

**Jonathan D. Victor, Ferenc Mechler, Ifije Ohiorhenuan, Anita M. Schmid and Keith P. Purpura**

*J Neurophysiol* 102:3414-3432, 2009. First published Oct 7, 2009; doi:10.1152/jn.00086.2009

**You might find this additional information useful...**

---

This article cites 49 articles, 18 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/102/6/3414#BIBL>

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

<http://jn.physiology.org/cgi/content/full/102/6/3414>

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

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

---

This information is current as of January 2, 2010 .

# Laminar and Orientation-Dependent Characteristics of Spatial Nonlinearities: Implications for the Computational Architecture of Visual Cortex

Jonathan D. Victor, Ferenc Mechler, Ifije Ohiorhenuan, Anita M. Schmid, and Keith P. Purpura

Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, New York

Submitted 29 January 2009; accepted in final form 30 September 2009

**Victor JD, Mechler F, Ohiorhenuan I, Schmid AM, Purpura KP.**

Laminar and orientation-dependent characteristics of spatial nonlinearities: implications for the computational architecture of visual cortex. *J Neurophysiol* 102: 3414–3432, 2009. First published October 7, 2009; doi:10.1152/jn.00086.2009. A full understanding of the computations performed in primary visual cortex is an important yet elusive goal. Receptive field models consisting of cascades of linear filters and static nonlinearities may be adequate to account for responses to simple stimuli such as gratings and random checkerboards, but their predictions of responses to complex stimuli such as natural scenes are only approximately correct. It is unclear whether these discrepancies are limited to quantitative inaccuracies that reflect well-recognized mechanisms such as response normalization, gain controls, and cross-orientation suppression or, alternatively, imply additional qualitative features of the underlying computations. To address this question, we examined responses of V1 and V2 neurons in the monkey and area 17 neurons in the cat to two-dimensional Hermite functions (TDHs). TDHs are intermediate in complexity between traditional analytic stimuli and natural scenes and have mathematical properties that facilitate their use to test candidate models. By exploiting these properties, along with the laminar organization of V1, we identify qualitative aspects of neural computations beyond those anticipated from the above-cited model framework. Specifically, we find that V1 neurons receive signals from orientation-selective mechanisms that are highly nonlinear: they are sensitive to phase correlations, not just spatial frequency content. That is, the behavior of V1 neurons departs from that of linear–nonlinear cascades with standard modulatory mechanisms in a qualitative manner: even relatively simple stimuli evoke responses that imply complex spatial nonlinearities. The presence of these findings in the input layers suggests that these nonlinearities act in a feedback fashion.

## INTRODUCTION

Understanding the behavior of neurons in primary visual cortex (V1) is important both as a first step in understanding central visual processing and also because of V1's status as a model system for understanding cortical computation in general. Ideally, one might hope to encapsulate this understanding in a predictively accurate yet concise mathematical model of individual neural responses to arbitrary visual stimuli. Although progress toward this goal has been made, attainment is elusive (Olshausen and Field 2004).

There is general agreement on a starting framework—a “new standard model” (Rust et al. 2005). This model takes the classical model of an oriented linear filter followed by a simple nonlinearity as a building block, combines these linear–nonlinear (“LN”) cascades in parallel, and adds modulatory mechanisms: gain controls, including normalizations (Albrecht and

Geisler 1991; Geisler and Albrecht 1992; Heeger 1992, 1993) and cross-orientation interactions (Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002). However, models of this structure provide an incomplete account (e.g., 50% of the variance is explained; David and Gallant 2005) of the responses of V1 neurons to complex, natural stimuli (see also (Felsen et al. 2005; Touryan et al. 2005; reviewed in Carandini et al. 2005).

It is not clear whether the shortcomings of current models merely reflect quantitative errors that would be corrected by improving the parametric account of known modulatory mechanisms or, alternatively, whether important qualitative aspects of the computations carried out by V1 remain to be identified. This is an important question—at the heart of understanding the design principles of cortical visual processing. Addressing it directly, however, is difficult: one would need to model the modulatory influences at a sufficient level of detail so that one could specify how each would respond to arbitrary images. To do this would require measurement of a multitude of parameters for each modulatory mechanism (e.g., its spatial-frequency tuning, orientation tuning, localization, and dynamics) on a cell-by-cell basis, all within the confines of the practical limits of single-unit recording.

Recognizing the importance of the problem but the impracticality of this direct approach, we developed an alternative experimental strategy that focuses on the identification of qualitative behaviors that distinguish classes of models, rather than detailed parametric modeling. This is an analog of a familiar idea: responses to sinusoidal stimuli allow one to detect the presence of nonlinear behavior, without having to model the linear behavior. Here, rather than attempt to measure modulatory influences, we use a stimulus class that effectively neutralizes them. This allows us to ask whether a cascade model accounts for what remains. If it does not, then we can conclude that a computational element is missing from the modulated-cascade picture and begin to characterize the missing component.

As described in the following text, we find that something is missing and this missing component indicates that V1 extracts orientation not just by spatial frequency content, but also by phase correlations, i.e., spatial correlations of order three and above. (The term “phase correlation” is synonymous with spatial correlations of order three and above. This is because third- and higher-order spatial correlations correspond to correlations across multiple Fourier components, whereas second-order correlations reflect the power of individual Fourier components, independent of their relative phase.) Moreover, we find that these orientation signals are available to modulate the behavior of neurons at the V1 input, suggesting the presence of nonlinear feedback from orientation-tuned neurons.

Address for reprint requests and other correspondence: J. D. Victor, Department of Neurology and Neuroscience, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065 (E-mail: jdvicto@med.cornell.edu).

*Background: two-dimensional Hermite functions*

To implement the above-mentioned strategy, we used a set of visual stimuli based on two-dimensional Hermite (TDH) functions (Victor et al. 2006), shown in Fig. 1. As we outline here, the mathematical properties of TDH functions (Victor and Knight 2003) allow us to ask questions about the structure of neural computations. This is because they determine the *qualitative* aspects of how stimuli based on these functions interact with the filters and gain controls that are the elements of many computational models.

An important property of TDH functions is that they can be grouped into overlapping orthonormal basis sets, Cartesian and polar, that are precisely equated for power and spatial frequency content. Each basis set can thus be used to map receptive fields (RFs) under conditions that eliminate the confounding influences of gain controls sensitive to total power, such as a global contrast normalization. Thus even in the presence of such global modulatory influences, RF maps determined with Cartesian and polar basis sets should be identical.

TDH functions have a second property that is more subtle and crucial to bypassing the effects of modulatory phenomena that are *not* global—i.e., those sensitive to power within a restricted range of frequencies, orientations, or locations. This second property reflects a fundamental difference between TDH functions and Gabor functions (to which some TDH functions bear a superficial resemblance). The Fourier transform of a Gabor function is a pair of Gaussian blobs in spatial frequency space, indicating that a Gabor filter is sensitive to a narrow range of spatial frequencies. In contrast, the Fourier transform of a TDH function (either Cartesian or polar) is the TDH function *itself*, indicating that it contains a broad range of spatial frequencies. Consequently, one antici-

pates that Gabor filters—and, by inference, mechanisms driven by local spatial frequency content in general—will be relatively unable to distinguish between Cartesian and polar TDH stimuli. This is admittedly an informal argument, but as shown in the following text (Figs. 6–8), it is borne out by a range of numerical simulations.

In addition to their analytic utility, we mention another motivation for the use of TDH functions as visual stimuli. Since they have intermediate complexity—more complex than gratings and spots, but less complex than natural scenes—they may tap into neural mechanisms that are recruited by natural scenes but not by the “simple” stimuli typically used to build models, such as gratings and unstructured noise.

METHODS

Our methods for animal preparation, visual stimulation, and recording have been previously described in detail (Aronov 2003; Mechler et al. 2002; Victor et al. 2006); we summarize them here. All animal procedures were performed in accordance with National Institutes of Health and local IACUC standards.

*Physiologic preparation*

Recordings were made under propofol and sufentanil anesthesia, under neuromuscular blockade. After atropine premedication (0.04 mg, administered intramuscularly [im]), initial anesthesia was induced with ketamine 10–20 mg/kg im (cats) or telazol 2–4 mg/kg im (macaques) followed by isoflurane masking for placement of an endotracheal tube, femoral vein and artery catheters, and a urethral catheter. During recording, anesthesia was maintained with propofol and sufentanil (mixture containing 10 mg/ml of propofol and 0.25 to 0.50  $\mu\text{g/ml}$  sufentanil, initially at 2  $\text{mg}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$  propofol, then titrated) and neuromuscular blockade was provided by vecuronium

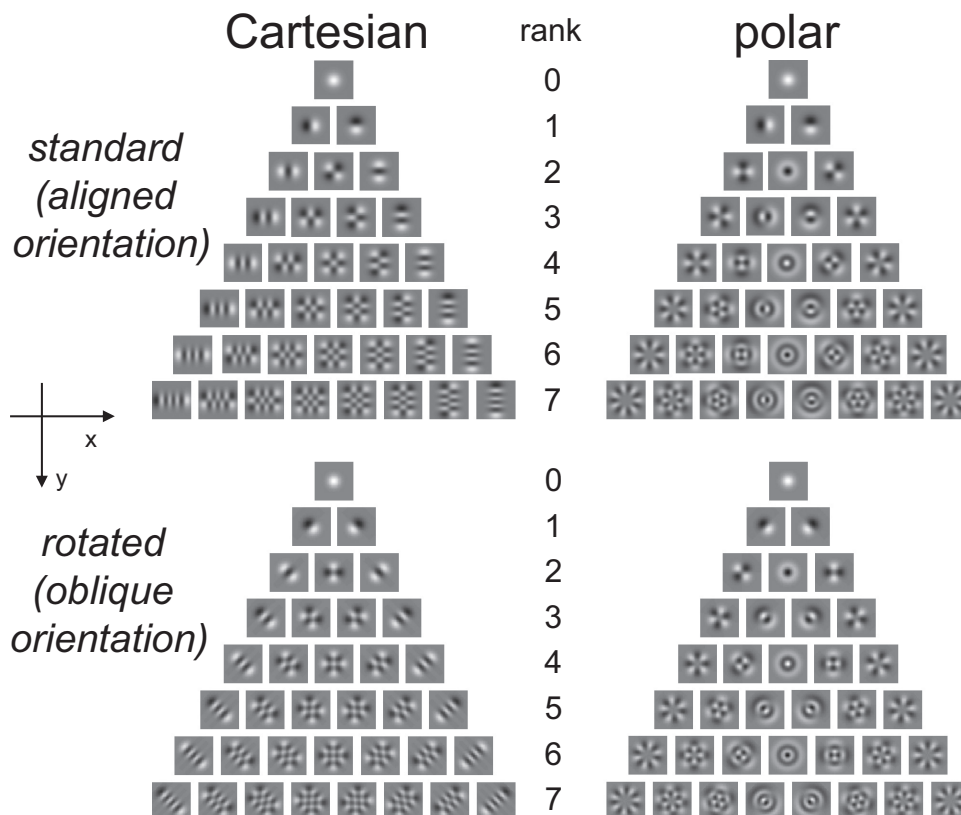


FIG. 1. The Cartesian and polar families of two-dimensional Hermite (TDH) functions. For *experiment 1*, only the “aligned orientation” stimuli were used; for *experiment 2* (the rotation experiment), these were interleaved with the stimuli in the lower half of the figure. Each family forms an orthonormal basis for 2-dimensional patterns, and the stimuli of each rank are equated in size, spatial frequency content, and power. For further details, see Victor et al. (2006).

0.25 mg/kg intravenous (iv) bolus, 0.25 mg·kg<sup>-1</sup>·h<sup>-1</sup> iv. Heart rate and rhythm, arterial blood pressure, body temperature, end-expiratory PCO<sub>2</sub>, arterial oxygen saturation, urine output, and electroencephalograms were monitored during the course of the experiment. Animal maintenance included intravenous fluids (lactated Ringer solution with 5% glucose, 2–3 ml·kg<sup>-1</sup>·h<sup>-1</sup>), administration of supplemental O<sub>2</sub> every 6 h, antibiotics (procaine penicillin G 75,000 U/kg im during surgery, gentamicin 5 mg/kg im daily if evidence of infection), application of 0.5% bupivacaine to wounds, and ocular instillation of atropine 1% and flurbiprofen 2.5% (and, for cats, Neo-Synephrine eye drops 10% to retract the nictitating membranes), dexamethasone (1 mg/kg im daily), and periodic cleaning of the contact lenses. With these measures, the preparation remained physiologically stable for 2 or 3 days (cats) and 4 or 5 days (macaques).

### Recording

A craniotomy was placed near P3, L1 (cats) or P15, L14 (macaques). For six macaques, an array of three tetrodes (Thomas Recording, Giessen, Germany), each coated with 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiI, Molecular Probes, Eugene, OR) to aid subsequent localization of their tracks, was inserted through a small durotomy. The tetrodes were in a "T" configuration: two tetrodes about 600 μm apart, placed just behind and parallel to the presumptive V1/V2 border, and a third tetrode 1,000 to 3,000 μm anteriorly, targeting V1 or V2. [For the two macaques and three cats reported in Victor et al. (2006), only one tetrode was used, targeting V1 or area 17.]

Each tetrode was advanced until spiking activity from one or more units was encountered. Regions of the RFs were hand-mapped and then centered on the display of a Sony GDM-F500 19-in. monitor (displaying a 1,024 × 768 raster at 100 Hz, 35 cd/m<sup>2</sup>), typically at a distance of 114 cm. Real-time spike-sorting software (Datawave Technologies) was then engaged to provide TTL pulses corresponding to the spikes of tentatively identified single units and to allow rapid, qualitative characterization of ocular dominance and grating responses via hand-controlled and computer-generated stimuli.

### Single-unit isolation

A single PC controlled the visual stimulus, logged discriminated event pulses corresponding to the single units tentatively identified by on-line cluster-cutting, and provided timing pulses for a Datawave spike sorting system that recorded spike event waveforms (32 samples at 0.04-ms resolution) for off-line analysis. Off-line spike sorting was performed with an in-house Matlab implementation (Reich 2000) of the methods of Fee et al. (1996) and Sahani et al. (1998). All the data in the following text are derived from these off-line spike sorts.

Usually, each tetrode yielded two to four simultaneously recorded neurons whose spikes were well isolated (signal-to-noise >2:1 and usually >3:1, distinctive shape via on-line spike sorting) and whose spike shapes across the tetrode were reliably discriminated. With the T-shaped recording geometry and the tetrode separations used, no single unit was detected by more than one of the tetrodes. Neurons recorded by different tetrodes usually had overlapping RFs (in V1 and/or V2) and often had similar orientation preferences.

Among the multiple spikes simultaneously recorded, one well-isolated spike on one tetrode was selected as the "target" neuron. We used the responses from these additional nontarget units for the experiments described in the following text, provided that their orientation tuning agreed with that of the target unit within 11.25° and their RFs were largely overlapping.

### Quantitative characterization of tuning to gratings

As described in the following text, tuning properties were determined for a target neuron and two-dimensional Hermite stimuli were

positioned and proportioned accordingly for the subsequent experiments. At approximately one third of recording sites, we repeated the quantitative characterizations for a second neuron with a different orientation preference on the same tetrode or a neuron on a second tetrode with a displaced RF, so it too could also serve as a "target" neuron.

Beginning with parameters determined by an initial hand characterization, computer-controlled stimulation paradigms were used to characterize the target neuron quantitatively with sine gratings. Orientation tuning was determined by the mean response ( $F_0$ ) and the fundamental modulated response ( $F_1$ ) to drifting gratings at orientations spaced in steps of 22.5° (or, for narrowly tuned units, 11.25°), presented at a contrast  $c = (L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$  of 0.5 or 1.0, with spatial and temporal frequency determined by the initial assessment. Following this, we sequentially determined spatial frequency tuning, temporal tuning, and the contrast response function, with each successive run replacing a parameter from the hand characterization by one determined quantitatively.  $F_1/F_0$  ratios quoted in the following text were determined from the response to a grating stimulus whose parameters were optimized in this fashion.

The position of the classical receptive field (cRF) of the target neuron was then determined from the poststimulus histograms (PSTHs) of the response to slowly moving bars. The size of the classical RF was determined from responses to a drifting grating (all parameters optimized) presented in discs of increasing diameter and annuli with a large outer diameter and decreasing inner diameters. The effective diameter  $D$  of the RF of the target neuron (used to determine the size of the TDH patterns) was taken to be the smallest inner diameter of an annulus that did not produce a response that was significantly (2SE) different from the spontaneous activity. The set of annuli was chosen so that  $D$  was determined to within 0.5° or, for smaller RFs, ≤0.25°. For further details on determining RF size and position, see Victor et al. (2006).

### Two-dimensional Hermite functions

STIMULI. Following characterization and alignment of one or more target neurons, we recorded responses to patches whose spatial contrast was determined by a two-dimensional Hermite function (TDH) (see Fig. 1). TDH functions consist of a Gaussian envelope,  $\exp[-(x^2 + y^2)/4\sigma^2]$ , multiplied by a polynomial in the coordinates  $(x, y)$ . We used two families of TDH functions, Cartesian and polar. For the Cartesian functions, the multiplying polynomial is of the form  $X(x)Y(y)$ , so the resulting TDH has zero-crossings and lobes parallel and perpendicular to the  $x$ -axis. For the polar functions, the multiplying polynomial is of the form  $R(r)\cos(\mu\theta)$  or  $R(r)\sin(\mu\theta)$ , so the resulting TDH has zero-crossings and lobes that are circular or radial. Each set (Cartesian and polar) of TDH functions forms a hierarchy of successively more complex patterns. At the  $n$ th rank ( $n = 0, 1, 2, \dots$ ), there are  $n + 1$  functions, each characterized by a polynomial of degree  $n$ . The zeroth rank TDH is an ordinary Gaussian; we used TDHs of rank  $n \leq 7$ .

We set the spatial scale parameter  $\sigma$  of the Gaussian envelope at  $\sigma = D/10$ , where  $D$  was the diameter of the classical receptive field (cRF) of the target neuron as determined by responses to disks and annuli containing the optimal drifting grating. In these experiments,  $\sigma$  had the following ranges (in deg): cat area 17, 0.2 to 0.7 (mean 0.39); macaque V1, 0.08 to 0.5 (mean 0.20); macaque V2, 0.1 to 0.6 (mean 0.25).

By choosing  $\sigma$  in this fashion, stimuli had one, two, or three oscillations within a region of space that covered the cRF and was well matched to sample (in the Nyquist sense) the typical receptive fields of cortical simple cells (Ringach 2002), which have two or three lobes. The contrast profiles of the lowest-rank stimuli lay within the cRF, but the contrast profiles of the higher-rank stimuli (by design) extended beyond the cRF. Each of the TDH patterns has the same total power  $[\iint f(x, y)^2 dx dy]$ . Contrast was scaled by a common factor for all stimuli so that the maximum contrast across all TDH stimuli was

1. See Victor et al. (2006) for further details on the rationale for this choice and the properties of the Hermite functions.

The above-cited procedures served to standardize the position and size of the TDH stimuli. There is an unavoidable measurement error in the determination of cRF position, size, and orientation of the target neuron. In addition, the other simultaneously recorded neurons often had somewhat different RF parameters. However, we note that the analysis procedure and our conclusions do not depend on an exact alignment or sizing and this theoretical robustness is confirmed by the numerical simulations below (Figs. 6–8).

We carried out two kinds of experiments. In the first experiment (as in Victor et al. 2006), we measured responses to Cartesian and polar stimuli, as shown in the top portion of Fig. 1, with the  $x$ -axis aligned to the preferred orientation. This experiment was carried out in all 149 units. Stimuli were presented both with the polarity as shown in Fig. 1, and with inverted-contrast polarity. Ranks 0 to 7 (as shown in Fig. 1) constituted 144 stimuli (36 Cartesian stimuli, 36 polar stimuli, and their contrast-inverses). With the addition of four “blank” stimuli (mean luminance), in which the contrast was held at zero, and removal of rank-0 and rank-1 duplicates between Cartesian and polar stimuli, this amounted to 142 stimuli. Each stimulus was presented for 250 ms, followed by 250 ms of mean luminance, in randomized order (to minimize the effects of gradual changes in responsiveness or residual eye movements), for 8 to 16 blocks.

In the second experiment, we presented the TDH stimuli as described earlier and also after rotating them by  $45^\circ$  (bottom portion of Fig. 1). This experiment was carried out in 59 units (12 cat, 47 macaque). For the two macaques and three cats reported in Victor et al. (2006), aligned and oblique stimuli were presented in sequential blocks (12 cat units, 9 macaque units). For the remaining six macaques (38 units), stimuli were presented in interleaved blocks. Randomization and timing were the same as in the first experiment.

Note that when the Cartesian TDH stimuli are presented in the standard orientation, their elongated regions and zero crossings are parallel or perpendicular to the preferred orientation. However, when they are presented after a  $45^\circ$  rotation, their contours are all oblique to the preferred orientation. Rotation of the polar TDH stimuli has a very different kind of effect: some polar TDH stimuli are rotationally invariant and, for many others, rotation turns one polar TDH stimulus into another one.

**ANALYSIS.** Responses to two-dimensional Hermite functions are analyzed as detailed in Victor et al. (2006) and are summarized here. The analysis has two stages: first, characterization of a neuron’s responses to a single TDH family (Cartesian or polar) and, second, comparison of these characterizations across the two families.

To characterize the response to a single TDH family, we model the response as the sum of three components: a maintained discharge  $R_m$ , a linear component, whose spatial sensitivity is described by  $L(x,y)$ , and a full-wave-rectifying component, whose spatial sensitivity is described by  $E(x,y)$ . That is, the response to an arbitrary stimulus  $S$  is modeled as

$$R(S) = R_m + \int S(x,y)L(x,y)dxdy + |\int S(x,y)E(x,y)dxdy| \quad (1)$$

We emphasize at the outset that we are not advancing Eq. 1 as a reasonable model for cortical computations in general, but simply as a way to summarize and compare responses to different kinds of TDH stimuli, so that we can draw inferences from this characterization. There are two reasons that Eq. 1 is useful for this purpose. First, for the stimuli used here, the characterization is effectively *complete*—it fully specifies the responses to one TDH basis set and accounts for all of the explainable variance (see comment following Eq. 2). Additionally, it explicitly predicts the responses to another TDH basis set. This follows from the fact that the Cartesian and polar TDH basis sets have exactly the same span and each is presented in two contrast polarities.

Second, as we will see below, the characterization afforded by Eq. 1 is only minimally affected by normalization mechanisms, gain controls, and orientation-specific interactions that are appended to cascade models (Albrecht and Geisler 1991; Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002; Geisler and Albrecht 1992; Heeger 1992, 1993). This is not true for general basis sets, but it is true for the TDH stimuli because of their special spectral characteristics.

As previously discussed (Victor et al. 2006), Eq. 1 formalizes a model that is related to a standard linear–nonlinear (LN) cascade. This formulation allows for separate linear ( $L$ ) and nonlinear pathways ( $E$ ) with arbitrary spatial filters, but requires a specific form (full-wave rectification) for the nonlinearity. We mention that using a model-fitting procedure that relaxes the latter constraint (the maximally informative-dimension procedure; Sharpee et al. 2004) leads to very similar findings (Sharpee and Victor 2009), indicating that the assumption about the form of the nonlinearity is not responsible for the findings we describe below. However, as shown in the Supplemental Materials of Sharpee and Victor (2009), the present approach has the advantage that it is more robust.

Because the TDH functions within each family (Cartesian and polar) constitute an orthonormal basis, we can use a reverse-correlation approach (Ringach et al. 1997) to estimate the quantities  $L$  and  $E$  of Eq. 1. The procedure is as follows: for each TDH stimulus  $f_k(x,y)$  (here,  $k$  is a generic index across all basis functions within one family), we first determine a scalar response measure  $R(f_k)$ , by averaging the spike rate in the first 250 ms following stimulus onset. [In Victor et al. (2006), we showed that other choices of the response measure, such as the ON-transient, the OFF-response, the OFF-transient, or the first principal component, led to similar results and also that the time course of the response to Cartesian and polar TDH functions was similar.] Next, to separate the contributions of  $L$  and  $E$ , responses to each stimulus  $f_k(x,y)$  and its contrast-inverse  $-f_k(x,y)$  are combined by addition and subtraction

$$L_k = [R(f_k) - R(-f_k)]/2$$

$$|E_k| = \{[R(f_k) + R(-f_k)]/2\} - R_m \quad (2)$$

The quantities  $L_k$  and  $E_k$  represent the projections of  $L$  and  $E$  onto  $f_k(x,y)$ . This simple transformation from responses to filter coefficients implies that the model of Eq. 1, constructed from one basis set, accounts for all of the explainable variance of the responses to that basis set. The reason is that each pair of filter coefficients  $L_k$  and  $E_k$  is derived from sums and differences of average responses to a separate pair of stimuli,  $f_k(x,y)$  and its contrast-inverse  $-f_k(x,y)$ . Thus the average responses to each of those stimuli can be reconstructed from sums and differences of  $L_k$  and  $E_k$  (and the mean rate  $R_m$ ). See Victor et al. (2006) for a discussion of this point.

The filters  $L$  and  $E$  (e.g., Fig. 2) can be reconstructed by summing these projections

$$L(x,y) = \sum_k L_k f_k(x,y) \\ E(x,y) = \sum_k E_k f_k(x,y) \quad (3)$$

Note that Eq. 2 does not determine the sign of the projection  $E_k$ ; either  $E_k = +|E_k|$  or  $E_k = -|E_k|$  will result in a profile that leads to the same responses to  $f_k$  and  $-f_k$ . For illustration of the profile  $E$ , we choose this sign to maximize the similarity of  $E$  to  $L$ . For quantitative analysis, we choose measures that are independent of this sign. See Victor et al. (2006) for further details. We use  $L_{cart}$  and  $E_{cart}$  to denote the profiles estimated from the responses to the Cartesian stimuli and  $L_{polar}$  and  $E_{polar}$  to denote the profiles estimated from the responses to the polar stimuli. We display these estimated profiles as maps. Figure 2 shows examples for real neurons; Fig. 6 shows examples for model neurons.

In interpreting these maps, it is important to recognize that they represent projections of the receptive fields into the spaces spanned by

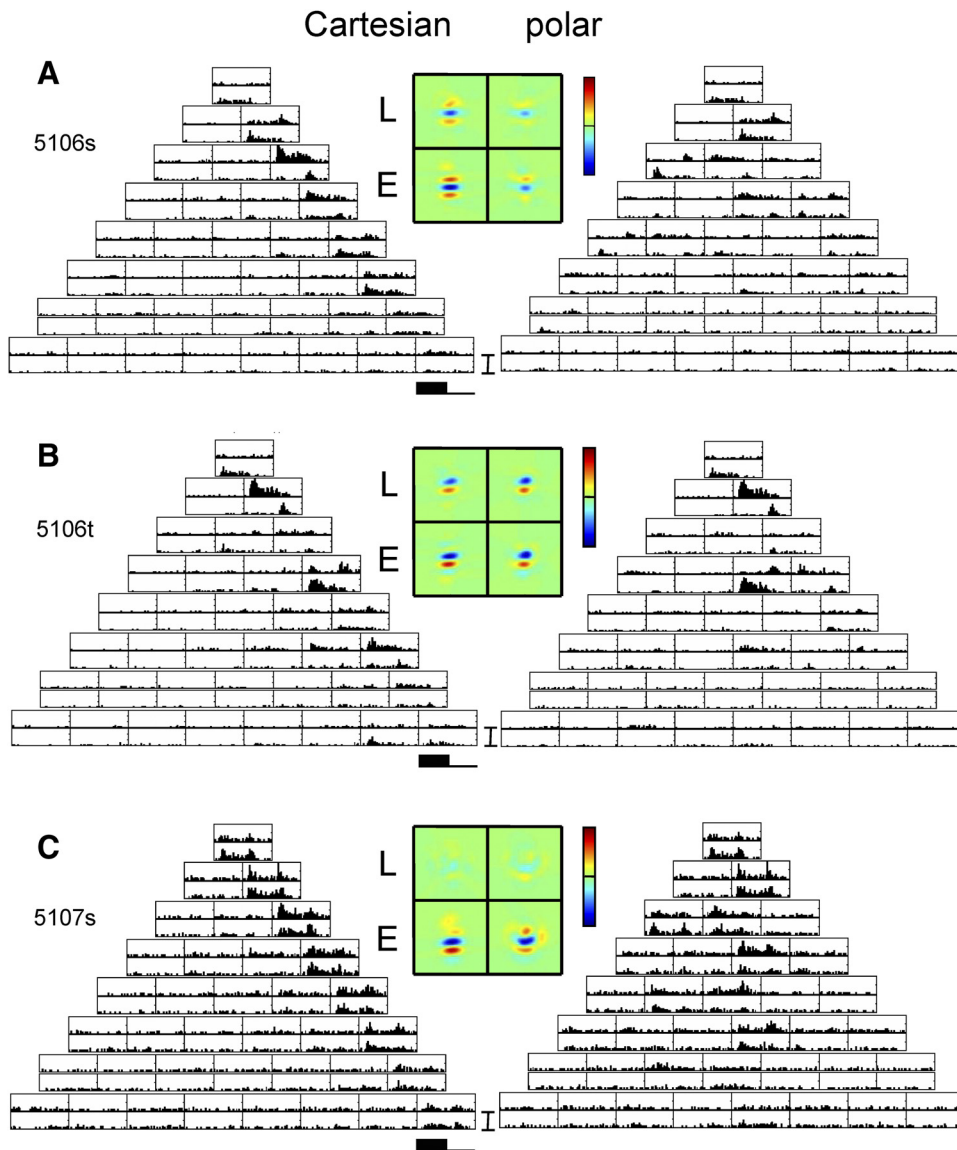


FIG. 2. Poststimulus histograms (PSTHs) and inferred sensitivity profiles for three macaque V1 Layer 4 neurons. Stimuli consisted of TDH functions (*left*, Cartesian stimuli; *right*, polar stimuli), each presented for 250 ms and followed by 250 ms of mean illumination. In each pair of histograms, the top histogram is the response to the stimulus shown in Fig. 1 and the bottom histogram is the response to the contrast-inverse of that stimulus. The four pseudocolor maps (sensitivity profiles) represent the spatial filters  $L_{cart}$ ,  $E_{cart}$ ,  $L_{polar}$ , and  $E_{polar}$  of Eq. 1. PSTH scale bar: 150 impulses/s (A), 150 impulses/s (B), and 100 impulses/s (C). Range for pseudocolor maps of sensitivity profiles:  $\pm 14$  impulses/s (A),  $\pm 16$  impulses/s (B), and  $\pm 12$  impulses/s (C). Maps of sensitivity profiles span  $5.4^\circ$  (A),  $5.4^\circ$  (B), and  $4.8^\circ$  (C). Stimuli were each presented 16 times.

the 36 basis functions. In a manner exactly analogous to the estimation of receptive field profiles from random checkerboard or grating stimuli, projection of the RF onto a finite basis set is expected to differ from the true RF, even for a strictly linear system, because the basis set is finite. For example, RFs estimated from random checkerboards have sharp edges at the pixel boundaries and RFs estimated from responses to a limited number of gratings have substantial ringing (Ringach 2002). Analogously, RFs estimated from a finite TDH basis are also distorted – they tend to be somewhat larger than the true RFs, and have modest ringing (see for example the numerical simulations, Fig. 6). For these reasons, we base our conclusions not on visual inspection of the RFs, but on quantitative comparisons between the Cartesian and polar RF estimates. Since the Cartesian and polar TDH functions span precisely the same subspace, the distorting effects of subspace projection cannot contribute to differences between the estimated RFs.

**INDICES.** The second stage of the analysis consisted of comparing RF profiles estimated from each family of stimuli (Cartesian and polar). If a neuron's responses were accurately described by linear–static nonlinear cascade—or, if the deviations from such a cascade do not materially affect the characterization provided by Eq. 1—then

either basis set should identify the same set of filters. That is, we would expect  $L_{cart} = L_{polar}$  and  $E_{cart} = E_{polar}$ .

This prediction can fail in two ways: Cartesian and polar filters may differ in shape or size (magnitude). To quantify these differences, we therefore use two indices (as in Victor et al. 2006). For shape, we use an index similar to a spatial correlation function

$$I_{shape} = \frac{\iint L_{cart}(x,y)L_{polar}(x,y)dxdy}{|L_{cart}||L_{polar}|} \quad (4)$$

For size (i.e., the overall preference for Cartesian vs. polar stimuli), we use an index that compares response magnitudes

$$I_{c-p} = \frac{(|L_{cart}|^2 + |E_{cart}|^2) - (|L_{polar}|^2 + |E_{polar}|^2)}{|L_{cart}|^2 + |E_{cart}|^2 + |L_{polar}|^2 + |E_{polar}|^2} \quad (5)$$

In the preceding equations

$$|L_{cart}|^2 = \iint |L_{cart}(x,y)|^2 dxdy \quad (6)$$

and similarly for  $L_{polar}$ ,  $E_{cart}$ , and  $E_{polar}$ .

$I_{shape}$  ranges between  $-1$  and  $1$ , where  $1$  indicates that Cartesian and polar filters have the same shape,  $0$  indicates that they are

uncorrelated, and  $-1$  indicates that they are equal and opposite. The size index  $I_{c-p}$  is 0 if Cartesian and polar stimuli lead to responses of the same size, 1 for a neuron that responds only to Cartesian stimuli, and  $-1$  for a neuron that responds only to polar stimuli. For the LN cascade neuron,  $I_{shape} = 1$  and  $I_{c-p} = 0$ ; departures from these values indicate departures from the cascade model predictions that  $L_{cart} = L_{polar}$  and  $E_{cart} = E_{polar}$ .

For the second experiment (in which stimuli are presented aligned to the preferred orientation and also oblique to it), we calculate the above-cited indices for both stimulus orientations and distinguish them by superscripts: i.e.,  $I_{shape}^{aligned}$ ,  $I_{shape}^{oblique}$ ,  $I_{c-p}^{aligned}$ , and  $I_{c-p}^{oblique}$ . We also determine the effect of stimulus orientation (aligned Cartesian vs. oblique Cartesian) on filter shape

$$I_{shape}^{aligned\_vs\_oblique} = \frac{\iint L_{cart}^{aligned}(x,y)L_{cart}^{oblique}(x,y)dx dy}{|L_{cart}^{aligned}| |L_{cart}^{oblique}|} \quad (7)$$

with  $I_{shape}^{aligned\_vs\_oblique} = 1$  if  $L_{cart}^{aligned}$  and  $L_{cart}^{oblique}$  have the same shape. (We consider only Cartesian responses for this measure, since changing orientation leaves many of the polar stimuli invariant or merely permutes them.)

Raw estimates of the indices  $I_{shape}$  and  $I_{c-p}$  were determined by substituting the profile estimates of Eq. 3 into the Eqs. 4 to 7. For  $I_{shape}$ , the null hypothesis is that there is no shape change; accordingly,  $P$  values are quoted with respect to deviations from 1. For  $I_{c-p}$ , the null hypothesis is that the sizes of the quantities compared in their numerators are equal; accordingly,  $P$  values are quoted with respect to deviations from 0. To compensate for the tendency of measurement error (response variability) to move these indices away from their null-hypothesis values, the raw (“plugin”) estimates were debiased via the jackknife (Efron 1982), with each replicate run considered a single observation. The jackknife procedure also yielded the quoted  $P$  values and confidence limits. The statistical calculations for  $I_{shape}$  were carried out following a Fisher transformation to normalize their distribution (Sharpee and Victor 2009; Victor et al. 2006). Correlations quoted are Pearson’s  $r$  and significance values are two-tailed. All calculations were performed in Matlab (The MathWorks, Natick, MA).

We emphasize that our motivation for the use of Eq. 1—and the indices derived from it—is to draw inferences about cortical computations, based on neural responses to TDH stimuli. Although we use the model to generate null hypotheses concerning the values of the indices, we do not intend to suggest that it is a model of cortical neuronal responses: it is evidently missing important elements such as response normalizations (Albrecht and Geisler 1991; Geisler and Albrecht 1992; Heeger 1992, 1993) and cross-orientation interactions (Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002). However, as shown below, the presence of these mechanisms does not detract from our ability to draw inferences concerning these indices, since they induce only minimal dependence of the fitted filters in Eq. 1 on the choice of basis set. The fundamental reason for this relates to mathematical properties of the TDH functions—each contains a wide and overlapping range of orientations and spatial frequencies; thus even gain controls and interactions tuned to particular spatial frequencies or orientations are surprisingly insensitive to the Cartesian versus polar distinction.

### Visual stimulus generation

Control signals for the cathode ray tube display were provided by a separate PC-hosted system optimized for OpenGL (Nvidia GeForce3 chipset) programmed in Delphi. [For the two macaques and three cats reported in Victor et al. (2006), bar and grating stimuli were generated by a PC-hosted VSG2/5 (8 Mb).] For presentation, TDH stimuli were discretized as limited by the display resolution. This typically meant  $\geq 64 \times 64$  display pixels across the stimulus, with

each display pixel subtending about 1 min. At the edge of each patch, stimulus contrast was reduced to  $< 1/256$  of its peak value.

Intensity linearization was separately performed for each display controller via Visualization Science Group (VSG, Burlington, MA) software or in-house software of comparable function.

### Histology

After all recordings, lesions were made by current passage (typically  $3 \mu\text{A} \times 5$  s, electrode negative) at three locations along each tetrode track, bracketing the recording sites. The animal was sacrificed and perfused with 4% paraformaldehyde solution in phosphate-buffered saline. Fluorescence microscopy of histologically unstained 40- $\mu\text{m}$  cryostatic sections was used to capture the Dil trace of the track. Light microscopy of the same sections after Nissl staining (Hevner and Wong-Riley 1990) was used to define laminar organization and locate the lesions.

### Numerical simulations

We carried out numerical simulations to determine the effect of gain controls and orientation-specific interactions on the characterization provided by Eqs. 2 and 3. Each simulation was carried out for 50 model neurons. Model neurons (see Fig. 6) had a Gabor spatial sensitivity profile  $G(x,y)$  whose parameters (Gaussian envelope, sinusoidal carrier) were determined by independent draws from the following distributions: envelope center position along each axis was drawn from a Gaussian with SD  $\sigma$ ; envelope  $1/e$  radius was drawn from a log-normal with geometric mean of  $\sigma$  and SD covering a factor of 1.25; envelope aspect ratio was drawn from a log-normal with geometric mean 1.5 and SD covering a factor of 1.25; envelope orientation was drawn from a Gaussian with SD  $\pi/16$  ( $11.25^\circ$ ) around the horizontal; carrier spatial frequency was drawn from a log-normal with geometric mean 0.3 cycles per  $1/e$  radius and SD covering a factor of 1.5; carrier orientation was drawn from a Gaussian with SD  $\pi/16$  around the envelope orientation; carrier spatial phase was drawn from a uniform distribution in  $[0, 2\pi]$ . This resulted in receptive fields that typically had one, two, or three lobes (rarely four), comparable to what is seen in V1 (Ringach 2002), with an orientation and center-position scatter similar to that of the receptive fields inferred for the recorded neurons (see Fig. 6). All Gaussian envelopes integrated to unity.

To model the generation of a normalization signal, we created a similar collection of 120 Gabor profiles  $G_i^{aux}(x,y)$  ( $i = 1, \dots, 120$ ). In the basic model, the parameter distributions matched those used for the model neuron, except that envelope orientation of neurons in the auxiliary pool was random. As described in RESULTS, we also created populations in which the distribution of locations and orientation of the auxiliary neurons were modified, to simulate different kinds of normalizations and cross-orientation interactions.

To determine the response of the model neuron with receptive field profile  $G(x,y)$  to a TDH stimulus  $S$ , we first convolved  $S$  with  $G$  and then applied a nonlinearity  $N$ . This yielded a “raw” response for the model neuron to the stimulus  $S$ , prior to any interactions between the model neuron and the pool of auxiliary neurons

$$g(S) = N[\iint S(x,y)G(x,y)dx dy] \quad (8)$$

Similarly, convolution of each auxiliary profile  $G_i^{aux}(x,y)$  with  $S$  yielded its “raw” response

$$g_i^{aux}(S) = N[\iint S(x,y)G_i^{aux}(x,y)dx dy] \quad (9)$$

$N$  was a half-wave rectifier  $N(u) = \max(u, 0)$  in most simulations; we also used a more nearly full-wave asymmetric rectifier  $N(u) = [|u| + \max(u, 0)]/2$  and an exponential nonlinearity  $N(u) = e^u$ .

Next, the raw responses of a randomly chosen pool of  $n_{aux}$  auxiliary neurons were combined to create a signal  $z(S)$  to provide normalization or cross-orientation interactions

$$z(S) = \frac{1}{n_{aux}} \sum_{i \in pool} N_{aux}[g_i^{aux}(S)] \quad (10)$$

Following Heeger (1993),  $N_{aux}$  was a half-squarer  $N_{aux}(u) = [\max(u/u_0, 0)]^2$ , but we also used full-squaring, half-wave rectification, and full-wave rectification. Here,  $u_0$  is the root-mean-squared raw response size across all TDH stimuli; this choice means that the quantities acted on by  $N_{aux}$  are on the order of 1. Pool sizes ranging from  $n_{aux} = 1$  to  $n_{aux} = 120$  were considered.

Finally, the modeled response  $R(S)$  to stimulus  $S$  was determined by allowing the pooled signal  $z(S)$  to interact with the raw response  $g(S)$  in a divisive fashion (Heeger 1993)

$$R(S) = \frac{g(S)}{1 + cz(S)} \quad (11)$$

The constant  $c$  was chosen by first identifying a value  $c_0$  for which the interaction in Eq. 11 typically attenuated the responses by a factor of 2, and then using  $c = c_0/2$ ,  $c = c_0$ ,  $c = 2c_0$ , and  $c = 3c_0$ , to provide a range of normalization and interaction strengths, up to a fourfold attenuation ( $c = 3c_0$ ). Along with the divisive interaction specified by Eq. 11, we also simulated models with a subtractive interaction

$$R(S) = \max[g(S) - cz(S), 0] \quad (12)$$

Note that the overall architecture of the model is feedforward (Albrecht and Geisler 1991) rather than feedback (Heeger 1993) and we do not consider response dynamics (cf. Heeger 1993). We do this simply to focus on the effects of various spatial configurations for the interactions, not to imply that feedback is absent. Thus in contrast to the feedforward model of Albrecht and Geisler (1991), the auxiliary pool contains oriented units. We elaborate on these issues in the DISCUSSION.

For each instance of the model (spatial configuration of the auxiliary pool, nonlinearity shape  $N$ , pool size  $n_{aux}$ , interaction strength  $c$ , and divisive vs. subtractive interaction), simulated responses  $R(S)$  were calculated for each TDH stimulus  $S = f_k$  and its negative  $S = -f_k$ . From these responses [ $R(f_k)$  and  $R(-f_k)$ ], and with  $R_m = 0$ , corresponding to the absence of a maintained discharge], Eqs. 2 to 5 were used to determine  $L$ - and  $E$ -filters and the various indices. Note that these simulations did not include response variability. Thus jackknife debiasing to compensate for response variability was applied to the neural data (and not to the simulations) so that the comparison between simulations and neural data would be a fair one.

## RESULTS

We present results from a total of 149 single units in 11 preparations (70 V1 units in 8 macaques, 45 V2 units in 3 of these animals, 34 area 17 units in 3 cats). All recordings were within  $5^\circ$  of the area centralis (cats) or fovea (macaques). This population consists of all spike-sorted units that had stable firing rates over the experimental period and had responses to TDH stimuli that could be reliably distinguished from the background, and includes five units that did not respond reliably to gratings (so their tuning functions and  $F_1/F_0$  ratios could not be determined). Approximately 20% of isolated neurons could not be driven by TDH stimuli and they are excluded from this analysis. A portion of these results (51 units: all 34 cat units and 17 of the V1 macaque units) have been previously reported (Sharpee et al. 2006; Victor and Mechler 2006; Victor et al. 2006).

As described earlier, the overall goal of this study was to identify and characterize nonlinear mechanisms in V1 that qualitatively differ from the behaviors expected from a LN cascade with modulatory influences. Our strategy to do this was to map RFs with two stimulus sets: the Cartesian and polar TDH functions. Since these functions are matched in contrast and spatial frequency content and constitute separate basis sets with identical spans, neurons that conform to the cascade model will yield identical RF maps. As reported previously (Sharpee et al. 2006; Victor and Mechler 2006; Victor et al. 2006), this expectation does not hold: RF maps constructed from the two basis sets often differ dramatically in shape and sensitivity.

In *experiment 1*, we extend this analysis to a much larger data set, to determine the laminar location of neurons that manifest these nonlinearities.

Then, we consider the effects of response normalizations and cross-orientation interactions. Since such mechanisms can be sensitive to power within a particular range of spatial frequencies and orientations, they might (in principle) lead to different RF maps with the two stimulus sets. However, as our simulations show, these mechanisms cannot account for the large changes in apparent RF maps seen in *experiment 1*.

*Experiment 2* determines what kinds of visual feature lead to the differences in apparent RF maps obtained from Cartesian and polar stimuli. We focus on two kinds of differences between the Cartesian and polar TDH stimuli. One distinction is that they differ as *classes*. That is, apparent differences in RF shape might be due to mechanisms sensitive to features generic to one class but not the other—such as curved contours, which are present in most of the polar TDHs but none of the Cartesian TDHs; or orthogonal contours, which are present in many of the Cartesian TDHs but none of the polar ones. Another kind of difference relates to characteristics specific to *individual* TDH stimuli—for example, the presence of an elongated contour at a specific orientation. Operationally, we can distinguish between these possibilities by comparing responses to Cartesian stimuli at two different orientations. We find that this manipulation alters the apparent RF profile, thus implying that the crucial feature is the orientation of individual contours, not the Cartesian versus polar distinction per se.

### *Experiment 1: prevalence and location of neurons sensitive to phase correlations*

This experiment was carried out on all 149 single units. We first present example responses and then a population summary.

**EXAMPLE RESPONSES.** Figure 2 shows example responses from a cluster of three macaque V1 neurons that were histologically localized to Layer 4b/4c $\alpha$ . Unit 5106s was a nondirectional complex cell ( $F_1/F_0 = 0.16$ ). For Cartesian stimuli (Fig. 2A, *left*), responses were almost entirely confined to the “Gabor-like” stimuli consisting of uninterrupted contrast bands aligned to its orientation preference (the stimuli along the right side of the array of Cartesian responses). In most cases, responses were similar when contrast was inverted (*top* and *bottom histograms* for each response block), but for some stimuli (e.g., the last stimulus in the *third row*), they were strikingly different: an ON response for one polarity and an OFF response for the



inverted polarity. Correspondingly, the RF maps inferred from these responses (see METHODS) consisted of Gabor-like profiles (Fig. 2A, middle, left two patches), with somewhat more power in the rectified pathway ( $E_{cart}$ ) than the linear pathway ( $L_{cart}$ ). The observation that the response is primarily but not exclusively carried by the rectified pathway is consistent with a cell that is mostly complex (ON-OFF), but not exclusively so (see Victor et al. 2006 for further discussion of this). For polar TDH responses (Fig. 2A, right), responses also showed a mixture of symmetrical and asymmetrical responses to stimuli and their contrast-inverses. However, the responses to the polar stimuli, overall, were only about half as large as the responses to Cartesian ones. Consequently, the sensitivity profiles  $L_{polar}$  and  $E_{polar}$  (Fig. 2A, middle, right two patches) had a lower amplitude than that of  $L_{cart}$  and  $E_{cart}$ . This difference is quantified by the index  $I_{c-p}$  (Eq. 5), which in this cell is  $0.50 \pm 0.06$ . It is significantly positive ( $P < 0.01$ ), corresponding to the evident overall preference for Cartesian stimuli. However, the shapes of the inferred profiles are nearly identical: the shape index  $I_{shape}$  (Eq. 4) is  $0.96 \pm 0.03$ . In sum, when studied with either family of TDH stimuli, this cell's responses are well described by a Gabor-like filter followed by partial rectification. Crucially, however, the responses to Cartesian stimuli are substantially larger than responses to polar stimuli.

The neuron in Fig. 2B, unit 5106t, recorded simultaneously with the neuron of Fig. 2A, had a different behavior: the shapes of the receptive field profiles inferred from the two stimulus sets were nearly identical ( $I_{shape}$ :  $0.96 \pm 0.02$ ) and the overall size of the response to Cartesian and polar stimuli were equal ( $I_{c-p}$ :  $0.03 \pm 0.05$ ). This unit shared the same orientation preference as that of the unit of Fig. 2A, but was directionally biased and more simple-like ( $F_1/F_0 = 1.0$ ). In keeping with its higher  $F_1/F_0$  ratio, most of the cell's TDH responses were strongly dependent on stimulus contrast polarity. Also in keeping with simple-like behavior, the overall sizes of the  $L$ - and  $E$ -filters were similar, corresponding to half-wave rectification (see Eq. 1 and Victor et al. 2006).

The neuron in Fig. 2C, unit 5107s, was recorded 100 microns further below the cortical surface. This cell was oriented and nondirectional and had an intermediate  $F_1/F_0$  ratio of 0.5. It responded to a greater number of TDH stimuli than the previous units and the responses were nearly identical for the two contrast polarities. Consequently, its sensitivity profiles (Fig. 2C, middle) were dominated by the nonlinear ( $E$ ) pathway and did not resemble Gabor patches as closely as the units of Fig. 2, A and B. Moreover, the shapes of the sensitivity profiles  $L_{cart}$  and  $L_{polar}$  differed: the shape index  $I_{shape}$  deviated substantially from unity ( $0.55 \pm 0.22$ ,  $P < 0.05$ ). Inspection of  $L_{cart}$  and  $L_{polar}$  show that the main change in RF shape is that the antagonistic regions were more evident in  $L_{polar}$  than that in  $L_{cart}$ . For this neuron, the overall sensitivity to Cartesian and polar stimuli were equal ( $I_{c-p}$ :  $0.08 \pm 0.08$ ).

In sum, we have examined the responses of three nearby Layer 4 neurons to the matched TDH basis sets. One neuron (Fig. 2B) had responses that fully conformed to the expectations of a linear-static nonlinear cascade, consisting of a Gabor-like spatial filter followed by partial rectification. The other two units did not: the unit of Fig. 2A responded preferentially to Cartesian stimuli ( $I_{c-p} > 0$ ) and the unit of Fig. 2C had a sensitivity profile whose shape changed ( $I_{shape} \approx 0.5$ ), depending on whether Cartesian or polar stimuli were used to

assay it. [Additional examples of individual-unit responses, including many units outside of Layer 4, are presented in the following text (Figs. 9–11) and also in Sharpee and Victor (2009) and Victor et al. (2006).]

**POPULATION SUMMARY.** To motivate our approach to population analysis, we consider the findings in Fig. 2 in more detail. Intuitively, one might expect that deviations from the predictions of the LN cascade are not so surprising, since the cascade model omits mechanisms such as cross-orientation suppression (Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002). These and similar mechanisms might be expected to alter the responses to one basis set or the other. For example, the greater sensitivity of the unit in Fig. 2A to Cartesian versus polar stimuli could arise from suppression of responses to stimuli that contain Fourier components that deviate from its preferred orientation; only the Cartesian basis set has stimuli that contain just the preferred orientation—the right-hand side of the Cartesian stimulus pyramid. For the polar stimuli, any stimulus with a Fourier component at the preferred orientation also contains a Fourier component at a nonpreferred orientation, thus suppressing its response. The net result would be that the apparent sensitivity to Cartesian stimuli would be larger, since only the Cartesian set has stimuli that contain just the preferred orientation. However, this intuition is not supported by quantitative analysis (see Figs. 6–8 in the following text). As we will see later, despite the presence of visually apparent oriented features in some of the TDH stimuli, all TDH stimuli are, in fact, broadband in orientation and spatial frequency. Because of this, mechanisms that are selective for the orientation of Fourier components are surprisingly unselective for one basis set versus the other.

This property of the TDH stimuli has the practical consequence that the two indices,  $I_{shape}$ , which quantifies apparent changes in RF shape, and  $I_{c-p}$ , which quantifies sensitivity differences, remain useful for the analysis of responses to Cartesian and polar TDH responses, even in the presence of normalizations and cross-orientation interactions. In the following text we describe the behavior of these indices in our data set and then consider their behavior in a range of models. Although we present only indices based on the ON-response, the findings we describe were substantially unchanged for indices  $I_{shape}$  and  $I_{c-p}$  based on other choices of a response measure, such as the size of the transient response (first 100 ms), the size of the OFF-response, or the size of the first principal component (as shown in Victor et al. 2006 for a subset of these data).

*Change in the shape of the sensitivity profile.* The index  $I_{shape}$  (Eq. 4) captures how the effective shape of the sensitivity profile of a single neuron depends on the basis set. To characterize this dependence across the population, we report the fraction of neurons that exhibited a dependence (i.e., neurons with  $I_{shape} < 1$  at some significance level) and the typical extent of this dependence, as quantified by the mean, median, and SD of  $I_{shape}$  across all neurons. Because of its relevance to circuitry, we focus on the difference between Layer 4 neurons and neurons in other layers.

Most units (41/70) in macaque V1 (Fig. 3, left column) showed a significant shape change between Cartesian and polar basis sets: for 11 units,  $I_{shape} < 1$  at  $0.01 \leq P < 0.05$ ; for 30 units,  $I_{shape} < 1$  at  $P < 0.01$ . This behavior was present in all layers, but it was more prevalent in the upper and lower layers (27/37 units) than in Layer 4 (14/33 units; Layer 4 vs. non-

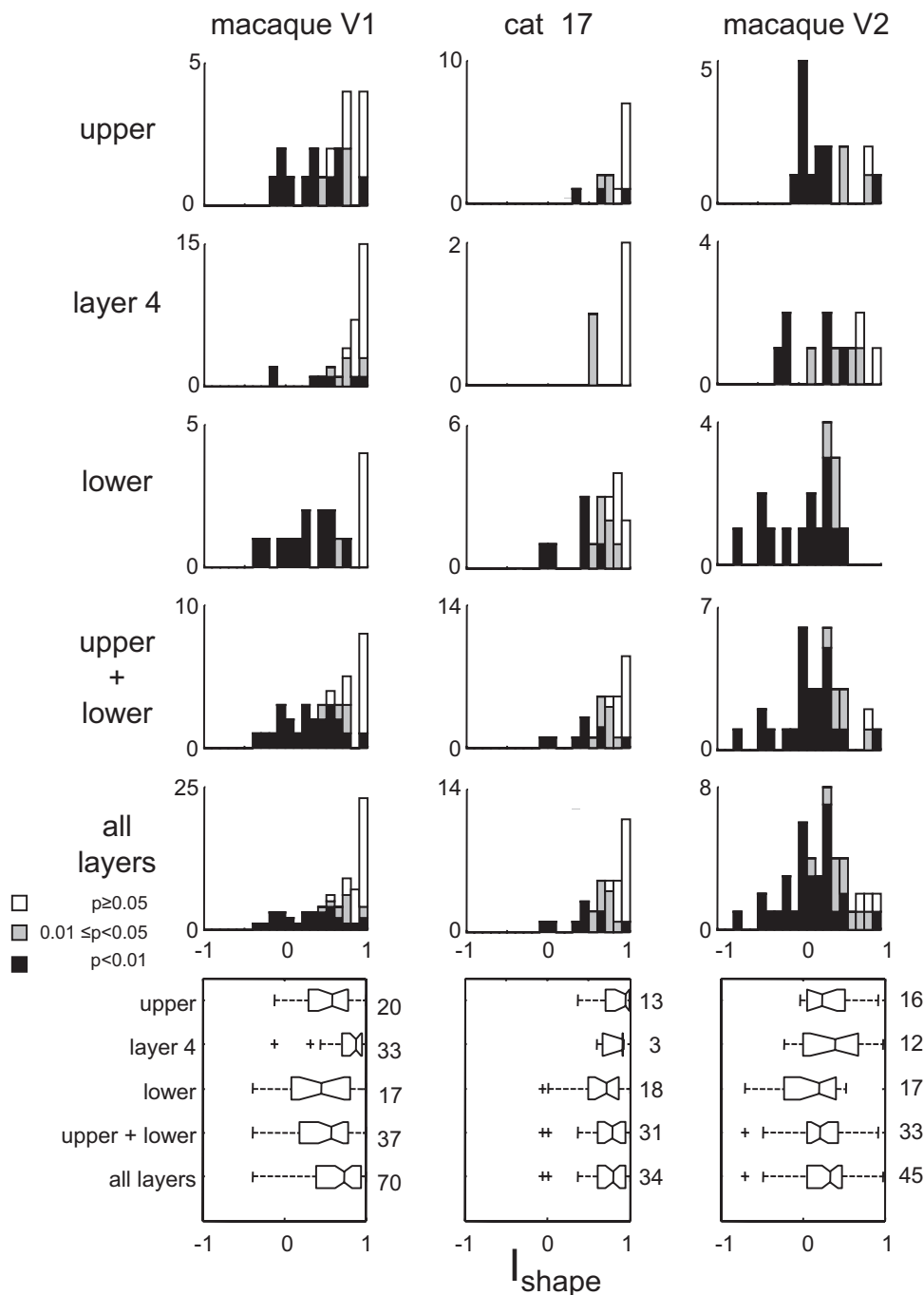


FIG. 3. Population summary of the dependence of the effective receptive field (RF) shape on basis set. The index  $I_{shape}$  (Eq. 4) measures the similarity between the RF sensitivity profile measured with Cartesian TDH functions, vs. with polar TDH functions (see METHODS). Values of  $I_{shape}$  significantly  $< 1$  indicate departure from the expected behavior of a linear–nonlinear (LN) cascade. Portions of the histograms (stacked) shaded black represent units for which  $I_{shape} < 1$  at  $P < 0.01$ ; portions shaded gray are significant at  $0.01 \leq P < 0.05$ ; unshaded portions correspond to  $P > 0.05$ . Boxplot notches indicate 95% confidence limits on median (located at the angle of the notch); boxplot size indicates interquartile range; “+” signs indicate outliers; numbers to the right of each plot indicate the number of units.

Layer 4 was significantly different at  $P < 0.02$ , two-tailed chi-squared [ $\chi^2$ ] test). There was no significant difference between supragranular layers (14/20 units) and infragranular layers (13/17 units). The extent of the dependence also was larger in non-Layer 4 units (mean  $\pm$  SD:  $0.48 \pm 0.39$ , median 0.58) than that in Layer 4 units (mean  $\pm$  SD:  $0.77 \pm 0.28$ , median 0.88; non-Layer 4 units, difference significant at  $P < 0.01$ , two-tailed Wilcoxon test). There was no difference between supragranular units (mean  $\pm$  SD:  $0.53 \pm 0.36$ , median 0.59) and infragranular units (mean  $\pm$  SD:  $0.43 \pm 0.43$ , median 0.46).

In cat area 17 (Fig. 3, middle column), the size and prevalence of shape changes were comparable to those of macaque

V1. There were insufficient data (only three Layer 4 units) to carry out a meaningful laminar analysis.

In macaque V2 (Fig. 3, right column), nearly all units showed a shape change between Cartesian and polar stimuli (42/45 units:  $I_{shape} < 1$  in 10 units at  $0.01 \leq P < 0.05$ , in 32 units at  $P < 0.01$ ) and the shape change was pronounced ( $I_{shape}$  was often close to 0: mean  $\pm$  SD,  $0.14 \pm 0.38$ , median 0.16). By both measures (prevalence and effect size), RF shape changes in V2 were significantly greater than those in V1 ( $P < 0.01$ ). This held whether the comparison was based on all V1 laminae or only its extragranular layers. However, in contrast to V1, we found no evidence for a laminar dependence of RF shape change in V2.

**Cartesian versus polar preference.** Figure 4 shows a parallel analysis of the index  $I_{c-p}$ , which compares the overall size of the response to the two families of TDH stimuli, with  $I_{c-p} > 0$  indicating a Cartesian preference and  $I_{c-p} < 0$  indicating a polar preference (Eq. 5). Because these two stimulus sets are matched for spatial frequency content, a linear-static nonlinear cascade would have responses characterized by  $I_{c-p} = 0$ .

Across the entire data set, 64/149 (43%) units had a significant deviation from  $I_{c-p} = 0$  (18 units at  $0.01 \leq P < 0.05$ , 46 units at  $P < 0.01$ ), whereas 85 units (57%) were consistent with  $I_{c-p} = 0$ . Overall preference for Cartesian and polar stimuli were similar in prevalence (36 units:  $I_{c-p} > 0$  at  $P < 0.05$ ; 28 units,  $I_{c-p} < 0$  at  $P < 0.05$ ) and magnitude

( $I_{c-p}$ : mean  $\pm$  SD  $0.07 \pm 0.36$ , median 0.00). A similar proportion of cells with a preference ( $P < 0.05$ ) for one basis set or the other was found in macaque V1 (20 units with  $I_{c-p} > 0$ , 12 units with  $I_{c-p} < 0$ , 38 units consistent with  $I_{c-p} = 0$ ), cat area 17 (6 units with  $I_{c-p} > 0$ , 7 units with  $I_{c-p} < 0$ , 21 units consistent with  $I_{c-p} = 0$ ), and macaque V2 (10 units with  $I_{c-p} > 0$ , 9 units with  $I_{c-p} < 0$ , 26 units consistent with  $I_{c-p} = 0$ ). Subdivision of the macaque V1 data set revealed no laminar differences (Layer 4: 10 units with  $I_{c-p} > 0$ , 3 units with  $I_{c-p} < 0$ , 20 units consistent with  $I_{c-p} = 0$ ; non-Layer 4: 10 units with  $I_{c-p} > 0$ , 9 units with  $I_{c-p} < 0$ , 18 units consistent with  $I_{c-p} = 0$ ). There was no correlation between  $I_{c-p}$  and  $I_{shape}$ .

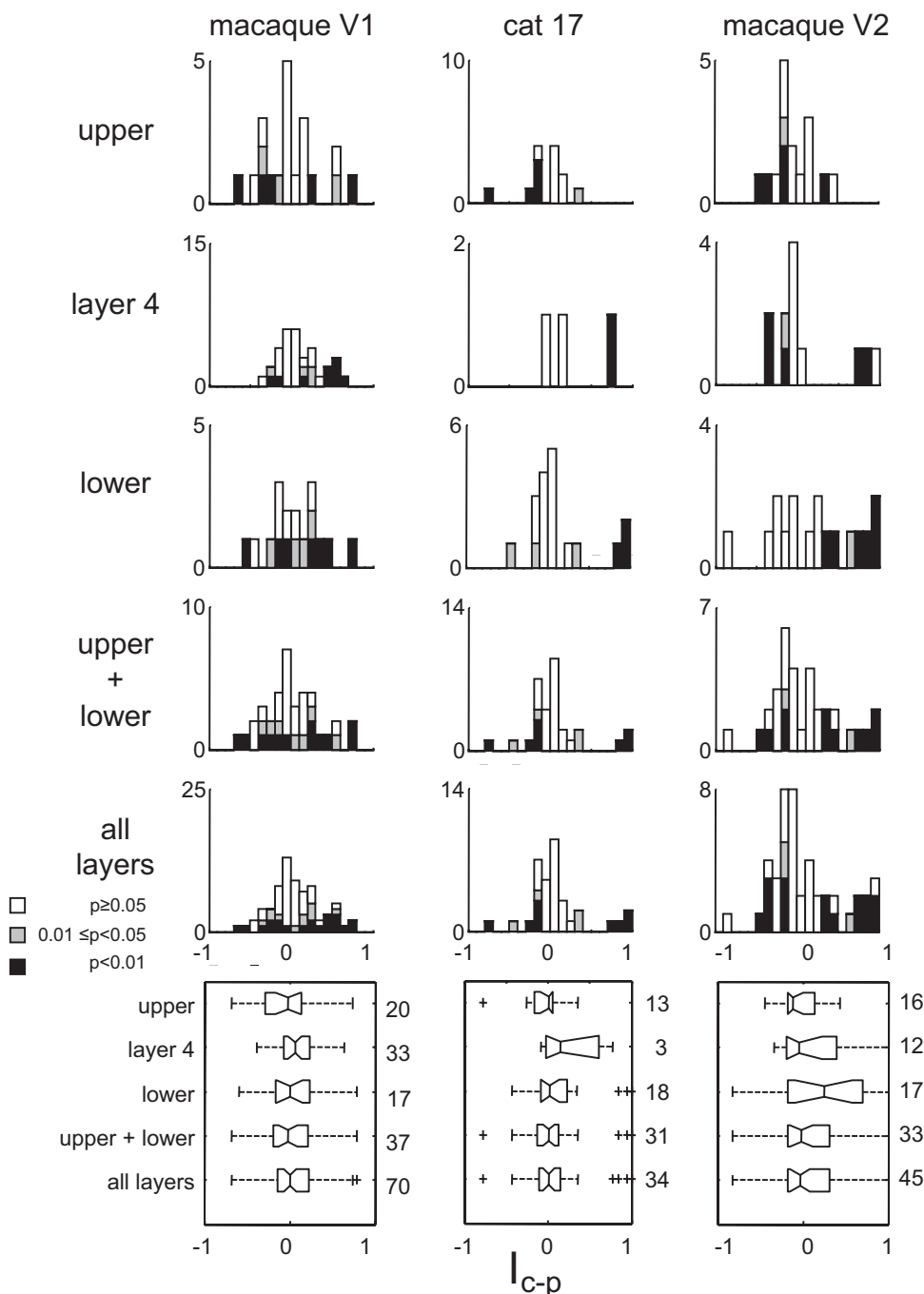


FIG. 4. Population summary of the dependence of overall response size on basis set, as measured by the index  $I_{c-p}$  (Eq. 5). Values of  $I_{c-p}$  significantly  $>0$  indicate an overall preference for Cartesian stimuli; values significantly  $<0$  indicate an overall preference for polar stimuli. Other plotting conventions as in Fig. 3.

**Correlation with simple versus complex behavior.** The distinction between simple and complex behavior is one of the fundamental emergent properties of striate cortex. Its genesis appears closely related to the balance between thalamic and intrinsic cortical inputs, although the extent to which this distinction is a categorical one governed by hierarchical anatomy (Hubel and Wiesel 1962, 1968; Skottun et al. 1991)—versus a continuum of behaviors governed by intrinsic cortical circuitry (Chance et al. 1999; Priebe et al. 2004; Tao et al. 2004)—remains the subject of debate (Abbott and Chance 2002). Recognizing that there is no consensus on how this distinction is best quantified (Kagan et al. 2002), we use the  $F_1/F_0$  ratio because it is readily quantified and captures the general tendency toward linear spatial summation ( $F_1/F_0$  large) and ON-OFF behavior ( $F_1/F_0$  near 0).

In macaque V1, we found (Fig. 5, *top left*) that the simple versus complex distinction correlated with the dependence of sensitivity profile shape on basis set ( $I_{shape}$ ). This correlation was statistically significant in a categorical analysis, in which complex cells are operationally defined by  $F_1/F_0 < 1$  (shape change in 35/50 complex cells and 6/15 simple cells,  $P < 0.05$ ,  $\chi^2$ ), and in a continuum analysis (correlation between  $I_{shape}$  and  $F_1/F_0$  across all cells was  $r = 0.34$ ,  $P < 0.01$ ,  $n = 65$ ). This correlation did not account for the finding that shape changes in V1 were more prominent outside of the input layer (Fig. 3). That is, the laminar differences persisted when the analysis was restricted to the 50 complex cells in macaque V1 (8/19 complex cells in Layer 4 with  $I_{shape} < 1$ , 22/31 complex cells not in Layer 4,  $P < 0.01$  by  $\chi^2$ ). For the 15 simple cells, the trend was in the same direction (3/10 simple cells in Layer 4 with  $I_{shape} < 1$ , 4/5 simple cells not in Layer 4,  $P \approx 0.1$  by  $\chi^2$ ).

A correlation between  $I_{shape}$  and the simple-versus-complex distinction was found in cat area 17 ( $r = 0.25$ ,  $P \approx 0.15$ ,  $n = 34$ ) and macaque V2 ( $r = 0.29$ ,  $P \approx 0.07$ ,  $n = 40$ ). The size of the correlation was similar to that seen in macaque V1, but did not reach statistical significance, likely because of the fewer number of units available for analysis.

In contrast to the behavior of  $I_{shape}$ , there was no correlation between the Cartesian versus polar preference ( $I_{c-p}$ ) and the

$F_1/F_0$  ratio (Fig. 5, *second row*: all  $|r| < 0.15$  and  $P > 0.3$ ), or between the number of units for which  $I_{c-p} \neq 0$ , and the simple-versus-complex categorical distinction ( $P > 0.3$  for macaque V1, cat area 17, and macaque V2).

#### Numerical simulations: gain controls and orientation-specific interactions

Here we show that normalizations driven by overall contrast (Geisler and Albrecht 1992; Heeger 1992, 1993) or interactions tuned to particular orientations (Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002) have only a minimal impact on our analysis and are thus unlikely to account for our findings. Our strategy is to carry out numerical simulations on several families of models that incorporate these mechanisms, to calculate RF maps from their responses just as for the recorded neurons, and to compare the behavior of the resulting indices  $I_{shape}$  and  $I_{c-p}$  with the above-cited findings.

We created a basic model neuron with a Gabor-shaped sensitivity profile, followed by half-wave rectification (METHODS, Eq. 8). We allowed the response of the model neuron to be influenced by a signal derived from the combined output of a population of similar auxiliary Gabor neurons (Eqs. 10 and 11). We then applied the same RF analysis procedures to the model neuron's responses to TDH functions as we applied to the recorded neurons. By varying the spatial parameters of the population of neurons that contributed to the modulatory signal, we determined the effect of multiple kinds of response normalizations and orientation-specific interactions.

The large profiles above each  $2 \times 2$  map set in Fig. 6 are the RF sensitivity profiles  $G(x,y)$  of a sample of simulated Gabor neurons in the absence of any normalization (see METHODS for details). The  $2 \times 2$  map sets below each large profile are the sensitivity profiles that are obtained by our analysis procedure, Eqs. 2 and 3, in the presence of a normalization circuit. Here, the normalization pool consisted of  $N_{aux} = 120$  randomly oriented neurons and the divisive normalization constant was  $c_0 = 36$  (Eq. 11). As can be seen from the figure, the inferred shapes of the Cartesian and polar linear filters,  $L_{cart}$  and  $L_{polar}$ , are very nearly

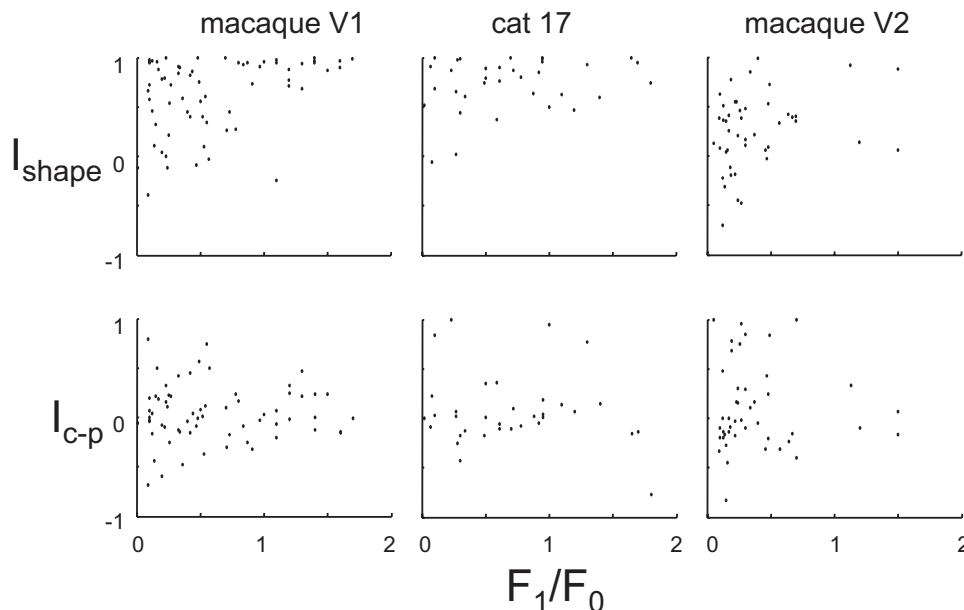


FIG. 5. Relationship of characteristics of TDH responses to the  $F_1/F_0$  ratio. *First row*: dependence of sensitivity profile  $L$  on basis set ( $I_{shape}$ , Eq. 4). *Second row*: Cartesian vs. polar preference ( $I_{c-p}$ , Eq. 5).

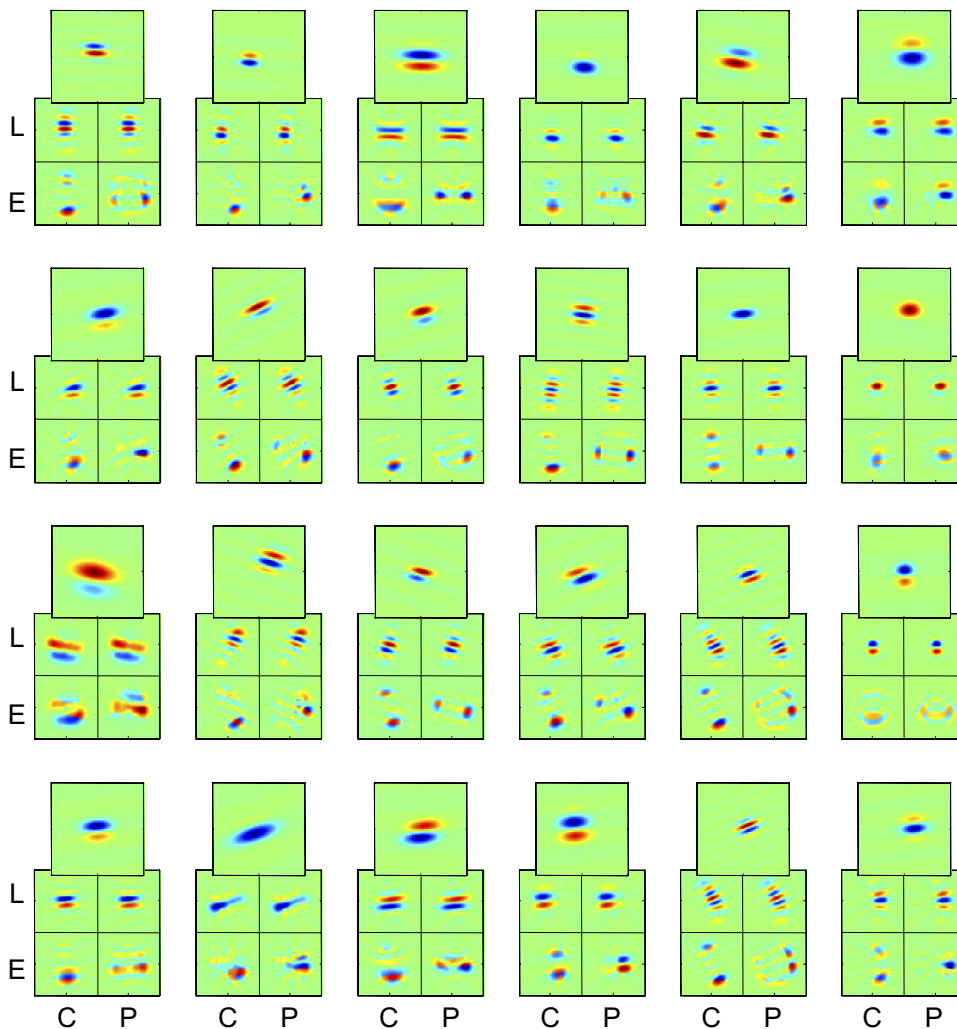


FIG. 6. Numerical simulation of the effect of a simple model of contrast normalization. In each set of profiles, the large profile is the receptive field sensitivity profile, i.e., a Gabor function  $G(x,y)$ . The  $2 \times 2$  map set underneath are the  $L$ - and  $E$ -filters estimated from simulated responses in which the output of this filter is divisively normalized by an isotropic (unoriented) local contrast signal, derived by combining the output of similar nearby Gabor filters. Note that the filters  $L_{cart}$  and  $L_{polar}$  estimated from responses to Cartesian and polar TDH stimuli are nearly identical. For further details, see METHODS.

identical to each other. Additionally, although the shapes of the filters  $E_{cart}$  and  $E_{polar}$  are somewhat different (as anticipated from the sign ambiguity discussed earlier in connection with Eqs. 2 and 3), the overall sensitivity to Cartesian and polar stimuli for each neuron is also very nearly identical.

We also note that although the shapes of the inferred filters ( $L_{cart}$  and  $L_{polar}$ ) are nearly identical to each other, they both differ somewhat from the shapes of the Gabor sensitivity profiles  $G(x,y)$  (the large profiles above each  $2 \times 2$  map set). This is expected, even in the absence of normalizations, because  $L_{cart}$  and  $L_{polar}$  estimate the *projection* of the neuron's sensitivity profile into the space spanned by the TDH functions. Just like the experimental estimates of  $L_{cart}$  and  $L_{polar}$  (e.g., Figs. 2, 9, 10, and 11), they tend to have more lobes than expected from cortical RFs. However, since Cartesian and polar basis sets span the same space, this projection does not interfere with testing the prediction that  $L_{cart} = L_{polar}$ , which holds to a very good approximation for this gain control model.

Figure 7 quantifies and extends this basic finding: that typical contrast normalization and cross-orientation interactions have very little effect on the sensitivity profile estimates that emerge from TDH stimulation. Each row shows the distribution of values of the index  $I_{shape}$  and the size index  $I_{c-p}$ , for the standard modulatory strength  $c = c_0$  and a higher strength  $c = 3c_0$ . The several rows correspond to six different kinds of normalizations

and cross-orientation interactions, defined in terms of the spatial configuration of the auxiliary neurons that generate the modulatory signal. Row A corresponds to the model used in Fig. 6; row B is a similar isotropic model, but the gain control population is spread over twice the distance; row C models oriented suppression from the orthogonal orientation within the RF (cross-orientation suppression); row D models iso-orientation suppression from the end-zones (end-stopping); row E models iso-oriented suppression from an annulus surrounding the RF (iso-orientation surround suppression); and row F is suppression in orientation sidebands from an annulus surrounding the RF. The second column of the figure displays the magnitude of the resulting modulatory effect, set at its standard value of  $c = c_0$ : responses to many TDH stimuli are typically attenuated by a factor of 2. The next two columns show that despite this large effect on responses to individual stimuli, the effect on receptive field maps is minimal: the resulting distribution of  $I_{shape}$  is very close to 1 and the distribution of  $I_{c-p}$  is close to zero. The final two columns show that when the strength of the modulatory interaction is tripled ( $c = 3c_0$ ), this finding persists:  $I_{shape}$  remains close to 1 and  $I_{c-p}$  remains close to zero. This is in marked contrast to the behavior of the physiological recordings (Figs. 3 and 4).

Figure 8 shows that the above-cited observations are robust with respect to the number of neurons in the auxiliary pool. As shown, even when there is only  $N_{aux} = 1$  neuron in the auxiliary

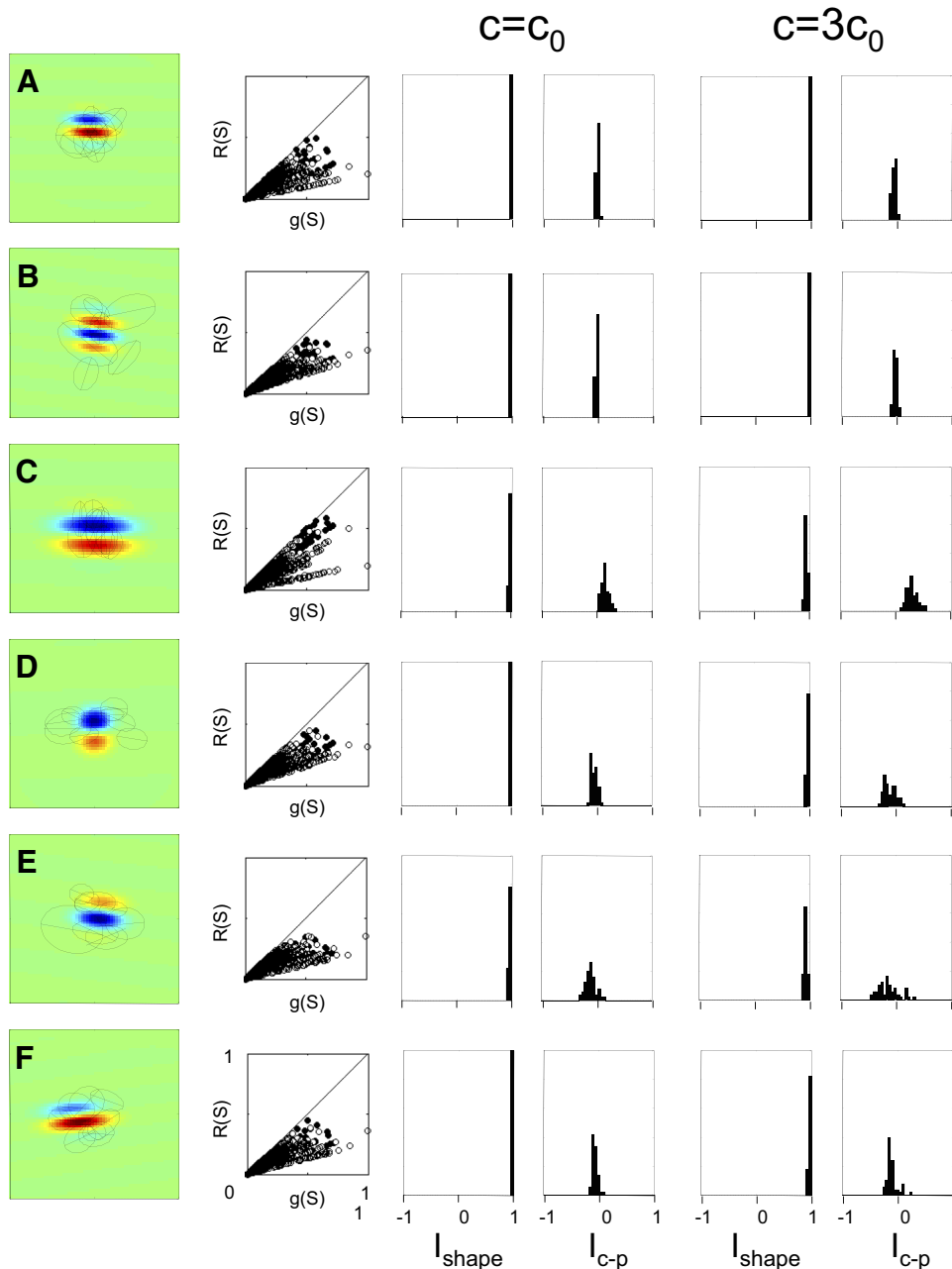


FIG. 7. Numerical simulation of the effect of several kinds of normalizations and orientation-specific interactions on RF indices. Characteristics of the auxiliary neurons that provide the normalization or interaction signal are as follows: *A*: randomly oriented, coextensive with the target neuron RF (center SD identical to RF envelope parameter  $\sigma$ ). *B*: randomly oriented, more widely distributed than the target neuron RF (center SD  $2\sigma$ ). *C*: as in *A*, but orthogonal orientations ( $SD\ 11.25^\circ$ ) to the modeled neuron. *D*: as in *A*, but same orientation as the modeled neuron and positioned, on average  $\pm\sigma$  away from its center (oriented end-zone inhibition) along the axis of RF orientation. *E*: as in *A*, but centers within a radius of  $\sigma$  excluded, and same orientation as modeled neuron (iso-orientation suppression from the nonclassical RF). *F*: as in *E*, but orientations differ from the modeled neuron by  $\pm 30^\circ$  (side-band suppression from the nonclassical RF). The 1st column shows examples of specific neurons simulated, with the modeled neuron's sensitivity profile in pseudo-color and typical auxiliary neurons diagrammed by the  $1/e$  level curves of their envelopes. (For clarity, only 8 auxiliary neurons are diagrammed;  $N_{aux} = 120$  were used in each simulation.) The 2nd column is a scattergram of the raw responses of the model neurons to each TDH stimulus [ $g(S)$ , Eq. 8] and its response following normalization [ $R(S)$ , Eq. 11], for Cartesian stimuli (filled) and polar stimuli (empty) for  $c = c_0$ , indicating that the normalization had a large effect on response size. The remaining 2 pairs of columns are histograms of the indices  $I_{shape}$  and  $I_{c-p}$ , for the standard normalization strength  $c = c_0$  (columns 3 and 4) and  $c = 3c_0$  (columns 5 and 6). Note that the bin width used in this figure (0.05) is half of the bin width used in Figs. 3 and 4 because the distributions of the indices derived from simulations are tightly clustered.

pool, the basic finding of Fig. 7 persists: the effective RF shape identified by Cartesian or polar stimuli is virtually unchanged and the sensitivity to Cartesian and polar stimuli is very nearly identical. Even in this extreme case, the behavior of these model neurons— $I_{shape}$  is nearly always  $>0.9$ —departs substantially from what was seen in the recorded neurons (Fig. 3).

The findings summarized in Figs. 7 and 8 were also seen with many other variations of the model (see METHODS for details), including 1) removal of the jitter of the distribution of orientations; 2) other kinds of nonlinearities for the individual neurons (asymmetric rectification and exponential); 3) other kinds of nonlinearities that pooled the auxiliary neurons' signals (squaring, half-wave, and full-wave rectification; see comments following Eq. 10); and 4) subtractive (Eq. 12), rather than divisive (Eq. 11), normalization. None of these models accounts for the apparent RF shape changes seen in the

recorded neurons: nearly all model neurons had  $I_{shape} > 0.9$ , but many of the recorded neurons had  $I_{shape} < 0.50$  (Fig. 3).

Although none of the model configurations led to substantial differences in apparent RF shape, some led to modest differences in sensitivity to Cartesian and polar stimulus sets (as quantified by the size index  $I_{c-p}$ ). For example, models that featured suppression from the orientation orthogonal to the preferred orientation (Figs. 7 and 8, row *C*) resulted in  $I_{c-p} > 0$  (greater sensitivity to Cartesian stimuli), qualitatively similar to the behavior seen for the unit of Fig. 2A. Conversely, models that featured suppression from orientations similar to the preferred orientation (Figs. 7 and 8, rows *D*, *E*, and *F*) resulted in  $I_{c-p} < 0$  (greater sensitivity to polar stimuli). Qualitatively similar behavior was also seen in recorded neurons (e.g., Fig. 11A). These behaviors make intuitive sense. A neuron with cross-orientation suppression will tend to respond optimally to stimuli that contain elongated regions

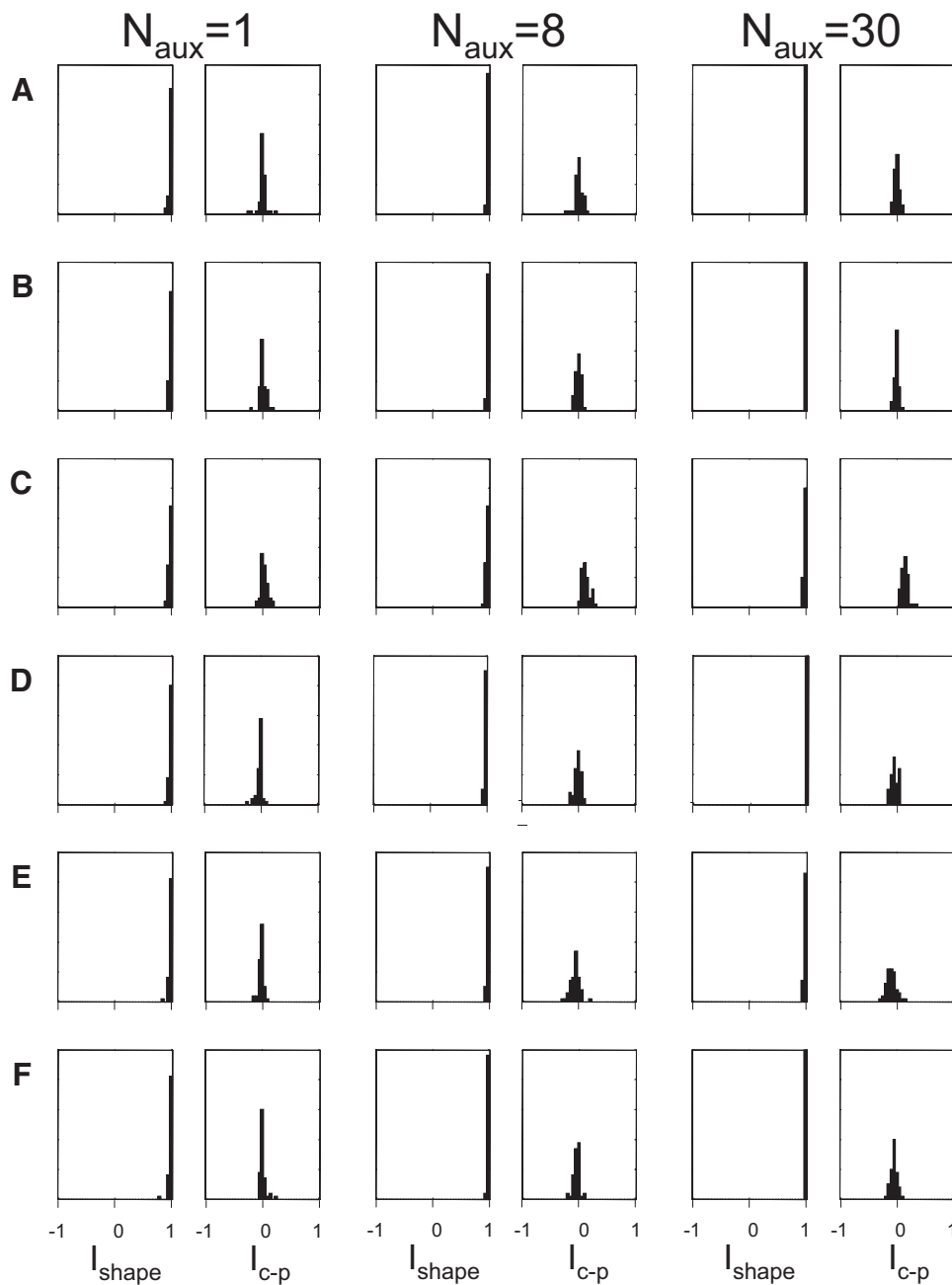


FIG. 8. Numerical simulation of the influence of  $N_{aux}$ , the number of neurons in the auxiliary pool. Pairs of columns shows histograms of the indices  $I_{shape}$  and  $I_{c-p}$  for  $N_{aux} = 1$ ,  $N_{aux} = 8$ , and  $N_{aux} = 30$ . Other simulation parameters are as in the first 4 columns of Fig. 7.

only in the preferred direction, and such stimuli are present only in the Cartesian basis set (row *C*). Conversely, neurons with iso-orientation suppression will respond least well to these stimuli and therefore have a preference for the polar basis set (rows *D*, *E*, and *F*).

However, none of the models accounts for the behavior of the recorded neurons. All modeled neurons had  $|I_{c-p}| < 0.5$ , whereas many of the recorded neurons (Fig. 4) were well outside this range.

#### Experiment 2: orientation dependence of receptive field nonlinearities

Above, we showed that most neurons in V1 and V2 manifest a spatial nonlinearity that alters their effective spatial sensitivity profile, depending on whether the sensitivity profile was deter-

mined with Cartesian or polar basis functions. This dependence on Cartesian versus polar context reflects an underlying nonlinearity in the neural computations and we now focus on characterizing it. The numerical simulations of Figs. 6–8 ruled out some possibilities; we next consider others.

We aim to determine the role played by orientation. One possibility is that orientation plays no role—i.e., that the observed dependence on context is independent of stimulus orientation and relates instead to differences between Cartesian and polar stimuli as *classes*. For example, only the polar stimuli have circular contours or rotational symmetry. The second possibility is that the orientation at which a stimulus is presented plays the critical role, not its Cartesian versus polar character.

To make this distinction, we compare responses to TDH functions *aligned* with the orientation preference (Fig. 1, *top*) and

responses to TDH functions rotated  $45^\circ$  so that their contours are *oblique* (Fig. 1, *bottom*). This manipulation does not influence attributes that are intrinsic to the Cartesian patterns, but markedly influences the orientations that they contain. We find that contour orientation, not the Cartesian versus polar distinction per se, is the crucial factor. Since we have shown earlier that orientation signals derived from spatial frequency content (extracted via pools of randomly positioned Gabor functions) cannot drive the substantial difference in the receptive fields identified by Cartesian and polar stimuli, we conclude that orientation must be extracted by a qualitatively different mechanism, such as nonlinearities that are sensitive to specific phase correlations (i.e., spatial correlations of order three and above).

These experiments were carried out in 59 of the units studied in *experiment 1* (29 macaque V1 units, 12 cat area 17 units, 18 macaque V2 units). As with *experiment 1*, we first present some example responses and then a population summary.

**EXAMPLE RESPONSES.** Figure 9 shows a unit recorded in Layer 4b of macaque V1. In keeping with its simple-like ( $F_1/F_0 = 1.5$ ) behavior, responses to most TDH stimuli were dependent on the stimulus polarity. For the TDH stimuli that were aligned to the preferred orientation (Fig. 9A), responses were generally larger for the Cartesian stimuli than those for the polar ones. The opposite was true for TDH stimuli that were oblique to the preferred orientation (Fig. 9B): responses to polar stimuli were large; responses to Cartesian stimuli were small or nonexistent. This observation is borne out by the index  $I_{c-p}$  (Eq. 5) of Cartesian versus polar preference: for aligned stimuli, there was a significant preference for Cartesian stimuli ( $I_{c-p}^{aligned}: 0.23 \pm 0.08, P < 0.05$ ), whereas for oblique stimuli the preference was strongly in the opposite direction ( $I_{c-p}^{oblique}: -0.50 \pm 0.12, P < 0.01$ ). For this unit, there was no significant change in RF shape; neither  $I_{shape}^{aligned}$  nor  $I_{shape}^{oblique}$  differed significantly from 1.

The unit of Fig. 10, recorded in Layer 4 of macaque V2, showed a behavior that contrasts in several respects to the V1 unit in Fig. 10. It had complex-like behavior ( $F_1/F_0 < 0.1$ ) and very similar responses to TDH functions of either polarity. It had at most a modest preference for Cartesian TDH functions ( $I_{c-p}^{aligned}: 0.36 \pm 0.80; I_{c-p}^{oblique}: -0.05 \pm 0.44$ , neither significantly different from 0). However, this unit showed a substantial change in the shape of its sensitivity profile when the oblique Cartesian set was presented ( $I_{shape}^{aligned}: 0.12 \pm 0.43; I_{shape}^{oblique}: 0.31 \pm 0.14$ , both  $P < 0.05$ ).

Sensitivity profiles are shown for three more units in Fig. 11. The unit of Fig. 11A is a V1 Layer 4 simple-like ( $F_1/F_0 = 1.1$ ) unit. This unit had a mild preference for polar stimuli under aligned conditions ( $I_{c-p}^{aligned}: -0.21 \pm 0.08, P < 0.05; I_{c-p}^{oblique}: -0.12 \pm 0.08, P > 0.05$ ) and no significant shape changes for either orientation. The two units of Fig. 11, B and C were recorded simultaneously and shared the same orientation preference, but one (55031s, Fig. 11B) was simple-like ( $F_1/F_0 = 1.2$ ) and directionally biased; the other (55031t, Fig. 11C) was complex-like ( $F_1/F_0 = 0.5$ ) and not directionally biased. Their responses to TDH functions had similar characteristics. For both units, responses were primarily polarity-independent for aligned and oblique stimuli, as seen by the larger contribution of the even-order pathway (*E*-filters) compared with the linear pathway (*L*-filters). Neither unit had a substantial preference for Cartesian or polar stimuli in either orientation ( $I_{c-p}^{aligned}$  and  $I_{c-p}^{oblique}$  not significantly different from 0). For the unit of Fig. 11B, there was a modest but statistically significant shape change for both aligned and oblique stimuli ( $I_{shape}^{aligned}: 0.77 \pm 0.10; I_{shape}^{oblique}: 0.80 \pm 0.08$ , both  $P < 0.05$ ). For the unit of Fig. 11C, a similar degree of shape change was present for aligned and oblique stimuli, but this was statistically significant only for aligned stimuli ( $I_{shape}^{aligned}: 0.40 \pm 0.13, P < 0.05; I_{shape}^{oblique}: 0.83 \pm 0.15, P > 0.05$ ).

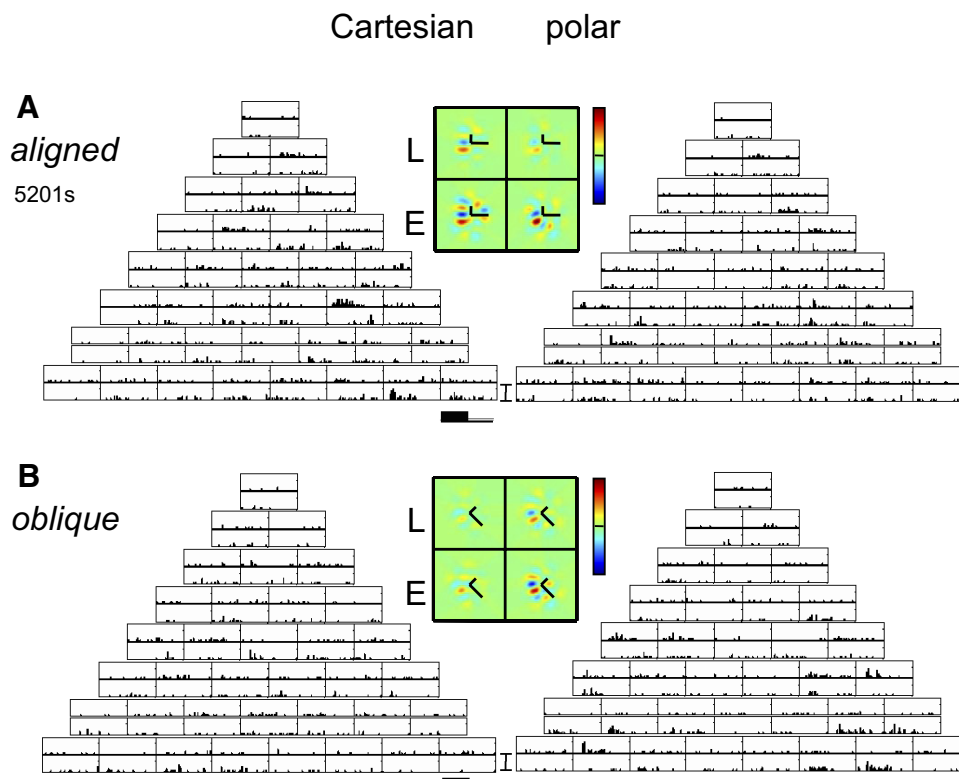


FIG. 9. PSTHs and inferred sensitivity profiles for a macaque V1 Layer 4 neuron, obtained from TDH functions aligned (A) and oblique (B) to the preferred orientation. PSTH scale bar: 75 impulses/s; range for pseudocolor maps of filters:  $\pm 5$  impulses/s. The long arm of the superimposed "L" indicates the *x*-axis of the stimulus; the short arm indicates the *y*-axis. Maps of sensitivity profiles span  $4.0^\circ$ . Other plotting conventions as in Fig. 2.



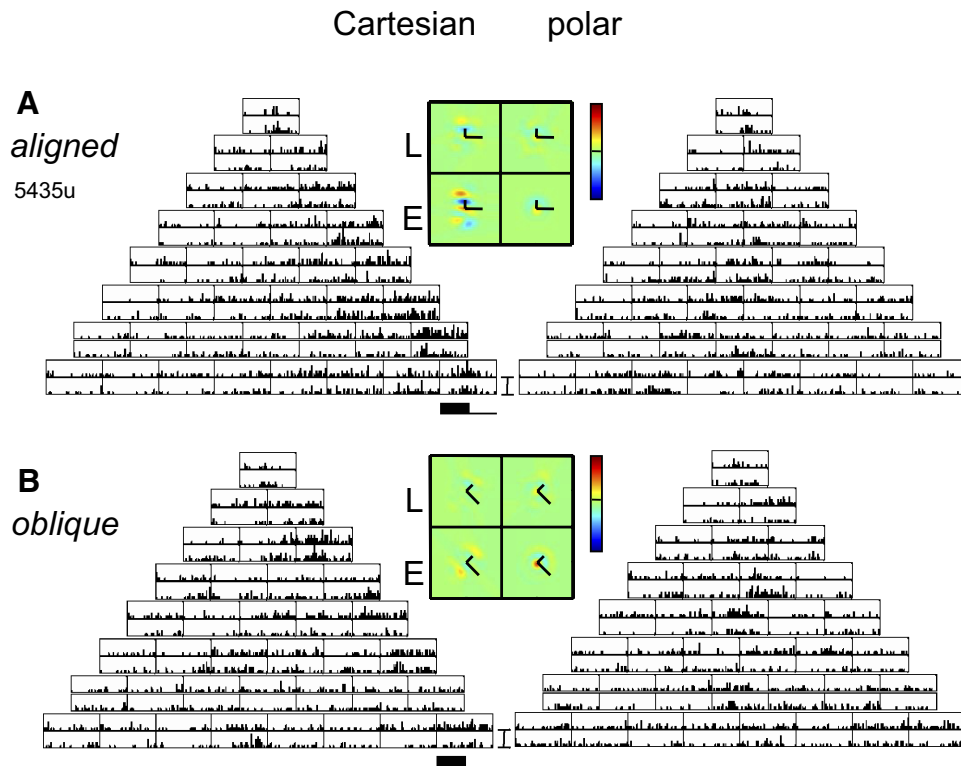


FIG. 10. PSTHs and inferred sensitivity profiles for a macaque V2 Layer 4 neuron, obtained from TDH functions aligned (A) and oblique (B) to the preferred orientation. PSTH scale bar: 50 impulses/s; range for pseudocolor maps of filters:  $\pm 10$  impulses/s. Maps of sensitivity profiles span  $4.8^\circ$ . Other plotting conventions as in Fig. 9.

POPULATION SUMMARY. The several examples above illustrate the substantial diversity we observe in the characteristics of neural responses to TDH functions. For many neurons, the apparent RF shape depends on whether it is assayed with Cartesian or polar functions. For another set of neurons, partially overlapping with the first set, there is an overall response preference for Cartesian or polar functions. These phenomena—both of which represent departures from the expected behavior of a linear-static nonlinear filter or an LN cascade modulated by gain controls such as divisive normalization or cross-orientation suppression—may be present, independently, for TDH stimuli aligned to the preferred orientation or oblique to it.

To focus on the question of interest—whether the nonlinearity that underlies the difference in apparent RF shape for Cartesian and polar stimuli is orientation sensitive—we make three com-

parisons of estimated RF shape: Cartesian versus polar estimates for aligned stimuli ( $I_{shape}^{aligned}$ ), Cartesian versus polar estimates for oblique stimuli ( $I_{shape}^{oblique}$ ), and Cartesian estimates for aligned and oblique stimuli ( $I_{shape}^{aligned\_vs\_oblique}$ ). (We do not compare estimates from polar stimuli for the two rotations, since many aligned and oblique polar stimuli are identical.)

Results are shown in Fig. 12. The difference between the sensitivity profile measured with the two orientations of Cartesian stimuli ( $I_{shape}^{aligned\_vs\_oblique}$ ) is more marked than the difference between Cartesian and polar stimuli ( $I_{shape}^{aligned}$  or  $I_{shape}^{oblique}$ ). This difference is highly significant in V1 and area 17 ( $P < 0.01$ ) and there is a trend in the same direction in V2. Since changing orientation of the elongated contours (but retaining their “Cartesian” character) has a greater effect than removing them altogether, the nonlinearity that drives the change in apparent recep-

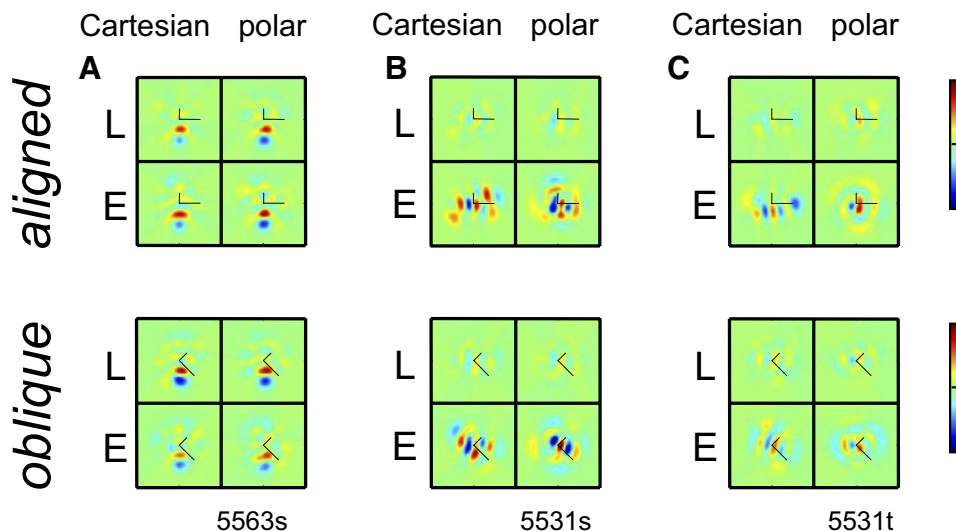


FIG. 11. Inferred sensitivity profiles for 3 macaque V1 neurons (A: layer 4; B and C: Layers 2/3) in response to aligned and oblique TDH functions. Range for pseudocolor maps of filters:  $\pm 10$  impulses/s (A),  $\pm 24$  impulses/s (B), and  $\pm 5$  impulses/s (C). Maps of sensitivity profiles span  $1.6^\circ$  (A),  $2.4^\circ$  (B), and  $2.4^\circ$  (C). Other plotting conventions as in Fig. 9.

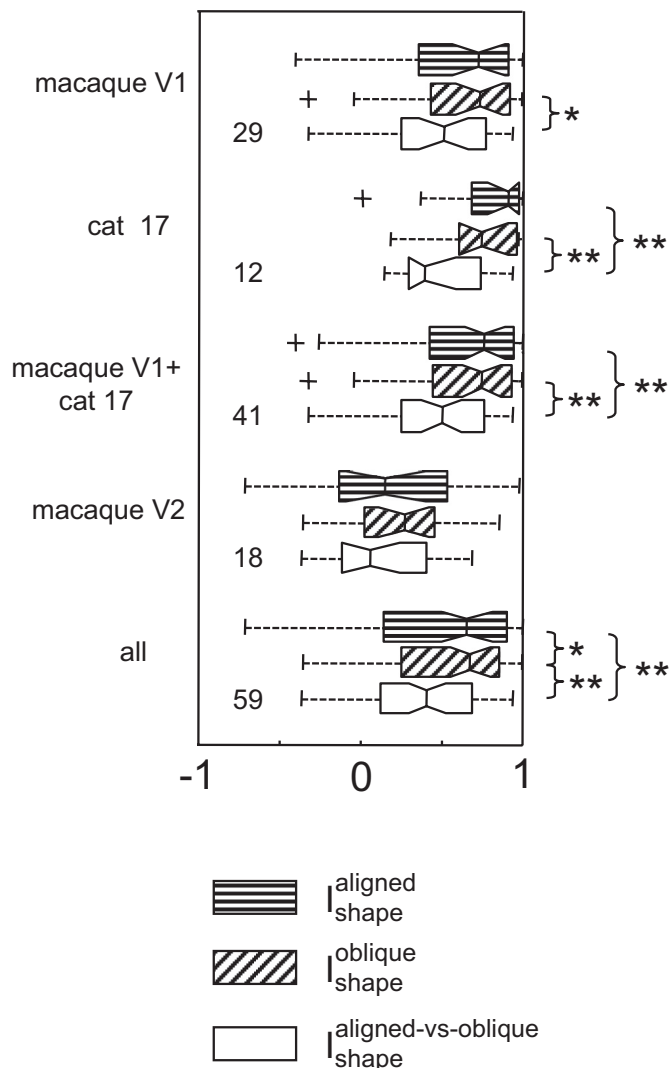


FIG. 12. Comparison of changes in the sensitivity profile for Cartesian vs. polar stimuli of a single orientation ( $I_{shape}^{aligned}$  and  $I_{shape}^{oblique}$ , Eq. 4, horizontal and oblique hatched symbols), and Cartesian stimuli in 2 different orientations ( $I_{shape}^{aligned\_vs\_oblique}$ , Eq. 7, open symbols). Within each set of boxplots, significance levels (2-tailed paired *t*-test) are indicated by \*\* $P < 0.01$  and \* $0.05 < P < 0.01$ . Other boxplot conventions as in Fig. 3.

tive field shape is sensitive to the orientation of the Cartesian stimuli, rather than to the Cartesian versus polar distinction per se.

We note that the range of shape changes is similar for aligned and oblique conditions and, on a neuron-by-neuron basis, the correlation between them is strong in macaque (V1:  $r = 0.79$ ,  $P < 0.001$ ,  $n = 29$ ; V2:  $r = 0.52$ ,  $P < 0.03$ ,  $n = 18$ ), with a similar trend in a smaller number of cat area 17 units ( $r = 0.28$ ,  $P \approx 0.4$ ,  $n = 12$ ). Thus if a neuron has a discrepancy between the sensitivity profiles for aligned Cartesian and polar TDH stimuli, then it will have a discrepancy of similar magnitude when studied with oblique stimuli, but the nature of the shape change typically differs.

We also did not find any difference between the overall sizes of the responses to aligned and oblique Cartesian stimuli. Such a difference might have been generated by a tuned cross-orientation suppression driven selectively by aligned or oblique Cartesian stimuli.

Finally, we mention that our finding (that orientation matters more than the Cartesian vs. polar distinction) is not in contradic-

tion to previous findings (Gallant et al. 1996; Hegde and Van Essen 2003; Mahon and De Valois 2001) concerning preference of visual neurons for Cartesian and polar gratings. Despite their superficial resemblance, TDH stimuli differ from the Cartesian and polar gratings used in those studies in an important way—TDH stimuli are matched for various low-level characteristics, as is required for our goal of testing models; this was not the case for the stimuli used in the earlier studies.

## DISCUSSION

The goal of this study was to further the understanding of the computations performed in primary visual cortex. To this end, we identified qualitative behavior that major *classes* of models must have in response to specific stimuli and tested whether this kind of behavior is present in V1 and V2 of macaque monkeys and area 17 of the cat. As described in the INTRODUCTION, our approach uses single-unit responses to stimuli based on TDH functions. Because of the way that the mathematical properties of TDH functions interact with typical building blocks of computational models, we can bypass stages of processing sensitive to luminance and contrast, to examine nonlinear mechanisms of spatial integration arising in local cortical circuits.

In our initial study with TDH stimuli (Victor et al. 2006), we showed that they elicited behavior that was qualitatively inconsistent with simple cascade models—that is, the apparent shape of the linear filter in the LN cascade depended on which basis set is used for the characterization. Subsequently (Sharpee and Victor 2009), we characterized this behavior as a contextual modulation and showed that the effect of this modulation was to change the strength of inputs from oriented subunits. This raises two further questions: 1) Which cells show this contextual modulation? 2) What stimulus features—and therefore what kinds of mechanisms—drive it? Here, we address these questions.

To determine which cells show this contextual modulation, we carried out a laminar analysis of the responses. This analysis showed that the contextual effect is present throughout V1, more prominently in the supragranular layers and infragranular layers than in Layer 4. Moreover, we found that the contextual effect is larger in V2 than that in V1 (Figs. 3 and 4). The progression of increasing contextual effect from V1 input layer to V1 intrinsic circuitry to V2 suggests that it is intrinsic to the goal of processing, not a “side effect” of early visual processing that is to be removed at later stages.

As a first step in identifying the mechanisms that modulate RF structure in the various TDH contexts, we considered modulatory mechanisms that have already been identified. These include gain controls driven by local contrast and/or spatial frequency content (Albright and Stoner 2002; Cavanaugh et al. 2002; Freeman et al. 2001; Ohzawa et al. 1985; Reid et al. 1992; Sceniak et al. 1999) and modulatory mechanisms sensitive to stimulus power in particular locations and/or at particular orientations (Albrecht and Geisler 1991; Allison et al. 2001; Bonds 1989; Carandini et al. 1998; Durand et al. 2007; Freeman et al. 2002; Geisler and Albrecht 1992; Heeger 1992, 1993).

To assess the possible roles of these mechanisms, we carried out several series of simulations, summarized in Figs. 6–8. We found virtually no influence of modulatory mechanisms on the apparent shape of cascade-model filters when assayed with Cartesian versus polar TDH functions (i.e.,  $I_{shape}$  remains near 1), and only a modest influence on the relative sensitivity to the

two basis sets (i.e.,  $I_{c-p}$  remains near 0). The simulations considered many scenarios for gain controls and spatial interactions, including influences from within and beyond the classical receptive field; isotropic and nonisotropic orientation distributions of the auxiliary neuron pool; and a range of cell numbers, interaction strengths, kinds of nonlinearities, and divisive versus subtractive influences.

The universal failure of these mechanisms to account for our findings might at first appear surprising, but it is an anticipated consequence of the properties of the TDH functions. In contrast to Gabor functions, each TDH function contains a *broad* range of spatial frequencies and orientations. (Since each TDH function is its own Fourier transform, its spectrum can be read off from the stimulus itself.) Because TDH functions are broadband, Gabor-based filters are surprisingly unselective for the Cartesian versus polar distinction—and this underlies the simulation results. A related consideration that helps to understand these simulation results is that cortical neurons are relatively broadly tuned to TDH stimuli (Figs. 2, 9, and 10). Thus many TDH functions enter into the basis-function expansion of the RF profile (Eq. 3)—implying that the ensemble properties of the stimulus sets are more relevant to the contextual modulation than the idiosyncratic luminance distribution of individual functions.

There are many variations on the theme of gain controls and cross-orientation interactions (MacEvoy et al. 2009) and a fully realistic cortical model would necessarily include parallel combinations of cascade-like units (Rust et al. 2005), dynamics, network interactions such as feedback (Chance et al. 1999; Heeger 1993), and spiking neurons (Tao et al. 2004, 2006). However, although such additions might provide greater realism and complexity, they do not alter the basic reason that the above-cited simulations fail to account for our findings: the requisite spatial selectivity is lacking.

What kinds of computations could generate a “context” signal that accounts for our findings? To address this, we consider the results of *experiment 2* in light of the above simulations. *Experiment 2* showed that “context” is determined by the orientation of the stimuli, not the Cartesian versus polar distinction per se. [Specifically, altering the orientation of the Cartesian stimuli by 45° induces a greater change in the apparent RF shape than switching from Cartesian to polar stimuli (Fig. 12).] So we can conclude that the context signal is, in fact, driven by orientation-sensitive mechanisms—but that these mechanisms are more selective than randomly positioned arrays of Gabor filters.

The implications for the nature of cortical computations become evident by casting the discussion in somewhat more abstract terms. Randomly positioned Gabor functions are sensitive to local power and spatial frequency content—i.e., the second-order statistics of the stimuli—and, in these respects, the various TDH ensembles are very nearly equated. To extract a consistent, orientation-dependent, contextual signal that discriminates among the two TDH stimulus sets, cortical mechanisms that extract the orientation of contours must therefore be sensitive to phase correlations, as proposed by Morrone and Burr (1988).

Although interactions of LN cascades with randomly positioned Gabor-like filters cannot carry out these computations, there are simple, biologically reasonable alternatives that can: interactions of Gabor-like filters across spatial scales (Morrone and Burr 1988). The critical requirement is that these filters must be *aligned* rather than randomly positioned, so that responses to oriented features are reinforced (Morrone and

Burr 1988). For example, an edge detector could be constructed by aligning Gabor-like filters of similar orientations but different spatial scales so that their zero-crossings coincide.

Previous studies of neural computations in V1 based on natural scenes (David et al. 2004; Felsen et al. 2005; Touryan et al. 2005) have also inferred the presence of mechanisms sensitive to phase correlations. By identifying the contribution of these mechanisms from responses to a library of designed stimuli, the present study adds to these results in two ways. First, because of the complexity of natural scenes, it is difficult to determine the essential qualities of stimuli that drive these nonlinear mechanisms; here we show that simple, local patterns suffice. Second, the above-cited studies could not rule out the possibility that gain controls or modulatory mechanisms might be the source of the observed nonlinear phenomema. Here, again because of the relative simplicity and controlled nature of the stimulus set, we were able to do so.

Certain features of our findings suggest an important role for intracortical feedback. As previously shown, context acts to modify the strength of *oriented* subunits (Sharpe and Victor 2009) within a receptive field. To this result, we add the current findings that the context signal itself is derived from orientation-sensitive mechanisms—and that contextual modulation is largest outside of Layer 4, but present in Layer 4 nevertheless. That is, Layer 4 neurons, whose subcortical inputs are presumably nonoriented, have RFs whose *oriented* inputs depend on context, implying a functional recurrence.

Based on the greater size of the context effect outside of Layer 4 and overall patterns of local circuit connectivity (Douglas et al. 1989; Kisvarday et al. 1986; Martin 2002; Thomson and Bannister 2003), we speculate that this recurrence primarily results from recurrent connections within supra- and infragranular layers and feedback of these signals to Layer 4. We are unable, however, to determine the anatomical substrate of this recurrence directly by tracking the latency of the context effects across cortical layers. This is because context effects are equally present in the first 25–50 ms as they are in later portions of the response (Victor et al. 2006) and signal-to-noise considerations prevent a substantially finer resolution.

### Conclusion

Our main motivation was to understand the computations carried out by V1 neurons. By analyzing neural responses to a specific set of designed stimuli, we identified a qualitative aspect of their behavior, changes in RF shape, driven by pattern context, and showed that the modulatory mechanism constitutes a functional recurrence of oriented signals. We showed that this behavior is not found in cascade models augmented by gain controls, response normalization, or cross-orientation interactions. The common denominator is that the latter processes extract orientation information based on local spatial frequency content, whereas our results indicate that sensitivity to phase correlations is required. Thus extraction of oriented features via phase correlation emerges as a qualitative aspect of V1 processing.

Natural scenes differ from traditional analytic stimuli in many ways—in addition to having local high-order spatial correlations, they are spatially extensive, cluttered, and often meaningful. Any (or all) of these differences might underlie why current computational models give an incomplete account of neural responses to natural scenes. Although studies with a restricted set of designed stimuli cannot replace analyses based on natural scenes, they do

have an important advantage. As shown here, the use of designed stimuli allowed us to isolate one factor—local spatial correlations—and to show that it alone suffices to require an expansion of our notion of what individual neurons do.

#### ACKNOWLEDGMENTS

Portions of this material were previously presented at the Society for Neuroscience (Victor et al. 2004; Victor and Mechler 2006).

#### GRANTS

This work was supported by National Eye Institute Grant R01 EY-09314 to J. Victor.

#### REFERENCES

- Abbott LF, Chance FS.** Rethinking the taxonomy of visual neurons. *Nat Neurosci* 5: 391–392, 2002.
- Albrecht DG, Geisler WS.** Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis Neurosci* 7: 531–546, 1991.
- Albright TD, Stoner GR.** Contextual influences on visual processing. *Annu Rev Neurosci* 25: 339–379, 2002.
- Allison JD, Smith KR, Bonds AB.** Temporal-frequency tuning of cross-orientation suppression in the cat striate cortex. *Vis Neurosci* 18: 941–948, 2001.
- Aronov D, Reich DS, Mechler F, Victor JD.** Neural coding of spatial phase in V1 of the macaque monkey. *J Neurophysiol* 89: 3304–3327, 2003.
- Bonds AB.** Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis Neurosci* 2: 41–55, 1989.
- Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC.** Do we know what the early visual system does? *J Neurosci* 25: 10577–10597, 2005.
- Carandini M, Movshon JA, Ferster D.** Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology* 37: 501–511, 1998.
- Cavanaugh JR, Bair W, Movshon JA.** Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88: 2530–2546, 2002.
- Chance FS, Nelson SB, Abbott LF.** Complex cells as cortically amplified simple cells. *Nat Neurosci* 2: 277–282, 1999.
- David SV, Gallant JL.** Predicting neuronal responses during natural vision. *Network* 16: 239–260, 2005.
- David SV, Vinje WE, Gallant JL.** Natural stimulus statistics alter the receptive field structure of V1 neurons. *J Neurosci* 24: 6991–7006, 2004.
- Douglas RJ, Martin KAC, Whitteridge D.** A canonical microcircuit for neocortex. *Neural Comput* 1: 480–488, 1989.
- Durand S, Freeman TC, Carandini M.** Temporal properties of surround suppression in cat primary visual cortex. *Vis Neurosci* 24: 679–690, 2007.
- Efron B.** *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Soc. for Industrial and Applied Mathematics, 1982.
- Fee MS, Mitra PP, Kleinfeld D.** Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* 69: 175–188, 1996.
- Felsen G, Touryan J, Han F, Dan Y.** Cortical sensitivity to visual features in natural scenes. *PLoS Biol* 3: e342, 2005.
- Freeman RD, Ohzawa I, Walker G.** Beyond the classical receptive field in the visual cortex. *Prog Brain Res* 134: 157–170, 2001.
- Freeman TC, Durand S, Kiper DC, Carandini M.** Suppression without inhibition in visual cortex. *Neuron* 35: 759–771, 2002.
- Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC.** Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76: 2718–2739, 1996.
- Geisler WS, Albrecht DG.** Cortical neurons: isolation of contrast gain control. *Vision Res* 32: 1409–1410, 1992.
- Heeger DJ.** Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181–197, 1992.
- Heeger DJ.** Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J Neurophysiol* 70: 1885–1898, 1993.
- Hegde J, Van Essen DC.** Strategies of shape representation in macaque visual area V2. *Vis Neurosci* 20: 313–328, 2003.
- Hevner RF, Wong-Riley MT.** Regulation of cytochrome oxidase protein levels by functional activity in the macaque monkey visual system. *J Neurosci* 10: 1331–1340, 1990.
- Hubel DH, Wiesel TN.** Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160: 106–154, 1962.
- Hubel DH, Wiesel TN.** Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195: 215–243, 1968.
- Kagan I, Gur M, Snodderly DM.** Spatial organization of receptive fields of V1 neurons of alert monkeys: comparison with responses to gratings. *J Neurophysiol* 88: 2557–2574, 2002.
- Kisvarday ZF, Martin KA, Freund TF, Magloczky Z, Whitteridge D, Somogyi P.** Synaptic targets of HRP-filled layer III pyramidal cells in the cat striate cortex. *Exp Brain Res* 64: 541–552, 1986.
- MacEvoy SP, Tucker TR, Fitzpatrick D.** A precise form of divisive suppression supports population coding in the primary visual cortex. *Nat Neurosci* 12: 637–645, 2009.
- Mahon LE, De Valois RL.** Cartesian and non-Cartesian responses in LGN, V1, and V2 cells. *Vis Neurosci* 18: 973–981, 2001.
- Martin KA.** Microcircuits in visual cortex. *Curr Opin Neurobiol* 12: 418–425, 2002.
- Mechler F, Reich DS, Victor JD.** Detection and discrimination of relative spatial phase by V1 neurons. *J Neurosci* 22: 6129–6157, 2002.
- Morrone MC, Burr DC.** Feature detection in human vision: a phase-dependent energy model. *Proc R Soc Lond B Biol Sci* 235: 221–245, 1988.
- Ohzawa I, Sclar G, Freeman RD.** Contrast gain control in the cat's visual system. *J Neurophysiol* 54: 651–667, 1985.
- Olshausen B, Field D.** What is the other 85% of V1 doing? In: *Problems in Systems Neuroscience*, edited by Sejnowski T, van Hemmen L. Oxford, UK: Oxford Univ. Press, 2004.
- Priebe NJ, Mechler F, Carandini M, Ferster D.** The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat Neurosci* 7: 1113–1122, 2004.
- Reich DS.** *Information Encoding by Individual Neurons and Groups of Neurons in the Primary Visual Cortex* (PhD thesis). New York: Rockefeller Univ., 2000.
- Reid RC, Victor JD, Shapley RM.** Broadband temporal stimuli decrease the integration time of neurons in cat striate cortex. *Vis Neurosci* 9: 39–45, 1992.
- Ringach DL.** Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *J Neurophysiol* 88: 455–463, 2002.
- Ringach DL, Sapiro G, Shapley R.** A subspace reverse-correlation technique for the study of visual neurons. *Vision Res* 37: 2455–2464, 1997.
- Rust NC, Schwartz O, Movshon JA, Simoncelli EP.** Spatiotemporal elements of macaque V1 receptive fields. *Neuron* 46: 945–956, 2005.
- Sahani M, Pezaris JS, Andersen RA.** On the separation of signals from neighboring cells in tetrode recordings. In: *Advances in Neural Information Processing Systems 10*, edited by Jordan MI, Kearns MJ, Solla SA. Cambridge, MA: MIT Press, 1998, p. 222–228.
- Sceniak MP, Ringach DL, Hawken MJ, Shapley R.** Contrast's effect on spatial summation by macaque V1 neurons. *Nat Neurosci* 2: 733–739, 1999.
- Sharpee T, Rust NC, Bialek W.** Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput* 16: 223–250, 2004.
- Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD.** Adaptive filtering enhances information transmission in visual cortex. *Nature* 439: 936–942, 2006.
- Sharpee TO, Victor JD.** Contextual modulation of V1 receptive fields depends on their spatial symmetry. *J Comput Neurosci* 26: 203–218, 2009.
- Skottun BC, De Valois RL, Grosfof DH, Movshon JA, Albrecht DG, Bonds AB.** Classifying simple and complex cells on the basis of response modulation. *Vision Res* 31: 1079–1086, 1991.
- Tao L, Cai D, McLaughlin DW, Shelley MJ, Shapley R.** Orientation selectivity in visual cortex by fluctuation-controlled criticality. *Proc Natl Acad Sci USA* 103: 12911–12916, 2006.
- Tao L, Shelley M, McLaughlin D, Shapley R.** An egalitarian network model for the emergence of simple and complex cells in visual cortex. *Proc Natl Acad Sci USA* 101: 366–371, 2004.
- Thomson AM, Bannister AP.** Interlaminar connections in the neocortex. *Cereb Cortex* 13: 5–14, 2003.
- Touryan J, Felsen G, Dan Y.** Spatial structure of complex cell receptive fields measured with natural images. *Neuron* 45: 781–791, 2005.
- Victor JD, Knight BW.** Simultaneously band and space limited functions in two dimensions, and receptive fields of visual neurons. In: *Springer Applied Mathematical Sciences Series*, Kaplan E, Marsden J, Sreenivasan KR. New York: Springer, 2003, p. 375–420.
- Victor JD, Mechler F.** Responses of cat and monkey striate cortical neurons to aligned and oblique two-dimensional Hermite function stimuli. *Soc Neurosci Abstr* 436.2, 2006.
- Victor JD, Mechler F, Repucci MA, Purpura KP, Knight BW.** Responses of neurons in cat and monkey striate cortex to two-dimensional Hermite function stimuli. *Soc Neurosci Abstr* 370.4, 2004.
- Victor JD, Mechler F, Repucci MA, Purpura KP, Sharpee T.** Responses of V1 neurons to two-dimensional Hermite functions. *J Neurophysiol* 95: 379–400, 2006.