

Estimating nonlinear receptive fields from natural images

Joaquín Rapela

Department of Electrical Engineering and Neuroscience
Graduate Program, University of Southern California,
Los Angeles, CA, USA



Jerry M. Mendel

Department of Electrical Engineering, University of
Southern California, Los Angeles, CA, USA



Norberto M. Grzywacz

Department of Biomedical Engineering and Neuroscience
Graduate Program, University of Southern California,
Los Angeles, CA, USA



The response of visual cells is a *nonlinear* function of their stimuli. In addition, an increasing amount of evidence shows that visual cells are optimized to process *natural images*. Hence, finding good nonlinear models to characterize visual cells using natural stimuli is important. The Volterra model is an appealing nonlinear model for visual cells. However, their large number of parameters and the limited size of physiological recordings have hindered its application. Recently, a substantiated hypothesis stating that the responses of each visual cell could depend on an especially low-dimensional subspace of the image space has been proposed. We use this low-dimensional subspace in the *Volterra relevant-space technique* to allow the estimation of high-order Volterra models. Most laboratories characterize the response of visual cells as a nonlinear function on the low-dimensional subspace. They estimate this nonlinear function using histograms and by fitting parametric functions to them. Here, we compare the Volterra model with these histogram-based techniques. We use simulated data from cortical simple cells as well as simulated and physiological data from cortical complex cells. Volterra models yield equal or superior predictive power in all conditions studied. Several methods have been proposed to estimate the low-dimensional subspace. In this article, we test projection pursuit regression (PPR), a nonlinear regression algorithm. We compare PPR with two popular models used in vision: spike-triggered average (STA) and spike-triggered covariance (STC). We observe that PPR has advantages over these alternative algorithms. Hence, we conclude that PPR is a viable algorithm to recover the relevant subspace from natural images and that the Volterra model, estimated through the Volterra relevant-space technique, is a compelling alternative to histogram-based techniques.

Keywords: natural images, nonlinear models, Volterra model, dimensionality reduction, Projection Pursuit Regression

Introduction

Nonlinearities are ubiquitous in the response of visual cells. Just in the primary visual cortex, we find rectification (Albrecht & De Valois, 1981; Movshon, Thompson, & Tolhurst, 1978), saturation (Dean, 1981; Maffei & Fiorentini, 1973), expansion (Albrecht & Hamilton, 1982; Sclar & Freeman, 1982), phase advance (Carandini & Heeger, 1994; Dean & Tolhurst, 1986), and cross-orientation inhibition (Bonds, 1989; Morrone, Burr, & Maffei, 1982), to mention a few. Therefore, several investigators looked for general methods to study these nonlinearities from the retina to the cortex. In theory, white noise could be an ideal stimulus to estimate nonlinear properties. This is because if a nonlinear system is stimulated with a white noise stimulus ensemble for a long enough time, then there is a finite probability that any given stimulus will appear, probing the nonlinear system thoroughly and efficiently. However, this rationale is weak for the visual system because the possible dimensions of stimuli space are too large (i.e., the number of possible images is immense). Moreover, data are noisy and limited in quantity in physio-

logical recordings. Hence, white noise does not provide enough signal-to-noise ratio to obtain accurate estimates of responses along all stimulus dimensions. Finally, some sensory neurons do not respond well to white noise. Thus, it would be better to use natural images to concentrate the power of the stimuli on the normal operating range of the cell.

More than 40 years ago, Barlow (1961) introduced the hypothesis that visual cells are optimized to process natural stimuli. Since then, several investigators have provided support to this natural-adaptation hypothesis (Dan, Attick, & Reid, 1996; Simoncelli & Olshausen, 2001; Srinivasan, Laughlin, & Dubs, 1982) and have shown that natural images emphasize features of responses not prominent when using synthetic stimuli (David, Vinje, & Gallant, 2004). However, only recently, a number of studies have begun to use nonlinear models and natural stimuli to characterize the response of visual cells. These studies obtained interesting findings, but the methods had limitations for investigators interested in general receptive fields. Some groups did not attempt to map receptive fields (Aggelopoulos, Franco, & Rolls, 2005; Dan et al., 1996; Guo, Robertson, Mahmoodi, & Young, 2005; Vinje & Gallant, 2000, 2002; Weliky, Fiser,

Hunt, & Wagner, 2003), whereas others limited their studies to linear (Ringach, Hawken, & Shapley, 2002; Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003; Theunissen et al., 2001; Willmore & Smyth, 2003) or second-order, nonlinear (Felsen, Touryan, Han, & Dan, 2005; Touryan, Felsen, & Dan, 2005) receptive fields. More general nonlinearities were first studied by modeling cell responses as a linear filter followed by a point nonlinearity (Chichilnisky, 2001; Nykamp & Ringach, 2002) or by fitting a priori models (David et al., 2004). The former models cannot capture nonlinear interactions between subregions of the receptive field and, thus, might not be sufficiently general to characterize responses of large classes of cells. Later, two methods that make no assumptions on the response-generation mechanism were introduced (Prenger, Wu, David, & Gallant, 2004; Sharpee, Rust, & Bialek, 2004; Sharpee et al., 2006). These methods are powerful but require delicate nonlinear optimizations that could be overwhelming to many investigators.

In an attempt to overcome some of the above limitations, we explored the Volterra model (Marmarelis, 2004). Because it does not rely on specific assumptions about the response-generation mechanism, this model can be applied to a broad class of cells. Moreover, the Volterra model has a long history in the study of nonlinear physiological systems. This model could be especially useful for visual cells because its first-order kernels are just like standard receptive fields. Therefore, the model can generalize them to the nonlinear domain. The large number of parameters in Volterra models of visual cells and limited physiological recordings have hindered its application to the visual system. In this article, we overcome this limitation using a recent substantiated hypothesis stating that the responses of each visual cell depend on an especially low-dimensional subspace of the image space (Bialek & de Ruyter van Steveninck, 2005; Brenner, Bialek, & de Ruyter van Steveninck, 2000; de Ruyter van Steveninck & Bialek, 1988; Rust, Schwartz, Movshon, & Simoncelli, 2005; Sharpee et al., 2004; Sharpee et al., 2006). Accordingly, we introduce the *Volterra relevant-space technique* to estimate Volterra models of visual cells from natural images.

Most laboratories characterize the response of visual cells as a nonlinear function on a one- or two-dimensional subspace of the image space (Chichilnisky, 2001; Felsen et al., 2005; Sharpee et al., 2004; Sharpee et al., 2006; Simoncelli, Paninski, Pillow, & Schwartz, 2004; Touryan, Lau, & Dan, 2002; Touryan et al., 2005). To estimate this nonlinear function, they construct one- or two-dimensional histograms of the response as a function of the projection of the input images based on this subspace. Then, they fit parametric functions to this histogram. Using simulated data from cortical simple cells and simulated and physiological data from cortical complex cells, we compare the Volterra model with histogram-based techniques for simple (Chichilnisky, 2001) and complex cells (Touryan et al., 2005).

Several methods, in different scientific disciplines, have been proposed to estimate the low-dimensional subspace

(de Boer & Kuyper, 1968; de Ruyter van Steveninck & Bialek, 1988; Friedman & Stuetzle, 1981; Helland, 1988; Li, 1991; Sharpee et al., 2004; Touryan et al., 2005). In this article, we test projection pursuit regression (PPR; Friedman & Stuetzle, 1981), a nonlinear regression algorithm. We compare PPR with two popular models used in vision: spike-triggered average (STA; de Boer & Kuyper, 1968) and the modification of spike-triggered covariance (STC) for natural images (Touryan et al., 2005).

We organize the rest of this article as follows: The **Theory** section describes the theoretical underpinnings of the modeling of cells' receptive fields. This section covers the **Volterra model** and the **Volterra relevant-space technique**. This theoretical section is followed by the **Methods** section, the **Results** section, and the **Discussion** section. Each theoretical material starts with an intuitive description of ideas so that readers can get the gist of the paper even if they skip its mathematical content. Then, those sections present the main mathematical results, leaving proofs and developments to appendices.

Theory

Volterra model

A useful way to think of the Volterra model is that it is a sequence of approximations. They are given by a sequence of filters, called kernels, with which to process the input. When Volterra models are applied to visual cells, the zeroth-order kernel represents the activity when the input is absent. The first-order Volterra kernel is the linear component of the response. For example, in the spatial domain, this kernel gives the response to small pulses of input at every position. Therefore, this kernel is a good representation of what people traditionally call the receptive field of a cell. The second-order kernel represents the weight that nonlinear interactions between two inputs (e.g., at two times or in two positions) have on the response. In general, the r th-order kernel represents the weight that nonlinear interactions between r inputs have on the response.

The development of Volterra models relies on the mathematical notion of the Volterra series introduced by Volterra (1930). In the time domain, the output $y(t)$ of a stationary, stable, causal system can be expressed in terms of its input signal $x(t)$ through its Volterra series expansion as

$$\begin{aligned}
 y(t) = & k_0 + \int_0^\infty k_1(\tau)x(t-\tau)d\tau \\
 & + \int_0^\infty \int_0^\infty k_2(\tau_1, \tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1d\tau_2 \\
 & + \int_0^\infty \int_0^\infty \int_0^\infty k_3(\tau_1, \tau_2, \tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3) \\
 & \times d\tau_1d\tau_2d\tau_3 + \dots
 \end{aligned} \tag{1}$$

A Volterra model is a truncated Volterra series; that is,

$$\begin{aligned}
 y_{Q,k}(t) &= k_0 + \int_0^\infty k_1(\tau)x(t-\tau)d\tau \\
 &+ \int_0^\infty \int_0^\infty k_2(\tau_1, \tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1d\tau_2 + \dots \\
 &+ \int_0^\infty \dots \int_0^\infty k_Q(\tau_1, \dots, \tau_Q)x(t-\tau_1)\dots x(t-\tau_Q) \\
 &\times d\tau_1\dots d\tau_Q + \varepsilon,
 \end{aligned} \tag{2}$$

where the subscript Q in $y_{Q,k}$ indicates that the Volterra series has been truncated at order Q , the subscript k reflects the dependence of the Volterra model on the particular choice of kernels $\{k_0, \mathbf{k}_1, \dots, \mathbf{k}_Q\}$, and ε stands for possible errors due to truncation and to noise in the data.

The Volterra model given in Equation 2 uses one-dimensional continuous inputs. However, the input to visual cells is multidimensional (one dimension of time, two dimensions of space, three dimensions of color, etc.), and discrete data are normally used to characterize their responses. Here, as a first approximation, we only use the spatial dimensions of the stimuli to estimate the response of visual cells, disregarding temporal interactions. As discussed in the Discussion section, the Volterra relevant-space technique can be extended to consider spatiotemporal interactions. The Volterra series used henceforth is

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= k_0 + \sum_{i,j=1}^N k_1(i, j)x(i, j) \\
 &+ \sum_{i_1, j_1, i_2, j_2=1}^N k_2(i_1, j_1, i_2, j_2)x(i_1, j_1)x(i_2, j_2) + \dots \\
 &+ \sum_{i_1, j_1, \dots, i_Q, j_Q=1}^N k_Q(i_1, j_1, \dots, i_Q, j_Q) \\
 &\times x(i_1, j_1)\dots x(i_Q, j_Q) + \varepsilon,
 \end{aligned} \tag{3}$$

where $\mathbf{x} \in \mathbb{R}^{N \times N}$ is the input image. Any Volterra model can be represented with kernels \mathbf{k}_r that are symmetric with respect to permutations in pairs of indices (i_p, j_p) . Hence, the number of parameters in Volterra kernels is less than their number of kernel coefficients. For example, for \mathbf{k}_2 , we have the symmetry $k_2(i_1, j_1, i_2, j_2) = k_2(i_2, j_2, i_1, j_1)$, which implies that the number of parameters is $\frac{(N^2+1)N^2}{2}$, whereas the number of kernel coefficients is N^4 . The number of parameters of the Volterra model in Equation 3, n_V , is

$$n_V = \frac{(N^2 + Q)!}{N^2! Q!}. \tag{4}$$

For example, a fourth-order Volterra model, $Q = 4$, using images of size 16×16 as input, contains $n_V = 186,043,585$ parameters. From the limited data that can be recorded in physiological experiments, we cannot estimate such a

large number of parameters. The Volterra relevant-space technique described in the next section addresses this problem.

Volterra relevant-space technique

Fortunately, one can reduce the dimensionality of Volterra models. This is because the response of each visual cell depends on a low-dimensional subspace of the image space. One reason for this low dimensionality is the high degree of correlation in natural images (Field, 1987; Ruderman & Bialek, 1994; Srinivasan et al., 1982). Another reason is that each neuron responds to particular features of the image, such as an edge at a given position with a given orientation. Here, we will show how to use this low-dimensional subspace to build a low-dimensional Volterra model and we will prove the equivalence between the original and low-dimensional Volterra models. Thus, we establish, for the first time, a mathematical link between the low-dimensional subspace and Volterra models.

The construction of the low-dimensional Volterra model begins by projecting the input images onto a low-dimensional subspace \mathcal{S} . If \mathbf{x} is an image and $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_L; \mathbf{b}_l \in \mathbb{R}^{N \times N}\}$ is a set of orthonormal basis functions spanning the space \mathcal{S} , then the projected image is

$$\Pi_{\mathcal{S}}(\mathbf{x}) = \sum_{l=1}^L \alpha_l(\mathbf{x})\mathbf{b}_l, \tag{5}$$

where

$$\alpha_l(\mathbf{x}) = \sum_{i,j=1}^N b_l(i, j)x(i, j) \tag{6}$$

is the projection coefficient of image \mathbf{x} onto the basis function \mathbf{b}_l . Using these coefficients, the low-dimensional Volterra model is

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= q_0 + \sum_{l=1}^L q_1(l)\alpha_l(\mathbf{x}) + \dots \\
 &+ \sum_{l_1, \dots, l_Q=1}^L q_Q(l_1, \dots, l_Q)\alpha_{l_1}(\mathbf{x})\dots \alpha_{l_Q}(\mathbf{x}) + \varepsilon.
 \end{aligned} \tag{7}$$

What are the conditions on the space \mathcal{S} for the low-dimensional Volterra model to be an equivalent representation of the original Volterra model in Equation 3? The answer is given by Proposition 1 (Appendix A). It states that *if a vector space \mathcal{S} is such that the response of the cell to any image equals the response to the projection of the image onto \mathcal{S} , then the original Volterra model can be rewritten as the low-dimensional Volterra model. A vector space \mathcal{S} satisfying this condition is called a relevant space. That is, \mathcal{S} is a relevant space if and only if*

$$y_{Q,k}(\mathbf{x}) = y_{Q,k}(\Pi_{\mathcal{S}}(\mathbf{x})) \forall \mathbf{x} \in X. \tag{8}$$

Any set of orthonormal basis function $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_L: \mathbf{b}_l \in \mathbb{R}^{N \times N}\}$ spanning a relevant space is called a set of *relevant dimensions*.

We are not seeking a relevant space such that the represented images look similar to the original ones. Instead, we are seeking a relevant space such that the response to the original images equals the response to the represented images. This equality is as in Equation 8. Therefore, for our purposes, a relevant-space representation does not depend only on the statistics of natural images, as in the case of image reconstruction. This representation depends also on the properties of the cell under study.

How much do we gain by using the relevant space to represent Volterra models? The number of parameters in the low-dimensional Volterra model, n_{LV} , is

$$n_{LV} = \frac{(L + Q)!}{L! Q!}. \tag{9}$$

As an example, for $L = 10$ relevant dimensions, a fourth-order ($Q = 4$) low-dimensional Volterra model will contain $n_{LV} = 1,001$ parameters. In contrast, the number of parameters in the original Volterra model (n_V) for a 16×16 image is 186,043,585 (Equation 4). Therefore, if a small relevant subspace can be identified, the Volterra relevant-space technique achieves enormous dimensionality reduction.

How are the coefficients, q_i , of the low-dimensional Volterra model (Equation 7), estimated? We seek coefficients that minimize the difference between true responses and those predicted by the Volterra model. To obtain a convenient method for this minimization, we begin by expressing Equation 7 as the linear Volterra equation

$$\mathbf{y} = \mathbf{A}\mathbf{q} + \varepsilon, \tag{10}$$

where, for K image–response pairs and M parameters, $\mathbf{y} \in \mathbb{R}^K$ is the response vector whose i th element, $y(\mathbf{x}_i)$, is the response of the cell to the i th image, $\mathbf{A} \in \mathbb{R}^{K \times M}$ is the data matrix, $\mathbf{q} \in \mathbb{R}^M$ is the vector of parameters, and $\varepsilon \in \mathbb{R}$ accounts for possible errors due to truncation and for noise in the data. For instance, for a second-order cell ($Q = 2$), the i th row of matrix \mathbf{A} , \mathbf{a}'_i , is

$$\mathbf{a}'_i = [1, \alpha_1(\mathbf{x}_i), \alpha_2(\mathbf{x}_i), \dots, \alpha_L(\mathbf{x}_i), \alpha_1(\mathbf{x}_i)^2, \alpha_1(\mathbf{x}_i)\alpha_2(\mathbf{x}_i), \dots, \alpha_1(\mathbf{x}_i)\alpha_L(\mathbf{x}_i), \alpha_2(\mathbf{x}_i)^2, \alpha_2(\mathbf{x}_i)\alpha_3(\mathbf{x}_i), \dots, \alpha_{L-1}(\mathbf{x}_i)\alpha_L(\mathbf{x}_i), \alpha_L(\mathbf{x}_i)^2], \tag{11}$$

and the parameters vector, \mathbf{q} , is

$$\mathbf{q} = [q_0, q_1(1), \dots, q_1(L), q_2(1, 1), q_2(1, 2), \dots, q_2(1, L), q_2(2, 2), q_2(2, 3), \dots, q_2(L - 1, L), q_2(L, L)]'. \tag{12}$$

By using the mean square error (*MSE*),

$$MSE(k_0, \mathbf{k}_1, \dots, \mathbf{k}_Q) = \frac{1}{n} \sum_{i=1}^n (y(\mathbf{x}_i) - y_{Q,k}(\mathbf{x}_i))^2, \tag{13}$$

as the goodness-of-fit criterion, we can apply any linear technique to estimate the parameters vector, \mathbf{q} . The use of linear techniques in the estimation of Volterra models removes the whiteness requirement of cross-correlation techniques. Therefore, we can use natural images to estimate Volterra models, without the need to whiten them or to modify image statistics in any way.

Having estimated the coefficients of the low-dimensional Volterra model, one approximates the true kernels, \mathbf{k}_i , in Equation 3, with

$$k_0 = q_0 \tag{14}$$

$$k_1(i, j) = \sum_{l=1}^L q_1(l) b_l(i, j) \tag{15}$$

...

$$k_Q(i_1, j_1, \dots, i_Q, j_Q) = \sum_{l_1, \dots, l_Q=1}^L q_Q(l_1, \dots, l_Q) \times b_{l_1}(i_1, j_1) \dots b_{l_Q}(i_Q, j_Q), \tag{16}$$

Proposition 2 (Appendix A) provides justification for these equations. This proposition proves that, if a cell can be represented by a Volterra model using true kernels, \mathbf{k}_i , then, for any set of true relevant dimensions $\{b_i\}$ (Equation 8), there exist coefficients q_0, \dots, q_Q , such that a Volterra model using the true kernels, \mathbf{k}_i , generates the same responses as a Volterra model using the approximated kernels in Equations 14–16. Thus, from a predictive perspective, these approximated kernels are indistinguishable from the true kernels.

The use of Equations 7–16 assumes a known order of the Volterra model, Q . Here, we selected the order Q by cross-validation, as indicated in the Volterra relevant-space technique section under the Receptive-field estimation section. In summary, the Volterra relevant-space technique uses the following procedure to estimate Volterra models from natural images:

1. Estimate the relevant space.
 2. Project the images in the estimated relevant dimensions (Equation 6), obtaining the coefficients α_i .
 3. Use the coefficients α_i to construct the data matrix (Equation 11) of the linear Volterra equation (Equation 10).
 4. Estimate the coefficients vector, q , of the linear Volterra equation by minimizing the *MSE* in Equation 13.
 5. Reconstruct the Volterra kernels using Equations 14–16.
- Code implementing this procedure along with the simulated data used in the Results section can be downloaded from <http://vpl.usc.edu/projects/nonlinearReceptiveFields/>.

After the relevant dimensions have been estimated, the remaining steps in the procedure (Steps 2–5) involve only linear operations and are therefore fast and easy to implement. In contrast, most implementations of histogram-based methods fit a nonlinear function to an estimated histogram. The selection of the optimal histogram bin size is an open question, and the nonlinear fit is slower and comparatively more difficult to implement than the linear operations required by the Volterra relevant-space technique.

Methods

Simulated data

Responses of the simulated simple-cell and complex-cell models described below were generated using 10×10 image patches extracted from calibrated natural images (van Hateren & van der Schaaf, 1998). Sample natural images are shown in Figure 1.

Simulated simple cell

The simple-cell data were generated by a model similar to those previously used to describe V1 simple cells (Jones, Stepnoski, & Palmer, 1987). The model consisted of a linear spatial filter followed by a sigmoidal rectification, $\text{rect}(v) = \text{saturation} / [1 + \exp(-\text{slope} \times (v - \text{threshold}))]$, and a

Poisson spike generator. We used a sigmoidal rectification instead of a halfway rectification because we wanted to use a model that was differentiable at every point. This differentiability allowed derivation of true Volterra kernels (see Appendix B). The simple-cell model is then given by

$$y(\mathbf{x}) = \text{rect}(\langle \mathbf{f}, \mathbf{x} \rangle), \quad (17)$$

where $\langle \dots \rangle$ represents the Euclidean inner-product operation and \mathbf{f} is the spatial filter. For the spatial filter, we used a two-dimensional Gabor function. It had a preferred orientation of 45 deg from vertical, a preferred spatial frequency of 2 cycles per receptive field, a spatial bandwidth of 1.6 octaves, and an even phase. For the sigmoidal rectification of the filtered image, we used slope = 5.0 and threshold = 1.0 and varied the saturation parameter (setting it to 458, 276, 92, and 46) to obtain different mean number of spikes per image (5, 3, 1, and 0.5, respectively). Because the responses of the simulated simple cell followed a Poisson distribution, the signal-to-noise ratio in the responses increased monotonically as a function of the mean number of spikes per image.

To assess the quality of the estimated relevant dimensions, we first derived analytically relevant dimensions for the simple cell. We then computed their projection onto the estimated relevant space. If the estimated relevant space embodies the true relevant space, then the analytically relevant dimensions should equal their projection onto the estimated space. In deriving the analytically relevant dimensions, we begin by observing that the first stage of the simple-cell model is a linear



Figure 1. Sample natural images.

filter. Hence, the response to any image will equal (up to the noise level) the response to the image projected in the span of the linear filter. Consequently, the linear filter spans the relevant space. After rescaling this filter to unit norm, it is the only relevant dimension for the simulated simple cell.

The derivatives of all orders of the simple-cell model in Equation 17 are nonzero. Hence, according to Equation B3, every kernel in Equation 3 is nonzero. This may cause a problem when the projections of the input images onto the filter \mathbf{f} (the coefficients $\alpha_i(\mathbf{x})$ in Equation 7) are large. In this case, truncation errors will become significant. However, in the opposite case, that is, when these projections are small, the problem is not serious. By using a model with kernels of sufficiently high order, these truncation errors will be negligible. To minimize the effect of truncation errors, we selected the amplitude of the filter \mathbf{f} in such a way that the absolute value of the largest filtered image was 1.

Simulated complex cell

The complex-cell data were generated by means of a spatial energy model (Adelson & Bergen, 1985). The input image was first filtered using two Gabor filters, in quadrature phase, that is, with a 90 deg phase relationship. These filters had a vertical preferred orientation, a preferred spatial frequency of 2 cycles per receptive field, and a spatial-frequency bandwidth of 1.6 octaves. The outputs of these filters were then squared and summed, generating the mean of a Poisson spike generator. The analytical expression of the complex-cell model is then

$$y(\mathbf{x}) = \langle \mathbf{f}_1, \mathbf{x} \rangle^2 + \langle \mathbf{f}_2, \mathbf{x} \rangle^2, \quad (18)$$

where \mathbf{f}_1 and \mathbf{f}_2 are the linear filters in quadrature phase. To vary the mean number of spikes in the responses, we scaled the magnitude of the Gabor filters, generating responses with 5.0, 3.0, 1.0, and 0.5 spikes/image.

As before, to assess the quality of the estimated relevant space, we derived analytically relevant dimensions for the complex-cell model. We then compared them with their projection onto the estimated relevant space. Due to the linear filters in the first stage of the complex-cell model, the response to any image will equal, up to the noise, the response to the image projected in the span of the filters. Consequently, the two linear filters span the relevant space. Because these filters are orthogonal, after rescaling them to unit norm, they constitute a set of relevant dimensions.

Again, to judge the quality of the estimated kernels, we compared them with the true kernels. We derived them for the complex-cell model using the procedure described in Appendix B. Only the second spatial derivatives with respect to \mathbf{x} of Equation 18, evaluated at zero, are different from zero. Hence, from Equation B3, only the second-order Volterra kernels of the complex cell are different from zero. Therefore, the complex cell can be represented with a second-order Volterra model.

Physiological data

The physiological data used here were recorded in the laboratory of Dr. Yang Dan, and the method used to obtain them were described by Touryan et al. (2005). Animal procedures to obtain these data were approved by the Animal Care and Use Committee at the University of California, Berkeley. Briefly, single-cell recordings were made in Area 17 of adult cats (2–6.5 kg) using tungsten electrodes. Cells were classified as simple if the receptive fields had clear ON and OFF subregions and if the ratio of the first harmonic to the DC component of the responses to an optimally oriented drifting sinusoidal grating was greater than 1. Natural images used in cell stimulation were selected at random from a database consisting of a variety of digitized movies (van Hateren & Ruderman, 1998), and the center patch (12×12 pixels) of each image was retained. Images were scaled so that they all had the same variance. These selected images had no temporal correlations and were presented at a frame rate of 24 Hz in an area slightly larger than the classical receptive field of the cell estimated by hand mapping. For analysis, Touryan et al. binned responses at 41.8 ms, the presentation frame rate, and the bin immediately following a frame was used as the response to that frame. They shared with us two to four repetitions of responses from four complex cells to 24,000 natural images. Below, we show the analysis of the cell for which we obtained the best predictions with both the Volterra and Touryan et al. models. For this cell, both models yielded better predictions when considering the spikes in the frame presentation bin as the cell response to that frame. We used the average response across four repetitions of the stimuli as the response variable to estimate the different models.

Data partitioning

To control for possible overfitting of the different receptive-field estimation models to the data used in their estimation, we partitioned the data into disjoint training and test data sets. Only the training data set was used to estimate model parameters. The test data set was reserved to evaluate the models' predictive power. By using different training and test data sets, the prediction results reflected the generalization ability of the different models to predict responses to novel natural images.

In addition, between 2 and 10 jackknifed data sets were generated by excluding different 10% segments of the training data set (Efron & Tibshirani, 1993). These jackknifed data sets were used to average out the noise from the estimated relevant dimensions and model parameters and to select model hyperparameters.

We used training sets of varying size (1,000, 3,000, 5,000, 10,000, and 15,000 image–response pairs). For the simulated and physiological data, we used test sets of size 4,500 and 4,000 image–response pairs, respectively.

Receptive-field estimation

Volterra relevant-space technique

Estimating the relevant dimensions. To develop low-dimensional Volterra models, one must estimate a set of relevant dimensions satisfying Equation 8. Methods for doing this rely on different assumptions about the statistics of the input, the properties of the response-generation function, or both. The methods of reverse correlation (de Boer & Kuyper, 1968), STC (de Ruyter van Steveninck & Bialek, 1988), and maximally informative dimensions (Sharpee et al., 2004) have been introduced in the field of neuroscience. These methods are powerful but have constraints about the dimensionality of the relevant space and the type of stimulus used or require an arduous high-dimensional nonlinear optimization that constrains the dimensionality of estimable relevant spaces. (An exception may be the modification of STC for natural images explained in the [STC for natural images](#) section.) Other interesting methods involve perception neural networks (Hertz, Palmer, & Krogh, 1991) and partial least squares (Geladi & Kowalski, 1986; Helland, 1988). The former suffers from the curse of dimensionality, and the latter relies on a linear assumption about the response-generation function. Other methods are sliced inverse regression and PPR, both of which were introduced in statistics. Sliced inverse regression (Li, 1991) makes no assumptions about the response-generation mechanism but relies on the assumption of a spherical distribution of the stimuli. In turn, PPR (Friedman & Stuetzle, 1981) makes no assumption about the distribution of the inputs and uses the following generalized linear model for the response-generation mechanism:

$$y(\mathbf{x}) = \bar{y} + \sum_{m=1}^{M_0} \beta_m \phi_m(\langle \mathbf{a}_m, \mathbf{x} \rangle) \text{ with}$$

$$E\{\phi_m(\langle \mathbf{a}_m, \mathbf{x} \rangle)\} = 0 \text{ and } E\{\phi_m^2(\langle \mathbf{a}_m, \mathbf{x} \rangle)\} = 1. \quad (19)$$

Provided M_0 is sufficiently large, any smooth function can be represented by Equation 19 (universal approximator property; Diaconis & Shahshahani, 1984). Another interesting property of PPR is that its estimation algorithm explicitly addresses the curse of dimensionality by separately estimating each term in the generalized linear model. In the following experiments, we evaluated PPR for the estimation of the relevant space. In [Appendix C](#), we outline the PPR algorithm (see Friedman, 1984, for further details). The estimated relevant dimensions for the low-dimensional Volterra model are computed by orthonormalizing the set of projection directions, $\{\mathbf{a}_1, \dots, \mathbf{a}_{M_0}\}$, in Equation 19. To understand why this procedure works, let \mathcal{S} denote the space spanned by the projection directions, $\mathcal{S} = \langle \mathbf{a}_1, \dots,$

$\mathbf{a}_{M_0} \rangle$. Let us also denote $\Pi_{\mathcal{S}}$ as the projection operator onto the space \mathcal{S} . Because $\langle \mathbf{a}_i, \mathbf{x} \rangle = \langle \mathbf{a}_i, \Pi_{\mathcal{S}}(\mathbf{x}) \rangle$,

$$\begin{aligned} y(\mathbf{x}) &= \bar{y} + \sum_{m=1}^{M_0} \beta_m \phi_m(\langle \mathbf{a}_m, \mathbf{x} \rangle) \\ &= \bar{y} + \sum_{m=1}^{M_0} \beta_m \phi_m(\langle \mathbf{a}_m, \Pi_{\mathcal{S}}(\mathbf{x}) \rangle) \\ &= y(\Pi_{\mathcal{S}}(\mathbf{x})). \end{aligned} \quad (20)$$

That $y(\mathbf{x}) = y(\Pi_{\mathcal{S}}(\mathbf{x}))$ implies that \mathcal{S} is a relevant space for y . Consequently, a set of relevant dimensions can be computed by orthonormalizing the set of projection directions $\{\mathbf{a}_1, \dots, \mathbf{a}_{M_0}\}$, as indicated above.

For the results reported here, we used the implementation of the PPR algorithm in R, an environment for statistical computing (Venables & Smith, 2002). This implementation selects initial values for the projection directions, \mathbf{a}_i , in Equation 19 by cross-correlating the input images with the residuals of the responses. Consequently, it might fail to estimate the relevant dimensions of cells for which the responses are completely uncorrelated with the input images. For these cases, one can use the implementation of Automatic Smoothing Spline Projection Pursuit (Roosen & Hastie, 1994) developed by Roosen (1994), which allows the specification of initial projection directions. However, for all the conditions reported below, there was enough correlation between images and responses for the R implementation of PPR to estimate good relevant dimensions.

The estimation of the projection directions requires two runs of the PPR algorithm. The first run estimates the number, L , of projection directions needed to characterize the cell. Then, the second run estimates the L projection directions. In the first run, for each jackknifed data set in our training set ([Data partitioning](#) section), we ran the PPR algorithm with parameters $M_0 = 1$ and $M = 6$. These parameters allowed assessing the predictive error of PPR models with one to six projection directions (see [Appendix C](#)). For each jackknifed data set, PPR generates a predictive-error-versus-number-of-projection-directions curve. The PPR predictive-error curves of different jackknifed data sets were averaged. This average generated a new curve that was used to select the number of projection directions. PPR predictive-error-versus-number-of-projection-directions curves follow a common pattern. Before a critical number of terms, the predictive error decreases substantially as we increase the number of terms. Increasing the number of terms beyond the critical value yields only small reductions in predictive error. We selected the number of projection directions as the largest number, between one and six, for which the error bars (Size 10 standard errors) did not intersect those of the previous number of projection directions.

After choosing the number of PPR projection directions, L , we proceeded to estimate the projection directions. We

ran the PPR algorithm, with parameters $M_0 = L$ and $M = 6$, for the N ($2 \leq N \leq 10$) jackknifed data sets, estimating N sets of L noisy projection directions. To attenuate the effect of noise, we averaged these sets of projection directions, as described in the [Averages of sets of relevant dimensions](#) section.

Note that PPR uses raw natural images as inputs, in contrast to spike-triggered techniques, which use spike-triggered images, that is, images weighted in proportion to their corresponding response.

Estimating the low-dimensional Volterra model. Having estimated the relevant dimensions, we constructed the low-dimensional Volterra model ([Equation 7](#)). We estimated its parameters, q_i , by minimizing the *MSE* between the responses of the cell model and the predictions of the low-dimensional Volterra model. As noted in the [Volterra relevant-space technique](#) section, the solution to [Equation 13](#) is a linear problem. We solved it using the pseudoinverse computed from the singular value decomposition (SVD; Mendel, 1995).

Selecting the order of Volterra model. For each training jackknife data set ([Data partitioning](#) section), we estimated low-dimensional Volterra models of different orders. We then evaluated the predictive power of these models using the 10% portion of the training data not included in the jackknife data set. The predictive power was the cross-correlation between cell responses and Volterra model predictions. We then generated a curve of predictive power versus model order. Curves of different data sets were averaged. The least order maximizing the average predictive power was selected as the order of the Volterra model.

Spike-triggered average

Suppose the responses of a cell are generated by a static nonlinearity on the projection of input images along a single relevant dimension. In other words, suppose $y(\mathbf{x}) = N(\langle \mathbf{h}, \mathbf{x} \rangle) + \varepsilon$, where $y(\mathbf{x})$ is the response of the cell to image \mathbf{x} , \mathbf{h} is a linear filter, $\langle \cdot, \cdot \rangle$ is the inner-product operation, and N is a static nonlinearity. When the images are Gaussian white noise, the filter \mathbf{h} can be estimated from pairs of images–responses by cross-correlating the inputs with the responses (Marmarelis & Marmarelis, 1978).

However, natural images are not white noise. For them, when the cell can be represented as a linear device, the cross-correlation between the stimuli and the responses equals the product of the autocorrelation of the inputs and the filter, i.e., $\mathbf{C}_{yx} = \mathbf{C}_{xx}\mathbf{h}$. Then, for nonwhite inputs and a linear cell, the filter \mathbf{h} can be recovered as $\mathbf{h} = \mathbf{C}_{xx}^{-1}\mathbf{C}_{yx}$. The autocorrelation matrix, \mathbf{C}_{xx} , for natural stimuli is nearly singular. Therefore, its true inverse tends to amplify noise. To avoid this problem, we regularized the autocorrelation matrix using the truncated SVD (Hansen, 1987) and computed the pseudoinverse (Ben-Israel & Greville, 1980) from this regularized matrix. The computation of the truncated SVD uses a threshold to decide how many singular values to include in the regularized matrix. For each condition in the studies reported below, we selected the optimal threshold by using k -fold cross-validation

(Efron & Tibshirani, 1993). We used the jackknifed data sets from the training data ([Data partitioning](#) section) to estimate different STA filters. To minimize the effect of noise, we used the average of these filters as the final STA estimate.

Histogram-based model for simple cells

The histogram-based model for simple cells characterizes the response to an image \mathbf{x} as a linear filter, \mathbf{h} , followed by a static one-dimensional nonlinearity (Chichilnisky, 2001). Precisely, $y(\mathbf{x}) = N(\langle \mathbf{h}, \mathbf{x} \rangle) + \varepsilon$. We estimated the linear filter using STA and the static nonlinearity by fitting the parameters of the same sigmoidal function used to generate the simulated data to a histogram of responses as a function of projections of input images in the estimated filter. We used histograms with equi-spaced bins and selected the bin size from the data range using the Sturges rule (Sturges, 1926).

STC for natural images

For Gaussian white noise stimuli, consider stimuli that elicit spikes (spike-triggering stimuli). The variance of the projections of these stimuli along the cell’s relevant dimensions should be higher or lower than the variance along nonrelevant dimensions. The dimensions with high or low variance correspond to the eigenvectors of the autocovariance matrix with high or low eigenvalues, respectively. For Gaussian stimuli, STC (de Ruyter van Steveninck & Bialek, 1988; Simoncelli et al., 2004) identifies these “extreme” eigenvectors as the relevant dimensions of a cell.

Touryan et al. (2005) have proposed a modification of STC for natural stimuli. This modification starts by whitening the natural stimuli. Denote by U the matrix of eigenvectors of the autocovariance of the stimuli (one eigenvector per column), and denote by λ_i its eigenvalues. The matrix of normalized eigenvectors is defined as

$$U_n = U \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\lambda_n}} \end{pmatrix}. \quad (21)$$

Then, the whitened natural images are $X_w = XU_n$. STC for natural images now performs a classical STC analysis on the whitened natural images, estimating a set of relevant dimensions, V_w (one relevant dimension per column). Finally, the estimated relevant dimensions of the cell, V (one relevant dimension per column), are $V = U_n V_w$.

The autocovariance matrix of natural images is nearly singular; hence, its last eigenvalues (λ_i , $i \gg 1$) will be very small and tend to amplify noise. To avoid this effect, we selected a threshold and set eigenvalues below this threshold to 1. In this way, only the eigenvectors corresponding to large eigenvalues are normalized in [Equation 21](#). With physiological and simulated data, we used different

threshold values for different cells and selected the value maximizing the predictive power. Best predictions were obtained using thresholds that normalized around 35% of the eigenvectors. Thus, for the results reported below, we used this criterion to select the threshold. The same criterion was used by Touryan et al. (2005). Multiple sets of STC relevant dimensions were estimated using the jackknifed data sets from the training data (Data partitioning section). We then averaged these sets (Averages of sets of relevant dimensions section) to obtain the final STC estimates.

Histogram-based model for complex cells

To reconstruct the nonlinear function from the relevant space by a histogram method, we use the technique proposed by Touryan et al. (2005). We construct two independent histograms for responses as a function of projections of images in each of the two estimated relevant dimensions. We then fit a power function, $f(p) = \beta|p|^\gamma + r_0$ to each histogram, where $f(p)$ is the response of the cell to a given image, p is the projection of the image onto a estimated relevant dimension, and β , γ , and r_0 are free parameters. The reconstructed nonlinearity is then the sum of the power functions fitted to each histogram, $y(\mathbf{x}) = f_1(\langle\phi_1, \mathbf{x}\rangle) + f_2(\langle\phi_2, \mathbf{x}\rangle)$, where $y(\mathbf{x})$ is the response of the cell to image \mathbf{x} , f_1 and f_2 are the power functions fitted to the histograms, ϕ_1 and ϕ_2 are the first and second relevant dimensions, and $\langle\cdot, \cdot\rangle$ is the inner-product operation.

Averages of sets of relevant dimensions

We wish to average N sets of L estimated relevant dimensions, when each set has been estimated from a different portion of the training data. All sets describe, with noise, a common relevant space. The goal is to estimate this common space. Because different sets might represent the common space in different coordinate systems, we cannot average them element by element. To perform the average, we instead collected the N sets into a matrix A . In A , columns $\{(i-1)L, \dots, iL\}$ correspond to the relevant dimensions of the i th set. The column space of A spans an L -dimensional relevant space (signal) and a noise space. We use the SVD to separate the column space of A into a signal and an orthogonal noise space (Scharf, 1991). Specifically, we compute the SVD of A , $A = U\Sigma U^T$, and use the first L left singular vectors (first L columns of matrix U) to represent the signal space. Thus, the first L columns of matrix U are the average of the N sets of L relevant dimensions.

Scaling kernels

To display kernels and to measure the goodness of fit of estimated kernels, we scaled them in proportion to their mean contribution to the response of the cell. The first-order

kernel was scaled by the mean absolute value of the stimuli. In turn, the second-order kernel was scaled by the mean absolute value of the autocorrelation. Precisely, by denoting the mean of z by \bar{z} , the scaled first- and second-order kernels are $k_1^s(i, j) = k_1(i, j)|\bar{x}(i, j)|$ and $k_2^s(i_1, j_1, i_2, j_2) = k_2(i_1, j_1, i_2, j_2)|\bar{x}(i_1, j_1)\bar{x}(i_2, j_2)|$, respectively.

Displaying kernels

In this article, we represent first- and second-order spatial Volterra kernel slices in three-dimensional (3D) spaces. We display 3D spaces as perspective and contour plots. First-order kernels, $k_1(i, j)$, have two independent variables: the dimensions of space and one dependent variable, the amplitude. The latter is represented on the vertical axis of the 3D plots, with i and j along the X - and Y -axes, respectively. Second-order kernels, $k_2(i_1, j_1, i_2, j_2)$, have four independent variables and one dependent variable. We display second-order kernel slices with respect to different reference positions. That is, we fix the reference position (i_1, j_1) and plot the value $k_2(i_1, j_1, i_2, j_2)$ on the vertical axis, with i_2 and j_2 on the X - and Y -axes, respectively. These plots show contributions to the response of interactions between individual pixels of the image and pixel $x(i_1, j_1)$. The kernels shown in the Results section have been scaled (Scaling kernels section). This allows the reader to appreciate visually the average contribution of a kernel to the cell response and to compare the contributions from different kernels. For predicting responses, on average, larger kernels are more relevant than smaller ones.

Goodness of fit of kernels

To measure the goodness of fit of estimated kernels to true kernels, we scaled them (Scaling kernels section), and computed their MSE . The MSE is a scale-sensitive measure. This implies that if true and estimated kernels have small amplitude, then the MSE between them will be small, independent of how well they correlate. For measuring errors between scaled kernels, this scale sensitivity is a good property. This is because scaled kernels with small amplitudes contribute little to the response, and thus it does not matter how well the estimated kernels correlate with the true kernel. Moreover, the MSE allows the comparison of errors across different kernels. This happens because the amplitude of scaled kernels indicates their mean contribution to the response. Therefore, a similar MSE for different scaled kernels indicates that they contribute similar errors to the response.

Volterra terms' contributions

To assess the importance of individual Volterra terms to the response, we decomposed Equation 3 as a sum of terms of increasing nonlinear order,

$$y(\mathbf{x}) = y_0 + y_1(\mathbf{x}) + \dots + y_Q(\mathbf{x}), \quad (22)$$

where

$$y_0(\mathbf{x}) = k_0, \quad (23)$$

$$y_1(\mathbf{x}) = \sum_{i,j=1}^N k_1(i,j)x(i,j), \quad (24)$$

...

$$y_Q(\mathbf{x}) = \sum_{i_1,j_1,\dots,i_Q,j_Q}^N k_Q(i_1,j_1,\dots,i_Q,j_Q)x(i_1,j_1)\dots x(i_Q,j_Q). \quad (25)$$

We wanted to know how relevant was each term relative to the others for the generation of the response. For this purpose, we defined the contribution of the i th term to the response $y(\mathbf{x})$ to image \mathbf{x} as $\text{contrib}_i(\mathbf{x}) = |y_i(\mathbf{x})| / \sum_{j=0}^Q |y_j(\mathbf{x})|$.

The definition of contribution is such that if for a given image the contribution of a specific term is large, then, if we set to zero the component of the response due to that term, the response is largely modified. On the other hand, if the contribution of a term is small, then the response is mostly unaffected when we set to zero the component of the response due to that term.

Results

We initially evaluated the Volterra relevant-space technique with models of cortical simple and complex cells. The use of simulated data allowed us to derive analytically true relevant dimensions and the true Volterra kernels of the cell models. We could then compare these true parameters with

those estimated from input–output data. The outcome of these comparisons appears in the True parameters sections under the [Simulated simple cell](#) section and the [Simulated complex cell](#) section. Next, we compared the Volterra model with two other well-known models that also use relevant dimensions. For the simulated simple cell, the comparison was with a model using STA to recover a single relevant dimension and histograms to estimate nonlinearity (Chichilnisky, 2001). For the simulated complex cell, the comparison was with a model using STC to recover the relevant space and histograms to estimate the nonlinearity (Touryan et al., 2005). In the [PPR versus STA](#) section and the [PPR versus STC](#) section, we present the results of comparing PPR with STA and STC, respectively. In turn, the Volterra model versus histogram-based model sections under the [Simulated simple cell](#) section and the [Simulated complex cell](#) section compare the Volterra model with the histogram-based models for the simulated simple and complex cells, respectively. Finally, we use physiological data from cortical complex cells to show the feasibility of our methods with real data ([Cortical complex cell](#) section).

Simulated simple cell

True parameters

We used PPR to estimate the relevant space of the simulated simple cell. For this estimation, we used a training set of 5,000 patches from natural images. [Figure 2a](#) shows the curve of predictive error versus the number of estimated relevant dimensions for the simple cell. Using two estimated relevant dimensions leads to a large reduction in prediction error compared with using only one dimension. In contrast, using more than two estimated relevant dimensions yield only small reductions in predictive error. (Here, we define large as the lack of intersection between neighboring error bars—Size 10 standard errors.) Based on these

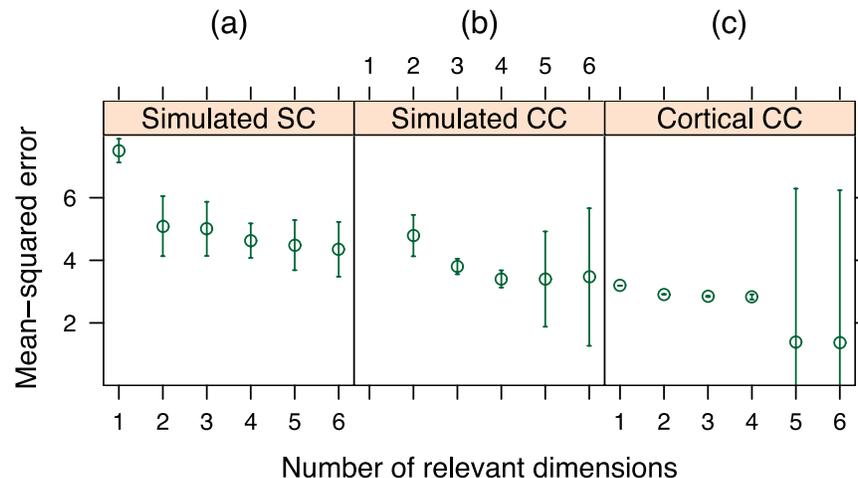


Figure 2. Goodness of fit of PPR models as a function of number of estimated relevant dimensions. (a) Simulated simple cell. (b) Simulated complex cell. (c) Cortical complex cell. The size of error bars is 10 standard errors. Based on these data, we use two, two, and three estimated relevant dimensions for the Volterra model of the simulated simple cell, simulated complex cell, and cortical complex cell, respectively.

data, we used the projection directions of a two-term PPR model to compute the estimated relevant dimensions for the Volterra model (Volterra relevant-space technique section under the [Receptive-field estimation](#) section).

Although we estimated two significant relevant dimensions from the data, as described in the [Simulated simple cell](#) section, the relevant space of the simple-cell model is spanned by a single relevant dimension. What matters for our method is that this space is part of the larger, estimated one. If this holds, then the estimation of the Volterra model, by minimization of [Equation 13](#), can prune nonrelevant estimated space. (This pruning could be done, e.g., by setting appropriate coefficients to zero.) We must thus ensure that the true relevant dimension is similar to its projection onto the estimated relevant space. The left column of [Figure 3](#) shows the analytically relevant dimension of the simple-cell model. The right column shows the projection of this dimension onto the estimated relevant space. Despite the effects of noise, the projection of the true relevant dimension onto the estimated relevant space captures well the preferred orientation, spatial frequency, and phase of the analytical one.

To measure the goodness of fit of the true relevant dimension and its projection onto the estimated relevant space,

we plot the coefficients of the true relevant dimension against the coefficients of its projection. We then fit a linear model to this plot and use the coefficient of determination (or r^2 statistic) of the fitted linear model as the goodness-of-fit measure. The r^2 statistic is the ratio between the reduction in variance achieved by the regression model (the total variance of the outputs minus the variance of the residuals) and the total variance of the outputs. Without a relationship between the true relevant dimension and its projection onto the estimated relevant space, the r^2 statistic would be 0. In contrast, if the true relevant dimension were identical to its projection, then the r^2 statistic would be 1. For the simple cell, the r^2 statistic for the analytically relevant dimension and its projection onto the estimated relevant space was .94.

Having estimated the relevant dimensions, we next estimated Volterra models of different orders using the same set of 5,000 image/response pairs used to estimate the relevant dimensions. We selected the order of the Volterra model as the least order maximizing the average predictive power (Volterra relevant-space technique section under the [Receptive-field estimation](#) section). [Figure 4a](#) shows the mean predictive power of simple-cell Volterra models as a function of their order. Second- and third-order Volterra

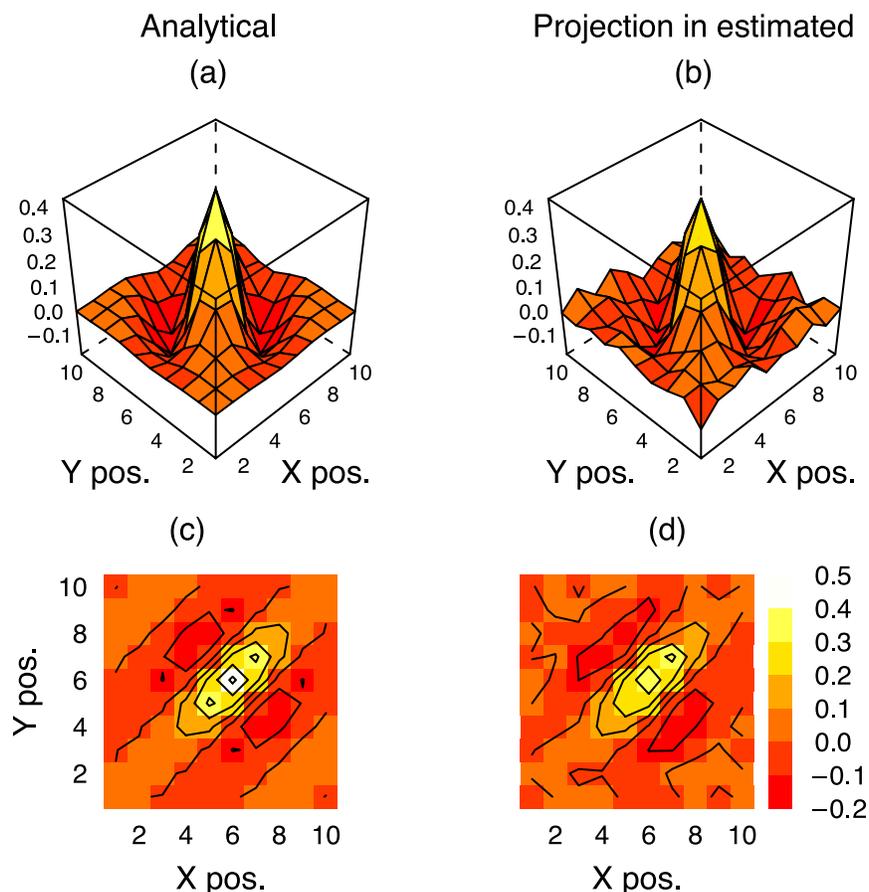


Figure 3. Simulated simple cell: relevant dimensions. (a, c) Analytical relevant dimension. (b, d) Projection of the analytical relevant dimension onto the estimated relevant space. (a, b) Perspective plot. (c, d) Contour plot. The analytical relevant dimension is accurately approximated by its projection onto the relevant space, $r^2 = .94$.

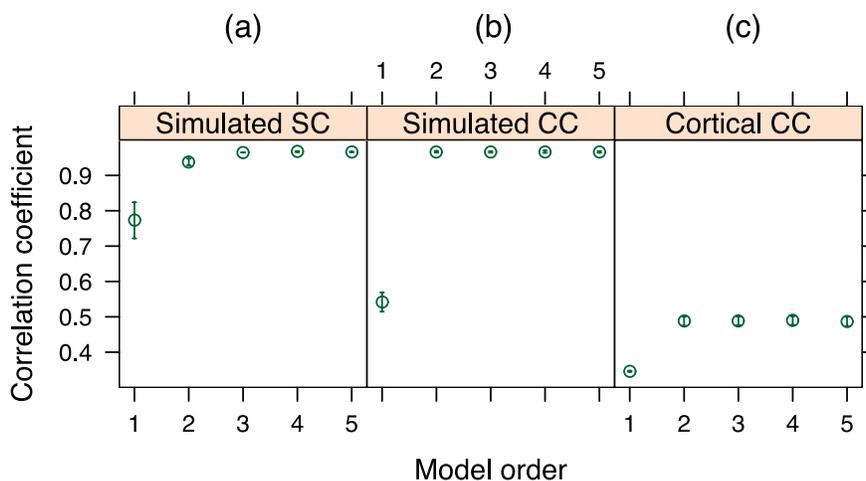


Figure 4. Correlation coefficients between cell responses and the Volterra model predictions versus the order of the Volterra model. (a) Simulated simple cell. (b) Simulated complex cell. (c) Cortical complex cell. The size of error bars is two standard errors. The order of the Volterra model was selected as the least order maximizing the average predictive power. We selected a third, second, and second order for the simulated simple cell, simulated complex cell, and cortical complex cell, respectively.

models lead to large improvements in cross-correlation (here, large is defined as the lack of intersection between neighboring error bars—Size 2 standard errors), illustrating the relevance of nonlinear contributions to characterize the simple-cell data. Volterra models of order higher than three do not lead to improvements in predictive power. Based on these data, we selected a third-order Volterra model.

How important are high-order nonlinearities in the description of responses? To answer this question, we defined a simple measure of the relative importance of each Volterra term to the overall response (**Volterra terms' contributions** section). We expected that the relative contribution to the response of different Volterra terms should be dependent on the response range. Higher responses might be more non-

linear and thus would require higher order Volterra terms. **Figure 5** shows the mean relative contributions of different Volterra terms to the response of the simple cell. We show the mean contributions for groups of small (first quartile), middle (second and third quartiles), and large (fourth quartile) responses. For small responses (**Figure 5a**), the most important terms are the low-order ones. In contrast, for medium (**Figure 5b**) and large (**Figure 5c**) responses, the situation reverses and high-order terms become dominant. Therefore, even for a cell often considered linear as the simple cell, high-order nonlinearities may matter, especially when describing responses to stimuli causing large responses. In the natural world, most responses may be small, as contrasts tend to be low (Balboa & Grzywacz, 2000;

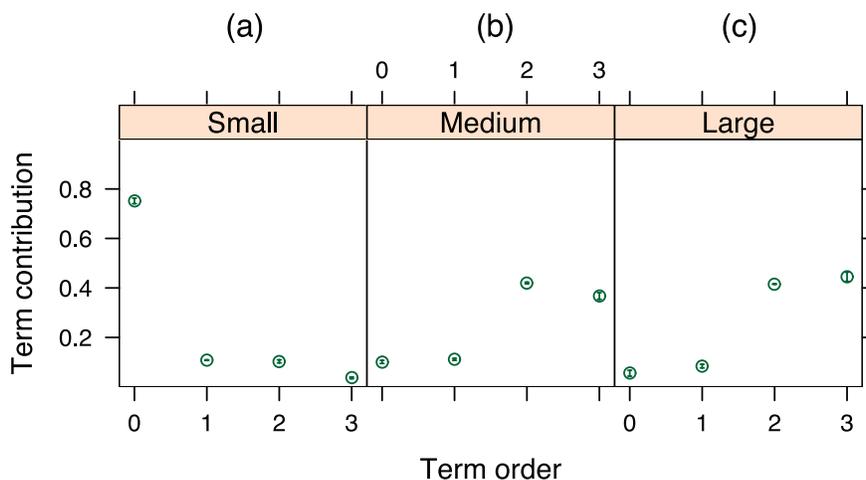


Figure 5. Simulated simple cell: mean relative contributions of Volterra terms to the response. Contributions have been averaged over different response ranges: (a) small responses (first quartile), (b) medium responses (second and third quartiles), and (c) large responses (fourth quartile). As the responses increase, high-order Volterra terms become more important.

Ruderman & Bialek, 1994; Tadmor & Tolhurst, 2000; Zhu & Mumford, 1997), but some responses will be high, as the distribution of natural intensities is kurtotic (Field, 1987).

To assess the quality of the estimated kernels, we compared them with the true kernels derived from the analytical expression of the simple-cell model (Appendix B). The left column in Figure 6 shows the first-order true kernel, whereas its estimation appears in the right column. Similarly, Figure 7 shows the second-order kernel slice with respect to Position (6, 6); this was the slice with largest error. The kernels have been scaled to reflect their mean contribution to the cell response (Displaying kernels section).

To measure the goodness of fit of estimated kernels to true kernels, we scaled them, in proportion to their mean contribution to the response, and computed their *MSE* (Goodness of fit of kernels section). For the simple cell, the *MSE* between scaled true and estimated kernels was $2.42\text{E}-03$ for the first-order kernel (Figure 6) and $1.83\text{E}-03$ for the second-order kernel slice with respect to Position (6, 6) (Figure 7). This slice had the largest *MSE* among all second-order kernel slices (the *MSE* averaged across all second-order kernel slices was $1.74\text{E}-04$). Therefore, despite the noise, the estimated kernels accurately approximate their true values.

Although, for the simple cell, the first-order kernel is similar to the second-order kernel slice with respect to Position (6, 6), their interpretation is different. The first-order kernel indicates that the natural image would elicit a response if it were bright in the yellow region and dark in the red region. In turn, the second-order kernel slice indicates that simultaneous bright stimuli anywhere in the yellow region and at Position (6, 6) would facilitate the response. However, having bright stimuli in the red regions and at Position (6, 6) at the same time would inhibit the response. In other words, first-order kernels describe how much a stimulus at a given position linearly contributes to the response. Second-order kernels describe nonlinear modulations of responses, through interactions between stimulus pairs at given positions.

PPR versus STA

So far, we showed that PPR is a good technique to estimate relevant dimensions that will be of use in the construction of Volterra models of simulated simple cells. Here, we compare PPR with STA, which is a technique often used to estimate the relevant dimension of simple cells. In particular, we compared the convergence

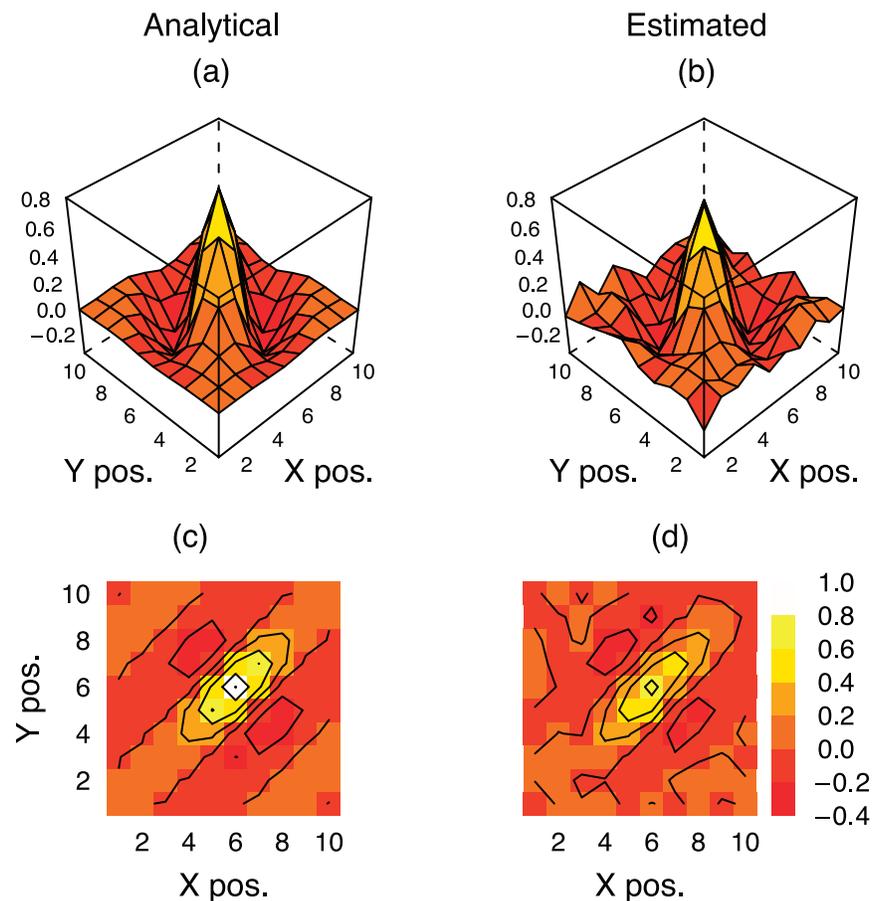


Figure 6. Simulated simple cell: first-order kernels. (a, c) True; (b, d) Estimated. (a, b) Perspective plot. (c, d) Contour plot. Despite the noise, the estimated kernel accurately approximates its true value, $MSE = 2.42\text{E}-03$.

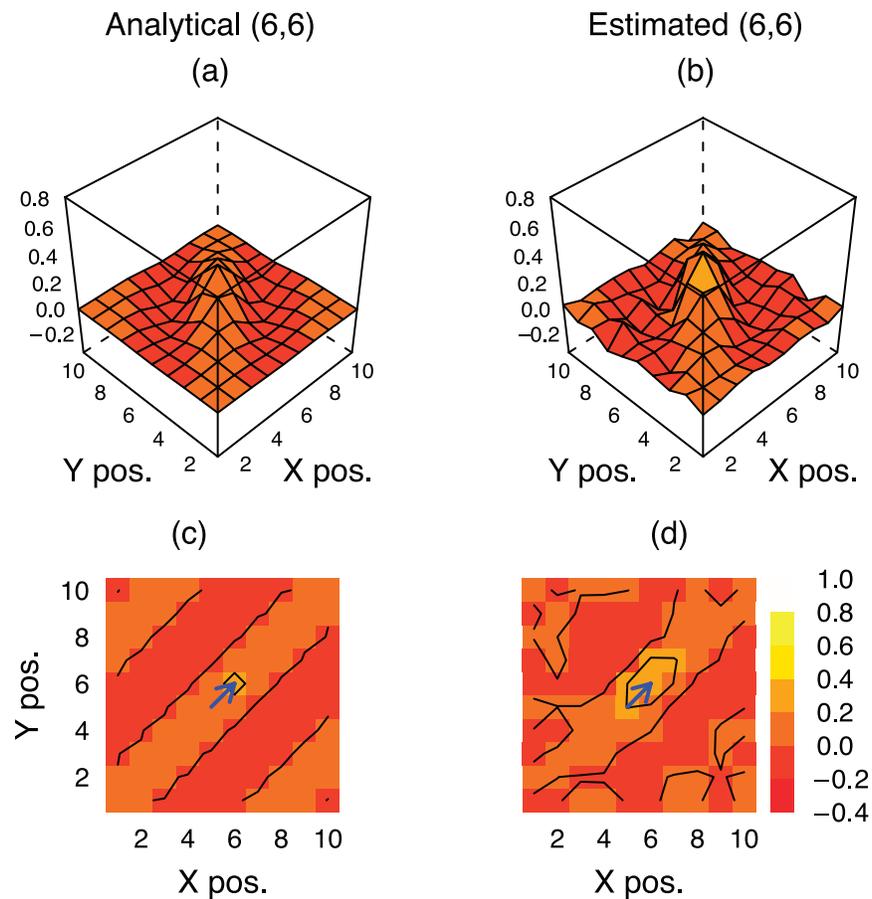


Figure 7. Simulated simple cell: second-order kernel slices with respect to Position (6, 6). (a, c) True; (b, d) Estimated. (a, b) Perspective plot. (c, d) Contour plot. Blue arrows in the contour plots indicate the reference position. Despite the noise, the estimated kernel accurately approximates its true value, $MSE = 1.83E-03$.

properties of both algorithms parametric on number of inputs and signal-to-noise ratio. Figure 8a plots the r^2 statistic for the regression between the coefficients of the true and estimated relevant dimensions as a function of the number of inputs used in the estimation. For this plot, the average response of the simulated simple cell was 5 spikes/image. Figure 8b plots the same r^2 statistic as a function of the number of spikes per image in the simulated responses, when using 5,000 inputs in the estimation. These figures show that the two PPR relevant dimensions better approximate the true relevant dimension of the simple-cell model than the relevant dimension estimated by STA (black vs. green curve). The superiority of PPR over STA is not due to the fact that PPR is using two relevant dimensions to approximate the true relevant dimension, although STA is using only one. PPR also outperforms STA when using one relevant dimension (red vs. green curves).

Volterra model versus histogram-based model

The True parameters section showed that Volterra models can represent the nonlinearities of simple cells well, starting from its estimated relevant dimension. We now compare the

Volterra model with another popular technique to represent nonlinearities, namely, a histogram-based model (Histogram-based model for simple cells section). To make our comparison independent from the method used to estimate the relevant space, we compare the predictive power of the Volterra model using the two PPR relevant dimensions, the Volterra model using the STA relevant dimension, the nonlinearity estimated from histograms using the first PPR relevant dimension, and the nonlinearity estimated from histograms using the STA relevant dimension. (We use only the first PPR relevant dimension with the histogram-based model because it is one dimensional. In other words, this model accepts only one relevant dimension as input.) The predictive power of these models as a function of the number of inputs (using a mean response of 5 spikes/image) is plotted in Figure 9a. A plot as a function of the number of spikes per image (using 5,000 images) appears in Figure 9b. Note that, as discussed in the Data partitioning section, the data used to evaluate the different models were different from those used to estimate their parameters. Thus, the predictive power measures are not inflated by overfitting effects.

The figures show that, when using the relevant dimensions estimated by PPR, the Volterra model predicts better than the nonlinearity estimated from histograms. This superiority

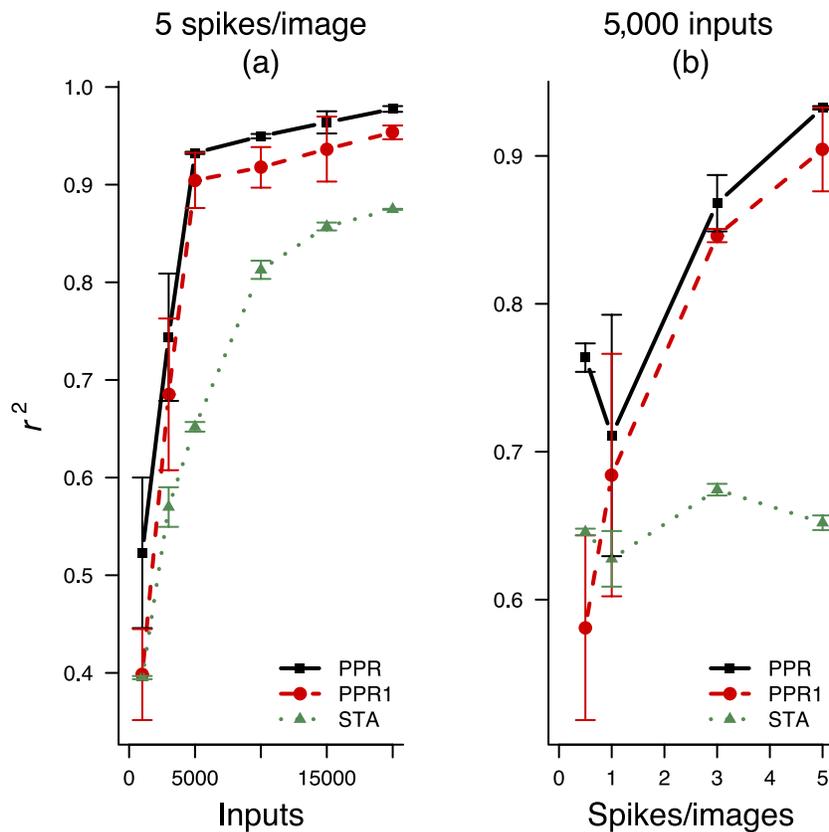


Figure 8. Simulated simple cell: comparison between PPR and STA. The figures show the r^2 statistic for the regression of the coefficients of the true and estimated relevant dimensions. This statistic appears as a function of (a) the number of inputs when the mean response of the simulated cell is 5 spikes/image and (b) the mean number of spikes in the responses when using 5,000 image–response pairs. Black curve: r^2 statistic between the true relevant dimension and its projection onto a space spanned by two PPR relevant dimensions. Red curve: r^2 statistic between the true relevant dimension and one PPR relevant dimension. Green curve: r^2 statistic between the true and STA relevant dimensions. The size of the error bars is two standard errors. PPR performs better than STA either when using one (red vs. green curve) or two (black vs. green) relevant dimensions to approximate the true relevant dimension.

holds for moderate to high number of inputs, that is, more than 1,000 image–response pairs, and for moderate to high signal-to-noise ratios, that is, more than 1 spike/image. When using the STA relevant dimension, the Volterra model performs at the same level as the nonlinear function estimated from histograms. This is significant because the simulated data accurately match the assumptions of the STA histogram model. Moreover, the parametric nonlinear function fitted to the histogram in the reconstruction of the nonlinear function ([Histogram-based model for simple cells](#) section) was the same sigmoidal function used to generate the simulated responses ([Simulated simple cell](#) section). Therefore, it is important that the Volterra model, making no assumptions on the response generation mechanism, is at least as good as the nonlinear function estimated from histograms.

Simulated complex cell

True parameters

As for simple cells, we also used PPR to estimate the relevant space of the simulated complex cell from 5,000 stimulus/response pairs. [Figure 2b](#) shows the mean error as a

function of the estimated number of relevant dimensions. Two is the largest number of estimated relevant dimensions for which models predict significantly better than models using fewer dimensions. Consequently, we computed the relevant dimensions for the Volterra model from a two-term PPR model.

The left columns of [Figures 10](#) and [11](#) show the first and second analytically relevant dimensions of the complex-cell model. In turn, the right columns show their projection onto the space spanned by the two estimated relevant dimensions. The analytically relevant dimensions are accurately approximated by their projection onto the estimated relevant space (first relevant dimension, $r^2 = .96$; second relevant dimension, $r^2 = .95$).

With these relevant dimensions we estimated Volterra models of different orders, using the same set of 5,000 stimulus/response pairs as for the estimation of the relevant dimensions. We selected the order of the Volterra model as the least order maximizing the average predictive power (Volterra relevant-space technique section under the [Receptive-field estimation](#) section). [Figure 4b](#) shows the mean predictive power of complex-cell Volterra models as

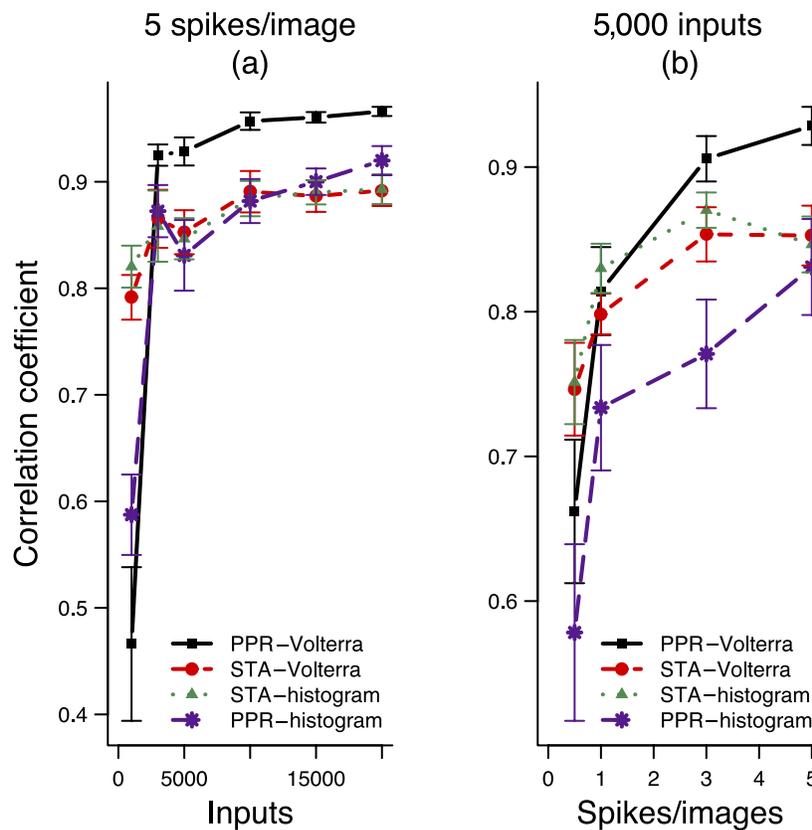


Figure 9. Simulated simple cell: predictions of the Volterra model and the nonlinearity estimated from histograms. This figure shows average correlation coefficients between simulated cell responses and model predictions. These coefficients appear as functions of (a) the number of inputs (using a mean response of 5 spikes/image) and (b) the number of spikes per image (using 5,000 image–response pairs). Black curve: Volterra model using the two PPR relevant dimensions. Red curve: Volterra model using the STA relevant dimension. Green curve: nonlinearity estimated from histograms using the STA relevant dimension. Blue curve: nonlinearity estimated from histograms using the first PPR relevant dimension. The size of the error bars is two standard errors. For sufficiently high number of inputs and signal-to-noise ratio, the Volterra model using the relevant dimensions estimated by PPR performs better than the nonlinearity estimated from histograms. When using the STA relevant dimension, both models perform equally.

a function of their order. Linear models predict the responses poorly due to a model mismatch. A linear model cannot accurately predict a quadratic response. Second-order and higher order Volterra models perform equally well. Hence, we used a second-order Volterra model to characterize the complex-cell data.

As for the simulated simple cell, we looked at the contributions of the various Volterra terms to the response. We again averaged the contributions across small (first quartile), medium (second and third quartiles), and large (fourth quartile) responses. These contributions are shown in Figure 12. The simulated complex cell is purely second order; thus, the second-order Volterra term should be the only one that contributes to the response. Accordingly, medium and large responses are dominated by the second-order term. However, the spontaneous, zeroth-order Volterra term has been spuriously fitted to a nonzero value. Its contribution is largest for small responses. Such spurious fits may occur because, at small responses, the Poisson distribution is highly kurtotic, admitting values that would be outliers in Gaussian distributions. Consequently, our least

squares fit, which is sensitive to outliers, might not be sufficiently robust at small responses. Nevertheless, these plots indicate that one should focus on the second-order kernel to interpret the responses of the simulated complex cell.

To assess the quality of the estimated kernels for the complex cell, we compared them with their true values. The left column in Figure 13 shows the first-order true kernel, whereas its estimation appears in the right column. Similarly, Figure 14 shows the second-order Volterra kernel slice with respect to Position (6, 6). The kernels have been rescaled to reflect their mean contributions to the cell response (Displaying kernels section).

To measure the goodness of fit of estimated kernels to true kernels, we scaled them, in proportion to their mean contribution to the response, and computed the *MSE* between them (Goodness of fit of kernels section). The *MSE* between scaled true and estimated kernels was $5.6E-04$ for the first-order kernel (Figure 13) and $1.42E-03$ for the second-order kernel slice with respect to Position (6, 6) (Figure 14). This slice was the one with maximum *MSE* among all slices (the mean *MSE* across all slices was

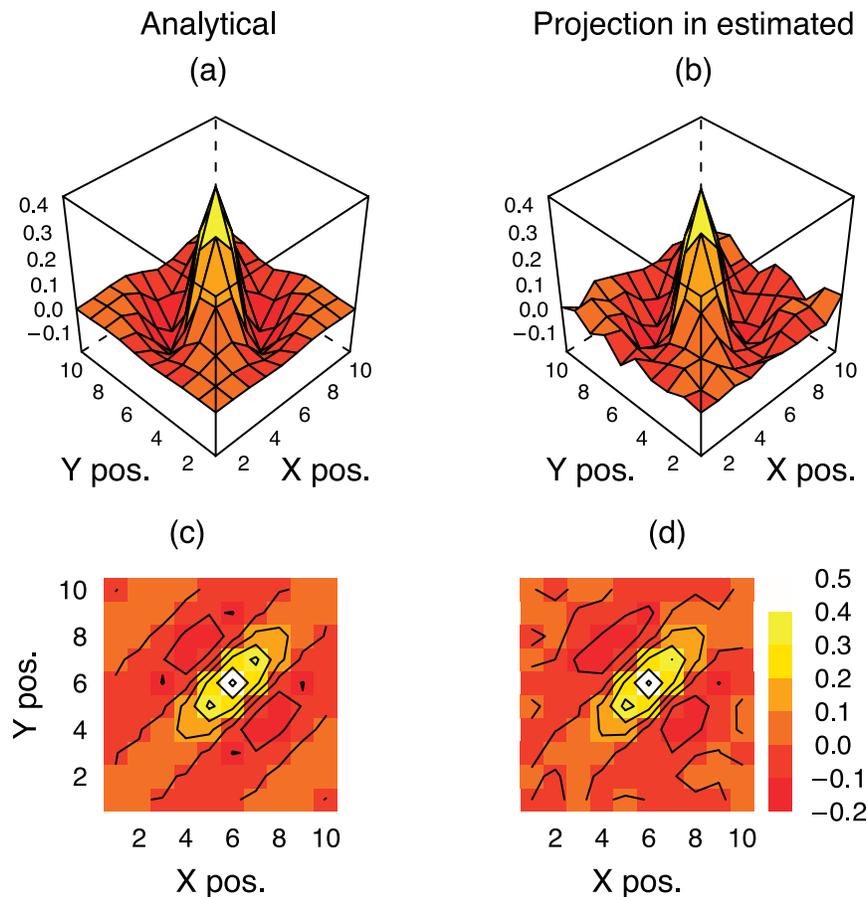


Figure 10. Simulated complex cell: first relevant dimension. (a, c) First analytical relevant dimension. (b, d) Projection of first analytical relevant dimension onto the estimated relevant space. (a, b) Perspective plot. (c, d) Contour plot. The analytical relevant dimension is well approximated by its projection onto the estimated relevant space, $r^2 = .96$.

$3.06E-04$). Thus, despite the noise, the estimated kernels accurately approximate their true values.

Note that for the complex cell, the first-order kernel should be zero. The estimated kernel is not zero and has structure. This is because, in the linear Volterra equation (Equation 10), small first-order coefficients, corresponding to a first-order Volterra term with small contribution (Figure 12), will be dominated by large coefficients, corresponding to spontaneous and second-order Volterra terms with large contributions. Then, in the regression procedure, the small first-order coefficients will tend to fit noise. Subsequently, in the reconstruction of the estimated first-order kernel (Equation 15), these noisy first-order coefficients will multiply relevant dimensions with clear structure, generating first-order kernels with spurious structure. Consequently, we should not try to interpret kernels corresponding to terms whose contributions to the response are very small.

PPR versus STC

The last section showed that PPR is a good technique to estimate relevant dimensions for use in the construction of Volterra models of simulated complex cells. Here, we com-

pare PPR with the modification of STC for natural images proposed by Touryan et al. (2005), henceforth referred as STC for natural images. These authors mentioned that natural images are highly variable in their global contrast, that is, variance. To avoid possible contrast adaptation of the cells, they normalized the natural images to have equal variance. Our simulated complex cell did not include adaptation mechanisms; thus, we did not need to normalize the images to have equal variance. When we used STC for natural images with simulated responses to nonequal-variance images, we obtained poor results, but when we used responses to equal-variance images, we obtained satisfactory results. To alleviate this problem, we equalized the variance of the natural images as a preprocessing step before computing STC. However, in these situations, we did not modify the simulated responses. The variance-equalization step was only necessary for the estimation of STC for natural images. Having estimated the relevant dimensions, the original, nonequal-variance natural images were used to recover the nonlinearity using histograms.

Figure 15a shows the average r^2 statistic as a function of the number of images for several conditions involving PPR and STC. Figure 15b plots the average r^2 statistic as a

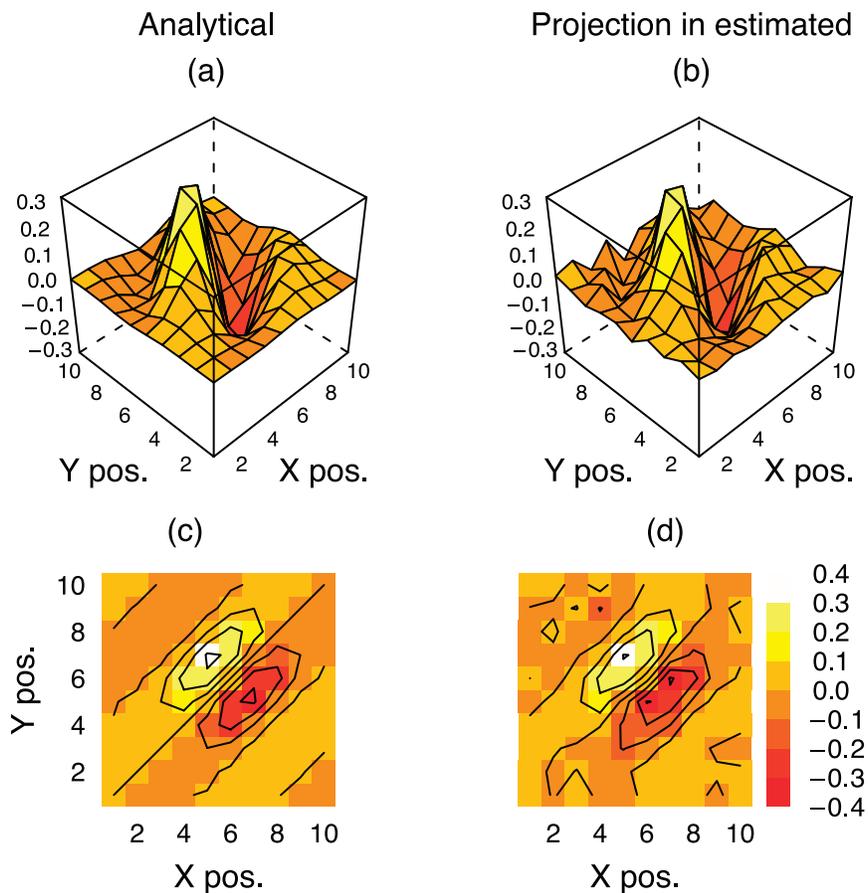


Figure 11. Simulated complex cell: second relevant dimension. (a, c) Second analytical relevant dimension. (b, d) Projection of second analytical relevant dimension onto the estimated relevant space. (a, b) Perspective plot. (c, d) Contour plot. The analytical relevant dimension is well approximated by its projection onto the estimated relevant space, $r^2 = .95$.

function of the average number of spikes in the responses. STC achieves its best performance with responses to natural images with equal variance (red curve), as in

Touryan et al. (2005). When the responses correspond to nonequal-variance images, STC for natural images performs poorly (blue curve) and the variance-equalization

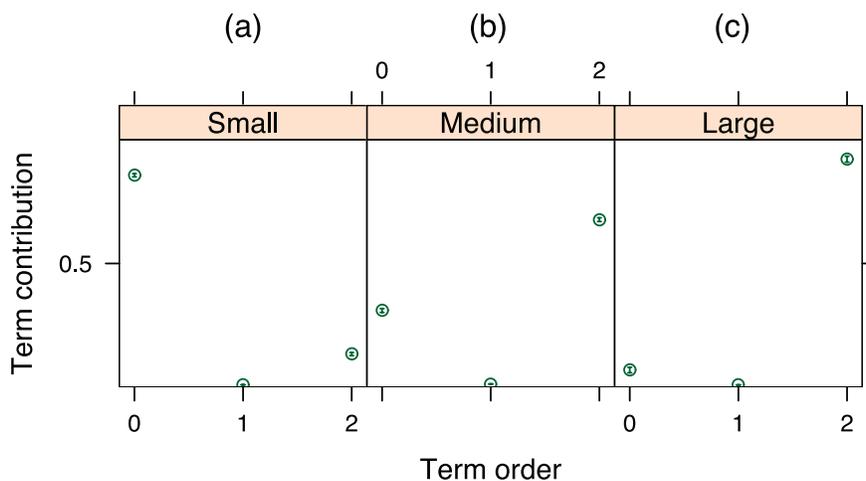


Figure 12. Simulated complex cell: mean relative contributions of Volterra terms to the response. Contributions have been averaged over different response ranges: (a) small responses (first quartile), (b) medium responses (second and third quartiles), and (c) large responses (fourth quartile). Medium and large responses of the purely quadratic complex-cell model are dominated by the second-order Volterra term. The contributions of the zeroth-order term reflect a spurious fit (see text).

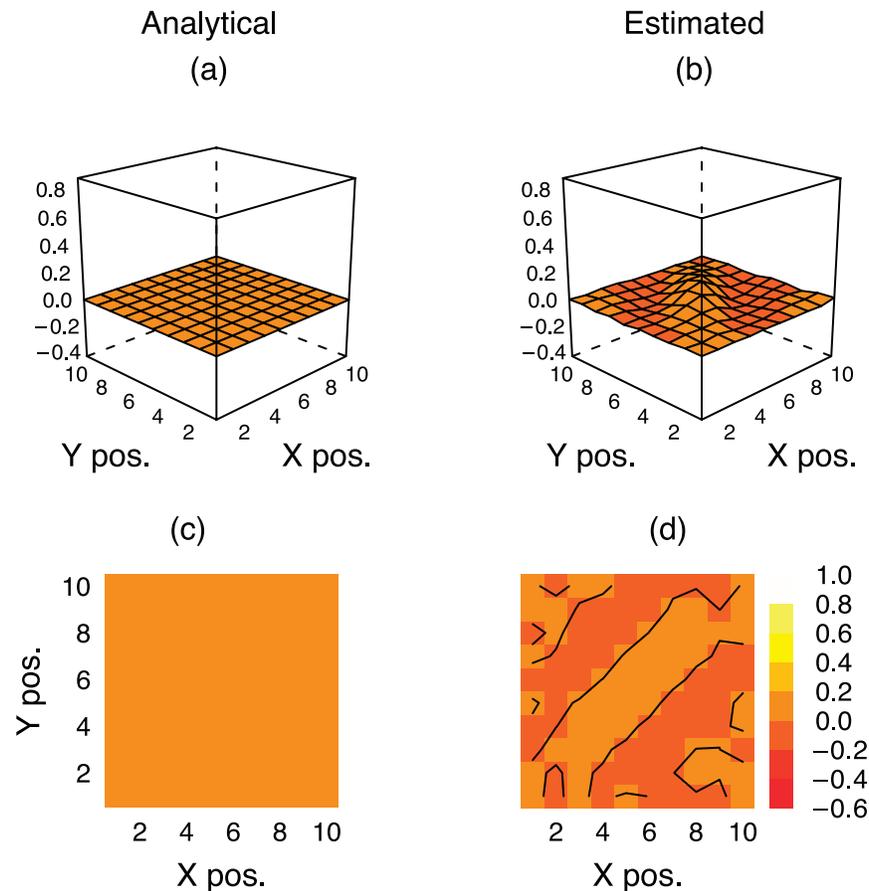


Figure 13. Simulated complex cell: first-order kernels. (a, c) True; (b, d) Estimated. (a, b) Perspective plot. (c, d) Contour plot. The kernels have been scaled to reflect their mean contribution to the cell response. Although the estimated but not the true kernel has some structure, the amplitude is low, making the estimation not much different from zero, $MSE = 5.6E-04$.

preprocessing step significantly improves its performance (green curve). With nonequal-variance data, PPR outperforms STC for natural images for any number of inputs or signal-to-noise ratios (black vs. green and blue curves). However, STC has an advantage in running time over PPR. The latter is a time-consuming iterative algorithm, requiring several estimations of smooth functions and least squares minimizations. STC is a fast algorithm that only requires the computation of covariance matrices and their spectral decomposition. A fast algorithm was essential for the online estimation of the relevant dimensions of a cell, an impressive task performed by Touryan et al. (2005).

Volterra model versus histogram-based model

The True parameters section showed that Volterra models can represent the nonlinearities of complex cells well, starting from its estimated relevant dimensions. Here, we compared the performance of the Volterra model with that of a complex-cell histogram-based model (Histogram-based model for complex cells section). To make the results independent from the method used to estimate the relevant dimensions, we compare the Volterra model using the rel-

evant dimensions estimated by PPR, the Volterra model using the STC relevant dimensions, the histogram-based model using the PPR relevant dimensions, and the histogram-based model using the STC relevant dimensions. Figures 16a and b plot the correlation coefficient between the simulated cell responses and the model predictions as a function of the number of inputs and the number of spikes in the simulated responses, respectively. The relevant dimensions estimated by PPR yield better predictions than those estimated by STC for both the Volterra (black vs. red curves) and the histogram-based models (blue vs. green curves). When using the PPR relevant dimensions, the Volterra model predicts responses slightly better than the histogram-based model (black vs. blue curves). With the STC relevant dimensions, both models perform similarly (red vs. green curves).

Cortical complex cell

Our results with simulated data from a cortical complex cell have shown that PPR can effectively estimate the cell's true relevant dimensions. Moreover, these results showed that the Volterra relevant-space technique can accurately

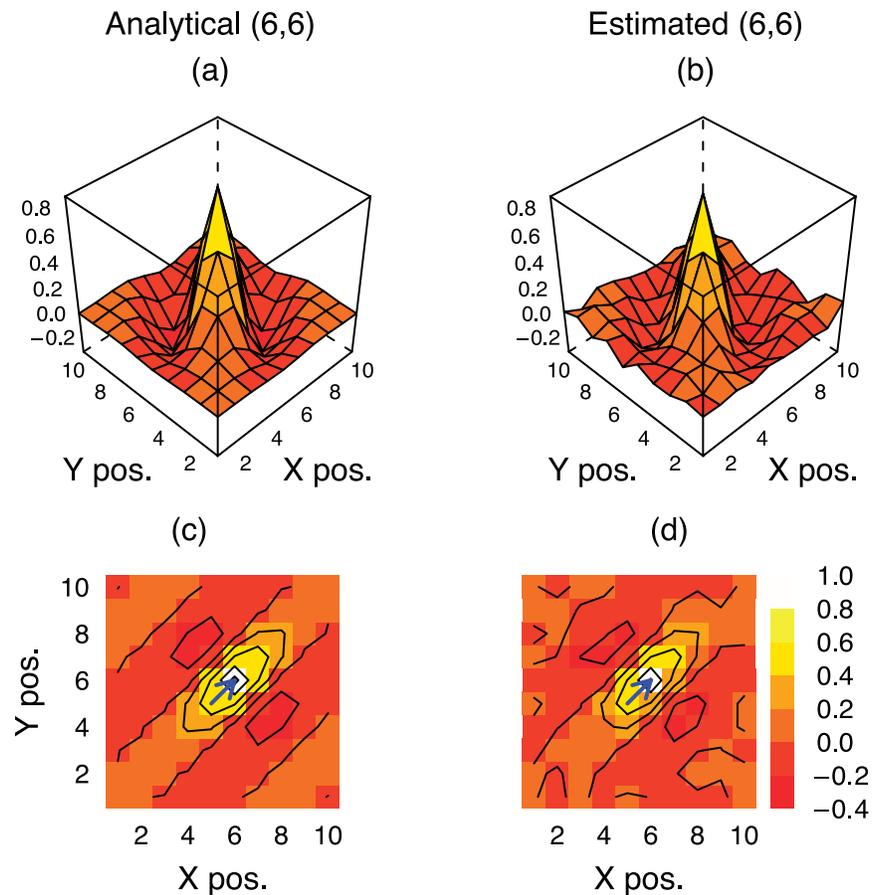


Figure 14. Simulated complex cell: second-order kernel slices with respect to Position (6, 6). (a, c) True; (b, d) Estimated. (a, b) Perspective plot. (c, d) Contour plot. Blue arrows in the contour plots indicate the reference position. Despite the noise, the estimated kernel accurately approximates its true values, $MSE = 1.42E-03$.

estimate the cell's true Volterra kernels. Finally, the results indicated that a Volterra model accurately predicts the simulated cell's responses. However, these results cannot tell whether the method is practical for real cells. Here, we test PPR, STC, the Volterra model, and the histogram-based model with physiological data from a cortical complex cell obtained by Touryan et al. (2005).

As for the simulated data, we began by estimating the mean predictive error of PPR models as a function of the number of estimated relevant dimensions as shown in Figure 2c. Based on this figure, we used three relevant dimensions for the Volterra model. The relevant dimensions estimated by PPR were noisy. To smooth them, we applied a low-pass filter. We then orthonormalized them. The resulting relevant dimensions used to estimate the Volterra model are shown in Figure 17. They have clear structure, showing a 45 deg orientation preference. These relevant dimensions yielded better predictive values than the original noisy relevant dimensions.

As before, we selected the order of the Volterra model as the least order maximizing the average predictive power. Figure 4c shows the mean predictive power of Volterra models as a function of their order. Based on these data, we selected a second-order Volterra model for the cortical com-

plex cell. As for the simulated complex cell, we determined the relative contributions of the various Volterra terms to the response of the cortical complex cell. We averaged the contributions across small (first quartile), medium (second and third quartiles), and large responses (fourth quartile). These contributions are shown in Figure 18. At all response levels, the term with most relative contribution was the spontaneous one, indicating that the complex cell had a strong background activity. Also, relative contributions from the first-order term are negligible, as for the simulated complex cell (Figure 12). As the response level increased, relative contributions from the spontaneous term became smaller, whereas contributions from the second-order term increased.

The left column of Figure 19 shows the estimated first-order kernel, and the right column shows the estimated second-order kernel slice with respect to Position (5, 9). This was the second-order kernel slice with the largest value. Both kernels have been scaled to reflect their mean contribution to the cell response. Because contributions from the first-order Volterra term are negligible (Figure 18), as in the case of the simulated complex cell, the first order is probably spurious (see discussion at the end of the **True parameters** section). Thus, we refrain from interpreting

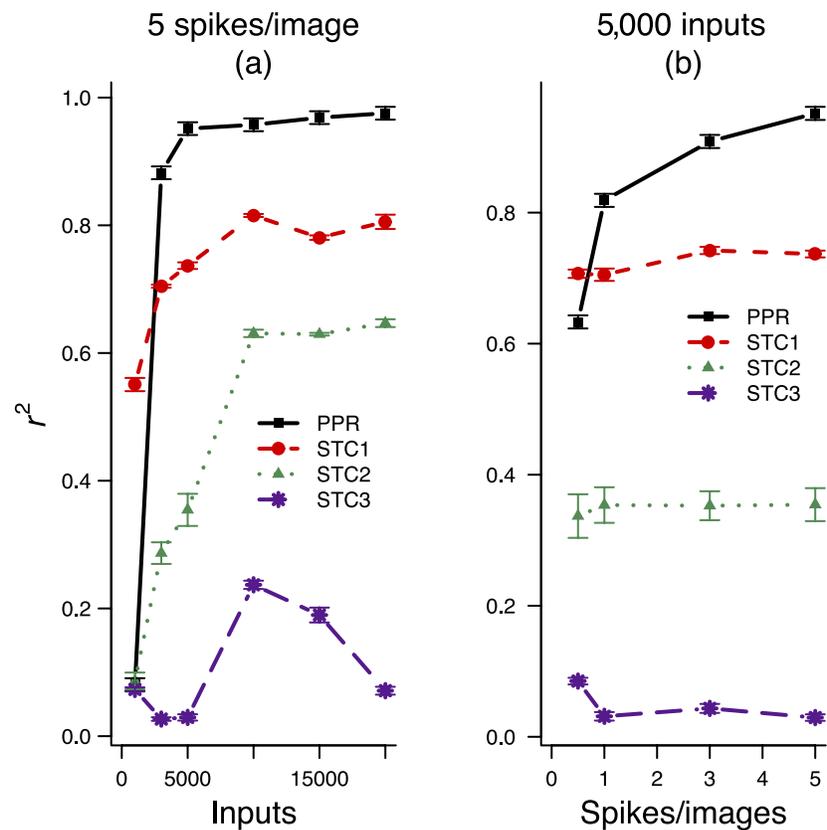


Figure 15. Simulated complex cell: comparison between PPR and STC for natural images at various conditions. These conditions were the following: red curve: STC1, equal-variance responses; green curve: STC2, nonequal-variance responses with variance-equalization preprocessing; blue curve: STC3, nonequal-variance responses without variance-equalization preprocessing. The black curve is for PPR with non-equal variance responses. The figure shows average r^2 statistic for the two relevant dimensions of the simulated complex cell. This statistic appears as a function of (a) the number of inputs when the mean response is 5 spikes/image and (b) the number of spikes per image when using 5,000 image–response pairs. The size of the error bars is two standard errors. STC achieves its best performance for responses to equal-variance images. Variance-equalization preprocessing improves the performance of STC when using responses to nonequal-variance images. For responses to nonequal-variance images, PPR outperforms STC for any number of inputs and signal-to-noise ratios.

this kernel. The second-order kernel slice with respect to Position (5, 9) is positive everywhere, being larger in the vicinity of the reference position (blue arrow) and smaller at positions along the central diagonal. Thus, positive correlations between a point of light at the reference position and a point of light in its close surrounding will induce the strongest pairwise response facilitations. In turn, positive correlations between a point of light at the reference position and a point of light in the central diagonal (along the antipreferred orientation) will induce weaker facilitations.

Points outside and near the boundary of the receptive field should not modulate the response in any sense, either linearly or nonlinearly. For the linear kernel, we observe that points near the boundary of the receptive field are close to zero. This is what we would expect from a well-estimated linear kernel. However, the second-order kernel slice is positive everywhere. This illustrates a problem because it shows that points on the boundary of the receptive field nonlinearly modulate the cell response. This problem could be an artifact of oversmoothing the PPR relevant

dimensions. However, these possibly oversmoothed relevant dimensions lead to better predictions than relevant dimensions with less smoothing (including nonsmoothed relevant dimensions). An alternative explanation is that the nonlinear component of the receptive field could be larger than its linear component. The size of the images used in the physiological experiments could have been adequate to characterize the linear component of the receptive field, as illustrated in Figures 19a and c, but not large enough to characterize the nonlinear component (Figures 19b and d).

Because the cortical complex cell's responses are dominated by the second-order kernel, it is now presented in greater detail. Figure 20 shows estimated second-order kernel slices with respect to Positions (7, 7) and (9, 5). The former slice (Figure 20, left column) is nearly flat and contains the smallest pairwise facilitations among all second-order slices. For the latter slice (Figure 20, right column), facilitation is strongest for interactions between the reference position and points in its close vicinity or points in the brightest region on the top-left quadrant. Facilitations are

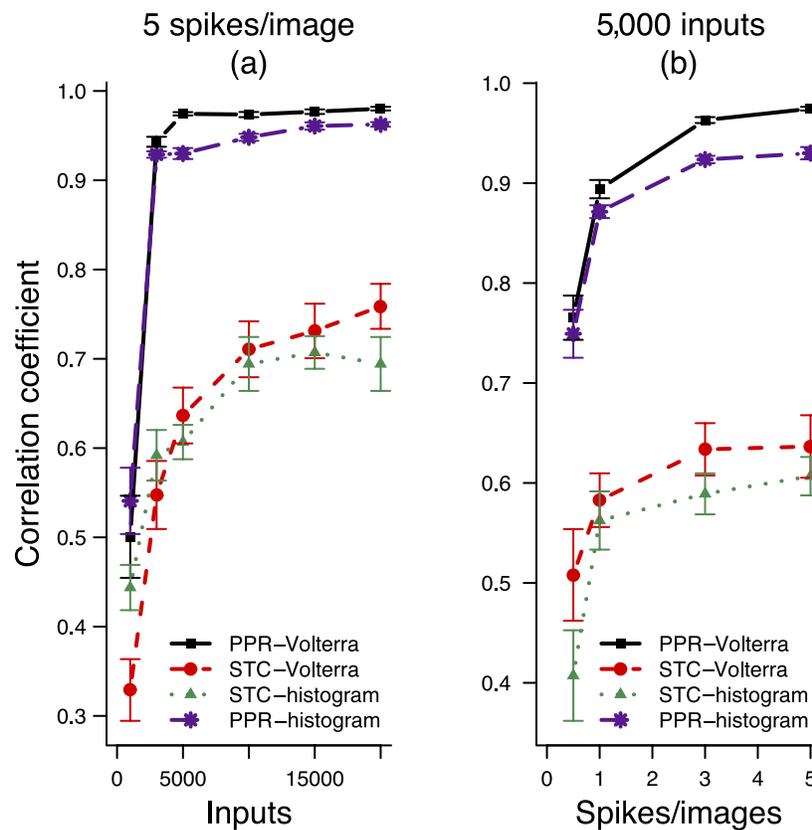


Figure 16. Simulated complex cell: predictions of the Volterra model and the nonlinearity estimated from histograms. This figure shows average correlation coefficients between simulated cell responses and model predictions. These coefficients appear as functions of (a) the number of inputs when the mean responses was 5 spikes/image and (b) the number of spikes per image when using 5,000 image–response pairs. Black curve: Volterra model using the PPR relevant dimensions. Red curve: Volterra model using the STC relevant dimensions. Green curve: nonlinearity estimated from histograms using the STC relevant dimensions. Blue curve: nonlinearity estimated from histograms using the PPR relevant dimensions. The size of the error bars is two standard errors. The relevant dimensions estimated by PPR yield better predictions than those estimated by STC. When using the PPR relevant dimensions, the Volterra model predicts responses slightly better than the histogram-based model.

weaker for interactions between the reference position and points along the central red band.

We next performed an STC analysis on the responses of the cortical complex cell and selected its first two relevant dimensions to describe the relevant space. As described in the [STC for natural images](#) section, we regularized the STC relevant dimensions by only normalizing the eigenvectors corresponding to the 35% largest eigenvalues. These relevant dimensions are shown in [Figure 21](#). These relevant dimensions show a cleaner structure than that of the PPR relevant dimensions in [Figure 17](#). However, the best relevant dimensions are those that lead to better predictions, which we will discuss next.

To compare the performance of the Volterra model with that of the complex-cell histogram model and to make the results independent of the algorithm used to estimate the relevant dimensions, we evaluated the predictive power of both models with PPR and STC relevant dimensions. [Figure 22](#) shows average cross-correlation coefficients between the predictions of the models and the complex-cell responses as a

function of the number of inputs. Note that, as discussed in the [Data partitioning](#) section, the data used to evaluate the different models were different from those used to estimate their parameters. Thus, the predictive power measures are not inflated by overfitting effects.

For any number of inputs, the Volterra models yield better predictions than the histogram-based model with both the PPR relevant dimensions (black vs. blue curves) and the STC relevant dimensions (red vs. green curves). For more than 1,000 inputs, the PPR relevant dimensions produced equal or better predictions than the relevant dimensions estimated by STC for both the Volterra (black vs. red curve) and histogram-based (blue vs. green curve) models. These results are not due to a faulty implementation of the STC algorithm or histogram-based models. The correlation coefficient of the histogram-based model estimated from 20,000 images using the STC relevant dimensions lies in the upper range of those reported in [Figure 6](#) of [Touryan et al. \(2005\)](#). Hence, of the models tested, we found that the best combination is to use Volterra models with PPR relevant dimensions.

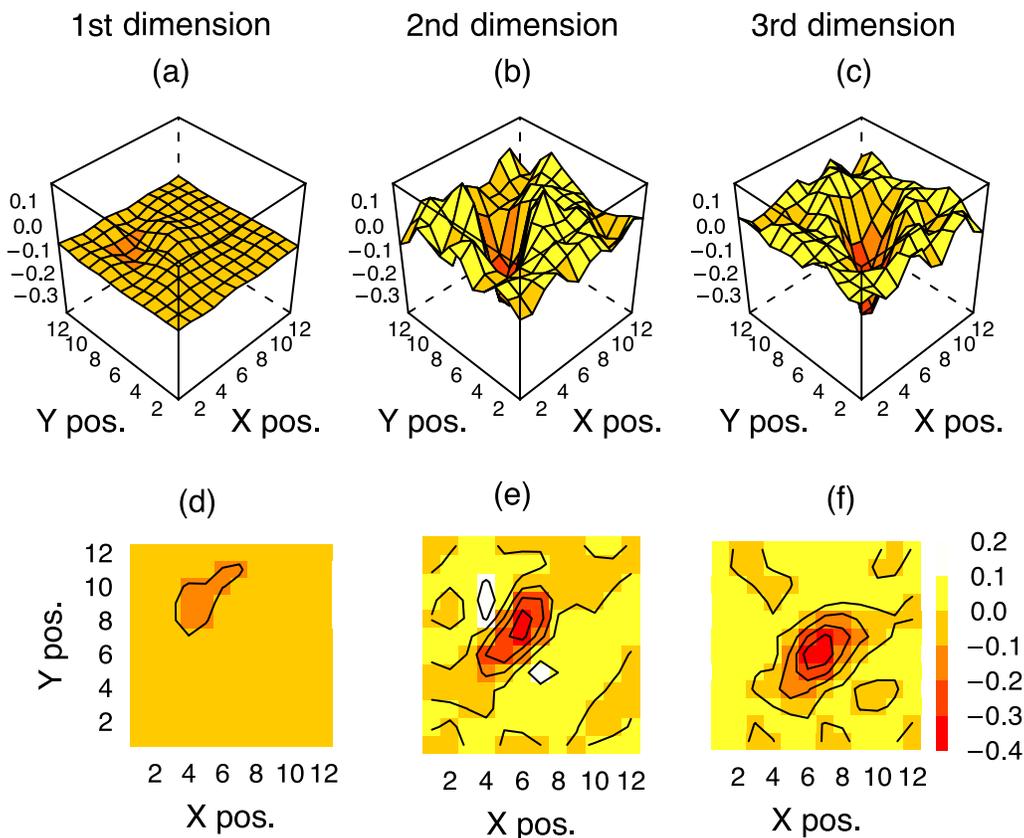


Figure 17. Cortical complex cell: PPR relevant dimensions. (a, d) First relevant dimension. (b, e) Second relevant dimension. (c, f) Third relevant dimension. (a, b, c) Perspective plot. (d, e, f) Contour plot. The relevant dimensions have a clear structure, showing a 45 deg orientation preference.

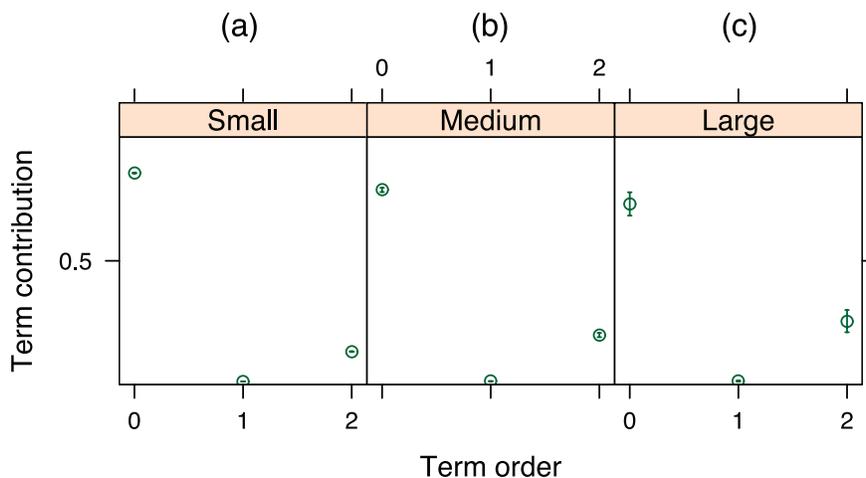


Figure 18. Cortical complex cell: mean relative contributions of Volterra terms to the response. Contributions have been averaged over different response ranges: (a) small responses (first quartile), (b) medium responses (second and third quartiles), and (c) large responses (fourth quartile). This complex cell has a large spontaneous activity. Relative contributions from the first-order term are negligible. As the response level increases, relative contributions from the spontaneous term become smaller, whereas relative contributions from the second-order term become larger.

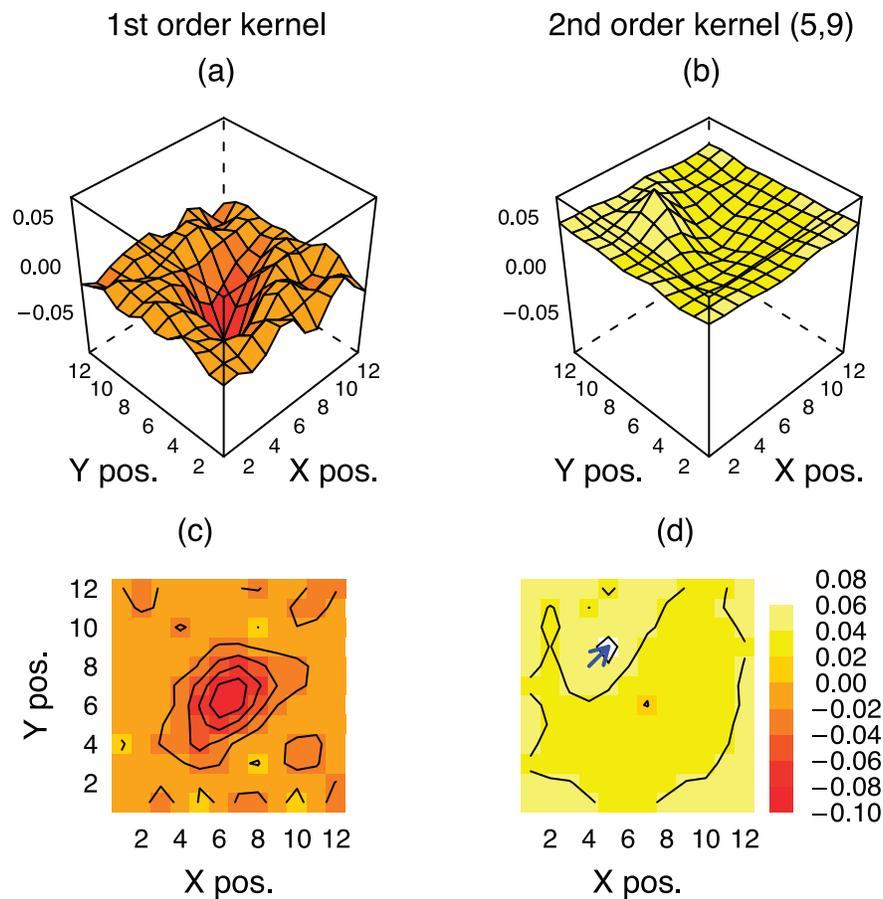


Figure 19. Cortical complex cell: first- and second-order kernels. (a, c) First order. (b, d) Second-order kernel slice with respect to Position (5, 9). (a, b) Perspective plot. (c, d) Contour plot. The blue arrow in (d) points to the reference position of the second-order kernel slice. Because contributions from the first-order Volterra term are negligible (Figure 18), as in the case of the simulated complex cell, the first order kernel is probably spurious. The second-order kernel slice indicates that positive correlations between a point of light at the reference position and a point of light in its close surrounding will induce the strongest pairwise response facilitations. In turn, positive correlations between a point of light at the reference position and a point of light in the central diagonal (along the antipreferred orientation) will induce weaker facilitations.

Discussion

Receptive fields are the main theoretical framework to represent functional properties of visual cells. In the most standard form, receptive fields tell how a cell responds when a point of light falls in a position of space (or time). Therefore, traditionally, one can think of receptive fields as spatial (or temporal) impulse responses; that is, we tend to think of receptive fields as linear filters. In addition, most stimuli used to obtain receptive fields are artificial, for example, dots of light, oriented bars, or sinusoidal gratings. Receptive fields obtained with such stimuli have been useful. However, visual cells are often nonlinear, and their functional properties may depend on the stimulus used to map them. In that case, we would like to map nonlinear receptive fields with natural images. How could receptive-field properties depend on the stimulus used to map them? One reason is that sensory systems adapt to the environ-

ment, changing their own functional properties (see Grzywacz & Balboa, 2002, and references therein). Another reason is that, as demonstrated by Proposition 2, by using a given stimuli set to characterize a cell, we can only observe the projection of the cell kernels onto the relevant space defined by the statistics of the stimulus set. In other words, the stimuli used to probe a cell constrain the functional properties that can be observed from its responses. We thus searched for a model that would allow us to map nonlinear receptive fields with natural stimuli. We explored the Volterra model, as it encompasses the receptive field as the first-order kernel and generalizes receptive fields to the nonlinear domain.

The large number of parameters in spatial Volterra models has hindered their application in the field of visual neuroscience. Fortunately, responses of visual cells appear to depend on a very low dimensional subspace of the image space (Volterra relevant-space technique section). Consequently, we could introduce the Volterra relevant-space

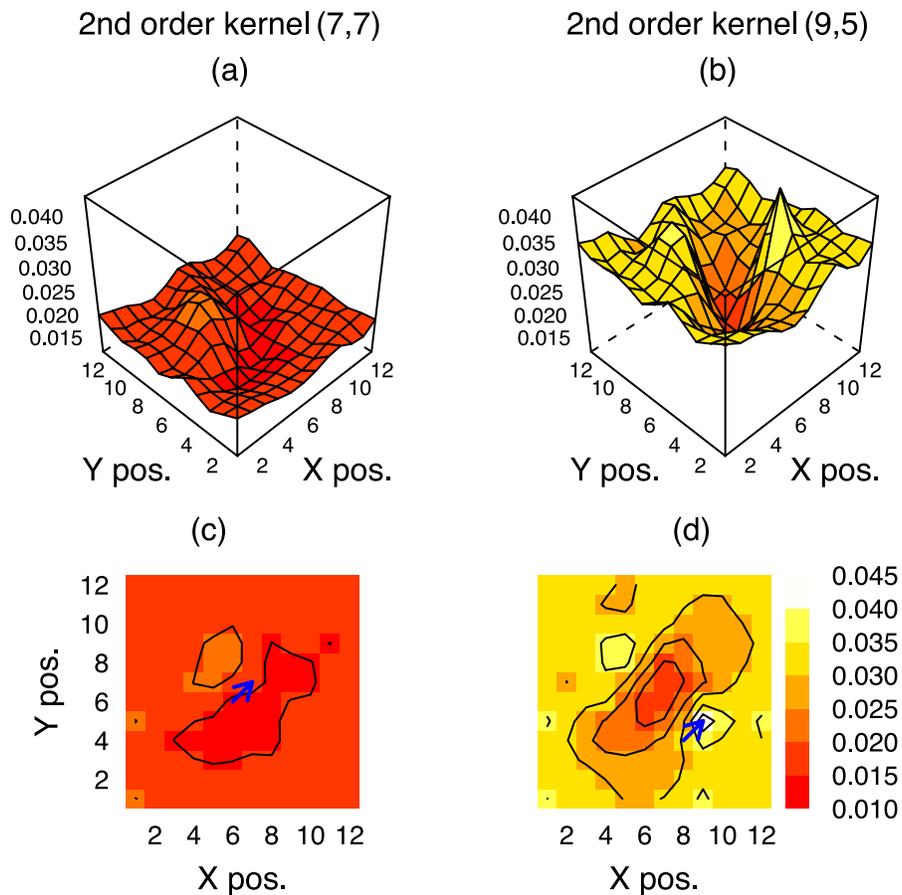


Figure 20. Cortical complex cell: second-order kernel slices estimated from PPR relevant dimensions. Slices correspond to different reference positions. (a, c) Position (7, 7); (b, d) Position (9, 5). Blue arrows in the contour plots indicate the reference positions. The slice with respect to Position (7, 7) (left column) is nearly flat and contains the smallest pairwise facilitations among all second-order slices. For the slice with respect to Position (9, 5) (right column), facilitation is strongest for interactions between the reference position and points in its close vicinity or points in the brightest region on the top-left quadrant. Facilitations are weaker for interactions between the reference position and points along the central red band.

technique for Volterra models. This representation allows the estimation of Volterra models from natural images.

To test the Volterra relevant-space technique, we used data from simulated cortical simple- and complex-cell models, for which we could analytically derive the true Volterra kernels. The complex-cell model, by being purely quadratic, can be exactly represented by a finite-order Volterra model. This is not the case for the simple-cell model, which has an infinite number of orders of nonlinearity. (However, when the projections of the input images onto the relevant dimensions are not too large, the Volterra model can be a good approximation for the simple-cell model. This would be true for natural images, which tend to have low contrasts; Balboa & Grzywacz, 2000; Ruderman & Bialek, 1994; Tadmor & Tolhurst, 2000; Zhu & Mumford, 1997.) Thus, the simple-cell model is a challenging test for the Volterra relevant-space technique. For both models, we showed that estimated kernels accurately approximate their true counterparts (Figures 6, 7, 13, and 14), illustrating that the Volterra relevant-space technique can accurately estimate Volterra models.

We also applied the Volterra relevant-space technique to physiological data from cortical complex cells. The estimated Volterra models yielded predictions of responses that were superior to those of a histogram-based model (Figure 22). This establishes this technique as a viable alternative for the estimation of nonlinear receptive fields with natural images.

An important step in the Volterra relevant-space technique is the estimation of relevant dimensions. We explored here three methods for this determination, namely, PPR, STA, and STC. For simulated data from cortical simple (Figure 8) and complex cells (Figure 15), we found that the relevant dimensions estimated by PPR better approximated the true relevant dimensions of the cell models. For physiological data from cortical complex cells, the relevant dimensions estimated by PPR yielded equal or better predictions than those estimated by STC (Figure 22). Furthermore, based on the simulated complex-cell data, PPR will probably succeed in estimating relevant dimensions for non-equal-variance responses, whereas STC will likely fail. In contrast, a disadvantage of PPR is its running time. PPR is

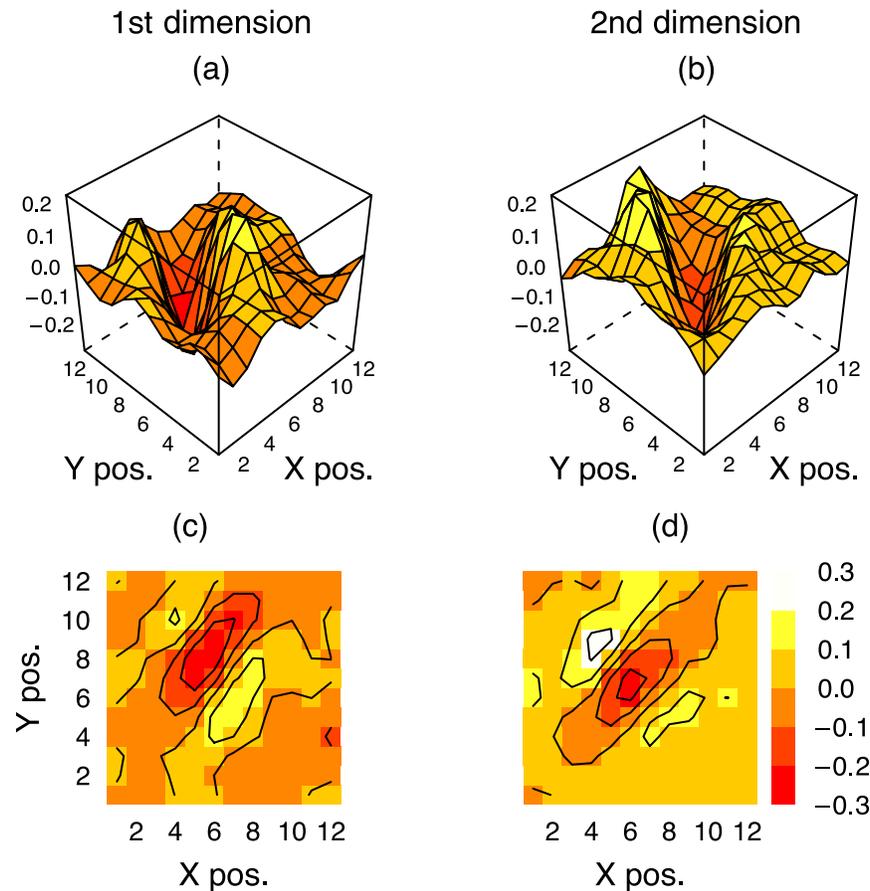


Figure 21. Cortical complex cell: STC relevant dimensions. (a, c) First relevant dimension. (b, d) Second relevant dimension. (a, b) Perspective plots. (c, d) Contour plots. These relevant dimensions show a cleaner structure than that of the PPR relevant dimensions in Figure 17. However, the best relevant dimensions are those that lead to better predictions (see text).

a time-consuming iterative algorithm, requiring several estimations of smooth functions and least squares minimizations. STA and STC are comparatively fast algorithms that only require matrix operations.

We also observed that the Volterra model yielded equal or superior predictions than histogram based models in all conditions studied (Figures 9, 16 and 22). A further advantage is that, after the relevant dimensions have been estimated, the Volterra relevant-space technique uses only linear operations. Thus, it is fast and easy to implement. In contrast, most implementations of histogram-based techniques require nonlinear optimizations that are slower and more subtle to implement.

From a prediction point of view, Volterra models are very accurate. However, a note of caution is pertinent regarding the interpretation of estimated Volterra kernels. Because Volterra models have a large number of parameters, physiological data are limited in size, and neurons are noisy; least squares fits to Volterra models can achieve similar predictive values with qualitatively different parameters. This could happen even after using different data sets to estimate a model and evaluate its predictive power. To minimize the effect of this problem, we have adopted several caution measures. First, we always use separate evaluation and test data

sets (Data partitioning section). Second, we divide our evaluation data set into jackknifed data sets (Efron & Tibshirani, 1993), estimate separate parameters (relevant dimensions and Volterra coefficients) for each jackknifed data set, and average our estimates. Third, we evaluate the relative contributions of Volterra terms (Figures 5, 12, and 18), and we do not try to interpret kernels of terms with low relative contributions (see justification in last paragraph of the True parameters section under the Simulated complex cell section). Fourth, whenever possible, we compare Volterra kernels estimated from different sets of relevant dimensions.

In this article, we have contrasted Volterra and histogram-based models. However, these two types of models are complementary, each describing the nonlinearity in a different way. For one- or two-dimensional relevant spaces, histogram-based models provide a succinct way of visualizing the nonlinearity. The nonlinearity would depend on the projection of the input images on one or two relevant dimensions. On the other hand, Volterra models can be estimated from much more than two dimensions and show more explicitly how individual pixels in the image contribute, linearly and nonlinearly, to generate responses. This is more in line with conventional models of receptive fields and thus may facilitate mechanistic interpretations for the

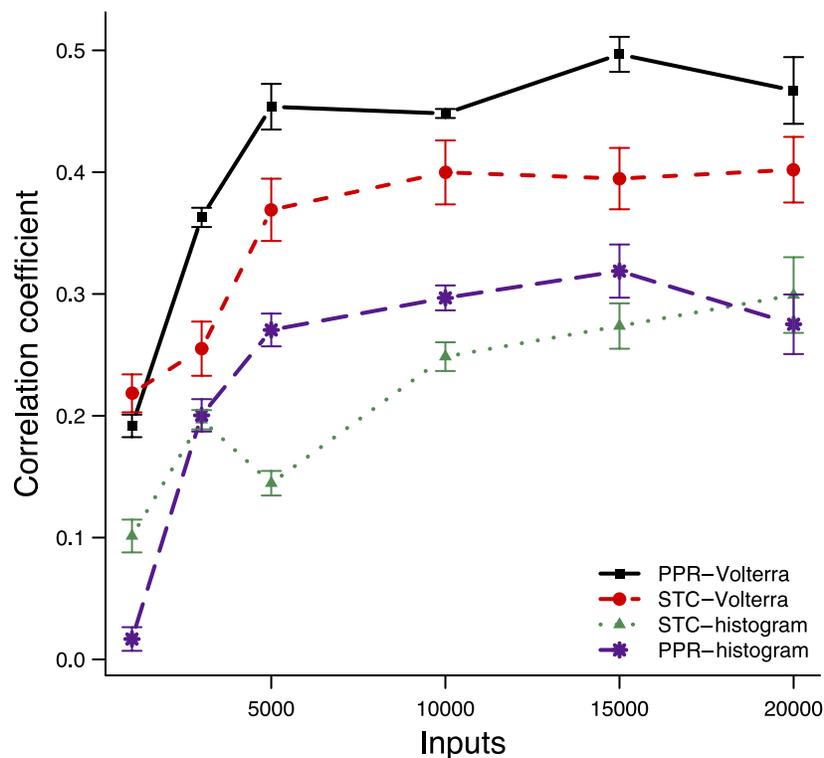


Figure 22. Cortical complex cell: predictions with two classes of models and two classes of estimated relevant dimensions. This figure shows average correlation coefficients between cell responses and model predictions. These coefficients appear as functions of the number of inputs. Black curve: Volterra model using the PPR relevant dimensions. Red curve: Volterra model using the STC relevant dimensions. Green curve: nonlinearity estimated from histograms using the STC relevant dimensions. Blue curve: nonlinearity estimated from histograms using the PPR relevant dimension. The size of the error bars is two standard errors. For any number of inputs, the Volterra models yield better predictions than the histogram-based model with both the PPR relevant dimensions (black vs. blue curves) and the STC relevant dimensions (red vs. green curves). For more than 1,000 inputs, the PPR relevant dimensions produced equal or better predictions than the relevant dimensions estimated by STC for both the Volterra (black vs. red curve) or histogram-based (blue vs. green curve) models.

cells. This mechanistic view of the nonlinearity could be of interest to many physiologists. For example, in the retina, starburst amacrine cells may produce a lateral nonlinear facilitation of the responses of certain ganglion cells. This facilitation would be mediated by acetylcholine (ACh; Amthor, Grzywacz, & Merwine, 1996; Tauchi & Masland, 1984). To validate this hypothesis, one could estimate Volterra models of the appropriate ganglion cells while blocking ACh and in control condition. The prediction would be that a nonlinear Volterra kernel would lose a central facilitatory zone while blocking ACh.

Relation to previous research

The Wiener series predated Volterra models for the characterization of nonlinear physiological systems. Wiener developed this series as an orthogonal representation of the Volterra series for white noise inputs (Wiener, 1958). Later, Lee and Schetzen (1965) introduced an efficient method (the cross-correlation technique) to estimate the Wiener

series. This technique allowed many investigators to use the Wiener series to study visual neurons (Emerson, Citron, Vaughn, & Klein, 1987; Livingstone & Conway, 2003; Marmarelis & Naka, 1972; Victor & Shapley, 1980). The Wiener series has several analytical and practical advantages over the Volterra series, and the cross-correlation technique greatly simplifies its estimation. However, the Wiener series has the following three major limitations for the characterization of visual cells from natural images: First, the orthogonality properties of the Wiener series and the cross-correlation technique apply only to Gaussian white noise inputs. Second, Wiener kernels lack the relatively simple interpretations of Volterra kernels. Third, the kernels estimated by cross-correlation have large estimation variance (Marmarelis, 2004).

These limitations motivated the development of other methods for the estimation of Volterra models. Their main problem is the exuberance of parameters; thus, dimensionality reduction techniques are needed to make the estimation of Volterra models feasible. The *kernels-expansion technique*, initially proposed by Wiener for the estimation of

Wiener kernels, has been the method of choice for the estimation of Volterra models (Amarocho & Brandstetter, 1971; Bose, 1956; Lee, 1964; Marmarelis, 1993; Watanabe & Stark, 1975). It reduces the number of parameters by compacting the representation of kernels and avoids the computation of cross-correlations, performing instead least squares fitting of the actual data to model parameters. The representation of Volterra kernels is compacted by expressing them as linear combinations of a few basis functions. One then uses the compact representation to build a low-dimensional Volterra model. Then, the kernels of the original Volterra model are reconstructed from the estimated parameters of the low-dimensional Volterra model.

The critical step in the dimensionality reduction of Volterra models for visual cells is the selection of basis functions. This is critical because they could bias the estimated kernels. For example, first-order Volterra kernels can only be linear combinations of the selected basis functions, as shown in Equation 15. Therefore, an improper selection of basis functions will impede a correct estimation of Volterra kernels. This error will be independent of the amount and quality of data used in the estimation procedure.

Marmarelis (1993) used Laguerre basis functions to analyze systems excited by one-dimensional, temporally modulated inputs. One rationale for selecting such basis functions is that discrete Laguerre basis functions are complete and orthonormal. Furthermore, they have an exponential time decay. Such decay approximates well the known linear kernels of many physiological systems with one-dimensional, temporally modulated inputs. However, it is not clear whether higher order Volterra kernels can be represented by similar basis functions as those for first-order Volterra kernels. Another issue is that, for every cell under study, the experimenter must select the number of Laguerre basis functions, that is, L in Equations 15 and 16, to represent the kernels. The experimenter must also select parameter α of the Laguerre basis functions. Marmarelis recommended a judicious selection of L and α based on trial and error. Later, Mitsis and Marmarelis (2002) proposed a method to learn parameter α from input–output data.

For systems excited by two-dimensional inputs, for example, visual cells, one-dimensional Laguerre basis functions cannot be used to represent their two-dimensional kernels. No known set of orthonormal two-dimensional basis functions fulfills the role that Laguerre functions have for systems with one-dimensional, temporally modulated inputs. Moreover, the number of parameters needed to characterize two-dimensional basis functions is larger than that required for one-dimensional basis functions. Tuning these parameters thus becomes an arduous process in the two-dimensional case. For systems excited by inputs varying in time and space, that is, 3D inputs, this problem is aggravated even further.

These difficulties in the selection of basis functions motivated our search for an alternate method to reduce the dimensionality of Volterra models. Instead of reducing the di-

dimensionality of the unknown kernels of the Volterra model, we thought that a better approach might be to reduce the dimensionality of the input images. This latter reduction leads to the Volterra relevant-space technique. The Volterra relevant-space technique has the following three advantages with respect to the kernels-expansion technique:

1. The effectiveness of the low-dimensional representation of Volterra models depends upon using a small set of basis functions. In the kernels-expansion technique, the assumption for effectiveness is that the kernels of the Volterra model can be represented using a small set of basis functions. This assumption cannot be tested experimentally a priori. In the Volterra relevant-space technique, the assumption is that, for a given cell, the relevant space is low-dimensional. As demonstrated in previous studies (de Ruyter van Steveninck & Bialek, 1988; Rust et al., 2005; Sharpee et al., 2006), this assumption has experimental support.
2. Optimization of Equation 8 allows the Volterra relevant-space technique to estimate the relevant dimensions from the data. Consequently, the Volterra relevant-space technique avoids the bias introduced in the estimates of kernels by an improper ad hoc selection of basis functions.
3. As described in the Volterra relevant-space technique section under the Receptive-field estimation section, the solution of Equation 8 is a well-studied problem in different fields of science. We can then use the expertise from these fields in the estimation of Volterra models.

Limitations of Volterra models

Although one can use Volterra models to study receptive fields, they have limitations in their applicability. We list the following three:

1. Convergence of the Volterra series requires certain smoothness conditions on the function being approximated (Palm & Poggio, 1977). This requirement is not severe because the firing rate of visual cells can be adequately approximated by smooth functions.
2. The response could depend on very high order Volterra terms. Consequently, even with the use of the Volterra relevant-space technique, we might not be able to estimate them from a limited amount of physiological data. However, because many cells in the early visual system have been adequately characterized with linear or quadratic models, this scenario does not seem to be a strong limitation. Furthermore, studies of statistics of natural images find that distributions of variables that modulate responses tend to have maxima at values that generate little response. For example, the distribution of contrasts tends to be maximal at zero (Balboa & Grzywacz, 2000; Ruderman & Bialek, 1994; Tadmor & Tolhurst, 2000; Zhu & Mumford, 1997). Therefore,

responses to natural images may be low, generating relatively little high-order nonlinearities.

3. A cell response may admit a Volterra representation, with all relevant terms of moderate order, but it may not depend on a low-dimensional subspace of the image space. In such a case, the Volterra relevant-space technique will not provide substantial dimensionality reduction to allow the estimation of high-order Volterra models. However, physiological experiments with some cells in the early visual system come to our aid (de Ruyter van Steveninck & Bialek, 1988; Rust et al., 2005; Sharpee et al., 2004), showing that the responses of these cells depend on such a low-dimensional subspace.

Extension to the temporal domain

Although we focus this article on spatial Volterra models, extending the Volterra relevant-space technique to include time is theoretically not a limitation. We just need to add the temporal dimension to the Volterra model (Equation 3), yielding

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= k_0(t) + \sum_{t=1}^{N_t} \sum_{j=1}^{N_s} k_1(i,j,t)x(i,j,t) + \dots \\
 &+ \sum_{t_1, \dots, t_Q=1}^{N_t} \sum_{i_1, j_1, \dots, i_Q, j_Q=1}^{N_s} k_Q(i_1, j_1, t_1, \dots, i_Q, j_Q, t_Q) \\
 &\times x(i_1, j_1, t_1) \dots x(i_Q, j_Q, t_Q) + \varepsilon.
 \end{aligned}
 \tag{26}$$

The Volterra relevant-space technique can also be extended to include the temporal dimension. However, the estimation of relevant dimensions in spatiotemporal models becomes a higher dimensional problem and, consequently, a more difficult one.

Appendix A: Proofs

Proposition 1

Given $y_{Q,k}(\mathbf{x}): \mathcal{H}^{N \times N} \rightarrow \mathcal{H}$, as in Equation 3. If a vector space \mathcal{S} , spanned by a set of basis functions $\{\mathbf{b}_1, \dots, \mathbf{b}_L: \mathbf{b}_l \in \mathbb{R}^{N \times N}\}$, fulfills Equation 8, then Equation 3 can be expressed as Equation 7, where q_0, \dots, q_Q are

$$q_0 = k_0 \tag{A1}$$

$$q_1(l) = \sum_{r,s=1}^N k_1(r,s)b_l(r,s) \tag{A2}$$

...

$$\begin{aligned}
 q_Q(l_1, \dots, l_Q) &= \sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_Q(r_1, s_1, \dots, r_Q, s_Q) \\
 &\times b_{l_1}(r_1, s_1) \dots b_{l_Q}(r_Q, s_Q).
 \end{aligned}
 \tag{A3}$$

Proof

From Equations 3, 5, and 8, observe that

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= y_{Q,k} \left(\sum_{l=1}^L a_l(\mathbf{x}) \mathbf{b}_l \right) \\
 &= k_0 + \sum_{r,s=1}^N k_1(r,s) \left(\sum_{l=1}^L a_l(\mathbf{x}) b_l(r,s) \right) + \dots \\
 &+ \sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_Q(r_1, s_1, \dots, r_Q, s_Q) \\
 &\times \left(\sum_{l_1, \dots, l_Q=1}^L a_{l_1}(\mathbf{x}) b_{l_1}(r_1, s_1) \dots a_{l_Q}(\mathbf{x}) b_{l_Q}(r_Q, s_Q) \right) + \varepsilon.
 \end{aligned}
 \tag{A4}$$

Switching the order of summations in Equation A4, we see that

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= k_0 + \sum_{l=1}^L \left(\sum_{r,s=1}^N k_1(r,s)b_l(r,s) \right) a_l(\mathbf{x}) + \dots \\
 &+ \sum_{l_1, \dots, l_Q=1}^L \left(\sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_Q(r_1, s_1, \dots, r_Q, s_Q) \right. \\
 &\left. \times b_{l_1}(r_1, s_1) \dots b_{l_Q}(r_Q, s_Q) \right) a_{l_1}(\mathbf{x}) \dots a_{l_Q}(\mathbf{x}) + \varepsilon.
 \end{aligned}
 \tag{A5}$$

Substituting Equations A1–A3 into Equation A5, we obtain Equation 7.

Proposition 2

Given $y_{Q,k}(\mathbf{x}): \mathcal{H}^{N \times N} \rightarrow \mathcal{H}$, as in Equation 3, assume that a vector space \mathcal{S} , spanned by a set of basis functions $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_L: \mathbf{b}_l \in \mathbb{R}^{N \times N}\}$, satisfies Equation 8. Consider the set projection coefficients of the kernels in the space \mathcal{S} given in Equations A1–A3 and the projected kernels in the space \mathcal{S} , k_i obtained by using these projection coefficients in Equations 14–16, then it is true that

$$y_{Q,k}(\mathbf{x}) = y_{Q,k}(\mathbf{x}) \forall \mathbf{x} \in X. \tag{A6}$$

Proof

From Equation 3,

$$\begin{aligned}
 y_{Q,k}(\mathbf{x}) &= k_0 \\
 &+ \sum_{i,j=1}^N k_1(i,j)x(i,j) + \dots \\
 &+ \sum_{i_1, j_1, \dots, i_Q, j_Q=1}^N k_Q(i_1, j_1, \dots, i_Q, j_Q) \\
 &\times x(i_1, j_1) \dots x(i_Q, j_Q) + \varepsilon.
 \end{aligned}
 \tag{A7}$$

Substituting Equations 14–16 into Equation A7, we see that

$$y_{Q,k}(\mathbf{x}) = q_0 + \sum_{i,j=1}^N \left(\sum_{l=1}^L q_1(l) b_l(i,j) \right) x(i,j) + \dots + \sum_{i_1, j_1, \dots, i_Q, j_Q=1}^N \left(\sum_{l_1, \dots, l_Q=1}^L q_Q(l_1, \dots, l_Q) b_{l_1}(i_1, j_1) \dots b_{l_Q}(i_Q, j_Q) \right) x(i_1, j_1) \dots x(i_Q, j_Q) + \varepsilon. \quad (\text{A8})$$

Substituting Equations A1–A3 into Equation A8, we then find that

$$y_{Q,k}(\mathbf{x}) = k_0 + \sum_{i,j=1}^N \left[\sum_{l=1}^L \left(\sum_{r,s=1}^N k_1(r,s) b_l(r,s) \right) b_l(i,j) \right] x(i,j) + \dots + \sum_{i_1, j_1, \dots, i_Q, j_Q=1}^N \left[\sum_{l_1, \dots, l_Q=1}^L \left(\sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_Q(r_1, s_1, \dots, r_Q, s_Q) b_{l_1}(r_1, s_1) \dots b_{l_Q}(r_Q, s_Q) \right) b_{l_1}(i_1, j_1) \dots b_{l_Q}(i_Q, j_Q) \right] x(i_1, j_1) \dots x(i_Q, j_Q) + \varepsilon. \quad (\text{A9})$$

Interchanging the order of summations, Equation A9 can be expressed as

$$y_{Q,k}(\mathbf{x}) = k_0 + \sum_{r,s=1}^N k_1(r,s) \left[\sum_{l=1}^L \left(\sum_{i,j=1}^N b_l(i,j) x(i,j) \right) b_l(r,s) \right] + \dots + \sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_Q(r_1, s_1, \dots, r_Q, s_Q) \times \left[\sum_{l_1=1}^L \left(\sum_{i_1, j_1=1}^N b_{l_1}(i_1, j_1) x(i_1, j_1) \right) b_{l_1}(r_1, s_1) \right] \dots \left[\sum_{l_Q=1}^L \left(\sum_{i_Q, j_Q=1}^N b_{l_Q}(i_Q, j_Q) x(i_Q, j_Q) \right) b_{l_Q}(r_Q, s_Q) \right] + \varepsilon. \quad (\text{A10})$$

Substituting Equation 6 into Equation A10 and using Equations 5 and 8, we see that

$$y_{Q,k}(\mathbf{x}) = k_0 + \sum_{r,s=1}^N k_1(r,s) \left(\sum_{l=1}^L \alpha_l(\mathbf{x}) b_l(r,s) \right) + \dots + \sum_{r_1, s_1, \dots, r_Q, s_Q=1}^N k_r(r_1, s_1, \dots, r_Q, s_Q) \left(\sum_{l_1=1}^L \alpha_{l_1}(\mathbf{x}) b_{l_1}(r_1, s_1) \right) \dots \left(\sum_{l_Q=1}^L \alpha_{l_Q}(\mathbf{x}) b_{l_Q}(r_Q, s_Q) \right) + \varepsilon = y_{Q,k} \left(\sum_{l=1}^L \alpha_l(\mathbf{x}) b_l(r,s) \right) = y_{Q,k}(\Pi_S(\mathbf{x})) = y_{Q,k}(\mathbf{x}). \quad (\text{A11})$$

Appendix B: True Volterra kernels

To simplify notation, we use one-dimensional vectors to represent input images, by concatenating the rows of their two-dimensional representation. Then, the Volterra representation of the response is given by

$$y_{Q,k}(\mathbf{x}) = k_0 + \sum_{i=1}^{N \times N} k_1(i) x(i) + \dots + \sum_{i_1, \dots, i_Q=1}^{N \times N} k_2(i_1, \dots, i_Q) x(i_1) \dots x(i_Q) + \varepsilon. \quad (\text{B1})$$

If the function describing the response model, $y_{Q,k}(\mathbf{x})$, is analytic on an open set $|\mathbf{x}| < R_0$ centered at zero with radius R_0 , then at each point \mathbf{x} in that set, $y_{Q,k}(\mathbf{x})$ has the Taylor-series representation (Brown & Churchill, 1996)

$$y_{Q,k}(\mathbf{x}) = y_{Q,k}(\mathbf{0}) + \sum_{i=1}^N \frac{\partial y_{Q,k}(\mathbf{0})}{\partial x_i} x(i) + \dots + \sum_{i_1, \dots, i_Q=1}^N \frac{1}{Q!} \frac{\partial^Q y_{Q,k}(\mathbf{0})}{\partial x_{i_1} \dots \partial x_{i_Q}} x(i_1) \dots x(i_Q) + \varepsilon. \quad (\text{B2})$$

By the uniqueness of Taylor series and comparing Equations B1 and B2, we derive the true Volterra kernels

$$k_0 = y_{Q,k}(\mathbf{0}) \\ k_1(i) = \frac{\partial y_{Q,k}(\mathbf{0})}{\partial x_i} \\ \dots \\ k_Q(i_1, \dots, i_Q) = \frac{1}{Q!} \frac{\partial^Q y_{Q,k}(\mathbf{0})}{\partial x_{i_1} \dots \partial x_{i_Q}}. \quad (\text{B3})$$

The functions describing the simple- and complex-cell models in Equations 17 and 18, respectively, are analytic. Hence, their analytic kernels can be calculated from Equation B3.

Appendix C: Projection pursuit regression

The PPR model represents the response of the system, y , to its input, $\mathbf{x} = (x_1, \dots, x_N)'$, using the nonlinear transformation given in Equation 19 (Friedman, 1984; Friedman & Stuetzle, 1981). The projection part of the name indicates that the input vector \mathbf{x} is projected onto the direction vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{M_0}$. In turn, the pursuit part indicates that an optimization technique is used to find a good set of direction vectors. The estimation of the PPR model parameters is a two-stage process. They comprise a forward stepwise procedure followed by a backward stepwise procedure. The PPR algorithm is controlled by two

parameters, M and M_0 . Parameter M controls the forward procedure, and parameter M_0 controls the backward procedure.

Forward stepwise procedure

For the forward procedure, an initial M -term model of the form given by Equation 19 is constructed. First, a trial direction \mathbf{a}_1 is used to compute the values $z_{i1} = \langle \mathbf{a}_1, \mathbf{x}_i \rangle$, $i = 1, \dots, n$, where $\mathbf{x}_i = (x_{i1}, \dots, x_{iN})'$ is the i th input and n is the number of inputs. Then, denoting the response of the system to input \mathbf{x}_i by y_i and $y_i^1 = y_i - \bar{y}$, a scatter plot of (y_i^1, z_{i1}) is constructed. This scatter plot is smoothed to obtain a first estimate ϕ_1 in Equation 19. Next, \mathbf{a}_1 is varied to minimize the sum of squares $\sum_{i=1}^n (y_i^1 - \phi_1(z_{i1}))^2$, where for each \mathbf{a}_1 in the optimization procedure, a new ϕ_1 is computed using the smoothing procedure. The result of this estimation, the pair (\mathbf{a}_1, ϕ_1) is then standardized according to Equation 19, and the corresponding value β_1 is calculated. We now have the approximation $y_i = \bar{y} + \beta_1 \phi_1(\langle \mathbf{a}_1, \mathbf{x}_i \rangle)$, $i = 1, \dots, n$. Next, we treat $y_i^2 = y_i - \bar{y} - \beta_1 \phi_1(z_{i1})$ as the response. We then fit a second term $\beta_2 \phi_2(z_{i2})$ to this modified response in the same manner that we fitted $\beta_1 \phi_1(z_{i1})$ to y_i^1 . This gives the approximation $y_i = \bar{y} + \beta_1 \phi_1(\langle \mathbf{a}_1, \mathbf{x}_i \rangle) + \beta_2 \phi_2(\langle \mathbf{a}_2, \mathbf{x}_i \rangle)$, $i = 1, \dots, n$. Continuing in this fashion, we arrive at the forward stepwise estimated model

$$y_i = \bar{y} + \sum_{m=1}^M \beta_m \phi_m(z_{im}) \quad i = 1, \dots, n. \quad (\text{C1})$$

Backward stepwise procedure

Next, models of decreasing order $m = M - 1, M - 2, \dots, M_0$ are fit in a backward stepwise manner. For each term in the model, the sum of squared residuals

$$\sum_{i=1}^n \left[y_i - \bar{y} - \sum_{l=1}^m \beta_l \phi_l(z_{il}) \right]^2 \quad (\text{C2})$$

is minimized with respect to β_l , \mathbf{a}_l , and ϕ_l . The initial values for these parameters are the solutions for the m most important values of the $m + 1$ terms in the previous order $m + 1$ model, where importance is measured by $|\beta_l|$.

Acknowledgments

We thank Drs. Yang Dan and Jon Touryan for their generosity in sharing their physiological data. We also thank the anonymous reviewers whose comments helped us enormously to improve the quality of our original submission. Thanks are also due to the following: Dr. Bartlett Mel and Mr. Bardia Behabadi for allowing us (and helping) to use their Linux cluster; the University of Southern California Center for High Performance Computing and Communi-

cations for the use of the Sun Computing Resource; Dr. Trevor Hastie for pointing us to Dr. Roosen's implementation of PPR; Mr. Xiwu Cao, Ms. Susmita Chatterjee, Dr. Eun Jin Lee, Dr. David Merwine, Dr. Mónica Padilla, and Mr. Jeff Wurfel for helpful discussions in our laboratory; Dr. Bosco Tjan for insightful and specially enjoyable interactions; Dr. Dong Song for earlier comments on the manuscript; and Mrs. Denise Steiner for administrative support. This work was supported by National Eye Institute Grants EY08921 and EY11170, and by National Science Foundation Grant EEC-0310723 to N.M.G.

Commercial relationships: none.

Corresponding author: Joaquín Rapela.

Email: rapela@usc.edu.

Address: 3641 Watt Way, University of Southern California, Los Angeles, CA 90089-90025.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, Optics and Image Science*, 2, 284–299. [PubMed]
- Aggelopoulos, N. C., Franco, L., & Rolls, E. T. (2005). Object perception in natural scenes: Encoding by inferior temporal cortex simultaneously recorded neurons. *Journal of Neurophysiology*, 93, 1342–1357. [PubMed] [Article]
- Albrecht, D. G., & De Valois, R. L. (1981). Striate cortex responses to periodic patterns with and without the fundamental harmonics. *The Journal of Physiology*, 319, 497–514. [PubMed] [Article]
- Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology*, 48, 217–237. [PubMed]
- Amorocho, J., & Brandstetter, A. (1971). Determination of nonlinear functional response functions in rainfall-runoff processes. *Water Resources Research*, 7, 1087–1101.
- Amthor, F. R., Grzywacz, N. M., & Merwine, D. K. (1996). Extra-receptive-field motion facilitation in on-off directionally selective ganglion cells of the rabbit retina. *Visual Neuroscience*, 13, 303–309. [PubMed]
- Balboa, R. M., & Grzywacz, N. M. (2000). The role of early lateral inhibition: More than maximizing luminance information. *Visual Neuroscience*, 17, 77–89. [PubMed]
- Barlow, H. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.
- Ben-Israel, A., & Greville, T. (1980). *Generalized inverses: Theory and applications*. Huntington, NY: Krieger.

- Bialek, W., & de Ruyter van Steveninck, R. R. (2005). *Features and dimensions: Motion estimation in fly vision*. Retrieved from <http://arxiv.org/abs/q-bio/0505003>
- Bonds, A. B. (1989). Role inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, *2*, 41–55. [[PubMed](#)]
- Bose, A. (1956). *A theory of nonlinear systems* (Tech. Rep. No. 309). Cambridge, MA: Research Laboratory of Electronics, MIT.
- Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, *26*, 695–702. [[PubMed](#)] [[Article](#)]
- Brown, J., & Churchill, R. (1996). *Complex variables and applications* (6th ed.). New York: McGraw-Hill, Inc.
- Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, *264*, 1333–1336. [[PubMed](#)]
- Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses. *Network*, *12*, 199–213. [[PubMed](#)]
- Dan, Y., Attick, J. J., & Reid, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *The Journal of Neuroscience*, *16*, 3351–3362. [[PubMed](#)] [[Article](#)]
- David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *The Journal of Neuroscience*, *24*, 6991–7006. [[PubMed](#)] [[Article](#)]
- de Boer, R., & Kuyper, P. (1968). Triggered correlation. *IEEE Transactions on Bio-medical Engineering*, *15*, 169–179. [[PubMed](#)]
- Dean, A. F. (1981). The relationship between response amplitude and contrast for cat striate cortical neurons. *The Journal of Physiology*, *318*, 413–427. [[PubMed](#)] [[Article](#)]
- Dean, A. F., & Tolhurst, D. J. (1986). Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex. *Experimental Brain Research*, *62*, 143–151. [[PubMed](#)]
- de Ruyter van Steveninck, R., & Bialek, W. (1988). Real-time performance of a movement-sensitive neuron in the blowfly visual system: Coding and information transfer in short spike sequences. *Proceedings of the Royal Society of London: Series B*, *234*, 379–414.
- Diaconis, P., & Shahshahani, M. (1984). On nonlinear functions of linear combinations. *SIAM Journal on Scientific Computing*, *5*, 175–191.
- Efron, B., & Tibshirani, R. (1993). *An introduction to bootstrap*. New York, NY: Chapman & Hall.
- Emerson, R. C., Citron, M. C., Vaughn, W. J., & Klein, S. A. (1987). Nonlinear directionally selective subunits in complex cells of a cat striate cortex. *Journal of Neurophysiology*, *58*, 33–65. [[PubMed](#)]
- Felsen, G., Touryan, J., Han, F., & Dan, Y. (2005). Cortical sensitivity to visual features in natural scenes. *Public Library of Science Biology*, *3*, e342. [[PubMed](#)] [[Article](#)]
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science*, *4*, 2379–2394. [[PubMed](#)]
- Friedman, J. (1984). *Smart user's guide* (Tech. Rep.). Stanford University, Department of Statistics.
- Friedman, J., & Stuetzle, W. (1981). Projection pursuit regression. *Journal of the American Statistical Association*, *76*, 817–823.
- Geladi, P., & Kowalski, B. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, *185*, 1–17.
- Grzywacz, N. M., & Balboa, R. M. (2002). A Bayesian framework for sensory adaptation. *Neural Computation*, *14*, 543–559. [[PubMed](#)]
- Guo, K., Robertson, R. G., Mahmoodi, S., & Young, M. P. (2005). Centre-surround interactions in response to natural scene stimulation in the primary visual cortex. *The European Journal of Neuroscience*, *21*, 536–548. [[PubMed](#)]
- Hansen, P. (1987). The truncated svd as a method for regularization. *BIT*, *27*, 534–553.
- Helland, I. (1988). On the structure of partial least squares regression. *Communications in Statistics: Simulation and Computation*, *17*, 584–607.
- Hertz, J., Palmer, R., & Krogh, A. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison-Wesley.
- Jones, J. P., Stepnoski, A., & Palmer, L. A. (1987). The two-dimensional spectral structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*, 1212–1232. [[PubMed](#)]
- Lee, Y. (1964). Contributions of Norbert Wiener to linear theory and nonlinear theory in engineering. In N. Wiener & Y. M. Lee (Eds.), *Selected papers of Norbert Wiener* (pp. 17–33). Cambridge, MA: MIT Press.
- Lee, Y., & Schetzen, M. (1965). Measurement of the wiener kernels of a nonlinear system by cross-correlation. *International Journal of Control*, *2*, 237–254.
- Li, K. (1991). Sliced inverse regression for dimension reduction (with discussion). *Journal of the American Statistical Association*, *86*, 316–342.
- Livingstone, M. S., & Conway, B. R. (2003). Substructure of direction-selective receptive fields in macaque V1. *Journal of Neurophysiology*, *89*, 2743–2759. [[PubMed](#)] [[Article](#)]

- Maffei, L., & Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser. *Vision Research*, *13*, 1255–1267. [PubMed]
- Marmarelis, P. Z., & Marmarelis, V. Z. (1978). *Analysis of physiological systems: The white-noise approach*. New York, NY: Plenum Press.
- Marmarelis, P. Z., & Naka, K. (1972). White-noise analysis of a neuron chain: An application of the Wiener theory. *Science*, *175*, 1276–1278. [PubMed]
- Marmarelis, V. Z. (1993). Identification of nonlinear biological systems using Laguerre expansions of kernels. *Annals of Biomedical Engineering*, *21*, 573–589. [PubMed]
- Marmarelis, V. Z. (2004). *Nonlinear dynamic modeling of physiological systems*. Hoboken, NJ: Wiley.
- Mendel, J. (1995). *Lessons in estimation theory for signal processing, communications, and control*. Upper Saddle River, NJ: Prentice Hall.
- Mitsis, G. D., & Marmarelis, V. Z. (2002). Modeling of nonlinear physiological systems with fast and slow dynamics: I. Methodology. *Annals of Biomedical Engineering*, *30*, 272–281. [PubMed]
- Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells: I. Neurophysiological evidence. *Proceedings of the Royal Society of London B*, *216*, 335–354. [PubMed]
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *The Journal of Physiology*, *283*, 53–77. [PubMed] [Article]
- Nykamp, D. Q., & Ringach, D. L. (2002). Full identification of a linear–nonlinear system via cross-correlation analysis. *Journal of Vision*, *2*(1), 1–11, <http://journalofvision/2/1/1/>, doi:10.1167/2.1.1. [PubMed] [Article]
- Palm, G., & Poggio, T. (1977). The volterra representation and the wiener expansion: Validity and pitfalls. *SIAM Journal on Applied Mathematics*, *33*, 195–216.
- Prenger, R., Wu, M. C., David, S. V., & Gallant, J. L. (2004). Nonlinear V1 responses to natural scenes revealed by neural network analysis. *Neural Networks*, *17*, 663–679. [PubMed]
- Ringach, D. L., Hawken, M. J., & Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision*, *2*(1), 12–24, <http://journalofvision/2/1/2/>, doi:10.1167/2.1.2. [PubMed] [Article]
- Roosen, C. (1994). *Implementation of Roosen and Hastie, 1994*. Retrieved from <http://lib.stat.cmu.edu/jcgs/roosen-hastie>
- Roosen, C., & Hastie, T. (1994). Automatic smoothing spline projection pursuit. *Journal of Computational and Graphical Statistics*, *3*, 235–248.
- Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters*, *73*, 814–817. [PubMed]
- Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, *46*, 945–956. [PubMed]
- Scharf, L. (1991). The SVD and reduced-rank signal processing. In R. J. Vaccaro (Ed.), *SVD and signal processing, II: Algorithms, analysis and applications* (pp. 3–31). New York, NY: Elsevier.
- Sclar, G., & Freeman, R. D. (1982). Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Experimental Brain Research*, *46*, 457–461. [PubMed]
- Sharpee, T., Rust, N. C., & Bialek, W. (2004). Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Computation*, *16*, 223–250. [PubMed]
- Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., & Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, *439*, 936–942. [PubMed]
- Simoncelli, E., Paninski, L., Pillow, J., & Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (3rd ed., pp. 327–338). Cambridge, MA: MIT Press.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216. [PubMed]
- Smyth, D., Willmore, B., Baker, G. E., Thompson, I. D., & Tolhurst, D. J. (2003). The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *The Journal of Neuroscience*, *23*, 4746–4759. [PubMed] [Article]
- Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London: Series B*, *216*, 427–459. [PubMed]
- Sturges, H. (1926). The choice of a class-interval. *Journal of the American Statistical Association*, *21*, 65–66.
- Tadmor, Y., & Tolhurst, D. J. (2000). Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. *Vision Research*, *40*, 3145–3147. [PubMed]
- Tauchi, M., & Masland, R. H. (1984). The shape and arrangement of the cholinergic neurons in the rabbit retina. *Proceedings of the Royal Society of London: Series B*, *223*, 101–119. [PubMed]

- Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E., & Gallant, J. L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network*, *12*, 289–316. [[PubMed](#)]
- Touryan, J., Felsen, G., & Dan, Y. (2005). Spatial structure of complex cell receptive fields measured with natural images. *Neuron*, *45*, 781–791. [[PubMed](#)] [[Article](#)]
- Touryan, J., Lau, B., & Dan, Y. (2002). Isolation of relevant visual features from random stimuli for cortical complex cells. *The Journal of Neuroscience*, *22*, 10811–10818. [[PubMed](#)] [[Article](#)]
- van Hateren, J. H., & Ruderman, D. L. (1998). Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society of London B*, *265*, 2315–2320. [[PubMed](#)]
- van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings: Biological Sciences/The Royal Society*, *265*, 359–366. Retrieved from <http://hlab.phys.rug.nl/imlib/index.html> [[PubMed](#)]
- Venables, W., & Smith, D. (2002). *An Introduction to R*. Network Theory Ltd.
- Victor, J., & Shapley, R. (1980). A method of nonlinear analysis in the frequency domain. *Biophysical Journal*, *29*, 459–483. [[PubMed](#)] [[Article](#)]
- Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287*, 1273–1276. [[PubMed](#)]
- Vinje, W. E., & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission in V1. *The Journal of Neuroscience*, *22*, 2904–2915. [[PubMed](#)] [[Article](#)]
- Volterra, V. (1930). *Theory of functionals and of integral and integro-differential equations*. New York: Dover Publications.
- Watanabe, A., & Stark, L. (1975). Kernel method for nonlinear analysis: Identification of biological control system. *Mathematical Biosciences*, *27*, 99–108.
- Weliky, M., Fiser, J., Hunt, R. H., & Wagner, D. N. (2003). Coding of natural scenes in primary visual cortex. *Neuron*, *37*, 703–718. [[PubMed](#)] [[Article](#)]
- Wiener, N. (1958). *Nonlinear problems in random theory*. New York: Wiley.
- Willmore, B., & Smyth, D. (2003). Methods for first-order kernel estimation: Simple-cell receptive fields from responses to natural scenes. *Network*, *14*, 553–577. [[PubMed](#)]
- Zhu, S., & Mumford, D. (1997). Prior learning and gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*, 1236–1250.