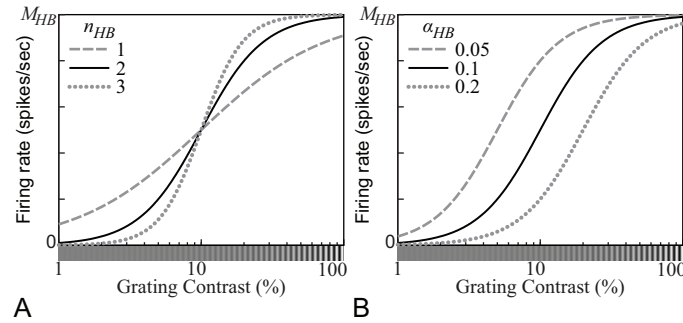


Figure 4: Contrast response functions (CFs) produced by the hyperbolic-ratio model (Equation 10). (A) The exponent parameter n_{HB} controls the slope of the CF as a function of the log-contrast of the stimulus. (B) The *semisaturation contrast* parameter α_{HB} controls the location of the CF. A stimulus with contrast α_{HB} elicits one-half of the saturation level M_{HB} . ($\alpha_{HB} = 0.1$ for panel A; $n_{HB} = 2$ for panel B.)

neurons well.

Equation 10 predicts that all CFs measured for the same neuron are multiplicatively scaled replicas of each other across the entire contrast range. That is, the ratio of the neuron’s responses to images with identical contrast but different templates $I_{T'}$ and $I_{T''}$ is invariant with respect to c :

$$\frac{R_{HB}(cI_{T'})}{R_{HB}(cI_{T''})} = \frac{M_{HB}(I_{T'})}{M_{HB}(I_{T''})}. \quad (11)$$

The empirical CFs (Fig. 3) of typical simple and complex cells in V1 are consistent with this prediction in many cases (e.g., Albrecht & Hamilton, 1982; Carandini et al., 1997; see Section 3.3.2 for further references). In particular, when $I_{T'}$ and $I_{T''}$ are gratings with different orientations, Equation 11 accounts for the approximate contrast invariance of the orientation tuning curves of typical V1 neurons (e.g., Sclar & Freeman, 1982; Skottun, Bradley, Sclar, Ohzawa, & Freeman, 1987; see Section 3.3.3 for further references). Note, however, that the spatial-frequency tuning curves of many V1 neurons have a slight but systematic dependence on contrast (e.g., Albrecht & Hamilton, 1982; Skottun et al., 1987; see Section 3.3.3 for further references and discussion).

Cortical neurons have a limited dynamic range and their firing rates saturate at high contrasts (e.g., Albrecht & Hamilton, 1982). It is important to note that this saturation is not simply an output nonlinearity because it occurs at a fixed contrast rather than at a fixed response level for different stimuli. Consider the CFs plotted on Figure 3C, for example. The responses to gratings whose frequency is 0.34 cpd saturate below 10 sps even though the same neuron can sustain firing rates above 30 sps when stimulated at its preferred frequency (0.75 cpd). Each neuron is characterized by an entire *family* of CFs—one for each stimulus template. This is why the saturation level M_{HB} depends on I_T in Equation 10. However, M_{HB} only acts as a scaling factor and all CFs within the family have the same shape. This shape only depends on c . It is for this reason that this type of divisive normalization is called a *contrast-set gain control* (or simply a *contrast gain control*).

The contrast gain control is a cortical mechanism and it should not be confused with the luminance gain control (Eq. 1) found in the retina. Because of the luminance gain control, the maximal contrast is $c_{max} = 1$ (or locally slightly greater for sharp-edged patches embedded within a larger field). The semisaturation contrast parameter in Equation 10 is in light-adapted units and should be set to a relatively small value ($\alpha_{HB} \ll 1$) so that the ratio $c^{n_{HB}} / (\alpha_{HB}^{n_{HB}} + c^{n_{HB}})$ is close to saturation for $c \approx 1$. The asymptotic limit $c \rightarrow \infty$ has no physiological interpretation.

The hyperbolic-ratio model requires the extraction of two distinct pieces of information about the stimulus. One is the contrast c , which is an intrinsic property of the image. The other is the degree of match between the input template I_T and the weighting function of the neuron. This match is a relational property that depends not only on the image but also on the WF. Both pieces of information are available in the image but they cannot be extracted by the application of a single filter. The stimulus drive extracted by a linear filter (Eq. 5) is a single number that confounds contrast information and degree-of-match information. If either of them is known, the other can be decoded from the stimulus drive. The two models discussed so far are complementary in this regard: The linear model characterizes stimulus selectivity, whereas the hyperbolic-ratio model describes dependence on contrast. When an arbitrary stimulus is presented, however, both pieces of information are unknown and multiple filters must be applied to resolve this ambiguity. The divisive normalization model pools the (half-squared) outputs of filters with diverse tuning preferences in order to estimate the intrinsic properties of the image. This pooled estimate is then used to normalize the stimulus drives to the individual units.

2.3 Divisive Normalization Model

The notion of a *normalization pool* is pivotal to the *divisive normalization (DNM)* model and sets it apart from the simpler models discussed above. The introduction of a normalization pool is motivated by three convergent lines of evidence. The first line comes from the experimental data on the contrast-set gain control outlined above (Fig. 3), coupled with the need to pool across filters with diverse tuning preferences to estimate the stimulus contrast. A second, related line comes from a priori considerations involving the so-called *noise-saturation dilemma* (Grossberg, 1988). Individual neuronal responses are noisy and they have a limited dynamic range. The brain needs to represent signals across very wide dynamic ranges. Hence the dilemma: “If the [activations of individual neurons] are sensitive to large inputs, then why do not small inputs get lost in internal system noise? If the [activations] are sensitive to small inputs, then why do they not all saturate at their maximum values in response to large inputs?” (Grossberg, 1988, p. 33). The proposed solution relies on pooled inhibition⁶ within a network of interacting neurons to normalize the individual responses relative to a dynamically adjustable

⁶The corresponding term in Grossberg (1973, 1988; Ellias & Grossberg, 1975) is *mass action*.

baseline. Such divisive normalization has been proposed (e.g., Carandini & Heeger, 2012) as a canonical type of neural computation in a wide variety of sensory modalities, brain regions, and species. The divisive nonlinearity in the hyperbolic-ratio model (Eq. 10) is an instantiation of this general principle.

The third and most direct line of evidence motivating the introduction of a normalization pool consists of experimental demonstrations of various broadly tuned suppressive effects in V1. In *cross-orientation suppression*, for example, the response to a grating (signal) is suppressed by another grating (mask) superimposed onto the signal within the neuron’s receptive field (e.g., DeAngelis, Robson, Ohzawa, & Freeman, 1992; Morrone, Burr, & Maffei, 1982; see Section 3.3.4 for further references and discussion). In *surround suppression*, the response is suppressed by masks presented outside the classical RF (e.g., Cavanaugh, Bair, & Movshon, 2002b; C.-Y. Li & Li, 1994; see Section 3.3.5).

Thus, modeling the responses of a single individual neuron requires filtering the input image with multiple linear filters that have diverse weighting functions. Let this diverse set be indexed by i , the filters in the normalization pool \mathcal{N} have preferences \mathbf{P}_i , and \mathbf{P}^* denote the tuning preference vector (Eq. 8) of the target neuron.

Heeger (1992b) combined all theoretical ideas introduced above into a single equation. His formulation of the divisive normalization model has been very influential (see Carandini & Heeger, 2012, for a recent review). In our notation, which is different from Heeger’s, this equation is

$$R_{DN}(I) = M_{DN} \frac{[E_{\mathbf{P}^*}(I)]^{n_{DN}}}{\alpha_{DN}^{n_{DN}} + \sum_i E_{\mathbf{P}_i}(I)^{n_{DN}}}, \quad (12)$$

where the summation in the denominator encompasses the normalization pool ($i \in \mathcal{N}$). The exponent n_{DN} and the semi-saturation contrast α_{DN} are free parameters analogous to their counterparts in the hyperbolic-ratio Equation 10 (Fig. 4). Note that the firing-rate parameter M_{DN} is a constant that does not depend on stimulus I . This parameter determines the stimulus drive $E_{\mathbf{P}^*}(I)$ to the target neuron, which can be either a simple or a complex cell (Eq. 9).

The *suppressive drive* $\sum_i E_{\mathbf{P}_i}(I)^{n_{DN}}$ in the denominator represents the aggregated inhibitory influence impinging on the target neuron. There is evidence that this inhibitory influence combines lateral inhibition from other neurons in V1, feedforward inhibition from and within the lateral geniculate nucleus (LGN), and feedback inhibition from higher cortical areas (e.g., Sengpiel, Baddeley, Freeman, Harrad, & Blakemore, 1998; Angelucci & Shushruth, 2013; see Section 4.3 for further references and a brief discussion). These sources have different temporal and spatial properties (e.g., Bair, Cavanaugh, & Movshon, 2003; see Sect. 4.3). This is an active research area that is beyond our present scope. Equation 12 models this suppressive drive as a sum of (exponentiated) homogenous terms $E_{\mathbf{P}_i}$, each of which is analogous to the stimulus drive $E_{\mathbf{P}^*}$ in the numerator.

It is important to emphasize that this formula is a *functional* description only. It entails no commitment about what neurophysiological mechanisms produce the suppressive effect. The formula itself suggests a three-stage sequence of linear filtering followed by exponentiation followed by divisive normalization, and this is indeed how Equation 12 is implemented on a computer. This sequential scheme, however, is not physiologically possible because the un-normalized intermediate terms $E_{\mathbf{P}_i}$ cannot be represented by substrates with a limited dynamic range such as membrane potentials or firing rates. This constraint is at the core of Grossberg’s (1988) noise-saturation dilemma. Instead, the normalization almost certainly involves dynamic inhibitory interactions within a recurrent network (e.g., Brosch & Neumann, 2014; Chance & Abbott, 2000; Ellias & Grossberg, 1975; Heeger, 1993; Kough & Poggio, 2008) in conjunction with other regulatory mechanisms (e.g., Carandini, Heeger, & Senn, 2002; Freeman, Durand, Kiper, & Carandini, 2002).

Various variants of Equation 12 have been used to account successfully for extracellular recordings (e.g., Heeger, 1992b; Carandini & Heeger, 2012), multi-electrode population recordings (e.g., Busse et al., 2009; Goris et al., 2009), and psychophysical data (e.g., Itti et al., 2000; Olzak & Thomas, 1999, 2003; To et al., 2010; Jesús & Laparra, 2010). Unfortunately, many of these applications use different mathematical formulations and idiosyncratic parameterizations. This practice makes it difficult to compare the results across studies despite the clear family resemblance of the model variants.

With the aim of consolidating this scattered literature, we propose that Equation 13 is used as the standard formulation of the divisive normalization model. *This is the main equation in the present article.* It is used in all of the simulations reported below. It was chosen on the basis of a theoretical analysis and the results of simulation experiments with several model variants that were compared informally on their ability to account for a comprehensive suite of empirical phenomena that are listed in Table 1 below. In our opinion, the following formulation achieves a good balance between flexibility and parsimony:

$$R(I) = M \frac{[\beta + k_n E_{\mathbf{P}^*}(I)]^{n_n}}{\alpha^{n_d} + k_d \sum_i w_i E_{\mathbf{P}_i}(I)^{n_d}}. \quad (13)$$

The firing-rate parameter M and the semisaturation contrast α have the same interpretation as their counterparts in Heeger’s (1992b) proposal (Eq. 12, Fig. 4). There are separate exponents n_n and n_d for the numerator and denominator, respectively. The *baseline* parameter β allows for non-zero responses when the stimulus drive $E_{\mathbf{P}^*}(I)$ is zero. The *maintained discharge* of the DNM neuron—its response to a blank stimulus (uniform gray field)—is $M[\beta]^{n_n}/\alpha^{n_d}$ sps. Note that the firing rate of an actual simple cell in V1 can be less than its maintained discharge when the inhibitory regions in the cell’s receptive field are stimulated by a light spot (Hubel & Wiesel, 1959). This property can be modeled by Equation 13, assuming $\beta > 0$, but not by Equation 12. Conversely, setting $\beta < 0$ within the scope of the half-wave rectification operator $[\cdot]$ amounts to setting a threshold

on the stimulus drive (the *iceberg effect*, Heeger, 1992a; Sceniak, Hawken, & Shapley, 2001; Tadmor & Tolhurst, 1989). The constraint that β must be fixed for a given neuron entails falsifiable predictions for the model (see Section 3.3.2 and Appendix C).

The *calibration constants* k_n and k_d are not free parameters. Conceptually, they are factored into the weights w_i in Equation 13 and the weighting function of the linear filter $E_{\mathbf{P}^*}$ (Eq. 5). Both constants are defined with respect to a single *calibration image* I_{cal} :

$$\left| \begin{array}{l} k_n = (E_{\mathbf{P}^*}(I_{cal}))^{-1} \\ k_d = (\sum_i w_i E_{\mathbf{P}_i}(I_{cal}))^{-1}. \end{array} \right. \quad (14)$$

The calibration image is chosen a priori as the grating whose frequency and orientation (and phase for a simple cell) match the preferences \mathbf{P}^* of the target neuron. The contrast of I_{cal} is 1 and its spatial extent is large enough to fill both the classical receptive field and the suppressive surround. Calibrating the model in terms of an explicit image is convenient because it establishes a standardized scale for the substantive parameters α and β —they can be interpreted as equivalent contrasts. To see why, consider the *calibration family* of gratings $\{cI_{cal}\}_{c \in [0,1]}$ that use I_{cal} as a template and sweep a range of contrasts c . For this special family, Equations 13 and 14 reduce to the following variant of the hyperbolic-ratio model (Eq. 10, see also Appendix C):

$$R(cI_{cal}) = M \frac{[\beta + c]^{n_n}}{\alpha^{n_d} + c^{n_d}}. \quad (15)$$

Then a positive baseline ($\beta > 0$) can be interpreted as the contrast of the counter-phase grating that cancels the maintained discharge of a simple cell, and a negative baseline ($\beta < 0$) as the contrast of the preferred (in-phase) grating that barely elicits a response.

To complete the specification of the divisive normalization model, we need to define the suppressive drive $\sum_i w_i E_{\mathbf{P}_i}(I)^{n_d}$ in the denominator of Equation 13. We need to specify three things: the weighting functions of the linear filters $E_{\mathbf{P}_i}$, the composition of the normalization pool \mathcal{N} encompassed by the sum, and the pooling weights w_i . We followed the common practice (e.g., Itti et al., 2000; Reynolds & Heeger, 2009) with respect to all three. First, we assume all linear filters have the same spatial-frequency bandwidth h_f (Eq. 4) and the same orientation bandwidth h_θ (Eq. 3). These h_f and h_θ are assumed for the stimulus drive $E_{\mathbf{P}^*}$ as well. Second, we assume the normalization pool tiles the space of frequencies and orientations. There is evidence (DeAngelis et al., 1992; DeAngelis, Freeman, & Ohzawa, 1994) that the suppressive effects are approximately invariant with respect to the phase of the mask grating. We model this by including only phase-invariant components E_C (Eq. 7) into \mathcal{N} . Third, we adopt the common simplifying assumption that the pooling weights w_i can be

separated⁷ into independent pooling kernels with respect to space, frequency, and orientation:

$$\sum_{i \in \mathcal{N}} w_i E_{\mathbf{P}_i}(I)^{n_d} = \sum_{X_i Y_i \in \mathcal{N}_{XY}} \sum_{F_i \in \mathcal{N}_F} \sum_{\Theta_i \in \mathcal{N}_\Theta} w_{X_i Y_i} w_{F_i} w_{\Theta_i} E_{C: X_i Y_i F_i \Theta_i}(I)^{n_d}, \quad (16)$$

where \mathcal{N}_{XY} is a grid of image locations, \mathcal{N}_F is a set of *frequency channels*, and \mathcal{N}_Θ is a set of *orientation channels*. This specification is constrained by empirical data on various forms of suppression (surveyed in Section 3.3) and by considerations of symmetry, parsimony, and computational efficiency.

The spatial pooling weights $w_{X_i Y_i}$ are defined by a radially symmetric 2D Gaussian kernel

$$w_{X_i Y_i} \propto \exp \left[-4 \ln 2 \frac{(X_i - X^*)^2 + (Y_i - Y^*)^2}{(h_R/F_i)^2} \right], \quad (17)$$

where X^* and Y^* are the coordinates of the center of the receptive field of the target neuron. The weights are defined up to a scaling factor and then calibrated by k_d in Equation 14. The diameter at half height of the kernel is proportional to the wavelength $1/F_i$ of each frequency channel. The *spatial pooling bandwidth* h_R (in number of cycles) is a free parameter common to all channels. Note that the overall *suppressive field* of the model is produced by a combination of two types of spatial summation. First, the individual components $E_{\mathbf{P}_i}$ perform summation within elliptical Gabor receptive fields whose spatial dimensions h_x and h_y also are proportional to the respective channel wavelengths (Eqs. 3 and 4). Second, after nonlinear rectification with exponent n_d , there is another summation across components centered on multiple locations $X_i Y_i$. Equation 17 defines the weighting function of the latter summation.

The frequency pooling weights w_{F_i} are defined by a Gaussian kernel along the log-frequency (octave) dimension:

$$w_{F_i} \propto \exp \left[-4 \ln 2 \frac{(\log_2 F_i - \log_2 F^*)^2}{h_F^2} \right], \quad (18)$$

where F^* is the preferred frequency of the target neuron. The *frequency pooling bandwidth* h_F (in octaves) controls the full width at half height of the kernel. This parameter is distinct from the bandwidth h_f of the weighting functions of the individual components $E_{\mathbf{P}_i}$. Because of the pooling, the frequency tuning of the suppressive effects is broader than h_f , in agreement with the data (DeAngelis et al., 1994; C.-Y. Li & Li, 1994).

Finally, the orientation pooling weights w_{Θ_i} are defined by a von Mises kernel

$$w_{\Theta_i} \propto \exp [\kappa_\Theta \cos 2(\Theta_i - \Theta^*)], \quad (19)$$

⁷This assumption ignores non-separable effects such as end stopping and contour integration, which require coordination among components with aligned orientations at neighboring locations (Kapadia, Ito, Gilbert, & Westheimer, 1995; Z. Li, 1998).

where Θ^* is the preferred orientation of the target neuron. The von Mises distribution is the circular analog of the normal distribution (Fisher, 1996). The dimensionless *concentration parameter* κ_Θ is intuitively similar to inverse variance. Equation 19 assigns $w_{max} \propto e^{\kappa_\Theta}$ to the preferred orientation and $w_{min} \propto e^{-\kappa_\Theta}$ to the orthogonal orientation. The two points whose height is halfway between these two extremes occur at orientations $\Theta^* \pm h_\Theta/2$, where

$$\cos(h_\Theta) = \frac{1}{\kappa_\Theta} \ln \left[\frac{e^{\kappa_\Theta} + e^{-\kappa_\Theta}}{2} \right]. \quad (20)$$

This equation establishes an invertible relationship in which h_Θ monotonically increases as κ_Θ decreases. The circular uniform distribution $w_{\Theta_i} \propto 1$ is the special case for $\kappa_\Theta = 0$, $h_\Theta = 90^\circ$. It is convenient to parameterize the DNM in terms of its *orientation pooling bandwidth* h_Θ (in degrees). Again, this parameter is distinct from the bandwidth h_θ of the weighting functions of the individual components $E_{\mathbf{P}_i}$. Because of the pooling, the tuning of cross-orientation suppression is broader than h_θ , in agreement with the data (DeAngelis et al., 1992; Morrone et al., 1982). Unfortunately, κ_Θ cannot be expressed as a closed-form function of h_Θ , but in practice, Equation 20 is easy to solve numerically.

Overall, our formulation of the divisive normalization model has ten free parameters: M , α , β , n_n , and n_d in Equation 13, the tuning bandwidths h_f and h_θ of the linear filters (Eq. 5), and the pooling bandwidths h_R , h_F , and h_Θ of the suppressive drive (Eq. 16). Five auxiliary constants ($h_{\tilde{x}}$, $h_{\tilde{y}}$, κ_Θ , k_n , and k_d) are calculated from the free parameters (cf. Table 2 below).

3 Simulation Experiments

The divisive normalization model (DNM, Eq. 13) was implemented as software and tested on a wide range of stimuli designed to replicate a comprehensive suite of published neurophysiological studies (Table 1).

3.1 Computational Implementation and Calibration

We developed software for Matlab (The MathWorks, 2015). This software takes a static grayscale image as input and produces a matrix of firing-rate responses for a population of DNM neurons. The neurons in this population have receptive fields at a single retinal location and they are tuned for a range of orientations and spatial frequencies. The details of this implementation are given in Appendix E. Briefly, the software provides tools for accomplishing two main computational tasks: constructing a DNM object for a given parameter set and calculating the responses of a model for a given input image.

A DNM object is a data structure that encapsulates the model’s parameters, weighting functions (WFs) for all linear filters, pooling weights for the suppressive drive, and various auxiliary information. Our implementa-

Table 1: Phenomena accounted for by the divisive normalization model (Equation 13) in our simulation experiments. RF = receptive field, CF = contrast response function.

N.	Phenomenon	Figure
1.	Size tuning: The receptive field (RF) has limited spatial extent	5
2.	The measured RF diameter increases as the grating contrast decreases	6
3.	The measured RF diameter decreases for non-preferred orientations	7
4.	The measured RF diameter decreases for non-preferred spatial frequencies	8
5.	The response decreases monotonically as the hole in an annular grating increases	9
6.	The response relationship with hole size is nearly invariant to stimulus contrast	9
7.	The contrast response function (CF) has a characteristic sigmoidal shape	3
8.	Supersaturation effect: The CF can slope downwards at very high contrasts	10
9.	The CF is affected by visual noise added to the grating stimulus	11
10.	The CF is scaled down for gratings with non-preferred orientations	3
11.	The CF is scaled down for gratings with non-preferred spatial frequencies	3
12.	The CF is affected by the size of the grating patch	12
13.	Orientation tuning: The response is maximal for the preferred orientation	2
14.	Spatial-frequency tuning: The response is maximal for the preferred frequency	2
15.	Orientation bandwidths become narrower as the size of grating patch increases	13
16.	Frequency bandwidths become narrower as the size of grating patch increases	13
17.	Orientation bandwidths tend to be invariant with respect to the grating contrast	14
18.	Frequency bandwidths tend to become narrower as the contrast decreases	14
19.	The spatial-frequency tuning function has a secondary peak for square gratings	15
20.	The CF shifts leftward for square gratings compared to sinusoidal ones	15
21.	Cross-orientation suppression: orientation tuning of the mask grating	16
22.	Cross-orientation suppression: spatial-frequency tuning of the mask grating	16
23.	Cross-orientation suppression: the mask contrast affects the CF	17
24.	Surround suppression: orientation tuning of the surrounding grating	18
25.	Surround suppression: spatial-frequency tuning of the surrounding grating	18
26.	Surround suppression: the contrast of an annular grating affects the CF	19
27.	Surround suppression: the orientation of an annular grating affects the CF	20
28.	Mapping the receptive field of a simple cell with a light spot	21
29.	Mapping the RF of a simple cell using the reverse correlation method	22
30.	Comparison of the mapped RF of a simple cell and its spatial-frequency tuning	23

tion used $N_{\Theta} = 12$ orientation channels spaced evenly at 15° increments around a circle. It used the following set of spatial-frequency channels: $\mathcal{N}_F = \{\sqrt{2}/2, 1, \sqrt{2}, 2, 2\sqrt{2}, 4, 4\sqrt{2}\}$ cyc/deg. The first and last channels in this set were auxiliary. They were included to contribute to the normalization of the five main channels in the middle. The WF of each channel was a Gabor function (Eq. 2). Two WFs were constructed per channel: one in sine ($\Phi = 90^{\circ}$) and one in cosine ($\Phi = 0^{\circ}$) phase (Fig. 1), for a total of $168 = 12 \times 7 \times 2$ Gabor patches. The sizes $h_{\tilde{x}}$ and $h_{\tilde{y}}$ of their Gaussian envelopes were calculated from the free parameters h_f and h_{θ} according to Equations 4 and 3. All images were rendered on a square 128×128 grid subtending 5.76 deg of visual angle. To improve efficiency, we also used a “small” 64×64 grid for some simulations that did not involve stimuli with extensive surrounds.

The calibration constants k_n and k_d were calculated according to Equation 14 with the aid of a calibration

Table 2: Parameters of the divisive normalization model (Equation 13). The values for the ten free parameters are compatible with typical neurophysiological measurements of representative simple and complex cells. Moreover, the model under this *standard parameterization* accounts qualitatively for the phenomena in Table 1. All simulation results in this article were produced with these parameter values unless explicitly stated otherwise. FWHH = full width at half height, WF = weighting function, cyc = wave cycles, oct = octaves = $\log[\text{cyc}/\text{deg}]$, Eq = Equation.

Parameter [units]	Eq.	Value
<i>Free parameters</i>		
Firing rate [spikes/sec]	13	$M = 40$
Semisaturation contrast [dimensionless]	13	$\alpha = 0.1$
Baseline (a.k.a. maintained discharge) [dimensionless]	13	$\beta = 0.02$
Stimulus-drive exponent	13	$n_n = 2$
Suppressive-drive exponent	13	$n_d = 2$
Orientation FWHH bandwidth of the Gabor WF [deg]	2, 3	$h_\theta = 40$
Spatial-frequency FWHH bandwidth of the Gabor WF [oct]	2, 4	$h_f = 1.5$
FWHH of the radial spatial-pooling kernel [cyc]	17	$h_R = 2.0$
Orientation pooling FWHH bandwidth [deg]	19, 20	$h_\Theta = 60$
Spatial-frequency pooling FWHH bandwidth [oct]	18	$h_F = 2.0$
<i>Constants calculated from the free parameters</i>		
FWHH size of the Gabor WF, perpendicular to grating [deg/cyc]	2, 4	$h_{\tilde{x}} \approx 0.92$
FWHH size of the Gabor WF, parallel to grating [deg/cyc]	2, 3	$h_{\tilde{y}} \approx 1.26$
Concentration parameter of the orientation pooling kernel	19, 20	$\kappa_\Theta \approx 1.22$
Stimulus-drive calibration constant [implementation-dependent]	13, 14	$k_n = 0.25$
Suppressive-drive calibration constant [implementation-dependent]	13, 14	$k_d \approx 0.011$
<i>Implementation specifications chosen a priori</i>		
Number of orientation channels, evenly spaced around the circle	16	$N_\Theta = 12$
Number of (main + auxiliary) spatial-frequency channels	16	$N_F = 5 + 2$
Spacing of the frequency channels [oct]		$\delta_F = 0.5$
Preferred orientation of the target neuron [deg]	14	$\Theta^* = 0$
Preferred frequency of the target neuron [cyc/deg]	14	$F^* = 2.0$
Size of a “large” (128×128) input image [deg]	5	5.76
Size of a “small” (64×64) input image [deg]	5	2.88

grating I_{cal} with unit contrast, vertical orientation ($\Theta^* = 0^\circ$), frequency $F^* = 2.0$ cyc/deg, phase $\Phi^* = 0^\circ$, and spatial extent covering the entire grid.

The calibrated model can be applied to an arbitrary grayscale input image. The computationally expensive operation is the calculation of the suppressive drive in Equation 16. The stimulus is convolved with each of the 168 Gabor filters to produce 84 phase-invariant suppressive terms $E_{\mathbf{P}_i}$ in Equation 13. The software uses the FFT algorithm (Fast Fourier Transform, Lathi, 2005) to compute these convolutions efficiently. See Appendix E for details. The suppressive drives of multiple simulated neurons can be computed as weighted linear combinations of the same $E_{\mathbf{P}_i}$ with different pooling kernels $w_{X_i Y_i}$, w_{F_i} , and w_{Θ_i} centered on different orientations and frequencies (Eqs. 16–19). Our software calculates the responses of 60 DNM complex cells and 240 DNM simple cells. The former are phase invariant (Eq. 7) and span the 12 orientations \times 5 frequencies

listed above (excluding the auxiliary frequencies). The latter vary also in phase: $\Phi = 0^\circ, 90^\circ, 180^\circ, \text{ and } 270^\circ$.

The channel with $\Theta^* = 0^\circ$ and $F^* = 2.0 \text{ cyc/deg}$ is singled out as the *target* and used to generate the DNM predictions in most figures below. Note, however, that it is only marginally more expensive to compute an entire *population code* of the input image (deCharms & Zador, 2000; Pouget, Dayan, & Zemel, 2003). Our software can be used as off-the-shelf front-end to larger models utilizing population codes. In fact, an earlier version of the software has been incorporated into such larger models (Petrov et al., 2005, 2006; Jacobs, 2009).

3.2 Standard Parameter Set

We propose the values listed in Table 2 as a standard parameterization of the DNM. These values were used to generate almost all of the DNM results in this article, with a few exceptions noted explicitly below. They are compatible with typical neurophysiological measurements of representative simple and complex cells, and with the phenomena in Table 1. Thus, these values are good defaults when the DNM is used as a building block for the construction of larger models of the visual system. The explicit reliance on a calibration image (Eq. 14) is designed to facilitate the reusability of parameter values across multiple applications.

A limitation inherent in the notion of standard parameterization must be acknowledged: Real V1 neurons have diverse properties that are impossible to subsume under a single parameter set. We reviewed single-cell recordings from different species (e.g., cats, new/old world monkeys, rabbits, rodents, and ferrets) obtained under different conditions (e.g., anesthesia vs. alertness, see Bereshpolova et al., 2011; Chen, Martinez-conde, Macknik, Swadlow, & Alonso, 2009; Disney, Aoki, & Hawken, 2007; Ecker et al., 2014; Goltstein, Montijn, & Pennartz, 2015; Niell & Stryker, 2010) and different experimental protocols (see Mukherjee & Kaplan, 1995; Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003, for examples). Furthermore, there is substantial variability within a sample of neurons recorded from a single animal under constant conditions. Clearly, model parameters need to vary substantially to accommodate this diversity. The standard set in Table 2 is proposed as an estimate of the central tendency of a broad distribution. Our goal is to produce a *qualitative* account of the phenomena in Table 1, rather than a *quantitative* fit of a specific data set.

Now, let us discuss briefly the ten free parameters in Table 2. The firing-rate parameter M converts the dimensionless ratio of Equation 13 into physiologically observable units (spikes/sec). This parameter plays no role in accounting for the qualitative patterns that are our focus here, but is indispensable for quantitative fitting of actual neuronal firing rates. The semi-saturation contrast α was discussed in Section 2.2 and illustrated in Figure 4B. Under the DNM calibration (Eq. 13), it is expressed in dimensionless units and can be interpreted as equivalent contrast. The proposed value ($\alpha = 0.1$) is consistent with empirical estimates obtained via hyperbolic-ratio fits to the contrast response functions of V1 neurons (Busse et al., 2009; Albrecht & Hamilton, 1982; Sclar

et al., 1990; Gardner, Anzai, Ohzawa, & Freeman, 1999a). The exponents n_n and n_d control the slope of the CF (Fig. 4A). Exponents greater than 1 are needed to account for the accelerating nonlinearity at low contrasts discussed in Appendix B. The standard value $n_n = n_d = 2$ is consistent with empirical estimates (e.g., Busse et al., 2009, p. 933; Albrecht et al., 2003, p. 752) and implements Heeger’s (1992a) half-squaring operator $[\cdot]^2$. Note that Equation 13 has separate exponents for the numerator and denominator. This additional flexibility is needed for quantitative fits of physiological (e.g., Carandini & Heeger, 2012) and psychophysical (e.g., Itti et al., 2000) data.

The baseline parameter β can be interpreted as equivalent contrast as discussed in Section 2.3. The interpretation depends on its sign. Negative values effectively impose a threshold on the stimulus drive, whereas positive values produce a *maintained discharge* in response to a blank stimulus (uniform gray field). Many real neurons in V1 emit spikes in the absence of external stimulation, although the spontaneous firing rates typically are quite low (Allison & Bonds, 1994; Hubel & Wiesel, 1959; Hubel, 1959; R. L. De Valois, Albrecht, & Thorell, 1982, 1982; Pettigrew, Nikara, & Bishop, 1968; Squatrito, Trotter, & Poggio, 1990). For example, Hubel and Wiesel (1959) recorded maintained discharges between 0.1 and 10 spikes/sec in V1 neurons of anesthetized cats. Under the standard parameterization, Equation 13 produces $M[\beta]^{n_n}/\alpha^{n_d} = 1.6$ spikes/sec. The β parameter is in the focus of the mathematical analyses below (see Eq. 21 and Appendix C) and several simulations (e.g., Fig. 10) that explore non-standard values.

The remaining five free parameters control various bandwidths. They have diverse units (listed in Table 2) and should be interpreted with care because their empirically observable analogs depend on complex interactions among the DNM components as discussed in Section 3.3.3 below. For instance, the orientation tuning bandwidth of the model as a whole is 31.8° under the standard parameters (Fig. 13 and 14). Note that this is smaller than the bandwidth $h_\theta = 40^\circ$ of the linear filtering stage and much smaller than the pooling bandwidth $h_\Theta = 60^\circ$ of the suppressive drive. Analogously, the frequency tuning bandwidth of the standard model as a whole is 1.11 oct, which is smaller than both $h_f = 1.5$ and $h_F = 2.0$ oct. The overall model bandwidths are within the reported ranges of neurophysiological measurements (R. L. De Valois, Albrecht, & Thorell, 1982; Kulikowski & Bishop, 1981; Movshon, Thompson, & Tolhurst, 1978a; Rose & Blakemore, 1974; Busse et al., 2009; Watkins & Berkley, 1974; Schiller, Finlay, & Volman, 1976b, 1976a). Finally, the standard FWHH of the radial spatial-pooling kernel (Eq. 17) covers $h_R = 2.0$ cycles of the neuron’s preferred frequency. This value is consistent with surround-suppression measurements (see Sections 3.3.1 and 3.3.5 for references).

3.3 Results

A systematic series of simulation experiments replicated the qualitative patterns characterizing the phenomena in Table 1. The simulated DNM patterns are plotted below alongside single-cell recording data from representative experiments. The physiological data were captured from the figures in their respective publications using PlotDigitizer (<http://plotdigitizer.sourceforge.net>) and replotted here in a unified format. All simulations used the standard parameters listed in Table 2 unless explicitly indicated otherwise.

3.3.1 Size tuning

Our first simulation measured the responses of a DNM complex cell as a function of the stimulus diameter. All stimuli were gratings with orientation $\Theta^* = 0^\circ$ and frequency $F^* = 2$ cpd that matched the cell's tuning preferences. The resulting *size tuning* function is plotted in Figure 5A for gratings with maximal contrast $c = 1$. The responses increased as stimulus size increased at first, reached a peak, decreased, and finally settled to an asymptote. This matches the qualitative pattern observed in single-cell recordings of V1 neurons (Gieselmann & Thiele, 2008; H. E. Jones, Grieve, Wang, & Sillito, 2001; Sengpiel et al., 1997; Schwabe et al., 2010).

The non-monotonic response pattern in Figure 5 indicates that the neuron's receptive field has limited spatial extent (Phenomenon 1 in Table 1). The diameter of the grating that induces the maximal response is often used in physiological studies to operationalize the size of the classical receptive field. For the DNM neuron, this *measured RF diameter* is MRFD $\approx 0.81^\circ$ (marked by the black triangle in Figure 5A). Note that it is narrower than the FWHH sizes of the elliptical contour of the weighting function of the model's linear stage ($h_{\hat{y}} = 1.26^\circ$ and $h_{\hat{x}} = 0.92^\circ$, Table 2, see also Fig. 1). Other parameter sets were also tested and the MRFD tended to expand as n_n , α , h_Θ , or h_F increased or β , n_d , h_θ , or h_f decreased. As a consistency check, note also that the measured asymptotic response rate (≈ 41.2 spikes/sec) matched the prediction of the hyperbolic-ratio Equation 15. This is because the gratings with very large diameters in this simulation became identical to the calibration image I_{cal} , thereby satisfying the condition for applicability of Equation 15.

The nonmonotonic response pattern in the model arises from the interplay between the numerator and denominator in Equation 13. As the stimulus diameter increases, the stimulus drive (Eq. 9) rises faster but saturates earlier than the suppressive drive (Eq. 16, Fig. 5C, Gieselmann & Thiele, 2008).

The measured RF diameters of real neurons depend on the parameters of the grating. The MRFD increases as the stimulus luminance contrast decreases (Phenomenon 2, Figure 6B, Schwabe et al., 2010; Sengpiel et al., 1997; Sceniak, Ringach, Hawken, & Shapley, 1999; Kapadia, Westheimer, & Gilbert, 1999; Cavanaugh, Bair, & Movshon, 2002a; Tailby, Solomon, Peirce, & Metha, 2007; Nienborg et al., 2013). Also, the MRFD decreases for gratings with non-preferred orientations (Ph. 3, Fig. 7B, Tailby et al., 2007). The DNM neuron reproduces

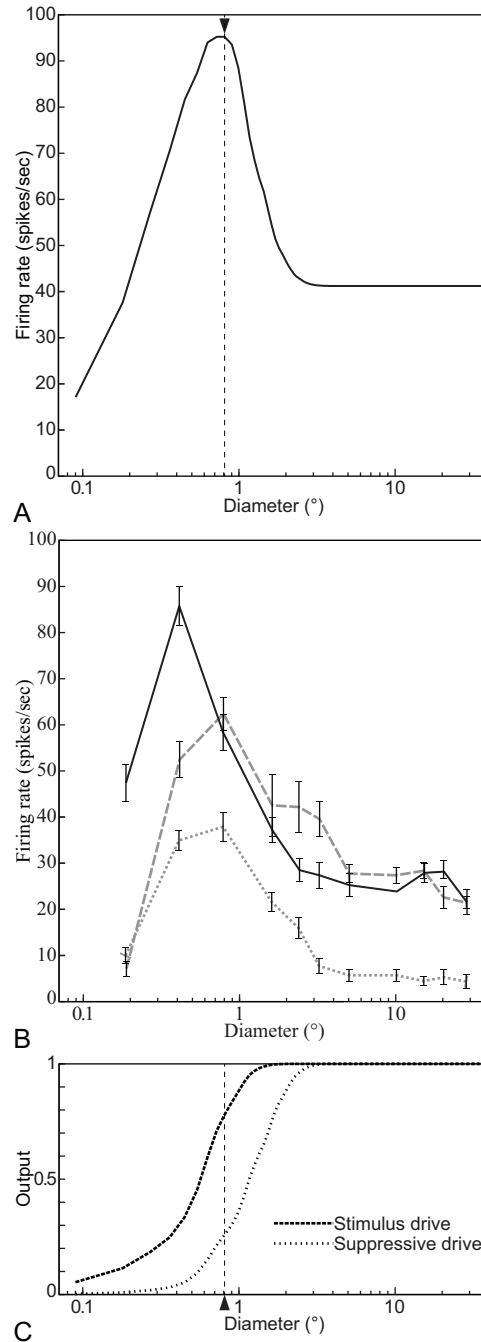


Figure 5: Size tuning functions of (A) a DNM complex cell with standard parameters, (B) three V1 complex cells (replotted from Schwabe et al., 2010, Fig. 2, anesthetized macaque, error bars = \pm SE), and (C) the stimulus- and suppressive-drive terms of the DNM equation. The *measured RF diameter* of the DNM neuron is 0.81° (indicated by the solid triangle in panel A). All stimuli were gratings with maximal contrast ($c = 1$) and orientation and frequency that matched the preferences of the respective neuron. (DNM = divisive normalization model. See Phenomenon 1 in Table 1.)

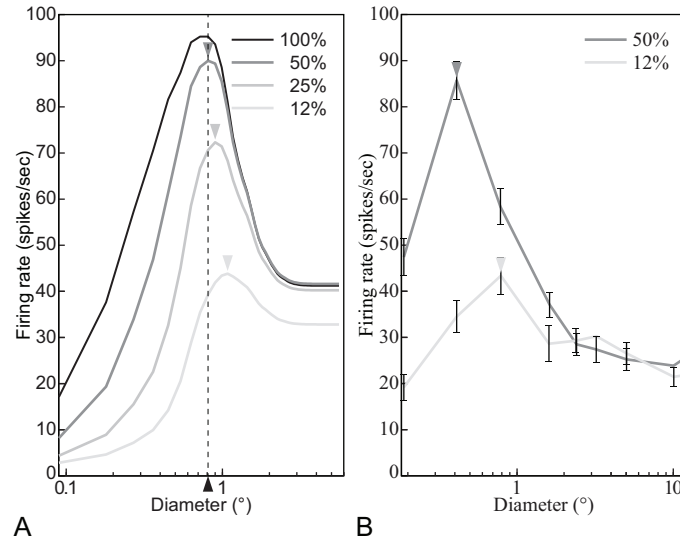


Figure 6: The size tuning function depends on the contrast of the stimulus grating. (A) DNM complex cell with standard parameters, (B) a V1 complex cell (replotted from Schwabe et al., 2010, Figure 2a, anesthetized macaque, error bars = \pm SE). The measured RF diameter (depicted by triangles) increases as the stimulus contrast (shown in the legend) decreases. Stimulus orientation and frequency matched the preferences of the respective neuron. (DNM = divisive normalization model. See Phenomenon 2 in Table 1).

both phenomena (Figures 6A and 7A). The simulation that produced Figure 7A was designed to emulate the method of Tailby et al. (2007). Specifically, the stimuli in the Opt condition (gray solid line) were gratings of the preferred orientation (0°). We determined the MRFD for the DNM neuron in the Opt condition (0.81° , vertical dashed line) and then determined the orientation tuning curve for gratings with this diameter (Fig. 13A). The orientation in the Ori Δ condition (15.9°) was determined by the half-height points of this (symmetric) tuning curve. Probing the DNM neuron with gratings at this sub-optimal orientation produced the dotted black line in Figure 7A. It replicates the qualitative pattern of the neurophysiological data in panel B.

The MRFD of the model is also affected by the spatial frequency of the grating. It decreases if the frequency is either lower or higher than the DNM neuron's preferred frequency, with stronger decreases for lower frequencies (Fig. 8A). The non-optimal stimulus frequencies ($Spf_L = -0.21$ oct, 0.86 cpd; $Spf_H = 1.52$ oct, 2.87 cpd) for this simulation were chosen at the half-height points of the model frequency-tuning curve (Fig. 13D; Tailby et al., 2007). The available recordings from real V1 neurons suggest that their MRFD tend to decrease as the stimulus frequency becomes higher than the preferred frequency, but no clear trend has been observed when the stimulus frequency becomes lower (Fig. 8B, Osaki, Naito, Sadakane, Okamoto, & Sato, 2011; see also Tailby et al., 2007). This qualitative pattern may be somewhat different than the DNM prediction. See Section 3.4 for further discussion of the relationships among stimulus size, frequency, and neuronal responses.

The size tuning of V1 neurons was also measured by using annular stimuli that overlay a circular grey "hole" in a larger circular grating with the neuron's preferred orientation and frequency (H. E. Jones et al.,

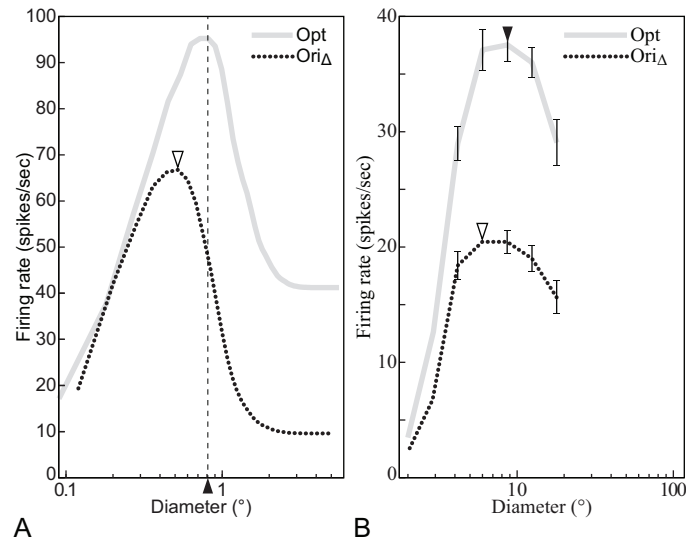


Figure 7: The size tuning function depends on the orientation of the stimulus grating. (A) DNM complex cell with standard parameters, (B) V1 complex cell (replotted from Tailby et al., 2007, Fig. 2b, anesthetized cat, error bars = \pm SE). The grating orientation in the Opt condition matched the preferences of the respective neuron. The orientation in the Ori Δ condition was determined by the half-height point of the orientation tuning function (Tailby et al., 2007). See text for details. All gratings had maximal contrast ($c = 1$) and their frequency matched the preference of the respective neuron. (DNM = divisive normalization model. See Phenomenon 3 in Table 1.)

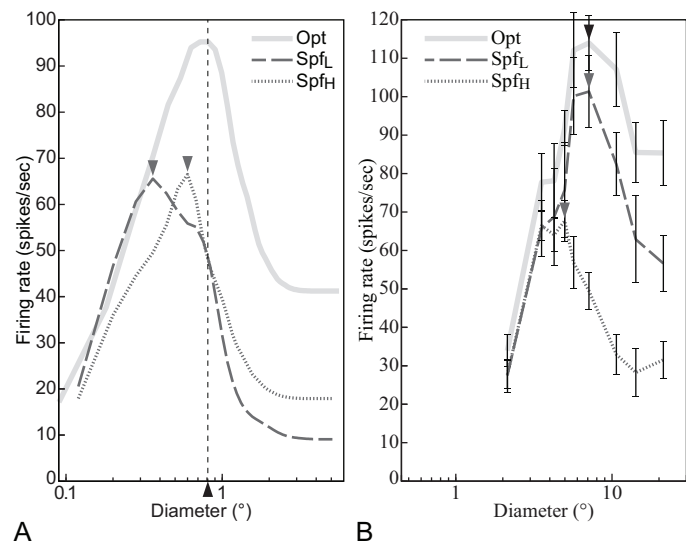


Figure 8: The size tuning function depends on the spatial frequency of the stimulus grating. (A) DNM complex cell with standard parameters, (B) a real V1 neuron (replotted from Osaki et al., 2011, Figures 2A and 4A, anesthetized cat, error bars = \pm SE). The stimulus frequency matched the preference of the respective neuron in the Opt condition, was lower in the Spf $_L$ condition, and higher in the Spf $_H$ condition. The side frequencies in (A) were determined by the half-height points of the DNM frequency tuning curve. See text for details. The data in (B) were collected with Spf $_L$ = 0.10 cpd, Opt = 0.20 cpd, and Spf $_H$ = 0.30 cpd (Osaki et al., 2011). All gratings had maximal contrast ($c = 1$) and their orientation matched the preference of the respective neuron. (DNM = divisive normalization model. See Phenomenon 4 in Table 1.)

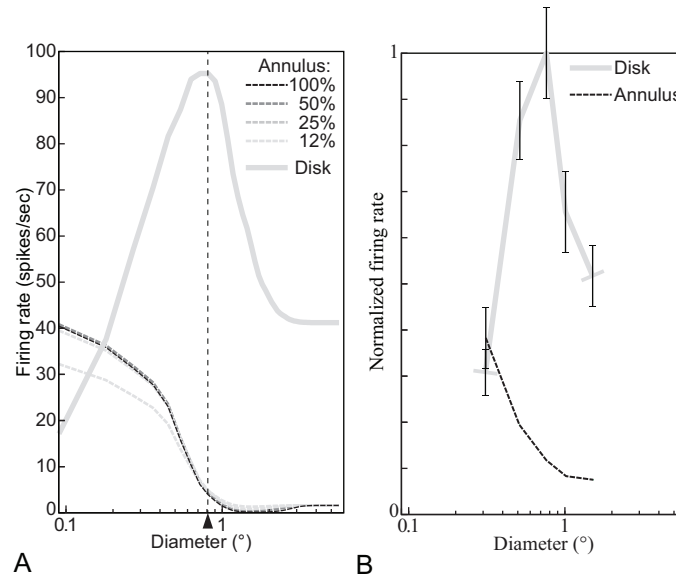


Figure 9: Responses to annular stimuli containing a circular grey “hole” inside a large circular grating. The X axis specifies the diameter of the hole for annular stimuli (dashed lines) and the outside diameter for disk stimuli (solid line, cf. Figure 5). (A) DNM complex cell with standard parameters. The legend indicates the contrast of the annular stimuli; the disk had 100% contrast. The measured RF diameter (0.81°) is indicated by a triangle. (B) V1 neuron (replotted from H. E. Jones et al., 2001, Fig. 1, anesthetized macaque, error bars = \pm SE). All stimuli are based on gratings whose frequency and orientation matched the preference of the respective neuron. See Phenomena 5 and 6 in Table 1.

2001; Sengpiel et al., 1997). The recordings from V1 neurons decreased monotonically as the diameter of the hole increased, and then leveled off at an asymptotic level that was similar to the spontaneous discharge of the cell (Fig. 9B, Ph. 5). The hole diameter for which the responses become nearly constant is similar to the RF diameter measured with disc stimuli as discussed above. The DNM neuron shows analogous trends (Figure 9A). Recall that when disc stimuli (with no holes) are used to measure the RF diameter, the latter depends on stimulus contrast for both real and model neurons (Fig. 6). When annular stimuli are used, however, the DNM predicts approximate invariance with respect to the contrast of the grating inside the annular envelope (Phenomenon 6).

3.3.2 Contrast sensitivity

The contrast sensitivity of the DNM neuron was measured using a grating whose orientation and spatial-frequency are those to which the neuron is tuned and its size is large enough (5.58°) to cover both the receptive field of the stimulus and the suppressive field of the model neuron. The model responses are plotted in Figure 10A (solid curve) as a function of the contrast (Phenomenon 7). The slope of the contrast sensitivity function becomes maximal when the contrast is at 9%, above 0 (see Sect. 2.2). It may be relevant to the psychophysical results that a differential threshold of a grating contrast becomes minimal above the contrast detection threshold (see Itti et al., 2000; Wilson, 1980 for a comparison between human- and model performance

on luminance contrast discrimination).

Note that the contrast sensitivity function of the virtual neuron slightly decreases if the contrast is higher than 50% with the standard parameter set (indicated by a solid curve and an open circle in Figure 10A). An analogous trend is also observed from some real neurons (C.-Y. Li & Creutzfeldt, 1984; Tyler & Apkarian, 1985; Somers et al., 1998; Peirce, 2007). This is called the *supersaturation effect* (Ph. 8). This effect could be emulated only weakly by the DNM neuron with the standard parameter set. Some V1 neurons, however, do show the supersaturation effect very clearly (Figure 10B, D). Mathematical analysis (Appendix C) indicates that the model produces this effect when the baseline parameter β exceeds the following critical value:

$$\beta > \frac{n_n}{n_d}(1 + \alpha^{n_d}) - 1 \quad (21)$$

This inequality suggests that the super-saturation effect becomes stronger as β increases, α decreases, and n_n/n_d decreases (Figure 10A, C; see also Sawada & Petrov, 2015, December, for supersaturation plots produced with a different parameter set).

When visual noise is added to the preferred grating of a real neuron, the neuron's responses tend to be scaled down by a multiplicative factor (Ph. 9, Carandini et al., 1997). This effect is observed in the model when the contrast of the tuned grating was high (e.g., 50, 25, and 12% in Figures 11A, B). When the contrast of the grating is zero or very low, adding noise increased the response of both real neurons (Squatrino et al., 1990) and the DNM neuron (Fig. 11A, B). Note that psychophysical results have shown that human performance on detecting a grating is improved by adding moderate amount of visual noise (e.g., 6% in Figure 11) to the grating (Moss, Ward, & Sannita, 2004; H. Sasaki et al., 2006; Simonotto et al., 1997). This is called *stochastic resonance*. The noise increased the response of the DNM neuron but decreased the slope of the contrast sensitivity function if the contrast of the grating was close to zero. Decreasing the slope is expected to increase the threshold of the model neuron for detecting the grating. However, if the output of one model neuron is relayed as excitatory input to another model neuron, it is possible that adding noise to the stimulus elevates the responses of the first neuron enough to trigger a response in the second neuron. In this way, moderate amounts of additive stimulus noise can improve the detection performance of the entire ensemble, despite the fact that the noise decreases the CF slope in the first stage.

The contrast sensitivity functions of real neurons were affected by the orientation (Phenomenon 10, Tolhurst & Dean, 1991; Carandini et al., 1997) and the spatial-frequency (Phenomenon 11, Albrecht & Hamilton, 1982; Carandini et al., 1997; Albrecht et al., 2003) of the stimulus grating. The CFs of real V1 neurons were multiplicatively scaled down if the orientation or the frequency of the grating were different from the neuron's preferred orientation or frequency (Fig. 3, Carandini et al., 1997). These trends are also observed in the model

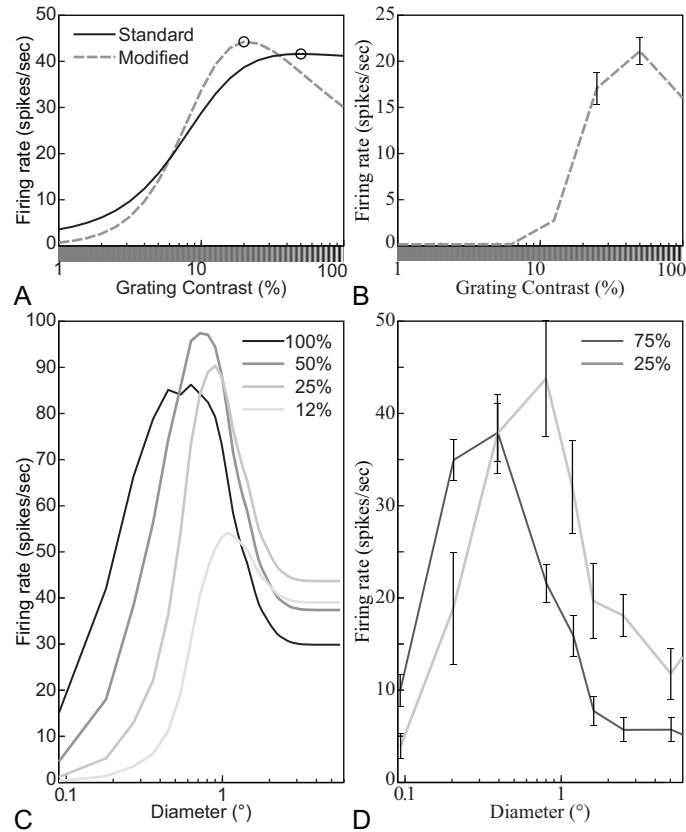


Figure 10: Examples of the *supersaturation effect* (Phenomenon 8 in Table 1). (A) Contrast sensitivity function of a DNM neuron with a standard (solid line) and a modified parameter set ($n_d = 2.35$, $\beta = 0$, $M = 30$, dashed line). (B) Contrast sensitivity function of a real neuron (replotted from Peirce, 2007, Fig. 2e, anesthetized macaque). (C) Size-tuning functions of the DNM neuron with the modified parameter set for different contrasts of a grating (shown in legend; see Fig. 6 for the standard-parameter analog). (D) Size-tuning functions of a V1 complex cell (replotted from Schwabe et al., 2010, Fig. 2C, anesthetized macaque, error bars = \pm SE). Stimulus frequency and orientation matched the DNM preferences in (A) and (C). The size of the grating patch was 2.88° for (A).

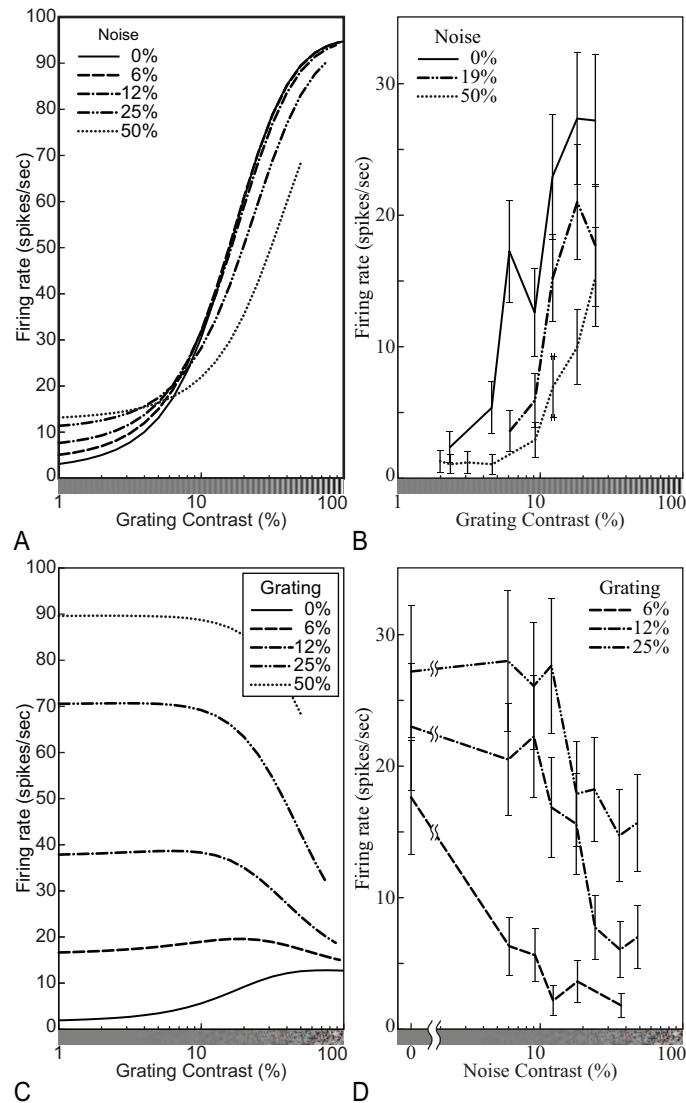


Figure 11: Responses of (A, C) the model neuron and (B, D) a V1 simple cell (replotted from Carandini et al., 1997, Fig. 14D and 14C, anesthetized macaque, error bars = \pm SE) to a grating with a random noise pattern superimposed. The graphs in (A, B) are plotted against the contrast of the grating, whereas those in (C, D) are plotted against the contrast of the noise. (A, C) The size of the stimuli was equal to the measured RF diameter (0.81°), and their orientation and frequency matched the tuning preferences of the model neuron. See Phenomenon 9 in Table 1.

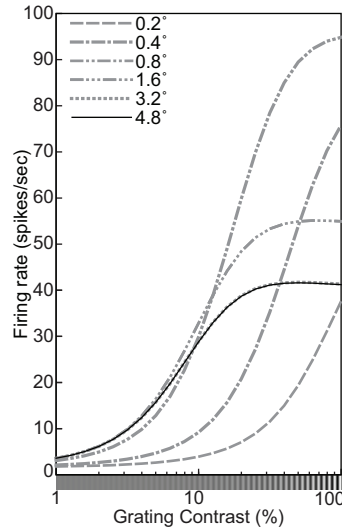


Figure 12: Contrast sensitivity functions of the model neuron for gratings with different diameters (indicated in the legend). The contrast, orientation, and spatial-frequency of the grating were 100%, 0° , and 2.0 cpd. See Phenomenon 12 in Table 1.

(Fig. 3B, D).

Our simulation experiments showed that the CF of the model neuron was also affected by the size of the grating patch (Fig. 12). As stimulus diameter increases up to 0.8° , the overall firing-rate increases, the CF slope at low contrasts becomes steeper, and the function becomes more saturating at high contrasts. Then, the asymptotic firing rate decreases while the slope at low contrasts is unchanged up to 3.2° . The function is unchanged even if stimulus diameter becomes larger than 3.2° . Note that this phenomenon was predicted by an analogous model of LGN neurons and was observed in a physiological experiment (Bonin, Mante, & Carandini, 2005). To the best of our knowledge, however, the phenomenon has never been observed in V1 neurons using grating patches as stimuli (see Schumer & Movshon, 1984). Note also that the supersaturation effect is observed from the DNM neuron if the stimulus diameter is 1.6° or larger.

3.3.3 Orientation and spatial-frequency tuning

The orientation and spatial-frequency tuning functions of the DNM neuron are shown in Figure 2 (Phenomena 13 and 14). The measured peaks of the tuning functions were consistent with the weighting function of the stimulus drive of the neuron. The measured bandwidths of the tuning functions were 31.8° in the orientation domain and 1.11 octaves in the frequency domain. Other parameter sets of the model were also tested and the measured bandwidths tended to become narrower as n_n , α , h_Θ , or h_F increase or β , n_d , h_θ , or h_f decrease. Note that the measured bandwidths were substantially narrower than those of the weighting function of the linear filter in Equation 5 ($h_\theta = 40^\circ$, $h_f = 1.5$ oct). This narrowing effect is attributed to the rectification and the expansive nonlinearity ($n_n > 1$) in the numerator of Equation 13.

Note that the bandwidths of the numerator of Equation 13 are a little narrower (29.2° and 1.04 oct) than those of the DNM neuron itself (31.8° and 1.11 oct). This occurs because the denominator of Equation 13, which represents suppression, tends to widen the tuning curves of the DNM neuron. The tuning functions of the denominator are unimodal and become maximal at the same orientation and frequency as those of the numerator. The divisive operation then has a widening effect, and becomes stronger as the denominator becomes more sharply tuned. This is why the bandwidths of the DNM neuron become wider as α , h_Θ , or h_F decrease or as n_d increases because they make the tuning functions of the denominator narrower. In particular, the h_Θ and h_F parameters (Table 2) allow an almost independent control of the widening of the orientation- and frequency tuning, respectively.

There is evidence that the orientation (Okamoto, Naito, Sadakane, Osaki, & Sato, 2009; C.-Y. Li & Li, 1994; see also Maffei & Fiorentini, 1976, for a report of a few exceptional neurons) and spatial-frequency (Osaki et al., 2011; Maffei & Fiorentini, 1976) bandwidths of real neurons become narrower when the size of the grating patch increases. The model simulated this effect well (Fig. 13A). This property of the model neuron is attributed to its stimulus drive. The bandwidths of the stimulus drive (and of the numerator) in the model also became narrower as the size of the patch increased (Fig. 13E, F). This suggests that the effect of the patch size on the tuning widths of the virtual neuron is *not* due to surround suppression. The weighting function of the linear stage of the model neuron is partially hidden by the surround suppression. This leads to a systematic underestimation of the measured receptive field (R. L. De Valois, Thorell, & Albrecht, 1985). Then, its measured bandwidths become wider if the grating patch just fills the underestimated RF. Note that the peak of the frequency tuning curve of real neurons tends to shift to a higher frequency as the size of the stimulus decreases (Teichert, Wachtler, Michler, Gail, & Eckhorn, 2007; Osaki et al., 2011). Note, however, that the model neuron did not show this trend. The peak of its frequency tuning curve was *not* affected by the size of the grating patch.

The orientation and spatial-frequency bandwidths of real neurons are affected by the contrast of the grating patches. As the contrast decreases, these bandwidths become slightly wider or narrower (or are invariant) depending on the individual neurons. Overall, the orientation bandwidths tend to be relatively invariant against the contrast (Phenomenon 17, Fig. 14A, B; Somers et al., 1995; Skottun et al., 1987; C.-Y. Li & Creutzfeldt, 1984; Sclar & Freeman, 1982; Anderson, Lampl, Gillespie, & Ferster, 2000; Alitto & Usrey, 2004; Troyer, Krukowski, Priebe, & Miller, 1998), but the spatial-frequency bandwidths tend to become narrower as the contrast decreases (Ph. 18, Fig. 14C, D; Sceniak et al., 2001; Albrecht & Hamilton, 1982; Skottun et al., 1987). These trends in the neurophysiological data could be emulated to some extent by the model neuron with a non-standard parameter set with smaller h_Θ (40°) and h_F (1.0 oct). Under the standard parameterization, the

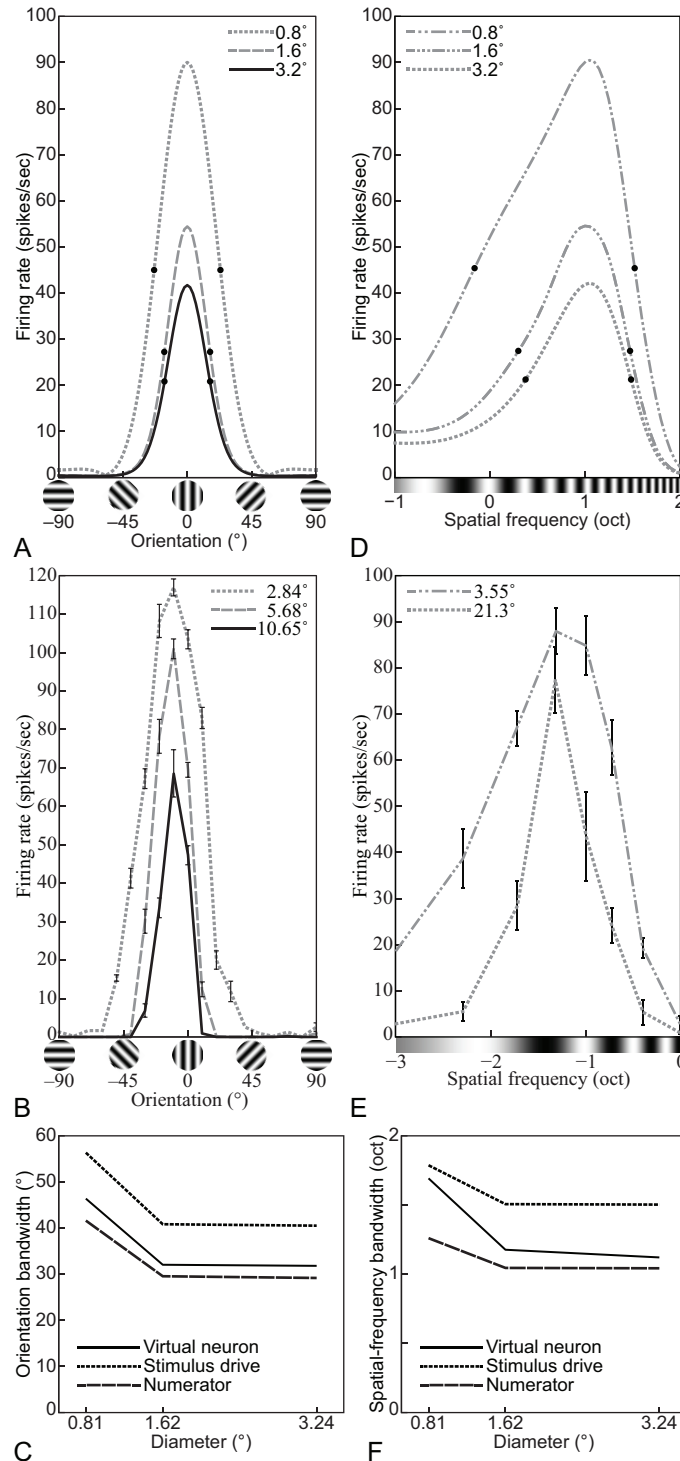


Figure 13: Orientation and spatial-frequency tuning functions of (A, D) the DNM neuron with standard parameters, (B) a V1 neuron (replotted from Okamoto et al., 2009, Fig. 1, anesthetized cat, error bars = \pm SE), and (E) a V1 simple cell (replotted from Osaki et al., 2011, Fig. 5B, anesthetized cat, error bars = \pm SE) for gratings with different diameters (indicated in legend). (C, F) The effect of the size on the bandwidths of the DNM neuron, its stimulus drive, and the numerator in Equation 13. The contrast of the grating was 100% for (A, D), its frequency was 2.0 cpd for (A), and its orientation was 0° for (D). (DNM = divisive normalization model. See Phenomena 15 and 16 in Table 1).

bandwidths of the DNM neuron in both domains became slightly wider as the contrast decreased (Figure 14). Note that these trends in the neurophysiological data are rather weak and also note that there are many neurons in V1 that show opposite trends (Sclar & Freeman, 1982; Sceniak et al., 2001; Alitto & Usrey, 2004; Kim, 2011).

The effect of stimulus contrast on the orientation- and frequency-bandwidths depends on the β parameter in the numerator and on the denominator of Equation 13. The bandwidths of the numerator of the model neuron become narrower as the contrast decreases if $\beta < 0$ (the *iceberg effect*, Heeger, 1992a; Sceniak et al., 2001; Tadmor & Tolhurst, 1989) and become wider if $\beta > 0$. The bandwidths of the model neuron are also widened by the denominator depending on stimulus contrast. Recall that the widening effect of the denominator becomes stronger as the denominator becomes more sharply tuned. The tuning function of the denominator widens as the contrast decreases and it becomes a constant (α^{nd}) at zero contrast. This contrast dependence of the widening effect of the denominator is consistent with the trend that the bandwidths of the model neuron become narrower as the contrast decreases.

Pollen and Ronner (1982) measured the spatial-frequency tuning of real neurons using sinusoidal and square gratings. They observed that, when measured with square gratings, the tuning function showed two peaks in the spatial-frequency domain (Phenomenon 19). The primary peak appeared at the neuron's preferred frequency and the secondary peak appeared at one-third of the preferred frequency (Figure 15B). The height of the secondary peak was between 0.6 and 0.8 times the height of the primary peak for most neurons. The secondary peak of some neurons was even as high as the primary peak. These observations partly agree with the predictions of the linear model (Section 2.1). According to Fourier-series decomposition, a square grating with frequency f_{sq} can be represented by a sum of sinusoidal gratings at the odd harmonics of f_{sq} with magnitudes proportional to reciprocals of the orders of the harmonics: $\sum_{j=0}^{\infty} (2j+1)^{-1} \sin(f_{sq}/(2j+1))$. Thus, if a neuron satisfying the assumption of the linear model is tuned to a sinusoidal grating with frequency f_{sin} , it is predicted that its tuning function to a square grating should show multiple local maxima in the spatial-frequency domain. The positions of these maxima are $f_{sin}/(2j+1)$ and their heights are $1/(2j+1)$ of the global maximum at f_{sin} , where j is an integer. However, the results of Pollen and Ronner (1982) showed some deviations from the linear predictions. Only the primary ($j=0$) and secondary peaks ($j=1$) of the tuning functions could be reliably identified reliably in the data, and the heights of the observed secondary peaks were higher than the predicted height. The DNM neuron with a non-standard parameter set ($h_f = 0.8, h_F = 0.4$) can emulate these results, including the discrepancy to some extent (Figure 15A). The tuning function of the model neuron shows the secondary peak at one-third of the tuned frequency and its height is about 0.9 of the height of the primary peak. On the other hand, the model neuron also showed the tertiary peak at one-fifth of the tuned frequency ($j=2$). Note also that the DNM neuron responds more strongly to the square grating than the sinusoidal

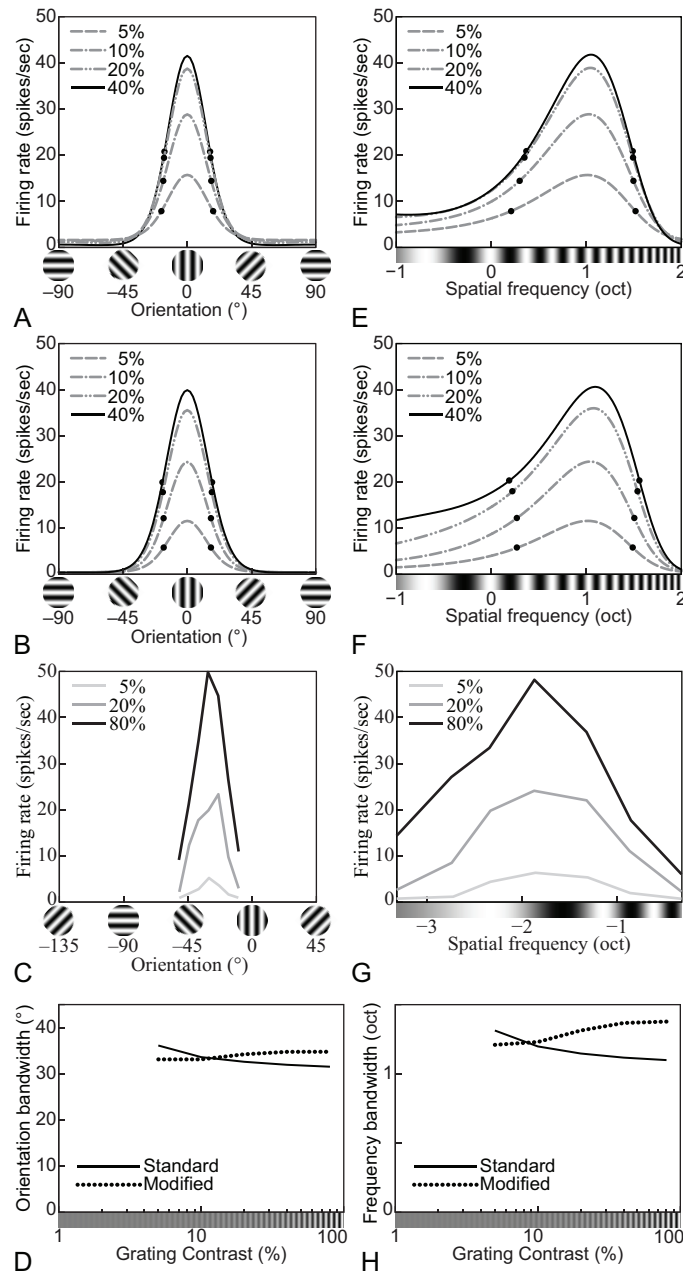


Figure 14: Orientation and spatial-frequency tuning functions for gratings with different contrasts (indicated in legend). (A, E) The DNM neuron with the standard parameter set; (B, F) The DNM neuron with a modified parameter set ($h_{\Theta} = 40^{\circ}$, $h_F = 1.0$ oct); (C) Orientation tuning of a simple striate cell (replotted from Skottun et al., 1987, Fig. 3A, anesthetized cat); (G) Frequency tuning of a simple striate cell (replotted from Skottun et al., 1987, Fig. 4A, anesthetized cat). (D, H) The effect of stimulus contrast on the bandwidths of the DNM neuron in the orientation (D) and frequency (H) domains for the two parameter sets (legend). (A, B, E, F) The size of the grating patch was 2.88° . The spatial-frequency of the grating was 2.0 cpd for (A, B) and its orientation was 0° for (E, F). (DNM = divisive normalization model. See Phenomena 17 and 18 in Table 1).

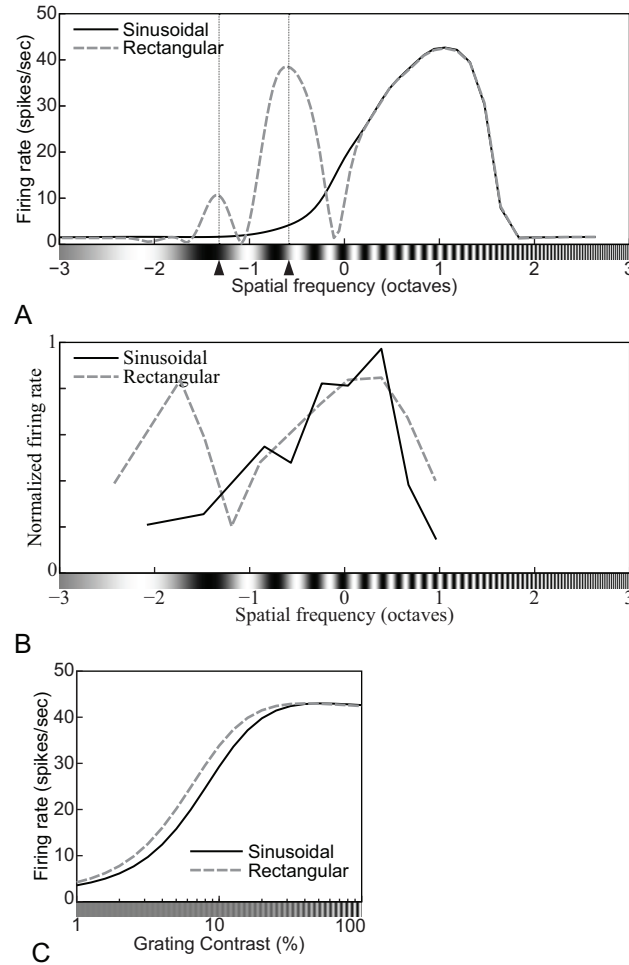


Figure 15: Spatial-frequency tuning functions for sinusoidal and square gratings: (A) of the model neuron with a modified parameter set ($h_f = 0.8$, $h_F = 0.4$) and (B) of a complex cell (replotted from Pollen & Ronner, 1982, Fig. 6E, anesthetized cat). Note the secondary peak at one-third of the tuned frequency for square gratings (dashed line). (C) The contrast sensitivity functions of the model neuron with the modified parameter set for sinusoidal and square gratings (indicated in legend). See Phenomena 19 and 20 in Table 1.

grating, especially when their contrast is moderately low (Figure 15C). The contrast sensitivity function of the DNM neuron with the square grating shifts horizontally to the left (Phenomenon 20). This is because the outputs of both stimulus- and suppressive drives are always larger to the square grating than to the sinusoidal grating with the same frequency and orientation (see Equation 5).⁸

3.3.4 Cross-orientation suppression

The responses of real neurons in V1 to a grating (signal) tend to be suppressed by another grating (mask) superimposed onto the signal grating within the neuron's measured RF. This *cross-orientation suppression* occurs in wider ranges of the orientation and spatial-frequency domains than the excitatory tuning of the neuron

⁸This prediction seems to be consistent with psychophysical results showing that human performance in contrast detection is better with a square grating than with a sinusoidal grating (Campbell & Robson, 1968). Note, however, that the square grating is expected to stimulate a wider variety of V1 neurons because a square grating is composed of multiple sinusoidal gratings, so these psychophysical results can be explained by properties of either single V1 neurons or by a population of V1 neurons.

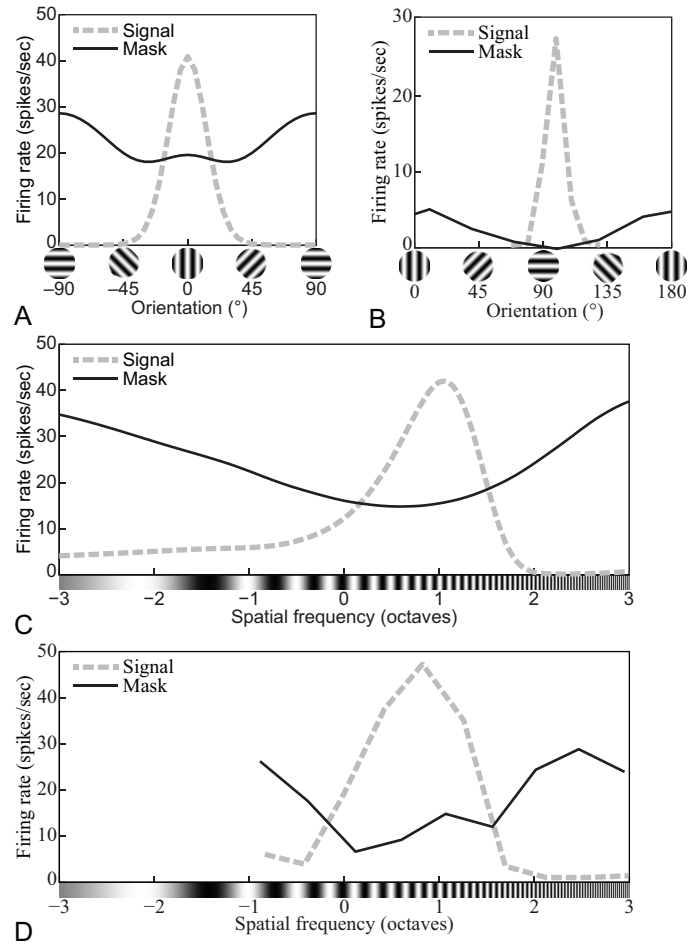


Figure 16: The effect of *cross-orientation suppression* on: (A, C) the orientation and frequency tuning functions of the model neuron with standard parameters, (B) the orientation tuning function of a complex cell (replotted from DeAngelis et al., 1992, Fig. 7CD, anesthetized cat), and (D) the frequency tuning function of a simple cell (replotted from DeAngelis et al., 1992, Fig. 3AB, anesthetized cat). In all panels, the standard tuning function (as probed with a non-masked grating) is plotted by the dashed gray line. The tuning function of orthogonal suppression is plotted by a solid black line. (A) The contrast and frequency of the signal grating were 20% and 1.0 oct. The contrast and frequency of the mask grating were 40% and 0.0 oct. The size of the stimuli was 2.88°. (C) The contrast and orientation of the signal grating were 20% and 0°. The contrast and orientation of the mask grating were 80% and 90°. The size of the stimuli was 2.88°. See Phenomena 21 and 22 in Table 1.

(DeAngelis et al., 1992; see also K. K. De Valois & Tootell, 1983). The strongest suppression occurs when the mask orientation is the same as the tuned orientation (Phenomenon 21, Morrone et al., 1982; DeAngelis et al., 1992, see also Bonds, 1989 for a mask-induced facilitation effect). These trends were captured by the DNM neuron (Figure 16A). As for spatial frequency, the most suppressive frequency of the mask grating does not necessarily coincide with the tuned frequency of the real neuron (Phenomenon 22, K. K. De Valois & Tootell, 1983; Morrone et al., 1982; DeAngelis et al., 1992). The most suppressive frequency of the DNM neuron (Figure 16C) seems to be lower than its tuned frequency (1.0 oct). One possible explanation of the latter result is that the diameter of the measured RF could have been too small for a grating with a spatial-frequency lower than the tuned frequency.

The mask grating affects the contrast response function (CF) of real neurons (Phenomenon 23, Carandini, 2004; Freeman et al., 2002; Morrone et al., 1982). Panel B of Figure 17 shows representative physiological data (Freeman et al., 2002) and panel A shows simulation results of a DNM neuron. The CF of the DNM neuron shifts horizontally right if the contrast of the mask is high enough ($\geq 12\%$). The suppression effect is negligible if the mask contrast is lower ($< 12\%$). This trend agrees with the physiological results. The responses are also plotted as functions of the mask contrast in panels C and D. In both the model and the data, the response functions shift upward and scale up vertically as the contrast of the signal grating increases.

3.3.5 Surround suppression

The responses of real neurons to a grating patch within the neurons's measured RF tend to be suppressed by an annular grating surrounding the RF. This *surround suppression* is strongest when the orientation and frequency of the annular grating are the same as those the neuron is tuned to (Phenomena 24 and 25, Nelson & Frost, 1978; C.-Y. Li & Li, 1994; Blakemore & Tobin, 1972; DeAngelis et al., 1994; Ozeki, Finn, Schaffer, Miller, & Ferster, 2009; Ozeki et al., 2004; Cavanaugh et al., 2002b; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995). The ranges of the surround suppression in both domains are wider than the respective tuned bandwidths. These trends were captured by the DNM neuron with standard parameters (Figure 18).

The contrast response function of real neurons is affected by the contrast of an annular grating surrounding the classical RF (Phenomenon 26, Cavanaugh et al., 2002a; Carandini, 2004; DeAngelis et al., 1994). Panel B of Figure 19 shows the CF of a complex cell in cat V1 with different contrasts of the annular grating (Carandini, 2004). Panel A shows the corresponding plot for a DNM neuron with standard parameters. Both empirical and simulated CFs shift horizontally to the right as the contrast of the annular grating increases. Note that the slope of the CF of the model neuron is not necessarily invariant against the annular grating. With a different parameter set ($n_n < n_d$), the CF of the DNM neuron became shallower as the contrast of the annular grating increased (Figure 19C). The same trend has been observed for some V1 neurons (Figure 19D, Carandini, 2004).

The effect of the annular grating also depends on its orientation for both real neurons (Figure 20B, Phenomenon 27, Cavanaugh et al., 2002b) and the model neuron (Figure 20A). Note that the cross-orientation suppression and the surround suppression of real neurons show analogous trends to one another, so they can be modeled jointly by the suppressive drive in Equation 13. Note that this similarity at the functional level does not necessarily imply similarity at the mechanistic level. In fact, physiological results indicate that the two types of suppression should be attributed to different mechanisms (see DeAngelis et al., 1992, 1994, for examples). We will discuss the difference between these two types of suppression in Section 4.3 below.

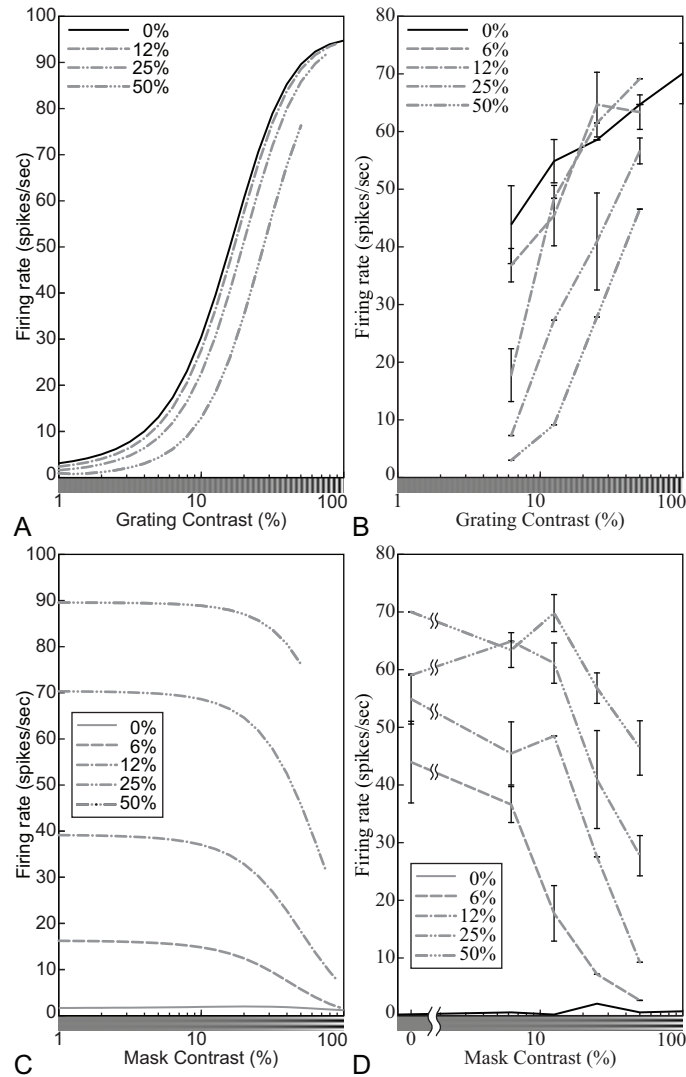


Figure 17: The effect of *cross-orientation suppression* on the contrast response functions of (A, C) the model neuron with standard parameters and (B, D) a V1 neuron (replotted from Freeman et al., 2002, Fig. 6, anesthetized cat). In (A, B), the responses of the model and real neurons are plotted against the contrast of the signal grating, with mask contrast shown in the legend. In (C, D), the responses are plotted against the contrast of the mask grating, with signal contrast in the legend. (A, C) The size of the stimuli was equal to the measured RF diameter (0.81°). The orientations of the signal and mask gratings were 0° and 90° , and their spatial-frequency was 2.0 cpd. See Phenomenon 23 in Table 1.

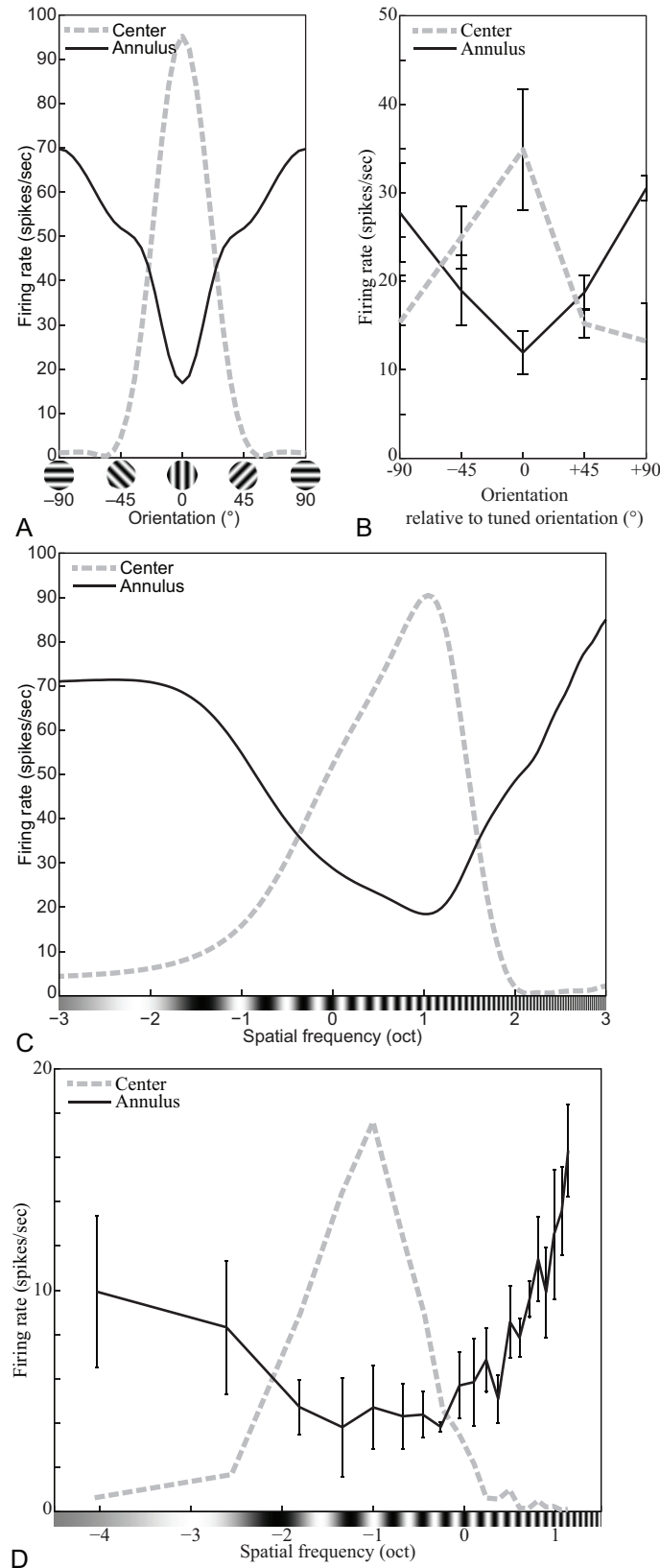


Figure 18: The effect of *surround suppression* on the response of (A, B) the model neuron with standard parameters, (C) a V1 complex cell (replotted from Cavanaugh et al., 2002b, Fig. 2B, anesthetized macaque, error bars = \pm SE), and (D) a V1 simple cell (replotted from C.-Y. Li & Li, 1994, Fig. 7C, anesthetized cat, error bars = \pm SE) in the orientation and the spatial-frequency domains. In all panels, the standard tuning function (as probed with no surrounding annulus) is plotted by the dashed grey line. The tuning function of the surround suppression is plotted by a solid black line. (A, C) The contrast of the center grating was 100% and its diameter was 0.81° (the measured RF). The contrast of the annulus grating were 100% and 5.76° . The frequency of the center and annular gratings was 2.0 cpd for (A). Their orientation was 0° for (C). See Phenomena 24 and 25 in Table 1.

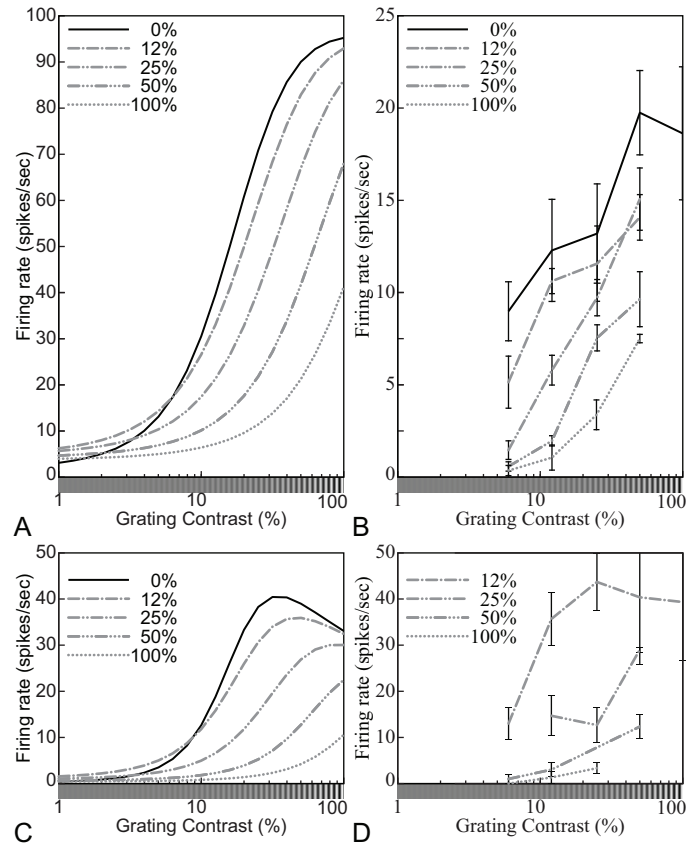


Figure 19: Surround-suppression effects on the contrast response functions for different contrasts of the annular grating. The contrast of the central disk is plotted on the horizontal axes, and the contrast of the annular surround is indicated in the legend. (A) The model neuron with the standard parameter set (cf. Table 2); (B) a complex cell in cat V1 (replotted from Carandini, 2004, Fig. 6); (C) the model neuron with a modified parameter set ($n_n = 2.8$, $n_d = 3.0$, $M = 10$); and (D) a complex cell in cat V1 (replotted from Carandini, 2004, Fig. 7). (A, C) The orientation and spatial-frequency of the center and annulus gratings were 0° and 2.0 cpd. The diameter of the center grating patch was 0.81° and it is equal to the measured RFs of the model neuron for both parameter sets. The size of the annulus grating was 5.76° . See Phenomenon 26 in Table 1.

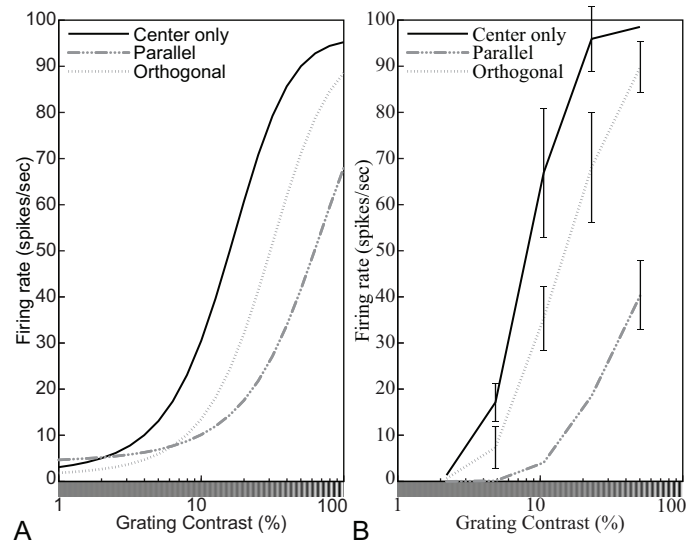


Figure 20: Surround-suppression effects on the contrast response functions for different orientations of the annular grating (indicated in the legend). (A) The DNM neuron with standard parameters and (B) a simple cell (replotted from Cavanaugh et al., 2002b, Fig. 5A, anesthetized macaque, error bars = \pm SE). (A) The orientation of the center grating was 0° . The spatial-frequency of the center and annulus gratings was 2.0 cpd. The size of the center grating patch was equal to the measured RF diameter (0.81°). The size of the annulus grating was 5.76° . (DNM = divisive normalization model. See Phenomenon 27 in Table 1.)

3.3.6 Receptive fields of simple cells

The RF of a real simple cell is composed of excitatory and inhibitory sub-regions. This composition has been mapped using local stimulus probes such as a light spot (Hubel & Wiesel, 1959; Volgushev, Vidyasagar, & Pei, 1996), a light bar (Andrews & Pollen, 1979), light and dark bars (Movshon et al., 1978b; Tadmor & Tolhurst, 1989; Kulikowski & Bishop, 1981; Kulikowski & Vidyasagar, 1986; Glezer, Tsherbach, Gauselman, & Bondarko, 1982), and the *reverse correlation* method (Nishimoto, Ishida, & Ohzawa, 2006; Ringach, 2002; Moore IV & Freeman, 2012; see also Gardner, Anzai, Ohzawa, & Freeman, 1999b). The classical RF of the DNM simple cell was also mapped using these methods (Figures 21 and 22, Phenomena 28 and 29). Note that a firing rate of a real simple cell becomes lower than its maintained discharge ($M[\beta]^{n_n}/\alpha^{n_d}$) when the light spot stimulates its inhibitory sub-regions (Hubel & Wiesel, 1959, Figure 21) or the light and dark bars stimulate its inhibitory and excitatory sub-regions respectively (Andrews & Pollen, 1979, Figure 23C). This effect can be observed with the DNM simple cell only when its maintained discharge in the absence of external stimulation is high enough to reveal the inhibitory effect of a light-spot probe. This occurs when $M[\beta]^{n_n}/\alpha^{n_d} \gg 0$.

Consider a hypothetical simple cell that acts as a purely linear filter of the visual stimulus. Under this linearity assumption (cf. Section 2.1), the bandwidths in the orientation and spatial-frequency domains can be derived from the 2D composition of its excitatory and inhibitory sub-regions in its RF in the image domain (Eq. 3 and 4; Graham, 1989; Lathi, 2005). However, the derived bandwidths of real simple cells tend to be

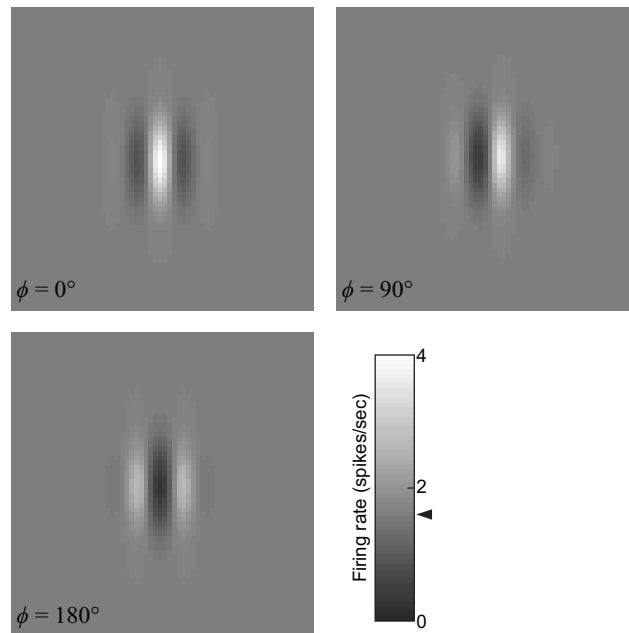


Figure 21: Receptive-field maps obtained with a light spot as a probe for three DNM simple cells with different phases ϕ . The grayscale levels indicate firing rates. The maintained discharge (the response to a uniform gray field) was the same for the three cells and is indicated by the triangle on the color bar. The size of the light-spot probe was $0.045^\circ \times 0.045^\circ$ (1 pixel). The stimulus background was uniform gray and the luminance of the spot was twice as high as the background gray. (DNM = divisive normalization model, standard parameters. See Phenomenon 28 in Table 1.)

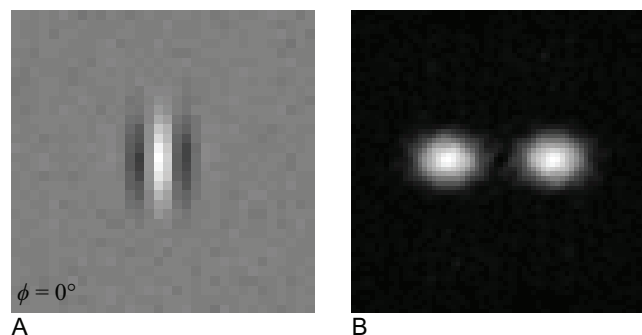


Figure 22: (A) The receptive field of a DNM simple cell obtained via the reverse correlation method. Brighter and darker regions represent excitatory and inhibitory regions, respectively. (B) The 2D Fourier spectrum energy distribution of the RF pattern in (A). The stimuli were 32×32 mosaics of random luminance spots (white noise). The size of the stimuli was 2.88° and that of the spots was 0.090° (2×2 pixels). (DNM = divisive normalization model, standard parameters. See Phenomenon 29 in Table 1.)

wider than these measured directly with gratings of various orientations and frequencies (Tadmor & Tolhurst, 1989; Gardner et al., 1999b; Nishimoto et al., 2006; Ringach, 2002; see also Section 3.3.3). These results suggest that real simple cells are non-linear. This discrepancy of the bandwidths is also observed with the model simple cell (Figures 22 and 23B). The derived bandwidths in the orientation and the frequency domains were 42.6° and 1.53 oct using the reverse correlation method and were wider than those measured with a grating (31.8° and 1.11 oct). This trend could be observed from the DNM cell if $n_n > 1$ (see Gardner et al., 1999a). Note that the derived bandwidths were rather close to the specified bandwidths of the stimulus drive (40° and 1.5 oct). Several other parameter sets (including $n_n < 1$) were tested and this similarity was observed reliably. Note that the stimulus probe itself could also affect the results. Recall that the RF could be measured with a bar as the probe. However, as the bars become wider, the derived bandwidth became wider and the tuned frequency derived was lower (Figure 23C). This underestimation of the tuned frequency was also observed in the physiological data (Tadmor & Tolhurst, 1989).

3.4 Summary and discussion

The results of our simulation experiments show that the model neuron based on the divisive normalization Equation 13 can account for many physiological phenomena (Table 1) with a standard parameter set (Table 2). A few other phenomena can also be accounted for, but they require customized parameter sets. Based on these simulation results and on mathematical analyses, we can make some falsifiable predictions that can be used to test the divisive normalization model (DNM).

All predictions involve probing a single neuron with multiple stimuli. The theoretical constraints stem from the fact that the model parameters must be fixed for each individual neuron. Thus, certain patterns are expected to occur together in the responses of a given neuron because they all depend on a single model parameter. In particular, the baseline parameter β in Equation 13 gives rise to interesting constraints. We have shown that the following three phenomena can be produced by the model only when β is sufficiently large:

(A) The inhibitory sub-regions of a simple cell probed with a single light spot (Hubel & Wiesel, 1959) can be observed only when $M[\beta]^{n_n}/\alpha^{n_d} \gg 0$. This condition ensures that the DNM simple cell has a substantial maintained discharge in the absence of external stimulation. This is necessary to reveal the inhibitory effect of a light-spot probe (Phenomenon 28, Figure 21).

(B) The supersaturation effect (Ph. 8, Fig. 10) can occur in the model only when $\beta > (1 + \alpha^{n_d})n_n/n_d - 1$ (Eq. 21). Recall from Section 3.3.2 that this effect refers to the non-monotonicity of the contrast sensitivity function of some V1 neurons. (Note that the majority of V1 neurons have monotonically increasing CFs.)

(C) Finally, the widening of the bandwidths in the orientation and in the spatial-frequency domains with

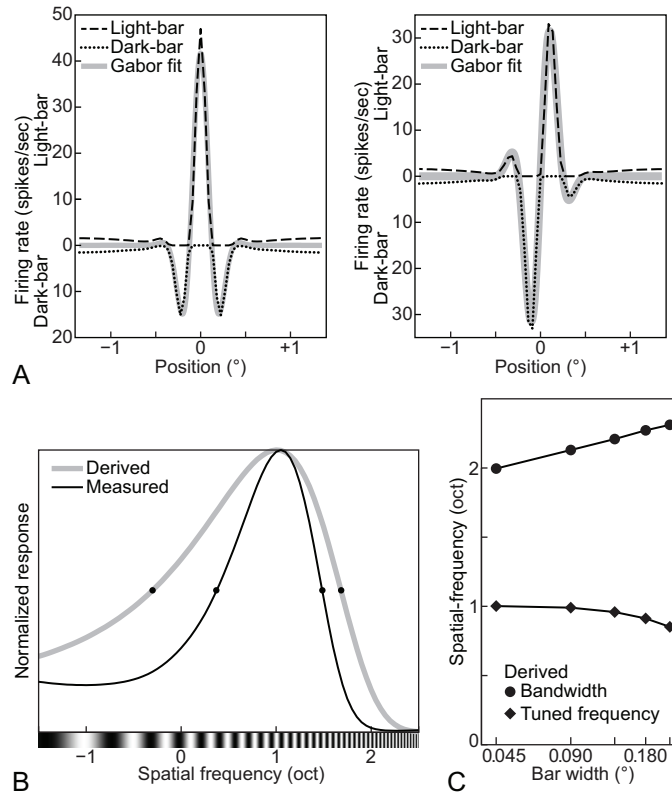


Figure 23: (A) A 1D distribution of the responses to light and dark bars parallel to the tuned orientation of the DNM simple cell. The abscissa of the graph represents the positions of the bars along an orientation perpendicular to the tuned orientation. The upper half of the graph is for the light bar and the lower half of the graph is for the dark bar. The best-fitting 1D Gabor pattern is superimposed (in gray) for comparison. The background of the stimulus was uniform gray and the luminance of the bars was 0 (dark) or twice as high as the background gray (light). The width of the bars was 0.045° (1 pixel). (B) The spatial-frequency tuning functions of the DNM simple cell derived (gray line) from the 1D pattern of its receptive field in (A) and measured (black line) using a grating. (C) The effect of the width of the bars on the tuned spatial-frequency derived from the measured RF as in (A). (DNM = divisive normalization model, standard parameters. See Phenomenon 30 in Table 1.)

decreasing stimulus contrast (Ph. 18, Fig. 14D, H, see also Ph. 17) can occur in the model only when $\beta > 0$.

Because of their dependence on a common parameter, the divisive normalization model predicts a correlation among these three phenomena if they are tested for each individual cell in a sample of V1 neurons recorded in a physiological experiment.

These interlocking patterns are methodologically important because they show that the divisive normalization model is constrained enough to be falsifiable even though it has 10 free parameters that give it the flexibility it needs to account for the diversity of V1 neurons. This novel theoretical result emerges from the examination of a comprehensive suite of phenomena within the framework of a single model with consistent parameters.

The model also provided alternative interpretations of a few phenomena observed in physiological studies. First, the tuning bandwidths of real simple cells are often narrower when measured directly in the orientation and the spatial-frequency domains compared to those derived from the mapped spatial patterns of the receptive fields (see Fig. 23). This departure from linearity has often been explained by the so-called *iceberg effect*, which occurs in our model when $\beta < 0$ (Carandini & Ferster, 2000; Sompolinsky & Shapley, 1997; Volgushev, Pernberg, & Eysel, 2000). The present simulations (Fig. 23) show that the model can produce such departures from linearity even when $\beta > 0$. Across its parameter range, the model produces departures from linearity due to the non-linear character of the numerator of the divisive-normalization equation (Gardner et al., 1999a).

Another phenomenon is the effect of stimulus size on the orientation- and frequency-bandwidths of real V1 neurons (see Fig. 13). The measured bandwidths depend on the size of the grating patch, and the size that maximizes the bandwidths is larger than the measured RF of the neuron. This occurs because RF mapping methods tend to underestimate the true extent of the weighting function (WF) of the linear filtering stage of the model neuron. The periphery of the WF of the model neuron is hidden by an annular region of the surround suppression, which narrows down the measured RF of the model neuron. This means that the stimulus patch must be larger than the measured RF to fill the true WF. Note that our simulations show that the surround suppression itself actually makes the bandwidths of the DNM neuron wider (see Section 3.3.3).

A few studies have shown that the size of the grating patch also bears some relationship to the preferred frequency of real neurons. The peak of the frequency tuning curve of real neurons tends to shift to a higher frequency as the size of the grating patch decreases (Teichert et al., 2007; Osaki et al., 2011). Also, the peak of the size tuning curve of the real neuron tends to shift to a smaller size if the spatial-frequency of the grating is higher than the tuned frequency of the neuron (Osaki et al., 2011). These trends were not observed with the DNM neuron used in the simulation experiments. Such trends, however, can be emulated by a model with some additional flexibility. For example, the suppressive drive of the DNM model can be modified so that the

suppression in the spatial-frequency domain changes depending on the eccentricity from the center of the RF. These trends can also be expected to be emulated by the modified DNM if the peak of the suppression in the spatial-frequency domain shifts to a higher frequency as eccentricity increases. Another possible modification involves the stimulus drive. Note that Naito and Sato (2015, August) proposed that if you want to model these trends, your model neurons should be composed of multiple Gabor filters. These filters can be tuned to different spatial-frequencies and the sizes of their WFs can be correlated with the tuned frequencies, i.e., smaller WFs are tuned to higher frequencies. Also, following the suggestions of Naito and Sato (2015, August), analogous modifications can be made to the DNM, composing its stimulus drive with multiple Gabor filters.

4 General Discussion

Our simulation experiments demonstrated that the divisive normalization model (DNM) can account for a comprehensive set of neurophysiological studies of both simple and complex cells in V1 (Table 1). Moreover, a mathematical analysis of Equation 13 predicts interdependence between certain observable phenomena. In the prior section, we explained how these predictions can be tested experimentally, which means that the standard formulation of the DNM specified in Section 2.3 is a falsifiable theory. If this formulation, especially Equation 13, is modified, the concrete mathematical results proven in the appendices will no longer apply. It seems plausible, however, that analogous results would hold for any formulation based on similar principles. In other words, our theoretical conclusions probably extend to the entire class of models based on a combination of linear filtering, half-wave rectification and squaring, and response normalization (Carandini et al., 2005). Our results ultimately rest on the fact that, although adjustable parameters are needed to accommodate the diversity across neurons, they must be fixed for each individual neuron. This fixedness gives rise to falsifiable constraints when a single neuron is probed with a judiciously chosen battery of stimuli.

With the standard parameter set in Table 2, the DNM qualitatively replicates most of the empirical phenomena listed in Table 1. Quantitative comparisons of the DNM with real neurons are also possible but, they would require quantitative estimation of all parameters of the DNM for the individual neurons from physiological data. The interpretability of such quantitative tests would depend on the precision of the estimation of the parameters. Finding the parameters that best fit the physiological data often involves the numerical optimization of an objective function. Depending on the data, it is possible that this function has multiple local minima that correspond to alternative parameter settings. When the parameter search procedure is trapped in a local optimum, it becomes difficult to compare the model predictions to empirical observations because different local minima would lead to different predictions. Such instability of the parameter estimation would jeopardize the status of the DNM even as a descriptive model unless the parameter search is guaranteed to find

the global optimum.

4.1 Correlation between the parameters of the model

The parameters of the model are conceptually independent from one another but physiological studies have shown correlations among some of them. These correlations can potentially decrease the number of independently adjustable parameters or limit the ranges of these parameters. This subsection briefly reviews some studies that suggest correlations among various DNM parameters.

The preferred spatial-frequency F^* of the divisive normalization model correlates with some parameters of the contrast response function (CF). Dean (1981) reported that the preferred spatial-frequencies of neurons in cat striate cortex correlates positively with their contrast threshold and negatively with the slope of their CF. Consider the hyperbolic-ratio model for simplicity (Eq. 10). Its mathematical analysis shows that the hyperbolic-ratio model can account for both correlation patterns above if the semisaturation contrast parameter α_{HB} correlates positively with the preferred frequency F^* across the population of model neurons (see Section 2.2). Some studies have shown that the bandwidths of V1 neurons tend to be narrower in neurons with higher preferred spatial frequencies (Kulikowski & Bishop, 1981; Kulikowski & Vidyasagar, 1986; R. L. De Valois, Albrecht, & Thorell, 1982; Yu et al., 2010). This effect can be emulated by the divisive normalization model if either β , n_d , or h_f decrease or α , n_n , or h_F increase as the tuned frequency F^* increases. (See Table 2 for the meaning of these symbols.) Note that the bandwidths in the spatial-frequency and orientation domains correlate with one another (Zhu, Xing, Shelley, & Shapley, 2010). Both of those bandwidths would be affected together if α , β , n_n , or n_d covary with F^* .

The receptive-field locations of neurons affect their orientation- and frequency tuning. The DNM can emulate this by introducing a correlation between the RF location parameters (x, y) and the tuning preference parameters Θ^* and F^* . First, as the retinal eccentricity of the positions increases, the tuned frequencies of real neurons decrease (Movshon et al., 1978b; Yu et al., 2010; Henriksson, Nurminen, Hyvärinen, & Vanni, 2008) and this correlation can be well represented well by the cortical magnification factor (Swindale, 1996; Daniel & Whitteridge, 1961; Tootell, Silverman, Switkes, & De Valois, 1982; Schwartz, 1980; Duncan & Boynton, 2003). Next, the observed distribution of tuned orientations has a preponderance of vertical or horizontal preferences for neurons with RFs in or near the fovea (R. L. De Valois, Yund, & Hepler, 1982; G. H. Henry, Dreher, & Bishop, 1974) and a preponderance of radial orientations for neurons in the visual periphery (Schall, Vitek, & Leventhal, 1986; see also Y. Sasaki et al., 2006, for fMRI results). The results of an fMRI study suggest that human V1 contains more vertically tuned neurons than horizontally tuned neurons (Yacoub, Harel, & Uğurbil, 2008). On the other hand, the number of the vertical and horizontal neurons found in cat striate

cortex are almost equal (B. Li, Peterson, & Freeman, 2003), whereas rat V1 apparently has a preponderance of neurons with horizontal preferences (Girman, Sauv e, & Lund, 1999). Note that the neurons tuned to vertical or horizontal orientations, and to high spatial-frequencies, have narrower bandwidths in the orientation domain when compared to neurons tuned to oblique orientations (Orban & Kennedy, 1981; B. Li et al., 2003).

4.2 The stimulus drive

The current formulation of the divisive normalization model specifies the weighting function of simple cells as a 2D Gabor filter (Eq. 5). There are, however, some other filters that can fit physiological data better than the Gabor filter does (see Stork & Wilson, 1990; Wallis, 2001, for reviews). All those models of the simple cell use a linear filtering stage (cf. Eq. 5) as their first processing step. This linear filtering predicts that the cell’s response will be maximal for a square wave grating with their tuned frequency, orientation, and phase.

Simple and complex cells are represented in qualitatively different ways in the current formulation of the DNM (see Section 2.1). Note, however, that it has been shown that these two types of neurons are better conceptualized as endpoints of a continuum rather than as a categorical distinction (Mechler & Ringach, 2002; Priebe, Mechler, Carandini, & Ferster, 2004; Martinez & Alonso, 2003; Chance, Nelson, & Abbott, 1999, 2000). Note that the original classification was based on 2D patterns of the classical RFs (Hubel & Wiesel, 1959, 1962, 1968; Bishop & Henry, 1972). Later proposals classified the neurons on the basis of the temporal modulation of their firing rates evoked by a drifting grating (e.g., Maffei & Fiorentini, 1973; Movshon et al., 1978c, 1978a; Andrews & Pollen, 1979). Simple cells tend to show greater modulation than complex cells. The classifications based on these two methods often agree with one another (R. L. De Valois, Albrecht, & Thorell, 1982; Skottun et al., 1991; Dean & Tolhurst, 1983; Mata & Ringach, 2005; Sengpiel et al., 1997; C. A. Henry & Hawken, 2013) but the agreement is not always complete (see Chen et al., 2009, for a review). There are always some neurons whose behaviors are intermediate between the “pure” types defined by either method (Van Kleef, Cloherty, & Ibbotson, 2010; Crowder, van Kleef, Dreher, & Ibbotson, 2007; Hietanen et al., 2013; Meffin, Hietanen, Cloherty, & Ibbotson, 2015; Mata & Ringach, 2005; Kagan, Gur, & Snodderly, 2002). These intermediate neurons would be difficult to emulate by models that make a categorical distinction between simple and complex cells. Also, the temporal modulation of the firing rates can be affected by other properties of the visual stimuli. For example, some complex cells show temporal modulation if the contrast of the drifting grating is low (Crowder et al., 2007; Van Kleef et al., 2010, see also C. A. Henry & Hawken, 2013). The temporal phase of the modulation also changes as a function of stimulus contrast (Albrecht, 1995). The temporal modulation of firing rate is also affected by stimulation of the surround regions of the neuron’s classical RF (Bardy, Huang, Wang, FitzGibbon, & Dreher, 2006). The effect of the surround stimulation on

the modulation is either facilitatory or suppressive depending on the individual neurons.

Neurons in V1 are tuned to gratings with particular orientations in static images and many of the neurons are selective to directions of drifting gratings (e.g., Livingstone, 1998; H. E. Jones et al., 2001). This *direction selectivity* could be modeled by a drifting 2D Gabor filter, that is, by a 2D Gabor filter whose phase changes over time (e.g., Heeger, 1993). This formulation of the direction selectivity predicts that the model responses would remain constant for a drifting grating whose orientation, spatial frequency, direction, and speed match the corresponding tuning preferences of the model neuron. Heeger (1993) showed that some other temporal properties of V1 neurons for drifting gratings can also be also emulated by integrating the drifting 2D Gabor filter and using it as the stimulus drive into the divisive normalization model. It is possible that many temporal properties of the neuronal responses to stimuli within their RFs can be well represented well by a relatively simple model that is based on divisive-normalization principles (see Albrecht et al., 2002, for an example).

4.3 The suppressive drive

The standard DNM formulation includes both surround- and cross-orientation suppression in a single *suppressive drive* (Eq. 16) for simplicity. Despite the fact that these two forms of suppression show some analogous properties, there is evidence that they should be attributed to different neurophysiological mechanisms (Sengpiel et al., 1998; Angelucci & Bullier, 2003). It has been suggested that cross-orientation suppression consists of a binocular factor arising from intracortical connections within V1 (B. Li, Peterson, Thompson, Duong, & Freeman, 2005; Endo, Kaas, Jain, Smith III, & Chino, 2000; Sengpiel & Vorobyov, 2005) and a monocular factor from thalamo-cortical connections between LGN and V1 (B. Li, Thompson, Duong, Peterson, & Freeman, 2006; DeAngelis et al., 1992; Priebe & Ferster, 2006; Smith, 2006; Freeman et al., 2002), but not from neurons in the LGN (Bauman & Bonds, 1991). It has also been suggested that surround suppression can come from LGN inputs to V1 (Webb, Dhruv, Solomon, Tailby, & Lennie, 2005), and from intracortical connections within V1 itself (Ozeki et al., 2009), and from feedback from higher cortical areas (Angelucci & Bullier, 2003; Nassi, Lomber, & Born, 2013; Bair et al., 2003; W. Li, Thier, & Wehrhahn, 2001; see also Angelucci & Bullier, 2003 for a review). In addition, the responses of some neurons are *facilitated* in a non-linear manner by stimuli outside their classical RFs (Fitzpatrick, 2000; C.-Y. Li & Li, 1994; Maffei & Fiorentini, 1976; Sillito et al., 1995; Cavanaugh et al., 2002b; Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998; H. E. Jones et al., 2001; Nelson & Frost, 1985), as well as by stimuli within the classical RFs (K. K. De Valois & Tootell, 1983). The facilitatory stimuli are different depending on the individual neurons. A judicious combination of surround suppression and surround facilitation can tune these neurons to more complex visual features (Grossberg & Mingolla, 1985; Fitzpatrick, 2000; Z. Li, 1998, 2000; Craft, Schütze, Niebur, & von der Heydt, 2007).

The various sources of suppression are expected to have different temporal properties. There is evidence that neuron-to-neuron signals propagate much faster along feedforward and feedback connections than along intracortical connections in V1 (Bringuier, Chavane, Glaeser, & Frégnac, 1999; Grinvald, Lieke, Frostig, & Hildesheim, 1994). This difference in temporal properties is expected to be larger for surround suppression because the temporal delay of the signal caused by the intracortical connection depends on the cortical distance, and the cortical distance depends on the distance between the corresponding points on the retina. The latency of the surround suppression becomes longer as the surround stimuli become further from the neurons' classical RFs (Bair et al., 2003). It has been shown that the feedback signal of the surround suppression from area MT is delayed less than 10 ms from the initial responses of the neurons (anesthetized macaque monkey: Hupé, James, Girard, Lomber, et al., 2001). The feedback signal from V2 and V3 is delayed by about 20 ms (alert macaque monkey: Nassi et al., 2013). The shorter latency of the surround suppression from MT can be attributed to a direct connection between the LGN and MT (Sincich, Park, Wohlgenuth, & Horton, 2004; Stepniewska, Qi, & Kaas, 1999).

In addition to the differences in their latencies, surround-suppression effects can be differentiated by their properties in different domains. Bair, Cavanaugh, and Movshon (2003) reported a negative correlation between strength and latency of the surround suppression across the neurons in their sample. The delay of the surround suppression from the initial response of the neurons could be as long as 30 ms for weak suppression, whereas it could be almost 0 ms (as early as the initial response) for strong suppression. Knierim and Van Essen (1992) tested the temporal properties of surround suppression while they controlled the orientation of the surround stimuli. Their results suggested that the surround suppression with a short delay (about 7 ms) was isotropic, but with a long delay (20 ms) it was selective to the orientation of the surround stimuli. Knierim and Van Essen argued that this isotropic surround suppression should be attributed to the neurons or to their processes in the LGN (or in the retina) because many of the subcortical neurons are insensitive to orientations of stimuli in their classical RFs or in their surround regions (see Bonin et al., 2005; Zaltsman, Heimel, & Van Hooser, 2015, for reviews).

Recall that the response of some neurons is facilitated by their surround stimuli (Fitzpatrick, 2000; C.-Y. Li & Li, 1994; Maffei & Fiorentini, 1976; Sillito et al., 1995; Cavanaugh et al., 2002b; Polat et al., 1998; H. E. Jones et al., 2001; Nelson & Frost, 1985). This facilitatory effect comes with a longer delay (40 ms) from the initial response than the facilitation produced by the suppressive effect (20 ms) (Hupé, James, Girard, & Bullier, 2001). It is worth pointing out that the delay of the suppressive effect (20 ms) observed by Hupé et al. (2001, anesthetized macaque monkey) is close to the delay observed in Nassi et al. (2013, alert macaque monkey). This suppressive effect was not affected by inactivating V2 (Hupé, James, Girard, & Bullier, 2001) but it was

weakened (or eliminated) by inactivating V2 and V3 (Nassi et al., 2013).

Some effects of the surround stimuli can be observed only by studying the temporal properties of the neuron's responses. Briggs and Usrey (2011) found that the initial response of the neurons becomes faster or slower by about 5 ms with the surround stimuli depending on the particular neuron. Also, the surround stimuli, themselves, can cause an excitatory signal and evoke firing responses in many neurons without any stimulation within the classical RFs with a very long latency (≥ 100 ms, W. Li et al., 2001; Rossi, Desimone, & Ungerleider, 2001). The latency of this excitatory effect from the surround stimuli is not affected by the distance of the surround stimuli from the RFs (Rossi et al., 2001). This constant latency of the excitatory effect suggests that this effect cannot be attributed to intracortical connections because the temporal delay of a signal mediated by these connections depends on the distance on the retina.

In sum, the surround suppression appears to be a combination of effects from different sources with different temporal and spatial properties. Cross-orientation suppression probably is similarly heterogeneous. It will be difficult to account for such complexity with a simple functional model.

4.4 Directions for future work

In this study, we considered only simple synthetic visual stimuli within a limited spatial context in a static (or steady-state) experimental condition. We did not assess the performance of the model under more ecologically-valid stimulation. There is a growing number of physiological studies that recorded responses from V1 neurons to photographs and to videos of real visual scenes (e.g., Touryan, Felsen, & Dan, 2005; Smyth et al., 2003), but Rust and Movshon (2005) warned about the interpretative difficulties inherent in the use of such complex stimuli. Furthermore, our review focused on relatively simple (first-order) properties of the stimuli, despite the fact that the responses of many real neurons can be affected by higher-order visual information including figure-ground organization (Hupé et al., 1998; Zhou, Friedman, & von der Heydt, 2000; Zhang & von der Heydt, 2010; Rossi et al., 2001), 3D context (Murray, Kersten, Olshausen, Schrater, & Woods, 2002; Murray, Boyaci, & Kersten, 2006), perceptual filling-in (Fiorani Jr., Rosa, Gattass, & Rocha-Miranda, 1992; Komatsu, 2006), and visual illusions (Ramsden, Hung, & Roe, 2001; see also Jancke, Chavane, Naaman, & Grinvald, 2004 for a brain imaging study). This means that no model can fully replicate the response of a neuron to natural stimuli unless such higher order information is taken into account. On the other hand, extracting such higher-order information from a retinal image is an open research problem. One practical way to circumvent this problem would be to prepare the higher-order information in advance and make it available to the model as its inputs from a top-down process.

The divisive normalization model (DNM) has also been used to fit physiological results of populations

of neurons in V1 (Busse et al., 2009; Goris et al., 2009). Note that modeling populations is different from modeling single cells, for the following reasons. First, a population may be able to process visual information better than individual neurons by forming a *population code* in which different neurons specialize in encoding different aspects of the input signal (deCharms & Zador, 2000; Pouget et al., 2003). For example, the neuronal population in V1 can systematically represent the second-order visual information while only a subset of the individual cells in V1 respond selectively to second-order stimuli (e.g., An et al., 2014). This complicates the derivation of model predictions about the population response because a given physiological phenomenon may have two possible explanations: one in terms of a population code and another in terms of the properties of individual neurons (e.g., Footnote 8 above). Another difficulty arises from the selective sampling of neurons in physiological experiments. It has been pointed out that recordings from some neurons are often excluded from the data set, producing a selection bias that can affect the experimental results (Olshausen & Field, 2005). So, the physiological results of the population can change depending on which neurons are included/excluded in the population. This can add extra parameters to the model of the population. On the other hand, the selection bias is less critical for single-cell recording because the model aims to account for the response properties of individual neurons. These differences suggest that different approaches may be required to model populations of neurons as opposed to modeling individual neurons. One fruitful area for future research is to explore the degree to which the divisive-normalization equation is applicable to both cases.

In conclusion, the divisive normalization model provides a useful functional characterization of the responses of simple and complex cells in primary visual cortex. It deserves to be designated as “the standard” model for many present purposes. We hope that the standard formulation proposed here, the standard parameter set, and the accompanying software implementation will facilitate future research based on this influential and successful model. Of course, the DNM will be supplanted and/or subsumed by future, more advanced models, just as it subsumed the “standard” linear model of the 1980s (Rust & Movshon, 2005). It seems particularly desirable to augment the DNM with mechanisms to account for the temporal dynamics of neuronal responses (e.g., Heeger, 1993; Brosch & Neumann, 2014). Accounting for this temporal dimension, however, poses interesting challenges. It is hard to see, for example, how a simple functional model can account for the temporal properties of surround suppression (and facilitation), which are attributable to several different sources as discussed above. Such an augmented model would need additional components and new adjustable parameters. Ideally, the components and parameters should have physiological interpretations that relate to physiological and/or anatomical data. Developing the DNM along those lines will bring it closer to a structural model.

5 Acknowledgement

The authors thank Kayla Higginbotham for her help in digitizing data from published figures.

6 Grants

This project was supported by NIH Research Grant R21 EY022745-01.

Appendix A A 2D Gabor filter and its tuning functions

Consider orientation and spatial-frequency tuning functions of a 2D Gabor filter (Equations 2 and 5) tuned to a spatial-frequency F (cycles/deg, cpd) and an orientation Θ ($^\circ$). The orientation tuning function of the 2D Gabor filter is cyclic and becomes maximal at Θ and $\Theta + 180^\circ$. The shapes of those two peaks are the same and the orientation bandwidth h_θ ($^\circ$) is defined as a width of each peak at its half-height. The orientation tuning function for a grating with spatial-frequency F is equivalent with a circular cross section of the 2D Fourier transform of the filter with a radius F . Note that a Fourier spectrum energy distribution of the 2D Gabor filter is a 2D Gauss distribution (see Figures 2.10 and 2.11 in Graham, 1989). Then, a relation between the orientation bandwidth and the bandwidth $h_{\tilde{y}}$ of the Gabor filter along θ can be written as follows (see Tables 2.2 and 2.4 in Graham, 1989):

$$h_{\tilde{y}} = \frac{4 \ln 2}{\pi F \frac{\pi h_\theta}{180}} = \frac{720 \ln 2}{\pi^2 F h_\theta} \quad (\text{A.1})$$

The spatial-frequency tuning function is unimodal and becomes maximal at F (cpd). Its bandwidth h_f (oct) is defined as a distance between two frequencies F_{low} and F_{high} (cpd) at the half-height of the function where $F_{low} < F_{high}$. Then, the relation between F , F_{low} , F_{high} , and h_f is:

$$\begin{cases} F_{low} + F_{high} = 2F \\ F_{high}/F_{low} = 2^{h_f} \end{cases} \quad (\text{A.2})$$

From Equation A.2, the bandwidth $F_{high} - F_{low}$ in cpd can be derived as follows:

$$F_{high} - F_{low} = 2F \frac{2^{h_f} - 1}{2^{h_f} + 1} \quad (\text{A.3})$$

Recall that the Fourier spectrum energy distribution of the 2D Gabor filter is the 2D Gauss function. The spatial-frequency tuning function of the filter for a grating with an orientation θ is equivalent with a straight cross section of the Gauss function along an orientation $\theta + 90^\circ$. Then, a relation between the frequency

bandwidth $F_{high} - F_{low}$ and the bandwidth $h_{\check{x}}$ of the Gabor filter for its tuned orientation Θ can be written as follows (see Tables 2.2 and 2.4 in Graham, 1989):

$$h_{\check{x}} = \frac{4 \ln 2}{\pi (F_{high} - F_{low})} \quad (\text{A.4})$$

From Equations A.3 and A.4, a relation between $h_{\check{x}}$ and h_f can be written as follows:

$$\begin{cases} h_f = \log_2 \frac{\pi F h_{\check{x}} + 2 \ln 2}{\pi F h_{\check{x}} - 2 \ln 2} \\ h_{\check{x}} = \frac{(2^{h_f} + 1) 2 \ln 2}{(2^{h_f} - 1) \pi F} \end{cases} \quad (\text{A.5})$$

Appendix B The Contrast Response Function of the hyperbolic-ratio model along the linear contrast axis

We analyze the contrast response function (CF) of the hyperbolic-ratio model (Equation 10) on a linear contrast axis. The CF is a function of the luminance contrast c of a grating stimulus. Note that the orientation and spatial-frequency of the grating are consistent with the tuning of the model and the size of the grating is large enough to fill the entire receptive field of the model.

The model becomes most sensitive to a change in luminance contrast c where the slope of the CF becomes maximal (see Itti et al., 2000; Wilson, 1980, see also Section 3.3.2). The maximal CF slope is:

$$\frac{(n_{HB} + 1)^{\frac{n_{HB}+1}{n_{HB}}} (n_{HB} - 1)^{\frac{n_{HB}-1}{n_{HB}}}}{4\alpha_{HB}n_{HB}} \quad (\text{B.1})$$

This maximum occurs for contrast

$$c^* = \alpha_{HB} \left(\frac{n_{HB} - 1}{n_{HB} + 1} \right)^{\frac{1}{n_{HB}}} \quad (\text{B.2})$$

if $n_{HB} > 1$ and for $c^* = 0$ otherwise (Figure 24). If $n_{HB} > 1$, the CF is convex downward around $c = 0$. This trend of the model CF can account for the shapes of CFs of real neurons around $c = 0$ (Albrecht & Hamilton, 1982; Albrecht et al., 2003).

Some other studies (e.g., Tadmor & Tolhurst, 1989; Dean, 1981; Heeger, 1992a) represent the shapes of the CFs of real neurons around $c = 0$ using a piecewise-linear model (Section 2.1) with a threshold β_{IB} :

$$R_{IB}(I) = M_{IB}[\beta_{IB} + E_{\mathbf{P}_i}(I)], \quad (\text{B.3})$$

where $\beta_{IB} < 0$. Note that CFs of real neurons often start to saturate at high contrasts (see Figure 4 for CFs of real neurons plotted on the log contrast axis). This trend can also be captured by the hyperbolic-ratio model

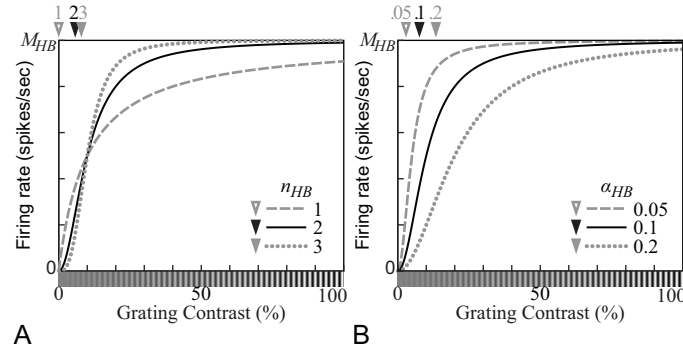


Figure 24: Contrast response functions (CFs) produced by the hyperbolic-ratio model (Equation 10) on the linear-contrast axis. Positions of the maximal slopes of the CFs are indicated by triangles. (A) The CFs with three different values of the exponent parameter n_{HB} . (B) The CFs with three different values of the *semisaturation contrast* parameter α_{HB} . A stimulus with contrast α_{HB} elicits one-half of the saturation level M_{HB} . ($\alpha_{HB} = 0.1$ for panel A; $n_{HB} = 2$ for panel B.)

(Figure 24, see also Figure 4 for the CFs of the hyperbolic-ratio model on the log contrast axis) but cannot be captured by the linear model.

Appendix C The shape of the contrast response function of the divisive normalization model

The shape of the contrast response function (CF) of the divisive normalization model (DNM) is mathematically analyzed in this appendix. Consider the DNM (Equation 13) and a grating $g(c)$ with contrast c , the model's tuned orientation and frequency (and phase for a simple cell), and spatial extent large enough to fill both the entire receptive and suppressive fields of the model. Then, the CF of the DNM (Equation 13) for $g(c)$ is:

$$R(g(c)) = M \frac{[\beta + c]^{n_n}}{\alpha^{n_d} + c^{n_d}} \quad (\text{C.1})$$

(Note that this equation is equivalent to Equation 15 in Section 2.3.) The first derivative of Equation C.1 with respect to c is:

$$\frac{d}{dc} R(g(c)) = \frac{(c^{n_d}(n_n - n_d) - \beta n_d c^{n_d-1} + n_n \alpha^{n_d})}{M^{-1}(\alpha^{n_d} + c^{n_d})^2 (\beta + c)^{1-n_n}} \quad (\text{C.2})$$

for $[-\beta] \leq c \leq 1$. Note that $dR(g(c))/dc > 0$ at $c = 0$ if $\beta > 0$ and at $c \rightarrow -\beta + 0$ if $\beta \leq 0$. This means the CF is increasing at low contrasts. The numerator of Equation C.2 is further differentiated:

$$\frac{d}{dc} (c^{n_d}(n_n - n_d) - \beta n_d c^{n_d-1} + n_n \alpha^{n_d}) = c^{n_d-2} n_d (c(n_n - n_d) - \beta(n_d - 1)) \quad (\text{C.3})$$

The derivative of the numerator becomes 0 at $c = 0$ and $c = c_{ex}$ where

$$c_{ex} = \beta \frac{n_d - 1}{n_n - n_d} \quad (\text{C.4})$$

Hence, the numerator has two local extrema: $n_n \alpha^{n_d}$ at $c = 0$ and

$$-\beta_{n_d} \frac{(n_d - 1)^{n_d - 1}}{(n_n - n_d)^{n_d - 1}} + n_n \alpha^{n_d} \quad (\text{C.5})$$

at $c = c_{ex}$. Besides, the denominator of Equation C.2 is always positive between $[-\beta]$ and 1, exclusive. Then, the possible shapes of the contrast sensitivity curves between $[-\beta]$ and 1 can be categorized into the following three types.

- (1) The sensitivity curve is unimodal and convex-upward if $dR(g(c))/dc < 0$ at $c = 1$. This condition is satisfied if β is large enough:

$$\frac{n_n}{n_d} (1 + \alpha^{n_d}) - 1 < \beta \quad (\text{C.6})$$

The CF is decreasing at high contrasts and must have a local maximum between $[-\beta]$ and 1, exclusive. Namely, the curve shows the super-saturation effect. This is the sufficient condition for the super-saturation effect for $g(c)$.

- (2) The CF is with both local maximum and minimum if $dR(g(c))/dc > 0$ at $c = 1$, $dR(g(c))/dc < 0$ at $c = c_{ex}$ (see Equation C.5), and $[-\beta] < c_{ex} < 1$. The local maximum appear between $[-\beta]$ and c_{ex} and the local minimum appears between c_{ex} and 1. Note that it may not be practically well-conditioned to judge those conditions based on physiological results. Many physiological studies have shown that the DNM can fit data well under an assumption that $n_n = n_d$. Then, $n_n - n_d$ can be too small compared with $n_d - 1$ to have stable estimates of c_{ex} and $dR(g(c))/dc$ at $c = c_{ex}$ (Equation C.5).

- (3) The CF is monotonically increasing otherwise.

Appendix D Contrast response function for a non-preferred stimulus

Consider the divisive normalization model (DNM, Equation 13) and a grating $g(c)$ with contrast c but with non-preferred orientation, spatial-frequency, or size. The stimulus drive (Equations 5 and 7) and the suppressive drive are still proportional to c between $0 \leq c \leq 1$ because $g(c) = cg(1)$. Then, the contrast response function

(CF) of the DNM (Equation 13, see also Equation C.1 above) for $g'(c)$ is:

$$R(g'(c)) = M \frac{[\beta + pc]^{n_n}}{\alpha^{n_d} + q^{n_d} c^{n_d}} \quad (\text{D.1})$$

where p and q are constants. Equation D.1 can be modified as follows:

$$R(g'(c)) = M \frac{[\beta + \frac{p}{q} 10^{(\log_{10} c + \log_{10} q)}]^{n_n}}{\alpha^{n_d} + 10^{n_d(\log_{10} c + \log_{10} q)}} \quad (\text{D.2})$$

Equation D.2 shows that a relation between plots of $R(g'(c))$ and $R(g(c))$ (Equation C.1) can be represented as a subset of affine transformation with c in a logarithmic scale and β sufficiently smaller than pc . The subset is translation $-\log_{10} q$ along the contrast axis and scaling $(p/q)^{n_n}$ along the response axis.

Appendix E Computational implementation of the divisive normalization model

Computational implementation of the divisive normalization model (DNM, Equation 13) allows the model to be applied to any visual stimulus used in physiological and also psychophysical studies. Its computational efficiency and practical usage should be also taken into consideration.

The suppressive drive of the DN model (Equation 16) is computationally demanding. It is represented as a weighted sum of $E_{C:X_i Y_i F_i \Theta_i}(I)^{n_d}$ (Equation 7) in the orientation, spatial frequency, and 2D retinal space domains, where $E_{C:X_i Y_i F_i \Theta_i}(I)$ is a square root of a sum of $E_{S:X_i Y_i F_i \Theta_i, 0^\circ}(I)^2$ and $E_{S:X_i Y_i F_i \Theta_i, 90^\circ}(I)^2$ (Equation 5). A part of the weighted sum of the suppressive drive in the 2D retinal space domain can be computed using 2D cross-correlation of a retinal image $I(x, y)$ (Equation 1) with two 2D Gabor filters $G_{X_i Y_i F_i \Theta_i \Phi_i}$ in Equation 2:

$$\sum_{X_i, Y_i} w_{X_i Y_i} E_{C:X_i Y_i F_i \Theta_i}(I)^{n_d} = \iint w_{X_i Y_i} \left((I(x, y) \otimes G_{X_i Y_i F_i \Theta_i, 0^\circ}(x, y))^2 + (I(x, y) \otimes G_{X_i Y_i F_i \Theta_i, 90^\circ}(x, y))^2 \right)^{\frac{n_d}{2}} dx dy \quad (\text{E.1})$$

where an operator \otimes represents a cross-correlation⁹. This is equivalent with computing $E_{C:X_i Y_i F_i \Theta_i}$ at every pixel of the input image. Then, the number of the 2D spatial sampling $N_X \times N_Y$ becomes equivalent with that of pixels in the filtered image.

⁹If the resolution of the image I is high enough, the following approximation also works:

$$\sum_{X_i, Y_i} w_{X_i Y_i} E_{C:X_i Y_i F_i \Theta_i}(I)^{n_d} \approx \iint w_{X_i Y_i} |I(x, y) \otimes G_{X_i Y_i F_i \Theta_i \Phi_i}(x, y)|^{n_d} dx dy$$

where Φ_i is arbitrary.

Note that the computational cost of the cross-correlation is an order of $O(N_X N_Y \log(N_X N_Y))$ and it has to be processed for $N_\Theta \times N_F$ times to compute the suppressive drive. Then, the total cost becomes $O(N_I N_\Theta N_F N_X N_Y \log(N_X N_Y))$ if the number of the input images is N_I . This process can be optimized using the convolution theory. Namely, the image filters in Equation 2 are applied to the input image in a 2D Fourier space:

$$I(x, y) \otimes G_{X_i Y_i F_i \Theta_i \Phi_i}(x, y) = \mathcal{F}^{-1}(\mathcal{F}(I(x, y))^* \mathcal{F}(G_{X_i Y_i F_i \Theta_i \Phi_i}(x, y))) \quad (\text{E.2})$$

where operators \mathcal{F} and \mathcal{F}^{-1} represent a Fourier transform and its inverse and $*$ represents a conjugation. The Fourier transforms of the image filters $G_{XYF\Theta\Phi}$ can be computed in advance and can be used for processing every image. Then, its total cost is $O(N_I N_X N_Y \log(N_X N_Y))$.

Note that Equations E.1 and E.2 become approximations because the sizes of I and $G_{XYF\Theta\Phi}$ are practically both finite. The approximation becomes better if I and $G_{XYF\Theta\Phi}$ are surrounded by sufficiently large regions with uniform gray. The size of I with the surrounding region should be the same as that of $G_{XYF\Theta\Phi}$ and be $2^{m_x} \times 2^{m_y}$ where m_x and m_y are positive integers. It allows the model to process the image efficiently using the Fast Fourier Transform algorithm in Equation E.2.

The process is further optimized so that the model can compute the suppressive drives of multiple “virtual neurons” to the single input image at once. The results of the image filtering process in Equation E.2 can be shared among the virtual neurons. It generates a set of the $N_\Theta \times N_F \times N_X \times N_Y$ different channels. The suppressive drives of the multiple virtual neurons can be computed from the same set of the channels with different distributions of the weights w_{Θ_i} , w_{F_i} , and $w_{X_i Y_i}$.

In the process of computing the suppressive drives, $N_\Theta \times N_F \times N_X \times N_Y \times 2$ different $E_{S:XYF\Theta\Phi}$ and $N_\Theta \times N_F \times N_X \times N_Y$ different $E_{C:XYF\Theta}$ are generated. They can be also used as the stimulus drives of the virtual neurons. Note that we practically need a limited number N_C , instead of $N_X \times N_Y$, of the 2D spatial positions of the virtual neurons, especially when the resolution of the input image is high. Also, not all N_F spatial-frequencies are appropriate for the stimulus drives because of the edge effect. The domain of spatial-frequency ranges between $\pm\infty$ but $E_{S:XYF\Theta\Phi}$ and $E_{C:XYF\Theta}$ can be sampled within its limited range. Consider virtual neurons whose tuned frequencies are on the upper (or lower) bound of the sampling range in the domain. They cannot be suppressed by the channels whose tuned frequencies are higher (or lower) than the tuned frequency of the neurons. The edge effect can be lightened by removing some neurons whose tuned frequencies are close to the bounds of the sampling range. Hence, there are $N_\Theta \times (N_F - \Delta_F) \times N_C \times 3$ virtual neurons generated in total: $N_\Theta \times (N_F - \Delta_F) \times N_C$ complex cells and $N_\Theta \times (N_F - \Delta_F) \times N_C \times 2$ simple cells (see Equation 7), where Δ_F is the number of the removed neurons for the edge effect.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America, A 2(2), 284–299.
- Albrecht, D. G. (1995). Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. Visual Neuroscience, 12, 1191–1210.
- Albrecht, D. G., De Valois, R. L., & Thorell, L. G. (1980). Visual cortical neurons: Are bars or gratings the optimal stimuli? Science, 207(4 Jan), 88–90.
- Albrecht, D. G., & Geisler, W. S. (1991). Motion sensitivity and the contrast-response function of simple cells in the visual cortex. Visual Neuroscience, 7, 531–546.
- Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2003). Nonlinear Properties of Visual Cortex Neurons: Temporal Dynamics, Stimulus Selectivity, Neural Performance. In L. Chalupa & J. Werner (Eds.), The visual neurosciences (pp. 747–764). Boston, MA: MIT Press.
- Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. Journal of Neurophysiology, 88(2), 888–913.
- Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. Journal of Neurophysiology, 48, 217–237.
- Alitto, H. J., & Usrey, W. M. (2004). Ferret Primary Visual Cortex. Journal of Neurophysiology, 91, 2797–2808.
- Allison, J. D., & Bonds, A. B. (1994). Inactivation of the infragranular striate cortex broadens orientation tuning of supragranular visual neurons in the cat. Experimental Brain Research, 101(3), 415–426.
- An, X., Gong, H., Yin, J., Wang, X., Pan, Y., Zhang, X., ... Wang, W. (2014). Orientation-cue invariant population responses to contrast-modulated and phase-reversed contour stimuli in macaque V1 and V2. Plos One, 9, (9):e10675. doi: doi:10.1371/journal.pone.0106753
- Anderson, J. S., Lampl, I., Gillespie, D. C., & Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. Science, 290(5498), 1968–1972. doi: 10.1126/science.290.5498.1968
- Andoni, S., Tan, A., & Priebe, N. J. (2013). The cortical assembly of visual receptive fields. In J. S. Werner & L. M. Chalupa (Eds.), The new visual neurosciences (pp. 367–379). Boston, MA: MIT Press.
- Andrews, B. W., & Pollen, D. A. (1979). Relationship between spatial frequency selectivity and receptive field profile of simple cells. Journal of Physiology, 287, 163–176.
- Angelucci, A., & Bullier, J. (2003). Reaching beyond the classical receptive field of V1 neurons: Horizontal or feedback axons? Journal of Physiology Paris, 97(2-3), 141–154. doi: 10.1016/j.jphysparis.2003.09.001
- Angelucci, A., & Shushruth, S. (2013). Beyond the classical receptive field: Contextual modulation in primary

- visual cortex. In J. S. Werner & L. M. Chalupa (Eds.), The new visual neurosciences (pp. 425–444). Boston, MA: MIT Press.
- Bair, W. (2005). Visual receptive field organization. Current Opinion in Neurobiology, 15(4), 459–64. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16023850> doi: 10.1016/j.conb.2005.07.006
- Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. Journal of Neuroscience, 23(20), 7690–7701. doi: 10.1523/JNEUROSCI.2320-03.2003
- Bardy, C., Huang, J.-Y., Wang, C., FitzGibbon, T., & Dreher, B. (2006). ‘Simplification’ of responses of complex cells in cat striate cortex: Suppressive surrounds and ‘feedback’ inactivation. Journal of Physiology, 574, 731–750. doi: 10.1113/jphysiol.2006.110320
- Bauman, L. A., & Bonds, A. B. (1991). Inhibitory refinement of spatial frequency selectivity in single cells of the cat striate cortex. Vision Research, 31(6), 933–944. doi: 10.1016/0042-6989(91)90201-F
- Ben-Yishai, R., Bar-Or, L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. Proceedings of the National Academy of Sciences, USA, 92, 3844–3848.
- Bereshpolova, Y., Stoelzel, C. R., Zhuang, J., Amitai, Y., Alonso, J.-M., & Swadlow, H. A. (2011). Getting Drowsy ? Alert / Nonalert Transitions and Visual Thalamocortical Network Dynamics. Journal of Neuroscience, 31(48), 17480–17487. doi: 10.1523/JNEUROSCI.2262-11.2011
- Bishop, P. O., & Henry, G. H. (1972). Striate neurons: Receptive field concepts. Investigative Ophthalmology, 11(5), 346–354.
- Blakemore, C., & Tobin, E. A. (1972). Lateral inhibition between orientation detectors in the cat’s visual cortex. Experimental Brain Research, 15, 439–440.
- Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Visual Neuroscience, 2, 41–55.
- Bonin, V., Mante, V., & Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. Journal of Neuroscience, 25(47), 10844–10856.
- Briggs, F., & Usrey, W. M. (2011). Distinct mechanisms for size tuning in primate visual cortex. Journal of Neuroscience, 31(35), 12644–12649. doi: 10.1523/JNEUROSCI.2268-11.2011
- Bringuiet, V., Chavane, F., Glaeser, L., & Frégnac, Y. (1999). Horizontal Propagation of Visual Activity in the Synaptic Integration Field of Area 17 Neurons. Science, 283(5402), 695–699. doi: 10.1126/science.283.5402.695
- Brosch, T., & Neumann, H. (2014). Computing with a canonical neural circuits model with pool normalization and modulating feedback. Neural Computation, 26, 2735–2789.
- Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity

- in visual cortex. Neuron, 64, 931-942.
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. Journal of Physiology, 197, 551-566.
- Carandini, M. (2004). Receptive fields and suppressive fields in the early visual system. In M. S. Gazzaniga (Ed.), The cognitive neurosciences III (pp. 313-326). Cambridge, MA: MIT Press.
- Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., ... Rust, N. C. (2005). Do we know what the early visual system does? Journal of Neuroscience, 25(46), 10577-10597.
- Carandini, M., & Ferster, D. (2000). Membrane potential and firing rate in cat primary visual cortex. Journal of Neuroscience, 20(1), 470-484.
- Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in visual cortex. Science, 264, 1333-1336.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. Nature Reviews Neuroscience, 13(1), 51-62.
- Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. Journal of Neuroscience, 17(21), 8621-8644.
- Carandini, M., Heeger, D. J., & Movshon, J. A. (1999). Linearity and gain control in V1 simple cells. In P. S. Ulinski, E. G. Jones, & A. Peters (Eds.), Cerebral cortex, vol. 13: Models of cortical circuits (pp. 401-443). New York: Kluwer Academic/Plenum Publishers.
- Carandini, M., Heeger, D. J., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. Journal of Neuroscience, 22, 100053-10065.
- Cauchy-Schwarz inequality. (n.d.). Encyclopedia of mathematics. Retrieved from http://www.encyclopediaofmath.org/index.php?title=Cauchy-Schwarz_inequality
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and Interaction of Signals From the Receptive Field Center and Surround in Macaque V1 Neurons. Journal of Neurophysiology, 88(5), 2530-2546. Retrieved from <http://jn.physiology.org/cgi/doi/10.1152/jn.00692.2001> doi: 10.1152/jn.00692.2001
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and Spatial Distribution of Signals From the Receptive Field Surround in Macaque V1 Neurons. Journal of Neurophysiology, 88(5), 2547-2556. Retrieved from <http://jn.physiology.org/cgi/doi/10.1152/jn.00693.2001> doi: 10.1152/jn.00693.2001
- Chance, F. S., & Abbott, L. F. (2000). Divisive inhibition in recurrent networks. Network: Computation in Neural Systems, 11, 119-129.

- Chance, F. S., Nelson, S. B., & Abbott, L. (2000). A recurrent network model for the phase invariance of complex cell responses. Neurocomputing, 32-33(October 1999), 339–344. doi: 10.1016/S0925-2312(00)00183-1
- Chance, F. S., Nelson, S. B., & Abbott, L. F. (1999). Complex cells as cortically amplified simple cells. Nature Neuroscience, 2(3), 277–282.
- Chen, Y., Martinez-conde, S., Macknik, S. L., Swadlow, H. A., & Alonso, J.-m. (2009). The linearity and selectivity of neuronal responses in awake visual cortex. Journal of Vision, 9(9), 12:1–17. doi: 10.1167/9.9.12.Introduction
- Cornsweet, T. N. (1970). Visual perception. New York: Harcourt College Publishers.
- Craft, E., Schütze, H., Niebur, E., & von der Heydt, R. (2007). A Neural Model of Figure-Ground Organization. Journal of Neurophysiology, 97(6), 4310–4326. Retrieved from <http://jn.physiology.org/cgi/doi/10.1152/jn.00203.2007> doi: 10.1152/jn.00203.2007
- Crowder, N. A., van Kleef, J., Dreher, B., & Ibbotson, M. R. (2007). Complex cells increase their phase sensitivity at low contrasts and following adaptation. Journal of Neurophysiology, 98(3), 1155–1166. doi: 10.1152/jn.00433.2007
- Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. The Journal of physiology, 159(2), 203–221. doi: 10.1113/jphysiol.1961.sp006803
- Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. Vision Research, 20, 847–856.
- Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America. A, Optics and image science, 2(7), 1160–1169. doi: 10.1364/JOSAA.2.001160
- De Valois, K. K., De Valois, R. L., & Yund, E. W. (1979). Responses of striate cortex cells to grating and checkerboard patterns. Journal of Physiology (London), 291, 483–505.
- De Valois, K. K., & Tootell, R. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. Journal of Physiology, 336, 359–376.
- De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. Vision Research, 22, 545–559.
- De Valois, R. L., & De Valois, K. K. (1988). Spatial vision. New York: Oxford University Press.
- De Valois, R. L., Thorell, L. G., & Albrecht, D. G. (1985). Periodicity of striate-cortex-cell receptive fields. Spatial Vision, 2, 1115–1123.
- De Valois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. Vision Research, 22, 531–544.

- Dean, A. F. (1981). The relationship between response amplitude and contrast for cat striate cortical neurones. Journal of Physiology, 318, 413–427.
- Dean, A. F., & Tolhurst, D. J. (1983). On the distinctness of simple and complex cells in the visual cortex of the cat. Journal of Physiology, 344, 305–325.
- DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. Journal of Neurophysiology, 71(1), 347–374.
- DeAngelis, G. C., Robson, J., Ohzawa, I., & Freeman, R. D. (1992). The organization of suppression in receptive fields of neurons in the cat's visual cortex. Journal of Neurophysiology, 68, 144–163.
- deCharms, R. C., & Zador, A. (2000). Neural representation and the cortical code. Annual Review of Neuroscience, 23, 613–647.
- Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. Journal of Physiology, 357, 219–240. doi: 10.1113/jphysiol.1984.sp015498
- Disney, A. A., Aoki, C., & Hawken, M. J. (2007). Gain Modulation by Nicotine in Macaque V1. Neuron, 56(4), 701–713. doi: 10.1016/j.neuron.2007.09.034
- Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. Neuron, 38(4), 659–671. doi: 10.1016/S0896-6273(03)00265-4
- Ecker, A., Berens, P., Cotton, R., Subramaniyan, M., Denfield, G., Cadwell, C., ... Tolias, A. (2014). State Dependence of Noise Correlations in Macaque Primary Visual Cortex. Neuron, 82(1), 235–248. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0896627314001044> doi: 10.1016/j.neuron.2014.02.006
- Ellias, S. A., & Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short-term memory of shunting on-center off-surround networks. Biological Cybernetics, 20, 69–98.
- Emerson, R. C., Korenberg, M. J., & Citron, M. C. (1992). Identification of complex-cell intensive nonlinearities in a cascade model of cat visual cortex. Biological Cybernetics, 66, 291–300.
- Endo, M., Kaas, J. H., Jain, N., Smith III, E. L., & Chino, Y. (2000). Binocular Cross-Orientation Suppression in the Primary Visual Cortex (V1) of Infant Rhesus Monkeys. Investigative Ophthalmology & Visual Science, 41(12), 4022–4031.
- Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. Annual Review of Neuroscience, 23, 441–471.
- Field, D. J., & Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. Proceedings of the Royal Society of London. Series B, 228, 379–400.
- Fiorani Jr., M., Rosa, M. G. P., Gattass, R., & Rocha-Miranda, C. E. (1992). Dynamic surrounds of receptive

- fields in primate striate cortex: A physiological basis for perceptual completion? Proceedings of the National Academy of Sciences, USA, 89, 8547–8551.
- Fisher, N. I. (1996). Statistical Analysis of Circular Data. Cambridge, MA: Cambridge University Press.
- Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. Current Opinion in Neurobiology, 10(4), 438–443. doi: 10.1016/S0959-4388(00)00113-6
- Freeman, T. C., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppression without inhibition in visual cortex. Neuron, 35, 759–771.
- Gardner, J. L., Anzai, A., Ohzawa, I., & Freeman, R. D. (1999a). Linear and nonlinear contributions to orientation tuning of simple cells in the cat's striate cortex. Visual Neuroscience, 16, 1115–1121.
- Gardner, J. L., Anzai, A., Ohzawa, I., & Freeman, R. D. (1999b). Linear and nonlinear contributions to orientation tuning of simple cells in the cat's striate cortex. Visual Neuroscience, 16(6), 1115–1121. doi: 10.1017/S0952523899166112
- Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. Visual Neuroscience, 14, 897–919. doi: 10.1017/S0952523800011627
- Gieselmann, M. A., & Thiele, A. (2008). Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque v1. European Journal of Neuroscience, 28, 447–459.
- Girman, S. V., Sauvé, Y., & Lund, R. D. (1999). Receptive field properties of single neurons in rat primary visual cortex. Journal of Neurophysiology, 82(1), 301–311.
- Glezer, V. D., Tsherbach, T. A., Gauselman, V. E., & Bondarko, V. M. (1982). Spatio-temporal organization of receptive fields of the cat striate cortex. Biological Cybernetics, 43, 35–49.
- Goltstein, P. M., Montijn, J. S., & Pennartz, C. M. a. (2015). Effects of Isoflurane Anesthesia on Ensemble Patterns of Ca²⁺ Activity in Mouse V1: Reduced Direction Selectivity Independent of Increased Correlations in Cellular Activity. Plos One, 10(2), e0118277. Retrieved from <http://dx.plos.org/10.1371/journal.pone.0118277> doi: 10.1371/journal.pone.0118277
- Goris, R. L. T., Wichmann, F. A., & Henning, G. B. (2009). A neurophysiologically plausible population code model for human contrast discrimination. Journal of Vision, 9(7), 15:1–22. doi: 10.1167/9.8.1004
- Graham, N. V. (1989). Visual pattern analyzers. New York: Oxford University Press.
- Graham, N. V. (1992). Breaking the visual stimulus into parts. Current Directions in Psychological Science, 1(2), 55–61.
- Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. Vision Research, 51, 1997–1430.

- Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. The Journal of Neuroscience, 14(5), 2545–2568.
- Grossberg, S. (1973). Contour Enhancement, Short Term Memory, and Constancies in Reverberating Neural Networks. Studies in Applied Mathematics, 52(3), 213–257.
- Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. Neural Networks, 1, 17–61.
- Grossberg, S., & Mingolla, E. (1985). Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading. Psychological Review, 92(2), 173–211.
- Heeger, D. J. (1992a). Half-squaring in responses of cat striate cells. Visual Neuroscience, 9, 427–443.
- Heeger, D. J. (1992b). Normalization of cell responses in cat striate cortex. Visual Neuroscience, 9, 181–197.
- Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. Journal of Neurophysiology, 70(5), 1885–1898.
- Henriksson, L., Nurminen, L., Hyvärinen, A., & Vanni, S. (2008). Spatial frequency tuning in human retinotopic visual areas. Journal of Vision, 8(10), 5:1–13. doi: 10.1167/8.10.5.Introduction
- Henry, C. A., & Hawken, M. J. (2013). Stability of simple/complex classification with contrast and extraclassical receptive field modulation in macaque V1. Journal of Neurophysiology, 109(7), 1793–1803. Retrieved from <http://jn.physiology.org/cgi/doi/10.1152/jn.00997.2012> doi: 10.1152/jn.00997.2012
- Henry, G. H., Dreher, B., & Bishop, P. O. (1974). Orientation specificity of cells in cat striate cortex. Journal of Neurophysiology, 37, 1394–1409.
- Herz, A. V. M., Gollisch, T., Machens, C. K., & Jaeger, D. (2006). Modeling single-neuron dynamics and computations: A balance of detail and abstraction. Science, 314, 80–86.
- Hietanen, M. A., Cloherty, S. L., van Kleef, J. P., Wang, C., Dreher, B., & Ibbotson, M. R. (2013). Phase sensitivity of complex cells in primary visual cortex. Neuroscience, 237, 19–28. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0306452213000675> doi: 10.1016/j.neuroscience.2013.01.030
- Hubel, D. H. (1959). Single unit activity in striate cortex of unrestrained cats. Journal of Physiology, 147, 226–238.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in cat's visual cortex. Journal of Physiology, 148, 574–591.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology, 160, 106–154.

- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology, 195(1), 215–243.
- Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. Proceedings of the Royal Society of London. Series B, 198(1130), 1–59. doi: 10.2307/77245
- Hupé, J.-M., James, A. C., Girard, P., & Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. Journal of Neurophysiology, 85(1), 146–163.
- Hupé, J.-M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. Journal of Neurophysiology, 85(1), 134–145. doi: 11152714
- Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature, 394(6695), 784–787. doi: 10.1038/29537
- Ibbotson, M. R., Price, N. S. C., & Crowder, N. A. (2005). On the division of cortical cells into simple and complex types: a comparative viewpoint. Journal of Neurophysiology, 93(6), 3699–3702. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15659524> doi: 10.1152/jn.01159.2004
- Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: Toward a unifying model. Journal of the Optical Society of America A, 17(11), 1899–1917. doi: 10.1364/JOSAA.17.001899
- Izhikevich, E. M. (2007). Dynamical systems in neuroscience: The geometry of excitability and bursting. Cambridge, MA: MIT Press.
- Jacobs, R. A. (2009). Adaptive precision pooling of model neuron activities predicts the efficiency of human visual learning. Journal of Vision, 9(4), 1–15.
- Jagadeesh, B., Wheat, H. S., & Ferster, D. (1998). Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. Science, 261, 1901–1904.
- Jancke, D., Chavane, F., Naaman, S., & Grinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. Nature, 428(6981), 423–426. doi: 10.1038/nature02396
- Jesús, M., & Laparra, V. (2010). Psychophysically Tuned Divisive Normalization Approximately Factorizes the PDF of Natural Images. Neural Computation, 22(12), 3179–3206. doi: 10.1017/CBO9781107415324.004
- Jones, H. E., Grieve, K. L., Wang, W., & Sillito, A. M. (2001). Surround suppression in primate V1. Journal of Neurophysiology, 86, 2011–2028.
- 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. doi: citeulike-article-id:

762473

- Jones, J. P., & Palmer, L. A. (1987b). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. Journal of Physiology, 58(6), 1187–1211.
- Kagan, I., Gur, M., & Snodderly, D. M. (2002). Spatial organization of receptive fields of V1 neurons of alert monkeys: comparison with responses to gratings. Journal of Neurophysiology, 88(5), 2557–2574. doi: 10.1152/jn.00858.2001
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. Neuron, 15, 843–856.
- Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. Proceedings of the National Academy of Sciences of the United States of America, 96(21), 12073–12078. doi: 10.1073/pnas.96.21.12073
- Kim, J.-N. (2011). The influence of contrasts on directional and spatial frequency tuning in visual cortex areas 17/18 of the cat. Korean Journal of Ophthalmology, 25(1), 48–53. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21350695> doi: 10.3341/kjo.2011.25.1.48
- Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. Journal of Neurophysiology, 67(4), 961–980.
- Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. Nature Reviews Neuroscience, 7, 220–231.
- Kough, M., & Poggio, T. (2008). A canonical neural circuit model for cortical nonlinear operations. Neural Computation, 20, 1427–1451.
- Kulikowski, J. J., & Bishop, P. O. (1981). Fourier analysis and spatial representation in the visual cortex. Experientia, 37, 160–163.
- Kulikowski, J. J., Marčelja, S., & Bishop, P. O. (1982). Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. Biological Cybernetics, 43, 187–198.
- Kulikowski, J. J., & Vidyasagar, T. R. (1986). Space and spatial frequency: analysis and representation in the macaque striate cortex. Experimental Brain Research, 64, 5–18.
- Lamme, V. A. F. (2003). Beyond the classical receptive field: Contextual modulation of V1 responses. In L. Chalupa & J. Werner (Eds.), The visual neurosciences (pp. 720–732). Boston, MA: MIT Press.
- Lathi, B. P. (2005). Linear systems and signals. Oxford University Press.
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. Journal of the Optical Society of America, 70(12), 1458–1471.
- Lehky, S. R., Sejnowski, T. J., & Desimone, R. (2005). Selectivity and sparseness in the responses of striate complex cells. Vision Research, 45(1), 57–73. Retrieved from <GotoISI>://WOS:000227767000006 doi:

10.1016/j.visres.2004.07.021

- Lennie, P., & Movshon, J. A. (2005). Coding of color and form in the geniculostriate visual pathway. Journal of the Optical Society of America, A, 22(10), 2013–2033.
- Li, B., Peterson, M. R., & Freeman, R. D. (2003). Oblique effect: A neural basis in the visual cortex. Journal of Neurophysiology, 90(1), 204–217. Retrieved from <http://jn.physiology.org/cgi/content/abstract/90/1/204> doi: 10.1152/jn.00954.2002
- Li, B., Peterson, M. R., Thompson, J. K., Duong, T., & Freeman, R. D. (2005). Cross-orientation suppression: monoptic and dichoptic mechanisms are different. Journal of Neurophysiology, 94, 1645–1650. doi: 10.1152/jn.00203.2005
- Li, B., Thompson, J. K., Duong, T., Peterson, M. R., & Freeman, R. D. (2006). Origins of Cross-Orientation Suppression in the Visual Cortex. Journal of Neurophysiology, 96, 1755–1764. doi: 10.1152/jn.00425.2006
- Li, C.-Y., & Creutzfeldt, O. (1984). The representation of contrast and other stimulus parameters by single neurons in area 17 of the cat. Pflügers Archiv - European Journal of Physiology, 401, 304–314.
- Li, C.-Y., & Li, W. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—classification and tuning properties. Vision Research, 34(18), 2337–2355. doi: 10.1016/0042-6989(94)90280-1
- Li, W., Thier, P., & Wehrhahn, C. (2001). Neuronal responses from beyond the classic receptive field in v1 of alert monkeys. Experimental Brain Research, 139, 359–371.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. Neural Computation, 10(4), 903–940. doi: 10.1162/089976698300017557
- Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. Spatial Vision, 13(1), 25–50. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10688228> doi: 10.1163/156856800741009
- Livingstone, M. S. (1998). Mechanisms of direction selectivity in macaque V1. Neuron, 20(3), 509–526. doi: 10.1016/S0896-6273(00)80991-5
- Maffei, L., & Fiorentini, A. (1973). The visual cortex as a spatial frequency analyzer. Vision Research, 13, 1255–1267.
- Maffei, L., & Fiorentini, A. (1976). The unresponsive regions of visual cortical receptive fields. Vision Research, 16, 1131–1139.
- Marr, D. (1982). Vision. Boston, MA: MIT Press.
- Martinez, L. M., & Alonso, J.-M. (2003). Complex receptive fields in primary visual cortex. Neuroscientist, 9(5), 317–331. doi: 10.1177/1073858403252732

- Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. Journal of the Optical Society of America, 70, 1297–1300.
- Mata, M. L., & Ringach, D. L. (2005). Spatial Overlap of ON and OFF Subregions and Its Relation to Response Modulation Ratio in Macaque Primary Visual Cortex. Journal of Neurophysiology, 93, 919–928. doi: 10.1152/jn.00668.2004.
- The MathWorks. (2015). MATLAB user's guide [Computer software manual]. Natick, MA: The MathWorks, Inc.
- Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. Vision Research, 42, 1017–1033. Retrieved from ISI:000175530500008 doi: 10.1016/S0042-6989(02)00025-1
- Meffin, H., Hietanen, M. A., Cloherty, S. L., & Ibbotson, M. R. (2015). Spatial phase sensitivity of complex cells in primary visual cortex depends on stimulus contrast. Journal of Neurophysiology. Retrieved from <http://jn.physiology.org/lookup/doi/10.1152/jn.00431.2015> doi: 10.1152/jn.00431.2015
- Moore IV, B. D., & Freeman, R. D. (2012). Development of orientation tuning in simple cells of primary visual cortex. Journal of Neurophysiology, 107(9), 2506–2516. doi: 10.1152/jn.00719.2011
- Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. i. neurophysiological evidence. Proceedings of the Royal Society, London (Biological Sciences), 216, 335–354.
- Moss, F., Ward, L. M., & Sannita, W. G. (2004). Stochastic resonance and sensory information processing: a tutorial and review of application. Clinical Neurophysiology, 115, 267–281.
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978a). Receptive field organization of complex cells in the cat's striate cortex. Journal of Physiology, 283, 79–99.
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978b). Spatial and temporal contrast sensitivity of the human eye. Journal of the Optical Society of America, 283, 101–120.
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978c). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. Journal of Physiology, 283, 53–77. doi: 10.1113/jphysiol.1978.sp012488
- Mukherjee, P., & Kaplan, E. (1995). Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling. Journal of Neurophysiology, 74(3), 1222–1243.
- Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. Nature Neuroscience, 9(3), 429–434. Retrieved from <http://www.nature.com/doifinder/10.1038/nn1641> doi: 10.1038/nn1641
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. Proceedings of the National Academy of Sciences of the

- United States of America, 99(23), 15164–15169. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=137561&tool=pmcentrez&rendertype=abstract> doi: 10.1073/pnas.192579399
- Naito, T., & Sato, H. (2015, August). Absolute and relative spatial frequency tuning in v1 neurons. Perception, 44, 87–87. (Paper presented at the 38th European Conference on Visual Perception (ECVP2015), Liverpool, UK.)
- Naka, K. I., & Rushton, W. A. H. (1966). S-Potentials from Luminosity Units in the Retina of Fish (Cyprinidae). Journal of Physiology, 185, 587–599.
- Nassi, J. J., Lomber, S. G., & Born, R. T. (2013). Corticocortical feedback contributes to surround suppression in v1 of the alert primate. Journal of Neuroscience, 33(19), 8504–8517.
- Nelson, J. I., & Frost, B. J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. Brain Research, 139, 359–365.
- Nelson, J. I., & Frost, B. J. (1985). Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. Experimental Brain Research, 61, 54–61.
- Niell, C. M., & Stryker, M. P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. Neuron, 65(4), 472–479. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0896627310000590> doi: 10.1016/j.neuron.2010.01.033
- Nienborg, H., Hasenstaub, A., Nauhaus, I., Taniguchi, H., Huang, Z. J., & Callaway, E. M. (2013). Contrast Dependence and Differential Contributions from Somatostatin- and Parvalbumin-Expressing Neurons to Spatial Integration in Mouse V1. Journal of Neuroscience, 33(27), 11145–11154. Retrieved from <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5320-12.2013> doi: 10.1523/JNEUROSCI.5320-12.2013
- 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, 3269–3280.
- Okamoto, M., Naito, T., Sadakane, O., Osaki, H., & Sato, H. (2009). Surround suppression sharpens orientation tuning in the cat primary visual cortex. European Journal of Neuroscience, 29, 1035–1046. doi: 10.1111/j.1460-9568.2009.06645.x
- Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding V1? Neural Computation, 17, 1665–1699.
- Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: Models and mechanisms. Vision Research, 39(2), 231–256.
- Olzak, L. A., & Thomas, J. P. (2003). Dual nonlinearities regulate contrast sensitivity in pattern discrimination

- tasks. Vision Research, 43, 1433–1442.
- Orban, G. A., & Kennedy, H. (1981). The influence of eccentricity on receptive field types and orientation selectivity in areas 17 and 18 of the cat. Brain Research, 208, 203–208.
- Osaki, H., Naito, T., Sadakane, O., Okamoto, M., & Sato, H. (2011). Surround suppression by high spatial frequency stimuli in the cat primary visual cortex. European Journal of Neuroscience, 33(5), 923–932. Retrieved from <http://doi.wiley.com/10.1111/j.1460-9568.2010.07572.x> doi: 10.1111/j.1460-9568.2010.07572.x
- Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., & Ferster, D. (2009). Inhibitory stabilization of the cortical network underlies visual. Neuron, 62(4), 578–592. doi: 10.1016/j.neuron.2009.03.028
- Ozeki, H., Sadakane, O., Akasaki, T., Naito, T., Shimegi, S., & Sato, H. (2004). Relationship between Excitation and Inhibition Underlying Size Tuning and Contextual Response Modulation in the Cat Primary Visual Cortex. Journal of Neuroscience, 24(6), 1428–1438.
- Peirce, J. W. (2007). The potential importance of saturating and supersaturating contrast response functions in visual cortex. Journal of Vision, 7(6:13), 1–15.
- Petrov, A. A., Doshier, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: an incremental reweighting model. Psychological Review, 112(4), 715–743.
- Petrov, A. A., Doshier, B. A., & Lu, Z.-L. (2006). Perceptual learning without feedback in non-stationary contexts: data and model. Vision Research, 46, 3177–3197.
- Pettigrew, J. D., Nikara, T., & Bishop, P. O. (1968). Responses to moving slits by single units in cat striate cortex. Experimental Brain Research, 6, 373–390.
- Pitt, M. A., Kim, W., Navarro, D., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. Psychological Review, 113(1), 57–83.
- Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature, 391(5 Feb), 580–584.
- Pollen, D. A., & Ronner, S. (1982). Spatial computation performed by simple and complex cells in the visual cortex of the cat. Vision Research, 22, 101–118.
- Pollen, D. A., & Ronner, S. F. (1983). Visual cortical neurons as localized spatial frequency filters. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13(5), 907–916.
- Popper, K. (1959). The logic of scientific discovery. London: Hutchinson.
- Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. Annual Review of Neuroscience, 26, 381–410.
- Priebe, N. J., & Ferster, D. (2006). Mechanisms underlying cross-orientation suppression in cat visual cor-

- tex. Nature Neuroscience, 9(4), 552–561. Retrieved from <http://www.nature.com/doi/10.1038/n1660> doi: 10.1038/n1660
- Priebe, N. J., Mechler, F., Carandini, M., & Ferster, D. (2004). The contribution of spike threshold to the dichotomy of cortical simple and complex cells. Nature Neuroscience, 7(10), 1113–1122. Retrieved from <http://www.nature.com/doi/10.1038/n1310> doi: 10.1038/n1310
- Ramsden, B. M., Hung, C. P., & Roe, A. W. (2001). Real and illusory contour processing in area V1 of the primate: A cortical balancing act. Cerebral Cortex, 11, 648–665.
- Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. Neuron, 61(2), 168–185. doi: doi:10.1016/j.neuron.2009.01.002
- Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. Journal of Neurophysiology, 88, 455–463.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? a comment on theory testing. Psychological Review, 107(2), 358–367.
- Rose, D., & Blakemore, C. (1974). An analysis of orientation selectivity in cat's visual cortex. Experimental Brain Research, 20(2), 1–17.
- Rossi, A. F., Desimone, R., & Ungerleider, L. G. (2001). Contextual modulation in primary visual cortex of macaques. Journal of Neuroscience, 21(5), 1698–1709. doi: 21/5/1698[pil]
- Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. Nature Neuroscience, 8(12), 1647–1650. Retrieved from <http://www.nature.com/doi/10.1038/n1606> doi: 10.1038/n1606
- Sasaki, H., Todorokihara, M., Ishida, T., Miyachi, J., Kitamura, T., & Aoki, R. (2006). Effect of noise on the contrast detection threshold in visual perception. Neuroscience Letters, 408, 94–97.
- Sasaki, Y., Rajimehr, R., Kim, B. W., Ekstrom, L. B., Vanduffel, W., & Tootell, R. B. H. (2006). The Radial Bias: A Different Slant on Visual Orientation Sensitivity in Human and Nonhuman Primates. Neuron, 51(5), 661–670. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16950163> doi: 10.1016/j.neuron.2006.07.021
- Sawada, T., & Petrov, A. A. (2015, December). A study of the role of the maintained-discharge parameter in the divisive normalization model of v1 neurons. Proceedings of the 8th International Conference on Bio-inspired Information and Communications Technologies. (Paper presented at the First International Workshop on Computational Models of the Visual Cortex, New York, NY.)
- Sceniak, M. P., Hawken, M. J., & Shapley, R. (2001). Contrast-dependent changes in spatial frequency tuning of macaque v1 neurons: effects of a changing receptive field size. Journal of Physiology, 88(3), 1363–1373.
- Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation

- by macaque v1 neurons. Nature Neuroscience, 2(8), 733–739.
- Schall, J. D., Vitek, D. J., & Leventhal, A. G. (1986). Retinal constraints on orientation specificity in cat visual cortex. Journal of Neuroscience, 6(3), 823–836.
- Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976a). Quantitative studies of single-cell properties in monkey striate cortex. III. Spatial frequency. Journal of Neurophysiology, 39(6), 1334–1551.
- Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976b). Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. Journal of Neurophysiology, 39(6), 1320–1333. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/825622>
- Schumer, R. A., & Movshon, J. A. (1984). Length summation in simple cells of cat striate cortex. Vision Research, 24(6), 565–571. doi: 0042-6989(84)90110-X[pii]
- Schwabe, L., Ichida, J. M., Shushruth, S., Mangapathy, P., & Angelucci, A. (2010). Contrast-dependence of surround suppression in macaque V1: Experimental testing of a recurrent network model. NeuroImage, 52, 777–792.
- Schwabe, L., Obermayer, K., Angelucci, A., & Bressloff, P. C. (2006). The role of feedback in shaping the extra-classical receptive field of cortical neurons: A recurrent network model. Journal of Neuroscience, 26(36), 9117–9129. doi: doi:10.1523/JNEUROSCI.1253-06.2006
- Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. Vision Research, 20, 645–669.
- Schwartz, E. L., Tootell, R. B., Silverman, M. S., Switkes, E., & De Valois, R. L. (1985). On the mathematical structure of the visuotopic mapping of macaque striate cortex. Science, 227, 1065–1066.
- Sclar, G., & Freeman, R. D. (1982). Orientation selectivity in cat's striate cortex is invariant with stimulus contrast. Experimental Brain Research, 46(2), 457–461.
- Sclar, G., Maunsell, J. H. R., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. Vision Research, 30, 1–10.
- Sengpiel, F., Baddeley, R. J., Freeman, T. C. B., Harrad, R., & Blakemore, C. (1998). Different mechanisms underlie three inhibitory phenomena in cat area 17. Vision Research, 38(14), 2067–2080. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698997004136> \delimitter"026E30F\$nhhttp://www.sciencedirect.com/science?{}_ob=MiamiImageURL{&}{}_cid=271122{&}{}_user=964000{&}{}_pii=S0042698997004136{&}{}_check=y{&}{}_origin=article{&}{}_zone=toolbar{&}{}_coverDate=31-Jul-1998{&}view=c{&}originContentFam doi: 10.1016/S0042-6989(97)00413-6
- Sengpiel, F., Sen, A., & Blakemore, C. (1997). Characteristics of surround inhibition in cat area 17.

Experimental Brain Research, 116, 216–228.

- Sengpiel, F., & Vorobyov, V. (2005). Intracortical origins of interocular suppression in the visual cortex. Journal of Neuroscience, 25(27), 6394–6400. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16000630>
doi: 25/27/6394[pii]10.1523/JNEUROSCI.0862-05.2005
- Shapley, R. M., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. Progress in Retinal Research, 3, 263–346.
- Shapley, R. M., & Lennie, P. (1985). Spatial frequency analysis in the visual system. Annual Review of Neuroscience, 8, 547–583.
- Shapley, R. M., & Xing, D. (2013). Local circuit inhibition in the cerebral cortex as the source of gain control and untuned suppression. Neural Networks, 37, 172–181.
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. Nature, 378, 492–496. doi: 10.1038/378492a0
- Simonotto, E., Riani, M., Seife, C., Roberts, M., Twitty, J., & Moss, F. (1997). Visual perception of stochastic resonance. Physical Review Letters, 78(6), 1186–1189.
- Sincich, L. C., Park, K. F., Wohlgenuth, M. J., & Horton, J. C. (2004). Bypassing V1: A direct geniculate input to area MT. Nature Neuroscience, 7(10), 1123–1128. doi: 10.1038/nn1318
- Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behavior. Journal of Physiology, 57(3), 773–786.
- Skottun, B. C., De Valois, R. L., Grosf, D. H., Movshon, J. A., Albrecht, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. Vision Research, 31(7/8), 1079–1086.
- Smith, M. A. (2006). Surround suppression in the early visual system. Journal of Neuroscience, 26(14), 3624–3625.
- Smith, M. A., Bair, W., & Movshon, J. A. (2004). Signals in macaque striate cortical neurons that support the perception of Glass patterns. Journal of Neuroscience, 22, 8334–8345.
- 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. doi: 23/11/4746[pii]
- Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. Journal of Neuroscience, 15(8), 5448–5465.
- Somers, D. C., Todorov, E. V., Siapas, A. G., Toth, L. J., Kim, D.-S., & Sur, M. (1998). A local circuit

- approach to understanding integration of long-range inputs in primary visual cortex. Cerebral Cortex, 8(3), 204–217. doi: 10.1093/cercor/8.3.204
- Sompolinsky, H., & Shapley, R. (1997). New perspectives on the mechanisms for orientation selectivity. Current Opinion in Neurobiology, 7, 514–522.
- Spitzer, H., & Hochstein, S. (1985). A complex-cell receptive-field model. Journal of Neurophysiology, 53, 1266–1286.
- Spitzer, H., & Hochstein, S. (1988). Complex-cell receptive field models. Progress in Neurobiology, 31(4), 285–309. doi: 10.1016/0301-0082(88)90016-0
- Squatrito, S., Trotter, Y., & Poggio, G. F. (1990). Influences of uniform and textured backgrounds on the impulse activity of neurons in area V1 of the alert macaque. Brain Research, 536(1-2), 261–270.
- Stepniewska, I., Qi, H.-X., & Kaas, J. H. (1999). Do superior colliculus projection zones in the inferior pulvinar project to MT in primates? European Journal of Neuroscience, 11(2), 469–480. doi: 10.1046/j.1460-9568.1999.00461.x
- Stork, D. G., & Wilson, H. R. (1990). Do Gabor functions provide appropriate descriptions of visual cortical receptive fields? Journal of the Optical Society of America. A, Optics and image science, 7(8), 1362–1373. doi: 10.1364/JOSAA.9.000334
- Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. Network: Computation in Neural Systems, 7, 161–247.
- Szulborski, R. G., & Palmer, L. A. (1990). The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. Vision Research, 30(2), 249–254.
- Tadmor, Y., & Tolhurst, D. J. (1989). The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning curve in simple cells of the cat's striate cortex. Visual Neuroscience, 3, 445–454.
- Tailby, C., Solomon, S. G., Peirce, J. W., & Metha, A. B. (2007). Two expressions of surround suppression in v1 that arise independent of cortical mechanisms of suppression. Visual Neuroscience, 24, 99–109.
- Teichert, T., Wachtler, T., Michler, F., Gail, A., & Eckhorn, R. (2007). Scale-invariance of receptive field properties in primary visual cortex. BMC Neuroscience, 8(Jun), 38:1–16. doi: 10.1186/1471-2202-8-38
- To, M. P. S., Lovell, P. G., Troscianko, T., & Tolhurst, D. J. (2010). Perception of suprathreshold naturalistic changes in colored natural images. Journal of Vision, 10(4), 12:1–22. doi: 10.1167/10.4.12
- Tolhurst, D. J., & Dean, A. F. (1991). Evaluation of a linear model of directional selectivity in simple cells of the cat's striate cortex. Vision Neuroscience, 6, 421–428.
- Tootell, R. B., Silverman, M. S., Switkes, E., & De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic

- organization in primate striate cortex. Science, 218(4575), 902–904.
- Touryan, J., Felsen, G., & Dan, Y. (2005). Spatial structure of complex cell receptive fields measured with natural images. Neuron, 45, 781–791. doi: 10.1016/j.neuron.2005.01.029
- Troyer, T. W., Krukowski, A. E., Priebe, N. J., & Miller, K. D. (1998). Contrast-invariant orientation tuning in cat visual cortex: Thalamocortical input tuning and correlation-based intracortical connectivity. Journal of Neuroscience, 18(15), 5908–5927.
- Tyler, C. W., & Apkarian, P. A. (1985). Effects of contrast, orientation and binocularity in the pattern evoked potential. Vision Research, 25, 755–766.
- Van Kleef, J. P., Cloherty, S. L., & Ibbotson, M. R. (2010). Complex cell receptive fields: evidence for a hierarchical mechanism. Journal of Physiology, 588(18), 3457–3470.
- Volgushev, M., Pernberg, J., & Eysel, U. T. (2000). Comparison of the selectivity of postsynaptic potentials and spike responses in cat visual cortex. European Journal of Neuroscience, 12(1), 257–263. doi: 10.1046/j.1460-9568.2000.00909.x
- Volgushev, M., Vidyasagar, T. R., & Pei, X. (1996). A linear model fails to predict orientation selectivity of cells in the cat visual cortex. Journal of Physiology, 496(3), 597–606.
- Wallis, G. (2001). Linear models of simple cells: correspondence to real cell responses and space spanning properties. Spatial Vision, 14(3-4), 237–260. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11817739>
- Watkins, D. W., & Berkley, M. A. (1974). The orientation selectivity of single neurons in cat striate cortex. Experimental Brain Research, 19(4), 433–46.
- Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. Journal of the Optical Society of America, A, 2(2), 322–342.
- Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C., & Lennie, P. (2005). Early and late mechanisms of surround suppression in striate cortex of macaque. Journal of Neuroscience, 25(50), 11666–11675.
- Webster, M. A., & De Valois, R. L. (1985). Relationship between spatial frequency and orientation tuning of striate-cortex cells. Journal of the Optical Society of America, 2(7), 1124–1132.
- Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. Biological Cybernetics, 38, 171–178.
- Yacoub, E., Harel, N., & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. Proceedings of the National Academy of Sciences of the United States of America, 105(30), 10607–10612.
- Yu, H.-H., Verma, R., Yang, Y., Tibballs, H. A., Lui, L. L., Reser, D. H., & Rosa, M. G. P. (2010). Spatial and temporal frequency tuning in striate cortex: Functional uniformity and specializations related to receptive

field eccentricity. European Journal of Neuroscience, 31, 1043–1062. doi: 10.1111/j.1460-9568.2010.07118

.x

- Zaltsman, J. B., Heimel, J. A., & Van Hooser, S. D. (2015). Weak orientation and direction selectivity in lateral geniculate nucleus representing central vision in the gray squirrel *Sciurus carolinensis*. Journal of Neurophysiology, 113(7), 2987–97. doi: 10.1152/jn.00516.2014
- Zhang, N. R., & von der Heydt, R. (2010). Analysis of the Context Integration Mechanisms Underlying Figure ?EGround Organization in the Visual Cortex. Journal of Neuroscience, 30(19), 6482– 6496. doi: 10.1523/JNEUROSCI.5168-09.2010
- Zhaoping, L. (2011). Neural circuit models for computations in early visual cortex. Current Opinion in Neurobiology, 21(5), 808–815.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. Journal of Neuroscience, 20(17), 6594–6611. doi: 10.1523/JNEUROSCI.2797-12.2013
- Zhu, W., Xing, D., Shelley, M., & Shapley, R. (2010). Correlation between spatial frequency and orientation selectivity in v1 cortex: implications of a network model. Vision Research, 50, 2261–2273.