SUBMITTED TO: Annual Computational Neuroscience Meeting
Alicante, Spain, 5-9 July 2003

Spike-triggered characterization of excitatory and suppressive stimulus dimensions in monkey V1 directionally selective neurons

Nicole C Rust[1], Odelia Schwartz[3], J Anthony Movshon[1,2] and Eero Simoncelli[1,2]

[1]Center for Neural Science, New York University, New York, NY
[2]HHMI, New York Unviersity, New York, NY
[3]Salk Institute, La Jolla, CA

Neurons in primary visual cortex (V1) are commonly characterized using linear models, or simple combinations of linear models. We examined this class of model directly by applying spike-triggered covariance analysis to the responses of monkey V1 neurons stimulated by binary white noise. The analysis extracts a low-dimensional subspace in which the axes define excitatory and suppressive influences upon the neural response. Across a population of cells, we reliably found more excitatory dimensions than are typically used to model V1 responses. We also found evidence for suppressive dimensions that were at least equal in number to the excitatory dimensions.

Within primary visual cortex (V1), a number of computations are performed. V1 is the first visual area to contain cells selective for the orientation and direction of a moving stimulus. It is also the first site where phase-invariant complex cells are found. Although much emphasis has been placed upon examining the details of these computations, the mechanisms underlying them are not fully characterized. Potential mechanisms can be constrained by models that accurately characterize the response properties of each cell type found within V1.

The motion-energy model (Adelson and Bergen, 1985) is a multi-stage model that describes the construction of directionally selective, phase-invariant units from oriented, phase-sensitive detectors. Specifically, this model describes directionally tuned simple cell responses as the linear combination of two linear filters, whereas complex cell responses are described as the nonlinear combination of two linear subunits in quadrature. Tests of the motion energy model have been hampered by the difficulty of adequately characterizing both excitation and suppression in the same experiment, due to low baseline firing rates and rectification. Furthermore, the second-order response properties of complex cells have hindered application of linear systems analysis to this prominent cell type. One recently developed extension to linear systems analysis, spike triggered covariance (de Ruyter and Bialek, 1998), has been shown to be capable of mapping the subunits of phase-insensitive complex cells (Touryan et al, 2002) as well as both suppressive and excitatory influences (Schwartz et al, 2001). In this article, we report a number of novel findings that arise from application of such techniques to analysis of responses in macaque primary visual cortex

Methods

We applied a spike triggered covariance (STC) analysis to data collected from isolated directionally tuned simple cells (n=12) and complex (n=11) cells within primary visual cortex (V1) of opiate anesthetized, paralyzed macaque monkeys (see Cavanaugh et al, 2002 for experimental details). Stimuli were extended temporal sequences in which each frame contained a set of parallel non-overlapping dark and bright bars with randomly assigned intensity. The orientation of the bars was aligned with the cell's preferred orientation and confined to the classical receptive field. The number of bars (8-32) was chosen based on the cell's preferred spatial frequency such that 4-8 bars fell within each spatial period and the whole array spanned the classical receptive field. Each frame had a duration of 10 msec.

We denote the set of stimuli that were presented over some fixed time interval preceding the nth spike as $S_n(x,t)$. In the conventional procedure of reverse correlation, one averages these stimuli over all spikes to obtain a linear kernel (the "spike-triggered average", STA):

$$STA(x,t) = \frac{1}{N}\sum_n S_n(x,t)$$

where N indicates the number of spikes. If one assumes that the neural response is generated by projection onto a single linear kernel, followed by a static nonlinearity and Poisson spike generation, the STA provides an unbiased estimate of this kernel (Chichilnisky, 2001; Paninski, 2002).

But if the neural response depends on more than a single axis within the stimulus space, the STA will provide an insufficient and possibly misleading description. A number of authors have suggested the natural extension of examining higher-order statistical properties (and in particular, the covariance) of the spike-triggered ensemble of stimuli (de Ruyter and Bialek, 1988; Arcas et al, 2001; Schwartz et al, 2001; Touryan et al, 2002). The idea is simple and intuitive: if the neural response is determined by a projection onto a low-dimensional subspace (within the space of all stimuli), an analysis of the spike-triggered covariance might allow us to recover this subspace.

In particular, consider first a stimulus direction to which the neuron is insensitive. Since the strength of the stimulus in that direction has no effect on the spiking behavior of the neuron, one expects the mean and variance (or any other statistic) of the spike-triggered stimulus ensemble along that direction to be essentially unchanged relative to that of the raw stimulus ensemble. Next consider a stimulus direction along which the cell's response grows monotonically. Here the mean of the spike-triggered stimulus ensemble will shift, relative to that of the raw ensemble. If the cell's response along this axis were symmetric (e.g., rectified, or squared), then one would expect the variance of the spike-triggered stimulus ensemble to increase. On the other hand, if the presence of of a stimulus component along the direction is suppressive, one expects the variance of the the spike-triggered stimulus ensemble to decrease.

In our analysis, we first compute the STA and project it out of the stimulus ensemble. Specifically, we compute the normalized (unit vector) STA, the nSTA, and define:

$$S_n'(x,t) = S_n(x,t) - (\sum_{x,t} S_n(x,t) \cdot nSTA(x,t)) \cdot nSTA(x,t)$$

This ensures that the axes obtained in the STC analysis will be orthogonal to the STA, and helps to avoid unwanted interactions between the STA and STC. We then compute the spike-triggered covariance:

$$COV(x_1, x_2, t_1, t_2) = \frac{1}{N_s - 1} \sum_n S'_n(x_1, t_1) S'n(x_2, t_2)$$

If one interprets this as a matrix, with parameter pairs $\{x_1,t_1\}$ and $\{x_2,t_2\}$ corresponding to the two indices, then this matrix captures the variance of a collection of samples in all possible directions within the space of stimuli. The surface swept out by a vector whose length is equal to the variance in its own direction is a hyper-ellipsoid. The set of orthogonal eigenvectors of the STC matrix, obtained by a principal components analysis (PCA), correspond to the fundamental axes of this ellipsoid, and the eigenvalues give the variances along these axes.

The relevant excitatory and inhibitory dimensions are identified with eigenvalues that deviate significantly from what is expected when the spikes are randomly correlated with the stimulus. These high- and low-variance axes, together with the STC, constitute a coordinate system for those aspects of the stimulus that are relevant to the neuron. The range of eigenvalues one expects when randomly selecting N samples of d-dimensional noise vectors depends on $\sqrt{(d/N)}$. The implication for these experiments is that mildly excitatory or suppressive axes will only be revealed when d/N is small enough. As a rough rule of thumb, we gather at least N = 100*d spikes.

Having determined this subspace coordinate system, we can examine the neural response function along axes, or pairs of axes within that subspace. Specifically, the firing rate for a given stimulus corresponds to quotient of the number of spikes observed for that stimulus and the number of times the stimulus was presented. As an estimate of this, we compute the quotient of the spike-triggered and raw stimulus histograms, as shown in Fig. 2. This is analogous to the usual method for computing the nonlinear response function associated with the STA (Chichilnisky, 2001).

Example cells

The STA calculated for an example simple and complex cell are shown in figure 1c. Space-time orientation in the STA of the simple cell suggests a preference for direction, confirmed by a grating direction tuning curve (not shown). In contrast, it is clear that the complex cell was poorly characterized by a single linear projection, as

evidenced by a relatively flat STA. Higher-order analyses are required to properly characterize the spatiotemporal properties of this cell type.

The results of the PCA applied to these two data-sets are shown in figure 1b. The eigenvector distribution shows that even after the STA had been removed, the relevant subspace for the simple cell included 3 additional excitatory dimensions and 6 suppressive dimensions; the resulting subspace for the complex cell contained 6 excitatory and 6 suppressive dimensions. Figure 1d displays the space-time and frequency structure of the first 6 excitatory and suppressive dimensions. For both cells, the excitatory and suppressive kernels were tuned for opposite directions of motion. The final panel of this figure (figure 1e) displays the sum of the significant spectra weighted by their corresponding eigenvalues.

It is important to note a few cautionary details regarding this analysis. Most importantly, the kernels (eigenvectors) should not be taken literally as physiologically instantiated mechanisms, since the analysis forces them to orthogonal. One should think of them as a representation of the set (subspace) of stimuli that the cell "cares about". More specifically, the labeling of any particular axis as "excitatory" or "suppressive" is based on whether the spike-triggered ensemble has a smaller or larger variance than the raw stimulus ensemble along that axis. If excitation and suppression coincide within a neuron along a common axis, the analysis might produce an axis of either label, or might even produce two orthogonal excitatory and suppressive axes. These concerns may be ameliorated by careful examination of the spike-triggered ensemble as projected onto the axes obtained in the analysis. We show examples of this in the next section.

Deviations from traditional models

Most models for simple cells postulate that the response is primarily based on a projection onto a single linear kernel. A spike-triggered covariance analysis applied to such a receptive field would reveal a single relevant dimension in the STA. In agreement with these models, the simple cell shown in figure 1 had distinct STA structure. However, not predicted by these models were the residual excitatory dimensions revealed by this analysis. To examine spiking properties along these axes, we plotted in figure 2a the joint two-dimensional projection of the stimuli preceding a spike (white dots) and the raw stimuli (black dots) onto the STA and the first eigenvector along with corresponding 1-dimensional histograms along the axes. Plots of the predicted spike rate at a given projection strength (figure 2b) show that only positive projections onto the STA caused increase in spiking whereas increasing the projection onto the first eigenvector or its inverse caused an increase in spike rate.

The linear kernel for a simple cell is usually assumed to be implemented with rectified inputs as at least one push-pull pair of inverted filters with one input having an excitatory influence and the other inhibitory. Due to the linearity produced by the perfect balance between push and pull, an arbitrary number of these pairs can combine to form a single, linear receptive field. For example, two space-time separable (orientation tuned) pairs displaced appropriately in space and time can combine to form a space-time inseparable (direction tuned) linear receptive field. The residual nonlinear behavior observed in the example simple cell could arise from the combination of two rectified subunits (push) without perfect complimentary (pull) balance. The spike triggered average of such a cell would be approximately the average of the two subunits and the residual nonlinear behavior produced by rectification would be resolved as an excitatory subunit by the PCA.

Figure 1. Spike triggered analysis applied to two directionally tuned V1 neurons. a) Eigenvalues calculated from a principle component analysis (PCA) of the covariance matrices for the same simple and complex cell. Gray lines denote the asymptotic cumulative eigenvector distribution expected given the stimulus dimensionality and number of spikes included in the analysis (Johnstone, 2000). Number of spikes collected for the simple cell: 42,053; for the complex cell: 79,048. Strong excitatory and suppressive eigenvalues are numbered. b) The spike triggered average (STA) for the same cells are shown in space-time coordinates and in the spatio-temporal frequency domain. The center of the latter plots corresponds to the zero (DC) frequency component. c) Excitatory and suppressive eigenvectors weighted according to their normalized eigenvalues and presented both in space-time coordinates and in frequency space. d) Excitatory and suppressive spectra weighted by their corresponding normalized eigenvalues before combination.

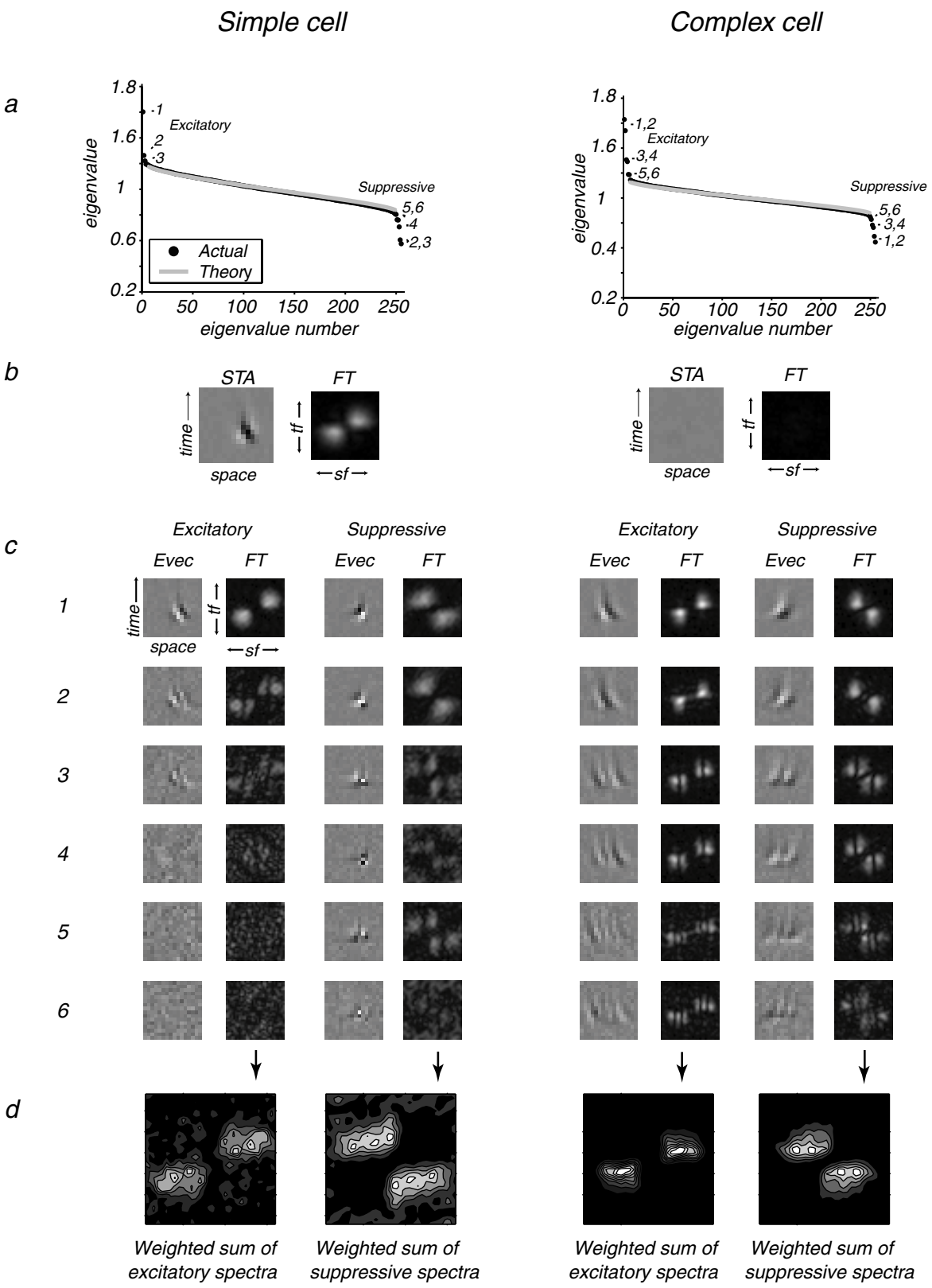
Simple cell
Complex cell
a
b
c
d
eigenvalue
eigenvalue number
Excitatory
Suppressive
Actual
Theory
STA
FT
time
space
tf
sf
Evec
1
2
3
4
5
6
Weighted sum of excitatory spectra
Weighted sum of suppressive spectra

Figure 1

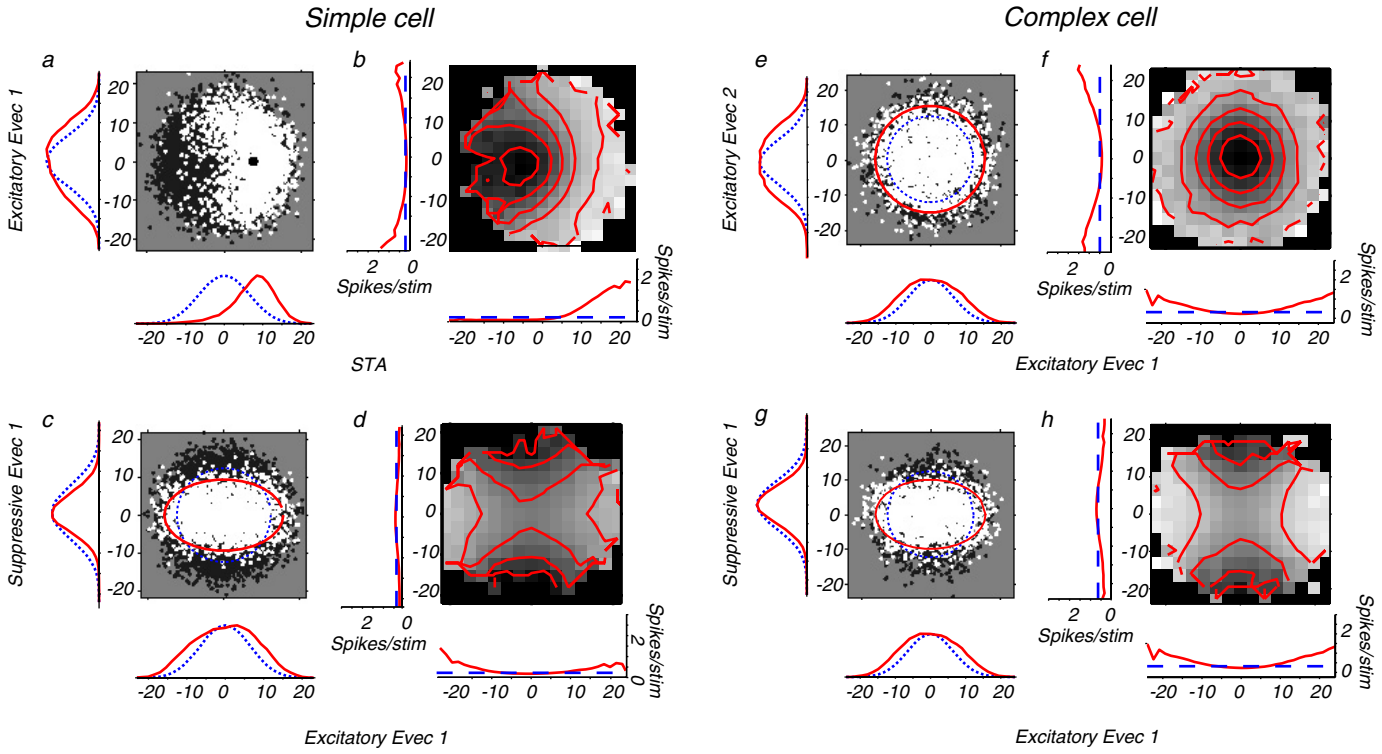

Figure 2 . Estimated response along selected axes of the recovered subspace for the example simple cell (a-d) and complex cell (e-h) in figure 1. a) Projections of all stimuli (black dots) and stimuli that elicited a spike (white dots) onto the STA and the strongest excitatory eigenvector (Evec). One-dimensional projections are displayed as histograms along the axes for both raw stimuli (blue dotted curve) and stimuli that elicited a spike (red solid curve). The large black dot indicates the mean of the spike triggered distributions. b) A two-dimensional histogram of the estimated spike rate at a given projection strength computed as the ratio between the spike-triggered versus raw stimulus distributions in a. In the one-dimensional projection, the dotted blue line indicates mean firing rate. Similar plots are also displayed for the following pairs of axes: c,d) simple cell, strongest excitatory and strongest suppressive; e,f) complex cell, first and second strongest excitatory; g,h) complex cell, strongest excitatory and strongest suppressive. In c, d, e, and g, the dotted blue curve indicates the variance of the stimulus; the red ellipse shows the variance of the spike triggered stimuli in each dimension as determined from the corresponding eigenvalues. The values of all axes are in units of percent relative to the maximum possible projection value.

The energy model for complex cells postulates that the response is based upon the squared projection onto a quadrature pair of filters. Each squared filter is assumed to be implemented physiologically by the sum of the responses of two rectified, sign-inverted units (push-push). Because the units are in quadrature, their phase cancels perfectly thus producing a flat STA. However, a spike-triggered covariance analysis applied to such a unit would resolve the relevant axes as two excitatory eigenvectors. These predictions are realized in the example complex cell (figure 1). But in addition a pair of relevant excitatory dimensions, this analysis reveals two additional pairs of excitatory axes not predicted by traditional models. Examination of the structure of the eigenvectors suggests that each subsequent pair has a derivative-like relationship to previous pairs. One potential mechanistic interpretation is that these vectors could represent multiple subunits of the cell arranged approximately in quadrature pairs and related by displacements in spatial position and spatiotemporal frequency. A second possible interpretation is that this behavior can be explained by deviations from the Linear-Nonlinear-Poisson (LNP) model assumed by the analysis (Arcas et al, 2001; Pillow and Simoncelli, 2003). We have collected as many as 200,000 spikes for some cells and have yet to find a clear suggestion of the upper limit on the number of significant axes revealed. Assuming that multiple subunits of a cell would be approximately equal in strength, the continued revelation of additional axes as more data is acquired is suggestive of a series expansion of some point nonlinearity not accounted for by the LNP model.

In both simple and complex cells, we found suppressive dimensions at least equal in number to the excitatory dimensions. Figure 2 d and h show the interactions between the strongest excitatory and suppressive axes revealed by STC for the simple and complex cell, respectively; in both cells increasing the projection along the suppressive dimension has a strong influence on the response to an otherwise excitatory stimulus. Although the structure of the suppressive eigenvectors appears to be tuned for the direction opposite the excitation in both example cells, for reasons described in the methods it is likely that the suppression observed here also exists in regions of excitation, consistent with models of gain control. Divisive influences that control gain and describe a variety of behaviors such as cross-orientation masking have been described in V1 (Carandini et al, 1997). The suppression could also be due in part to a subtractive rather than a divisive influence; subtractive influences are postulated to be involved in shaping receptive field properties such as direction selectivity (Adelson and Bergen, 1985). The degree to which the suppression observed here is subtractive or divisive will be determined through exploration of models within the relevant subspace for each cell.

The spatiotemporal tuning properties of the two example cells described thus far are representative of the majority of our data, including space/time oriented excitation with suppressive axes oriented in the opposite direction. We did observe some exceptions, including a directionally biased complex cell whose weak STA and excitatory eigenvector spectra suggest orientation rather direction tuning but whose suppressive spectra appear to be tuned for direction. We also observed directionally tuned excitation coupled with non-directional suppression.

Discussion

Using a dense noise stimulus and a spike-triggered covariance analysis, we have characterized a subspace of excitation and suppression for a population of directionally selective neurons in monkey V1. The strongest excitatory axes revealed by this analysis are predicted by standard V1 models, including a space-time oriented STA for simple cells and two space-time oriented eigenvectors in approximate quadrature for complex cells. In all cells tested, we also found evidence for the existence of weaker excitatory and suppressive axes that are not predicted by standard models of V1 neurons, including phase-invariant excitatory and suppressive axes in simple cells and behavior suggestive of an unaccounted for nonlinearity in both cell types.

Emerson et al (1992) applied a second-order analysis to two-bar interaction data obtained from cat V1 cells. They found evidence for the existence of two relevant excitatory dimensions in complex cells but no evidence of suppression in the direction opposite the excitation. Touryan et al (2002) performed experiments and analysis very similar to those reported here in cat V1 complex cells and typically found subspaces limited to two excitatory dimensions. One possible explanation for the additional excitatory and suppressive axes we report here is a difference in the V1 receptive field structure of the cat versus the monkey. Another possible difference is the number of spikes analyzed. We find that the weaker excitatory and suppressive axes are revealed only after a considerable number (20,000 – 50,000+) of spikes are collected.

One must be cautious when interpreting the results of spike triggered covariance, particularly with regard to the specific appearance of the relevant axes of the subspace determined for a cell. These recovered subspaces need to be considered in their entirety with special attention paid to the interactions between suppression and excitation within and between dimensions. Nevertheless, application of this technique allows for the first time determination of a multi-dimensional subspace of excitatory and suppressive influences for a cell from spiking responses. Future work is certain to reveal interesting and unexpected behaviors that occur within these unexplored subspaces.

References

Adelson EH, Bergen JR (1985) J Opt Soc Am A 2:284-299.
y Arcas BA, Fairhall AL, Bialek W (2001) NIPS Vol 13, pp 75-81.
Carandini M, Heeger DJ, Movshon JA (1997) J Neurosci 17:8621-8644.
Cavanaugh JR, Bair W, Movshon JA (2002) J Neurophysiol 88:2530-2546.
Chichilnisky EJ (2001) Network 12:199-213.
de Ruyter van Steveninick R, Bialek W (1988) Proc R Soc Lond B 234: 379-414.
Emerson RC, Bergen JR, Adelson EH (1992) Vision Res 32:203-218.
Johnstone IM (2000) Technical report. Department of Statistics. Stanford University. August.
Pillow JW and Simoncelli EP (2003) Neurocomputing. To appear.
Schwartz O, Chichilnisky EJ, Simoncelli E (2002) NIPS Vol 14, pp 269-276.
Touryan J, Lau B, Dan Y (2002) J Neurosci 22:10811-10818.