

2005-01

# A Theoretical Analysis of the Influence of Fixational Instability on the Development of Thalamocortical Connectivity

---

<https://hdl.handle.net/2144/2370>

*"Downloaded from OpenBU. Boston University's institutional repository."*

**A theoretical analysis of the influence of fixational  
instability on the development of  
thalamocortical connectivity**

**Antonino Casile and Michele Rucci**

**January, 2005**

**Technical Report CAS/CNS-2005-001**

Permission to copy without fee all or part of this material is granted provided that: 1. The copies are not made or distributed for direct commercial advantage; 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and / or special permission.

Copyright © 2005

Boston University Center for Adaptive Systems  
and  
Department of Cognitive and Neural Systems  
677 Beacon Street  
Boston, MA 02215

# A theoretical analysis of the influence of fixational instability on the development of thalamocortical connectivity

Antonino Casile and Michele Rucci

## Abstract

Under natural viewing conditions, the physiological instability of visual fixation keeps the projection of the stimulus on the retina in constant motion. After eye opening, chronic exposure to a constantly moving retinal image might influence the experience-dependent refinement of cell response characteristics. The results of previous modeling studies have suggested a contribution of fixational instability in the Hebbian maturation of the receptive fields of V1 simple cells (Rucci, Edelman, & Wray, 2000; Rucci & Casile, 2004). This paper presents a mathematical explanation of our previous computational results. Using quasi-linear models of LGN units and V1 simple cells, we derive analytical expressions for the second-order statistics of thalamocortical activity before and after eye opening. We show that, in the presence of natural stimulation, fixational instability introduces a spatially uncorrelated signal in the retinal input, which strongly influences the structure of correlated activity in the model.

## 1 Introduction

In the primary visual cortex (V1), distinct regions in the receptive fields of simple cells tend to receive afferents from either ON- or OFF-center neurons in the Lateral Geniculate Nucleus (LGN) (Hubel & Wiesel, 1962; Reid & Alonso, 1995; Ferster, Chung, & Wheat, 1996). It is a long-standing proposal that this pattern of connectivity originates from a Hebbian stabilization of synchronously firing geniculate afferents onto common post-synaptic targets, which is initially driven by endogenous spontaneous activity and later refined by visual experience (Stent, 1973; Changeux & Danchin, 1976; Miller, Erwin, & Kayser, 1999). This hypothesis is challenged by the substantially different structures of endogenous spontaneous activity and visually evoked responses. Already at the level of the retina, spontaneous activity appears to be correlated on a narrow spatial scale (Mastronarde, 1983), whereas external visual stimulation is known to be

characterized by broad spatial correlations (Field, 1987; Burton & Moorhead, 1987; Ruderman, 1994). This difference raises the question as to how the same activity-dependent mechanism of synaptic plasticity could account for both the initial emergence and the later refinement of V1 receptive fields.

A possible solution to this problem is represented by the fact that, after eye opening, the statistics of neural activity depend not only on the visual scene but also on the observer's behavior during the acquisition of visual information. Eye movements, in particular, with their direct impact on the sampling of visual information, may profoundly influence neural responses. Eye movements are a constant presence during natural viewing. In addition to saccades that relocate the direction of gaze every few hundred milliseconds, small fixational eye movements keep the eyes in constant motion even during the periods of fixation (Yarbus, 1967; Ratliff & Riggs, 1950). Recent neurophysiological studies have shown that fixational eye movements strongly affect the responses of geniculate (Martinez-Conde, Macknik, & Hubel, 2002) and cortical neurons (Gur, Beylin, & Snodderly, 1997; Martinez-Conde, Macknik, & Hubel, 2000; Leopold & Logothetis, 1998; Snodderly, Kagan, & Gur, 2001). Furthermore, experiments with kittens in which eye movements were prevented during the critical period have reported serious impairments in the maturation of characteristics of V1 neurons, such as orientation selectivity (Buisseret, Gary-Bobo, & Imbert, 1978; Gary-Bobo, Milleret, & Buisseret, 1986) and ocular dominance (Fiorentini, Maffei, & Bisti, 1979; Freeman & Bonds, 1979; Singer & Rauschecker, 1982).

In previous studies, we simulated the responses of LGN and V1 neurons to analyze the second-order statistics of thalamic (Rucci et al., 2000) and thalamo-cortical activity (Rucci & Casile, 2004) before and after eye opening. In these simulations, patterns of correlated activity were found to be consistent with a Hebbian maturation of simple cell receptive fields both in the presence of spontaneous activity and when images of natural scenes were scanned by eye movements, but not when the same images were examined in the absence of the retinal image motion produced by fixational instability. These results were highly robust. They were little affected by the precise characteristics of neuronal models and simulated oculomotor activity.

In this paper, to better understand the possible influence of fixational instability on visual development, we used quasi-linear models of LGN and V1 units to derive analytical expressions of the patterns of correlated activity. We show that the similarity between the statistics of thalamocortical activity present in our model before and after eye opening originates from a decorrelation of the retinal input operated by fixational instability.

## 2 Measuring correlated activity in a model of the LGN and V1

By definition, the strength of a Hebbian synapse is proportional to the level of correlation between the responses of pre- and post-synaptic elements. In this study, instead of explicitly modeling synaptic changes, we analyzed the compatibility of Hebbian synaptic plasticity and the spatial organization of simple cell receptive fields by directly examining the second-order statistics of neural activity in a model of the cat LGN and V1.

To determine whether a simple cell  $\eta$  would establish stronger connections with afferents from either ON- or OFF-center geniculate units at a location  $\mathbf{x}$  within its receptive field, we evaluated the correlation difference map:

$$R_\eta(\mathbf{x}) = \langle \eta_{\mathbf{x}_\eta}(t) [\alpha_{\mathbf{x}_\alpha}^{\text{ON}}(t) - \alpha_{\mathbf{x}_\alpha}^{\text{OFF}}(t)] \rangle_{\mathcal{I}, t}$$

where  $\eta(t)$  is the activity of the cortical neuron,  $\alpha^{\text{ON}}(t)$  and  $\alpha^{\text{OFF}}(t)$  are the responses of the ON- and OFF-center geniculate cells,  $\mathbf{x}_\eta$  and  $\mathbf{x}_\alpha$  are the positions of receptive field centers ( $\mathbf{x} = \mathbf{x}_\eta - \mathbf{x}_\alpha$ ), and the average is evaluated over time  $t$  and an ensemble of input images  $\mathcal{I}$ .

A positive value of  $R_\eta(\mathbf{x})$  implies that the simple cell response is more strongly correlated with the response of an ON-center (rather than an OFF-center) geniculate unit with receptive field at relative separation  $\mathbf{x}$ . The opposite holds for a negative value of  $R_\eta(\mathbf{x})$ .  $R_\eta$  can be seen as  $\eta$ 's receptive field predicted by Hebbian synapses. To preserve and refine the spatial organization of  $\eta$ 's receptive field,  $R_\eta$  needs to be positive at locations that correspond to ON subregions and negative in correspondence of the OFF subregions.

### 2.1 Modeling cell responses

We modeled the responses of simple cells in V1 and non-lagged ON- and OFF-center X cells in the LGN. The responses of both cortical and geniculate neurons were evaluated on the basis of the convolution between the visual input  $I(\mathbf{x}, t)$  and the cell spatiotemporal kernel  $k(\mathbf{x}, t)$ . For both LGN and V1 units we assumed a space-time separable kernel  $k(\mathbf{x}, t) = s(\mathbf{x})h(t)$ , where  $s(\mathbf{x})$  and  $h(t)$  represent the spatial and temporal components. Cell responses were obtained by rectifying the convolution output using a threshold  $\Theta$ . That is,  $\alpha(t) = k_\alpha(\mathbf{x}, t) \star I(\mathbf{x}, t) - \Theta$  if  $k_\alpha(\mathbf{x}, t) \star I(\mathbf{x}, t) > \Theta$ ,  $\alpha(t) = 0$  otherwise.

Spatial receptive fields of simple cells were modeled by means of Gabor filters:

$$s_\eta(\mathbf{x}) = A_\eta \cos([\omega_\eta \ 0] \cdot \mathbf{x} + \phi) e^{-\frac{\mathbf{x}^T \mathbf{R}(\rho)^T \sigma_\eta \mathbf{R}(\rho) \mathbf{x}}{2}}$$

where  $A_\eta$  is the amplitude,  $\sigma_\eta = \begin{pmatrix} \sigma_{\eta_x}^{-2} & 0 \\ 0 & \sigma_{\eta_y}^{-2} \end{pmatrix}$  is the covariance matrix of the Gaussian,  $\omega_\eta$  and  $\phi$  are the angular velocity and phase of the plane wave, and  $\mathbf{R}$  is a rotation matrix

that introduces the angle  $\rho$  between the major axis of the Gaussian and the plane wave. Parameters were adjusted to model 10 simple cells following neurophysiological data from Jones and Palmer (1987a, see their Table 1).

Spatial kernels of a geniculate units were modeled as differences of Gaussians:

$$s_\alpha(\mathbf{x}) = A_{\text{cnt}} e^{-\frac{\mathbf{x}^T \mathbf{x}}{2\sigma_{\text{cnt}}^2}} - A_{\text{srn}} e^{-\frac{\mathbf{x}^T \mathbf{x}}{2\sigma_{\text{srn}}^2}}$$

where the subscripts indicate contributions from the receptive field center (*cnt*) and surround (*srn*). Kernel parameters followed neurophysiological data from Linsenmeier, Frishman, Jakiela, and Enroth-Cugell (1982) to model ON-center cells with receptive fields located between  $5^\circ$  and  $15^\circ$  of visual eccentricity. At each angle of visual eccentricity, spatial receptive fields of modeled OFF-center cells were equal in magnitude and opposite in sign to those of ON-center units, *i.e.*  $s_\alpha^{\text{ON}} = -s_\alpha^{\text{OFF}}$ .

Since many neurons in the LGN and V1 possess similar temporal dynamics (Alonso, Usrey, & Reid, 2001), for both cortical and geniculate units, the temporal element  $h(t)$  was modeled as a difference of two gamma functions (DeAngelis, Ohzawa, & Freeman, 1993a; Cai, DeAngelis, & Freeman, 1997):

$$h_\alpha(t) = h_\eta(t) = k_1 \Gamma(t, t_1, c_1, n_1) - k_2 \Gamma(t, t_2, c_2, n_2)$$

where  $\Gamma(t, t_0, c, n) = \frac{[c(t-t_0)]^n e^{-c(t-t_0)}}{n^n e^{-n}}$ . Following data from Cai et al. (1997), temporal parameters were:  $t_1 = t_2 = 0$ ,  $n_1 = n_2 = 2$ ,  $k_1 = 1$ ,  $k_2 = 0.6$ ,  $c_1 = 60\text{s}^{-1}$ ,  $c_2 = 40\text{s}^{-1}$ .

Previous studies in which cell responses were simulated during free-viewing of natural images have shown that the second-order statistics of thalamocortical activity produced by this model are insensitive to the level of rectification (Rucci et al., 2000; Rucci & Casile, 2004). To probe into the origins of our previous simulation results, in this study we focused on the specific case of no rectification for simple cells and rectification with zero threshold for geniculate units. This assumption enables correlation difference maps to be expressed as the product of linear geniculate and cortical kernels:

$$R_\eta(\mathbf{x}) = \langle \eta_{\mathbf{x}_\eta}(t) [\alpha_{\mathbf{x}_\alpha}^{\text{ON}}(t) - \alpha_{\mathbf{x}_\alpha}^{\text{OFF}}(t)] \rangle_{\mathcal{I}, t} = \langle \eta_{\mathbf{x}_\eta}(t) \alpha_{\mathbf{x}_\alpha}(t) \rangle_{\mathcal{I}, t}$$

where  $\alpha(t) = \alpha^{\text{ON}}(t) - \alpha^{\text{OFF}}(t) = k_\alpha^{\text{ON}}(\mathbf{x}, t) \star I(\mathbf{x}, t)$ . While this choice of rectification parameters simplified the mathematical analysis of this paper, our previous simulation data ensure that results remain valid for a wide range of thresholds.

### 3 Thalamocortical activity before eye opening

To establish a reference baseline, we first examined the structure of thalamocortical activity immediately before eye opening. Experimental evidence indicates that many of the response features of V1 cells are already present at the time of eye opening (Hubel & Wiesel, 1963; Blakemore & van Sluyters, 1975). Computational studies have shown

that correlation-based mechanisms of synaptic plasticity are compatible with the emergence of simple cell receptive fields in the presence of endogenous spontaneous activity (Linsker, 1986; Miller, 1994; Miyashita & Tanaka, 1992).

For simplicity, we restricted our analysis to the two-dimensional case of one spatial and one temporal dimension, by considering sections of the spatial receptive fields. The receptive fields of simple cells were sectioned along the axis orthogonal to the cell preferred orientation. For LGN cells, we considered a section along a generic axis crossing the center of the receptive field. Results are, however, general and can be directly extended to the full 3D space-time case.

In the presence of spontaneous retinal activity, levels of correlation between the responses of thalamic and cortical units can be estimated by means of linear system theory:

$$R_\eta(x) = \mathcal{F}^{-1}\{C_{SA}(\omega_x, \omega_t)K_\eta(\omega_x, \omega_t)\overline{K_\alpha(\omega_x, \omega_t)}\}|_{t=0} \quad (1)$$

where  $C_{SA}$  is the power spectrum of spontaneous activity in the retina,  $\mathcal{F}^{-1}$  indicates the operation of inverse Fourier transform, and  $K_\alpha(\omega_x, \omega_t)$  and  $K_\eta(\omega_x, \omega_t)$  are the Fourier transforms of LGN and V1 kernels.

Under the model assumption of space-time separability of cell kernels, Eq. 1 gives:

$$R_\eta(x) = T\mathcal{F}^{-1}\{S_{SA}(\omega_x)S_\eta(\omega_x)\overline{S_\alpha(\omega_x)}\} \quad (2)$$

where we also assumed space-time separability of the power spectrum of spontaneous retinal activity.  $T$  is a multiplicative factor equal to  $\int_{-\infty}^{\infty} H_{SA}(\omega_t)H_\eta(\omega_t)\overline{H_\alpha(\omega_t)}d\omega_t$ , and  $S_{SA}(\omega_x)$ ,  $H_{SA}(\omega_t)$ ,  $S_\alpha(\omega_x)$ ,  $H_\alpha(\omega_t)$ ,  $S_\eta(\omega_x)$  and  $H_\eta(\omega_t)$  are, respectively, the spatial and temporal components of the power spectrum of spontaneous retinal activity and of the Fourier transforms of LGN and V1 kernels.

Data from Mastrorarde (1983) show that retinal spontaneous activity is characterized by narrow spatial correlations. These data are accurately interpolated by Gaussian functions. Least squares interpolations of levels of correlation between ganglion cells at different separations produced Gaussians with amplitude  $A_{SA} = 13.9$  independent of the cell eccentricity and standard deviation  $\sigma_{SA}$  that ranged from  $0.18^\circ$  at eccentricity  $5^\circ$  to  $0.35^\circ$  at  $25^\circ$ . Use in Eq. 2 of a Gaussian approximation for retinal spontaneous activity gives, after some algebraic manipulations, an analytical expression for the structure of correlated activity:

$$R_\eta(x) \propto \widehat{A}e^{-\frac{x^2}{2\widehat{\sigma}^2}} \cos(\widehat{\omega}x + \phi) + \widetilde{A}e^{-\frac{x^2}{2\widetilde{\sigma}^2}} \cos(\widetilde{\omega}x + \phi) = \widehat{R}_{\eta\alpha}(x) + \widetilde{R}_{\eta\alpha}(x) \quad (3)$$

where the parameters are given by:

$$\left\{ \begin{array}{l} \widehat{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{\text{cnt}}^2 + \sigma_{SA}^2} \\ \widehat{\omega} = \frac{\sigma_\eta^2}{\widehat{\sigma}^2}\omega_\eta \\ \widehat{A} = \frac{A_\eta A_{\text{cnt}} A_{SA} \sigma_{\text{cnt}} \sigma_\eta \sigma_{SA} 2\pi}{\widehat{\sigma}} e^{-\frac{\sigma_\eta^2 \omega_\eta (\widehat{\omega} - \omega_\eta)}{2}} \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \widetilde{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{\text{srn}}^2 + \sigma_{SA}^2} \\ \widetilde{\omega} = \frac{\sigma_\eta^2}{\widetilde{\sigma}^2}\omega_\eta \\ \widetilde{A} = \frac{A_\eta A_{\text{srn}} A_{SA} \sigma_{\text{srn}} \sigma_\eta \sigma_{SA} 2\pi}{\widetilde{\sigma}} e^{-\frac{\sigma_\eta^2 \omega_\eta (\widetilde{\omega} - \omega_\eta)}{2}} \end{array} \right. \quad (4)$$

Substitution of cell receptive field parameters in Eq. 4 yields  $\hat{A} \gg \tilde{A}$  at all considered angles of visual eccentricity. Thus, the second term of Eq. 3 can be neglected, and correlation difference maps are described by Gabor functions:

$$R_\eta(x) \approx \hat{R}_\eta(x) = \hat{A} e^{-\frac{x^2}{2\hat{\sigma}^2}} \cos(\hat{\omega}x + \phi) \quad (5)$$

Since also the spatial receptive fields of modeled V1 units are represented by Gabor functions, patterns of correlated activity can be directly compared to the structure of receptive fields. Apart from a scaling factor, a Gabor function is defined by two parameters: the width  $\sigma$  of the Gaussian, and the spatial frequency  $\omega$  of the plane wave. Thus, the similarity between correlation difference maps and cortical receptive fields can be quantified by directly comparing the values of these parameters by means of the two ratios  $r_\omega = \hat{\omega}/\omega_\eta$  and  $r_\sigma = \hat{\sigma}/\sigma_\eta$ . The closer these two ratios are to 1 the higher is the similarity between patterns of correlation and the spatial structure of simple cell receptive fields.

Fig. 1 compares the correlation difference maps given by Eq. 5 to the receptive fields of modeled V1 units. Since the precise location of the receptive fields of recorded cells were not reported by Jones and Palmer (1987b, 1987a), we estimated the patterns of correlation that each modeled V1 unit would establish with LGN cells located at various angles of visual eccentricity. Fig. 1 (a) shows an example for one of the modeled V1 units. The patterns of correlated activity measured at both  $5^\circ$  and  $15^\circ$  of visual eccentricity closely resembled the receptive field profile of the cortical cell. Fig. 1 (b) shows the mean values of the two ratios  $r_\omega$  and  $r_\sigma$  evaluated over all the 10 modeled V1 cells as a function of the visual eccentricity of geniculate units. Both ratios were close to 1 at all eccentricities indicating a close matching between the patterns of correlated activity and the receptive fields of all simulated cells. The average values of the two indices of similarity were  $\bar{r}_\sigma = 1.08 \pm 0.02$  and  $\bar{r}_\omega = 0.86 \pm 0.04$  respectively. Thus, in the model, the structure of thalamo-cortical activity present immediately before eye opening was compatible with a Hebbian maturation of simple cell receptive fields.

It is important to notice that the similarity between receptive fields and correlation difference maps shown in Fig. 1 originated from the narrow spatial correlations of spontaneous activity. Fig. 2 illustrates the results obtained when no spatial correlation was present at the level of the retina. When retinal spontaneous activity was modeled as white noise ( $C_{SA}(\omega_x, \omega_t) = 1$  in Eq. 1), correlation difference maps were again given by the sum of two terms  $\hat{R}_{\eta\alpha}(x)$  and  $\tilde{R}_{\eta\alpha}(x)$ , as in Eq. 3, with parameters:

$$\left\{ \begin{array}{l} \hat{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{\text{cnt}}^2} \\ \hat{\omega} = \frac{\sigma_\eta^2}{\hat{\sigma}^2} \omega_\eta \\ \hat{A} = \frac{A_\eta A_{\text{cnt}} \sigma_{\text{cnt}} \sigma_\eta \sqrt{2\pi}}{\hat{\sigma}} e^{-\frac{\sigma_\eta^2 \omega_\eta (\hat{\omega} - \omega_\eta)}{2}} \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \tilde{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{\text{srn}}^2} \\ \tilde{\omega} = \frac{\sigma_\eta^2}{\tilde{\sigma}^2} \omega_\eta \\ \tilde{A} = \frac{A_\eta A_{\text{srn}} \sigma_{\text{srn}} \sigma_\eta \sqrt{2\pi}}{\tilde{\sigma}} e^{-\frac{\sigma_\eta^2 \omega_\eta (\tilde{\omega} - \omega_\eta)}{2}} \end{array} \right. \quad (6)$$

As before, substitution of cell parameters into Eq. 6 gives  $\hat{A} \gg \tilde{A}$  so that  $R_\eta(x) \approx \hat{R}_\eta(x)$ . Similar to the previous case, correlation difference maps and cortical receptive

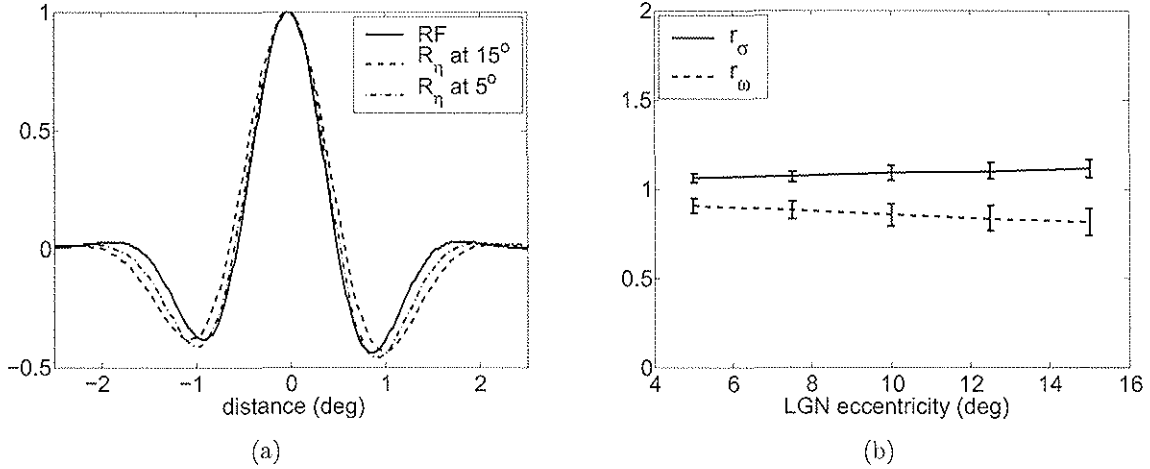


Figure 1: Comparison between the spatial organization of simple cell receptive fields and the structure of thalamo-cortical activity immediately before eye opening. (a) Results for one of the 10 modeled simple cells. The correlation difference maps ( $R_\eta$ ) measured between the considered simple cell and arrays of geniculate units located around  $5^\circ$  and  $15^\circ$  of visual eccentricity are compared to the profile of the receptive field (RF). (b) Comparison between the parameters of the Gabor functions that represented receptive fields and patterns of correlated activity. The two curves show the mean ratios  $r_\omega = \hat{\omega}/\omega_\eta$  and  $r_\sigma = \hat{\sigma}/\sigma_\eta$  evaluated over 10 modeled simple cells. Error bars represent the standard deviation. The  $x$  axis marks the angle of visual eccentricity of the receptive fields of geniculate units.

fields were highly similar in the absence of input spatial correlations. As shown in Fig. 2, both  $r_\sigma$  and  $r_\omega$  were close to 1 for all modeled V1 units independent of the visual eccentricity of geniculate cells. Mean ratios were  $\bar{r}_\sigma = 1.03 \pm 0.007$  and  $\bar{r}_\omega = 0.94 \pm 0.01$ . These results, together with those of Fig. 1, indicate that a regime of stochastic retinal activity with narrow spatial correlations is compatible with a Hebbian maturation of cortical receptive fields.

## 4 Thalamocortical activity after eye opening

After eye opening, the assumption of narrow spatial correlation in the visual input is no longer valid. Luminance values in natural scenes are correlated over relatively large distances, as revealed by the power-law profile of the power spectrum of natural images (Field, 1987; Burton & Moorhead, 1987; Ruderman, 1994).

Although the structure of natural visual stimulation is known to be non-Gaussian, a

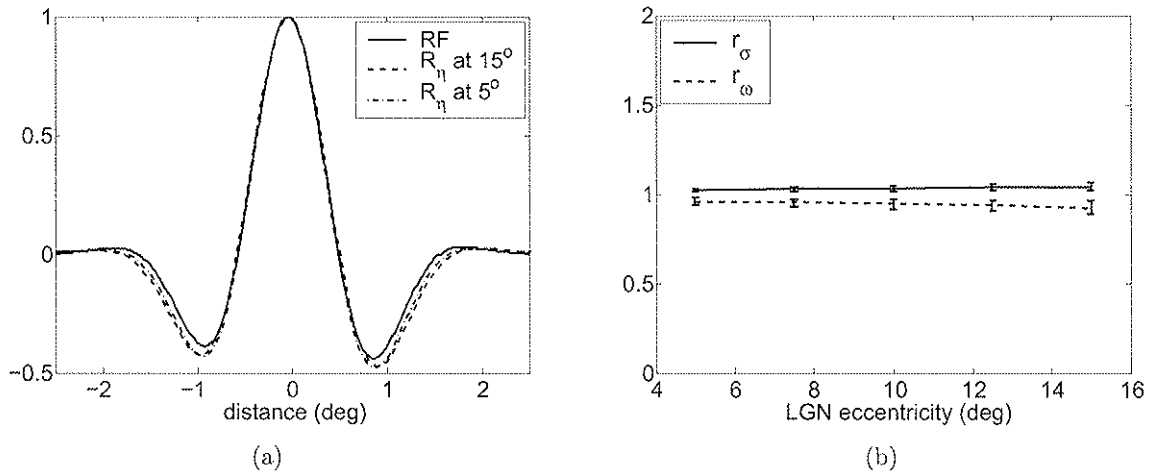


Figure 2: Comparison between the spatial organization of simple cell receptive fields and correlation difference maps measured in the absence of spatial correlation in retinal activity. The layout of the panels and the graphic notation are the same as in Fig. 1.

first understanding of the impact of these broad spatial correlations on thalamocortical activity can be gained by increasing in Eq. 4 the extent  $\sigma_{SA}$  of correlated activity in the retina. As shown by Fig. 3, in the presence of broad input correlations patterns of correlated activity did not match the spatial organization of simple cell receptive fields. In particular, a clear mismatch was present in correspondence of the secondary lobes in the receptive fields, where levels of correlation predicted strong input from LGN units with the wrong polarity (see Fig. 3 (a)). As shown in Fig. 3 (b), both  $r_\sigma$  and  $r_\omega$  diverged from 1 for increasing values of  $\sigma_{SA}$ , producing correlation difference maps that were wider and possessed lower cut-off frequencies than corresponding receptive fields. Thus, in the model, the second-order statistics of thalamocortical activity measured in the presence of broad spatial correlations was not compatible with a Hebbian maturation of simple cell receptive fields.

The impact of the broad correlations of natural scenes on the structure of thalamocortical activity can be analyzed by following a similar approach. In this case, correlation difference maps were given by:

$$R_\eta(\mathbf{x}) = C\mathcal{F}^{-1}\{\mathcal{N}(\omega)S_\eta(\omega)\overline{S_\alpha(\omega)}\} \quad (7)$$

where  $\mathcal{N}(\omega)$  is the power spectrum of natural images and  $C$  is a multiplicative factor equal to  $H_\alpha(0)H_\eta(0)$ . Fig. 4 compares the receptive fields of modeled simple cells to the correlation difference maps obtained by solving numerically Eq. 7. The power spectrum  $\mathcal{N}(\omega)$  was estimated from a set of 15 natural images (van Hateren & van der Schaaf, 1998). Its radial mean was best interpolated by  $\overline{\mathcal{N}}(\omega) \propto \omega^{-2.02}$ , which is

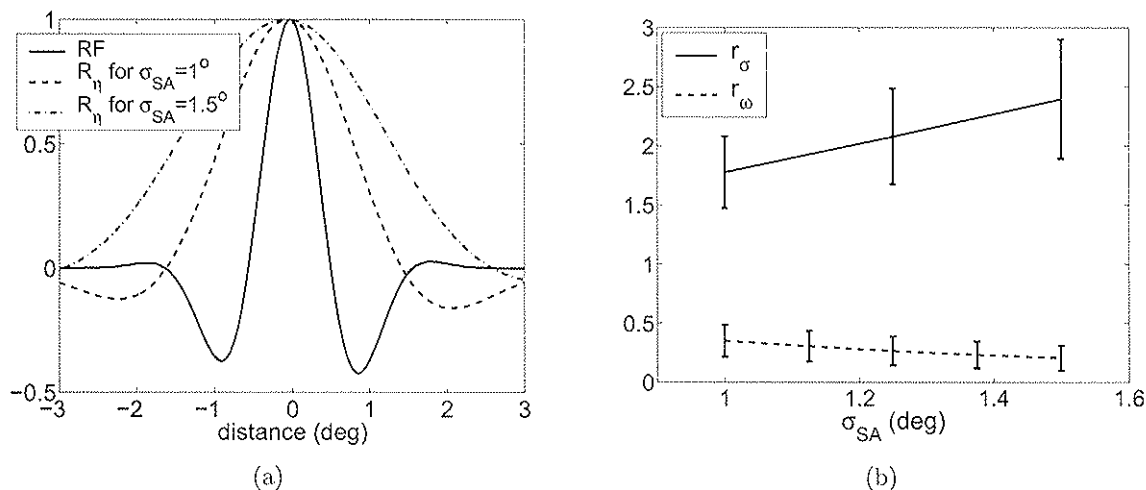


Figure 3: Comparison between the spatial organization of simple cell receptive fields and the structure of thalamo-cortical activity produced by retinal activity with broad spatial correlations. (a) Results for one of the 10 modeled simple cells. The correlation difference maps ( $R_\eta$ ) estimated from Eqs. 4 and 5 for two different values of  $\sigma_{SA}$  ( $\sigma_{SA} = 1^\circ$  and  $\sigma_{SA} = 1.5^\circ$ ) are compared to the profile of the simple cell receptive field (RF). (b) Comparison between the parameters of the Gabor functions representing receptive fields and patterns of correlated activity. The two curves show the mean ratios  $r_\omega = \hat{\omega}/\omega_\eta$  and  $r_\sigma = \hat{\sigma}/\sigma_\eta$  evaluated over 10 modeled simple cells. Error bars represent the standard deviation. The  $x$  axis represents the spatial extent of correlation in retinal activity. Parameters of LGN units simulated an eccentricity of  $10^\circ$ .

in agreement with previous measurements (Field, 1987; Ruderman, 1994). Since, in this case, an analytical expression of  $R_\eta(\mathbf{x})$  was not available, we compared correlation difference maps and receptive fields by means of their correlation coefficient  $r_{RF}$ . This index measures the similarity of two patterns. It varies between -1 and 1, with +1 indicating perfect matching and -1 perfect mirror symmetry.

Similar to the case of broad Gaussian input correlations, when images of natural scenes were passively presented to the model, patterns of correlated activity did not match the receptive fields of simple cells. An example for one of the 10 modeled simple cells is shown in Fig. 4 (a), which compares the profile of the cell receptive field to sections of the correlation difference maps measured at  $5^\circ$  and  $15^\circ$  of visual eccentricity. Similar to Fig. 3 (a), in correspondence of the side lobes of the receptive field, levels of correlation predicted stabilization of afferents from geniculate cells with the wrong polarity (ON- instead of OFF-center). Fig. 4 (b) shows average results obtained over the entire population of simulated simple cells. In addition to the mean correlation

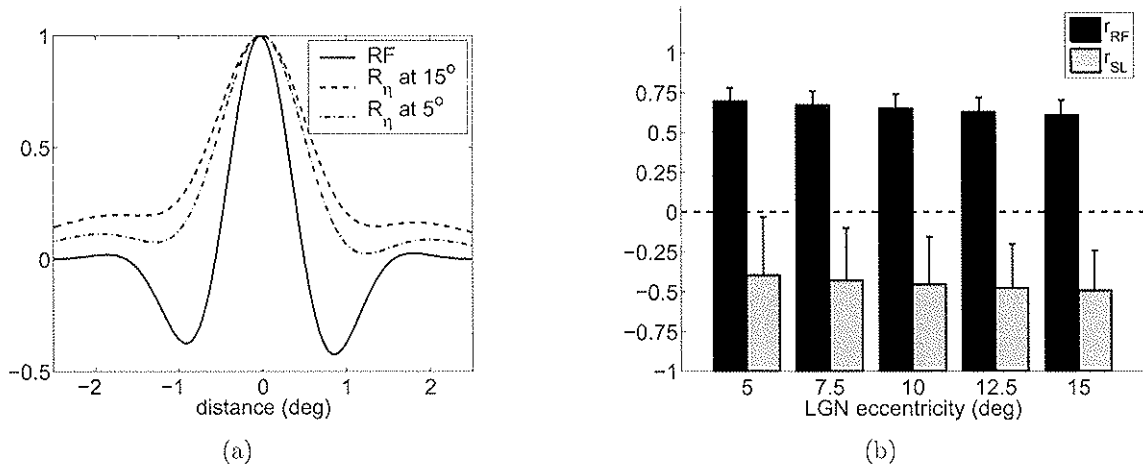


Figure 4: Comparison between the spatial organization of simple cell receptive fields and the structure of thalamo-cortical activity in the case of static presentation of natural images. (a) Results for one of the 10 modeled simple cells. The two correlation difference maps ( $R_\eta$ ) measured between the considered simple cell and arrays of geniculate units located around  $5^\circ$  and  $15^\circ$  of visual eccentricity are compared to the profile of the simple cell receptive field (RF). (b) Average matching across the 10 modeled V1 units. Bars indicate the matching between correlation difference maps and cortical receptive fields evaluated both over the entire receptive field ( $r_{RF}$ ) and only in correspondence of the secondary lobes ( $r_{SL}$ ) (see text for details). The  $x$  axis represents the angle of visual eccentricity of simulated geniculate units. Vertical lines represent the standard deviation.

coefficient  $r_{RF}$ , a second, more specific correlation coefficient index,  $r_{SL}$ , quantified the similarity between receptive fields and correlation difference maps only over the secondary lobes of cell receptive fields. The results obtained for geniculate units with parameters that replicated the average characteristics of cells located at various angles of visual eccentricity are ordered on the  $x$  axis. At all considered eccentricities, a clear mismatch was present between correlation maps and receptive fields. Average correlation coefficients were  $\bar{r}_{RF} = 0.65 \pm 0.03$  over the entire receptive fields, and  $\bar{r}_{SL} = -0.45 \pm 0.04$  in correspondence of the secondary lobes. That is, contrary to the case of retinal spontaneous activity, the structure of correlated activity measured in the presence of natural images was not compatible with a Hebbian refinement of the receptive fields of simple cells.

The results of Fig. 4 were obtained in the absence of eye movements. Under natural viewing conditions, however, the retinal image is always in motion as small movements of the eyes, head and body prevent maintenance of a steady direction of gaze.

Results from previous computational studies have shown a strong influence of fixational instability on the structure of correlated activity in models of the LGN and V1 (Rucci et al., 2000; Rucci & Casile, 2004). To examine the origins of this influence, in this paper we modeled fixational instability by means of a two-dimensional ergodic process  $\boldsymbol{\xi}(t) = [\xi_x(t), \xi_y(t)]^T$ . For simplicity, we assumed zero moments of the first-order ( $\langle \xi_x(t) \rangle = 0$  and  $\langle \xi_y(t) \rangle = 0$ ) and uncorrelated movements along the two axes ( $R_{\xi_x \xi_y}(t) = 0$ ).

By means of Taylor expansion, the luminance profile  $I(\mathbf{x})$  of a natural image in the neighborhood of a generic point  $\mathbf{x}_0$  can be approximated as :

$$I(\mathbf{x}) \approx I(\mathbf{x}_0) + \nabla I(\mathbf{x}_0)^T \cdot [\mathbf{x} - \mathbf{x}_0]$$

Thus, if the average area covered by fixational instability is sufficiently small, the input to the retina during visual fixation can be approximated by:

$$\begin{aligned} I_r(\mathbf{x}, t) &\approx I(\mathbf{x}) + \nabla I(\mathbf{x})^T \cdot \boldsymbol{\xi}(t) = \\ &= I(\mathbf{x}) + \frac{\partial I(\mathbf{x})}{\partial x} \xi_x(t) + \frac{\partial I(\mathbf{x})}{\partial y} \xi_y(t) \end{aligned}$$

Using this approximation, we can estimate the responses of cortical and geniculate cells during visual fixation:

$$\begin{aligned} \eta_{\mathbf{x}_0}(t) &\approx k_\eta(\boldsymbol{\mu}, \tau) \star I_r(\boldsymbol{\mu}, \tau) \Big|_{(\mathbf{x}_0, t)} = k_\eta(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_0, t)} + \\ &\quad + k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{(\mathbf{x}_0, t)} + k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau) \Big|_{(\mathbf{x}_0, t)} = \\ &= \eta_{\mathbf{x}_0}^S(t) + \eta_{\mathbf{x}_0}^D(t) \\ \alpha_{\mathbf{x}_1}(t) &\approx k_\alpha(\boldsymbol{\mu}, \tau) \star I_r(\boldsymbol{\mu}, \tau) \Big|_{(\mathbf{x}_1, t)} = k_\alpha(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_1, t)} + \\ &\quad + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{(\mathbf{x}_1, t)} + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau) \Big|_{(\mathbf{x}_1, t)} = \\ &= \alpha_{\mathbf{x}_1}^S(t) + \alpha_{\mathbf{x}_1}^D(t) \end{aligned} \tag{8}$$

where  $\mathbf{x}_0$  and  $\mathbf{x}_1$  are the locations of receptive fields centers and

$$\begin{cases} \eta_{\mathbf{x}_0}^S(t) = k_\eta(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_0, t)} \\ \eta_{\mathbf{x}_0}^D(t) = k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{(\mathbf{x}_0, t)} + k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau) \Big|_{(\mathbf{x}_0, t)} = \\ \quad = \eta_{\mathbf{x}_0}^{D_x}(t) + \eta_{\mathbf{x}_0}^{D_y}(t) \\ \alpha_{\mathbf{x}_1}^S(t) = k_\alpha(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_1, t)} \\ \alpha_{\mathbf{x}_1}^D(t) = k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{(\mathbf{x}_1, t)} + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau) \Big|_{(\mathbf{x}_1, t)} = \\ \quad = \alpha_{\mathbf{x}_1}^{D_x}(t) + \alpha_{\mathbf{x}_1}^{D_y}(t) \end{cases}$$

That is, cell responses can be decomposed into a *static* component with non-zero mean ( $\eta^S$  and  $\alpha^S$ ) and a zero-mean *dynamic* component determined by fixational instability ( $\eta^D$  and  $\alpha^D$ ).  $\eta^{D_x}$ ,  $\alpha^{D_x}$ ,  $\eta^{D_y}$  and  $\alpha^{D_y}$  are the contributions to cell responses generated by the instability of visual fixation along the  $x$  and  $y$  axes.

Given this decomposition, also correlation difference maps can be expressed as the sum of static and dynamic terms:

$$R_\eta(\mathbf{x}) = R_\eta^S(\mathbf{x}) + R_\eta^D(\mathbf{x}) \quad (9)$$

Indeed, from our assumptions on the statistical moments of fixational instability it follows that only three of the nine terms obtained by direct multiplication of the responses  $\eta_{\mathbf{x}_0}(t)$  and  $\alpha_{\mathbf{x}_1}(t)$  have non-zero means. The first one of these terms is given by:

$$\begin{aligned} & \langle \eta_{\mathbf{x}_0}^S(t) \alpha_{\mathbf{x}_1}^S(t) \rangle_{\xi, \mathcal{I}, t} = \\ & = \left\langle \left( k_\eta(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_1, t)} \right) \left( k_\alpha(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_0, t)} \right) \right\rangle_{\xi, \mathcal{I}, t} = \\ & = \left( s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N(\boldsymbol{\mu}) \Big|_{(\mathbf{x}_1 - \mathbf{x}_0)} \right) \left\langle \int_{-\infty}^{\infty} h_\alpha(\tau) d\tau \int_{-\infty}^{\infty} h_\eta(\tau) d\tau \right\rangle_t = \\ & = C s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N(\boldsymbol{\mu}) \Big|_{\mathbf{x}} \end{aligned}$$

where  $N(\mathbf{x})$  is the autocorrelation function of natural images. Since this term depends only on the static components of cell responses, it represents the correlation difference map that would be obtained in the absence of fixational instability (see Eq. 7).

The second term is given by:

$$\begin{aligned} & \langle \eta_{\mathbf{x}_0}^{D_x}(t) \alpha_{\mathbf{x}_1}^{D_x}(t) \rangle_{\xi, \mathcal{I}, t} = \\ & \left\langle \left( k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{\mathbf{x}_1, t} \right) \left( k_\alpha(\boldsymbol{\mu}, t) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau) \Big|_{\mathbf{x}_0, t} \right) \right\rangle_{\xi, \mathcal{I}, t} = \\ & = \left( s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N_x(\boldsymbol{\mu}) \Big|_{\mathbf{x}_1 - \mathbf{x}_0} \right) \left\langle h_\eta(\tau) \star h_\alpha(-\tau) \star R_{\xi_x \xi_x}(\tau) \Big|_{\tau=0} \right\rangle_t = \\ & = D s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'_x(\boldsymbol{\mu}) \Big|_{\mathbf{x}} \end{aligned} \quad (10)$$

where  $N'_x(\boldsymbol{\mu})$  is the autocorrelation function of the first component of the gradient of natural images (the derivative along the  $x$  axis).  $D$  is a constant equal to  $\int_{-\infty}^{\infty} H_\eta(\omega_t) \overline{H_\alpha}(\omega_t) \mathcal{R}_{\xi_x \xi_x}(\omega_t) d\omega_t$  where  $\mathcal{R}_{\xi_x \xi_x}(\omega_t)$  indicates the Fourier transform of  $R_{\xi_x \xi_x}(t)$ .

By using a similar procedure, we obtain the third term:

$$\langle \eta_{\mathbf{x}_0}^{D_y}(t) \alpha_{\mathbf{x}_1}^{D_y}(t) \rangle_{\xi, \mathcal{I}, t} = D s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'_y(\boldsymbol{\mu}) \Big|_{\mathbf{x}}$$

where  $N'_y(\boldsymbol{\mu})$  is the autocorrelation function of the second component of the gradient of natural images (the derivative along the  $y$  axis).

By adding these three terms and defining  $N'(\boldsymbol{\mu}) = N'_x(\boldsymbol{\mu}) + N'_y(\boldsymbol{\mu})$ , we obtain:

$$R_\eta(\mathbf{x}) = Cs_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N(\boldsymbol{\mu})|_{\mathbf{x}} + \\ + Ds_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'(\boldsymbol{\mu})|_{\mathbf{x}} = R_\eta^S(\mathbf{x}) + R_\eta^D(\mathbf{x}) \quad (11)$$

which proves Eq. 9.

Eq. 11 shows that fixational instability adds a contribution  $R_\eta^D(\mathbf{x})$  to the correlation map  $R_\eta^S(\mathbf{x})$  obtained with presentation of the same stimuli in the absence of retinal image motion. Whereas in the absence of fixational instability, levels of correlation depend on the autocorrelation function of the stimulus  $N(\mathbf{x})$  (or, equivalently, its power spectrum  $\mathcal{N}(\boldsymbol{\omega})$ , the term  $R_\eta^D(\mathbf{x})$  introduced by the jittering of the eye depends on the autocorrelation function of the gradient of the stimulus,  $N'(\mathbf{x})$  (or equivalently its power spectrum  $\mathcal{N}'(\boldsymbol{\omega})$ , the ‘‘dynamic’’ power spectrum).

Fig. 5 compares  $\mathcal{N}(\boldsymbol{\omega})$  and  $\mathcal{N}'(\boldsymbol{\omega})$  for the case of images of natural scenes. Whereas  $\mathcal{N}(\boldsymbol{\omega})$  followed, as expected, a power law with exponent approximately equal to -2, the dynamic power spectrum  $\mathcal{N}'(\boldsymbol{\omega})$  was almost flat up to a cut-off frequency of about 10 cycles/deg. That is, in the presence of natural images, fixational instability adds an input signal that discards spatial correlations.

The ‘whitening’ of the dynamic power spectrum is a direct consequence of the scale-invariance of the second-order statistics of natural images. It has a simple explanation in the frequency domain. Since the Fourier transform of the two partial derivatives  $\frac{\partial I(\mathbf{x})}{\partial x}$  and  $\frac{\partial I(\mathbf{x})}{\partial y}$  are respectively proportional to  $\omega_x I(\boldsymbol{\omega})$  and  $\omega_y I(\boldsymbol{\omega})$ , the two power spectra  $\mathcal{N}'_x(\boldsymbol{\omega})$  and  $\mathcal{N}'_y(\boldsymbol{\omega})$  are proportional to  $\omega_x^2 \mathcal{N}(\boldsymbol{\omega})$  and  $\omega_y^2 \mathcal{N}(\boldsymbol{\omega})$ . Thus,  $\mathcal{N}'(\boldsymbol{\omega}) = \mathcal{N}'_x(\boldsymbol{\omega}) + \mathcal{N}'_y(\boldsymbol{\omega}) \propto |\boldsymbol{\omega}|^2 \mathcal{N}(\boldsymbol{\omega})$ . For images of natural scenes,  $\mathcal{N}(\boldsymbol{\omega}) \propto |\boldsymbol{\omega}|^{-2}$  (Field, 1987). Therefore, the product  $|\boldsymbol{\omega}|^2 \mathcal{N}(\boldsymbol{\omega})$  produces a dynamic power spectrum  $\mathcal{N}'(\boldsymbol{\omega})$  with uniform spectral density. That is, fixational instability represents an optimal decorrelation strategy for visual input with power spectrum that declines as  $|\boldsymbol{\omega}|^{-2}$ .

To summarize, Eq. 11 shows that in the presence of the self motion of the retinal image that occurs during natural viewing, the second-order statistics of thalamo-cortical activity depends both on the configuration of the stimulus and on how its retinal projection changes during visual fixation. This latter component is determined by the dynamic power spectrum  $\mathcal{N}'(\boldsymbol{\omega})$  (and not by the power spectrum of the stimulus  $\mathcal{N}(\boldsymbol{\omega})$ ), a spectrum that discards broad spatial correlations.

We have already shown in Fig. 4 that the patterns of correlated activity  $R_\eta^S(\mathbf{x})$  measured with presentation of images of natural scenes without fixational instability did not match the receptive fields of modeled simple cells. Fig. 6 analyzes the contribution of fixational instability to the structure of correlated activity (the term  $R_\eta^D(\mathbf{x})$  in Eq. 11). In this case, for all modeled simple cells correlation difference maps closely resembled the spatial organization of receptive fields irrespective of the eccentricity of simulated geniculate units (compare Fig. 6 with Fig. 4). The mean matching index was  $\bar{r}_{RF} = 0.98 \pm 0.002$  over the entire receptive fields and  $\bar{r}_{SL} = 0.92 \pm 0.004$  over the secondary lobes. That is, each simple cell established strong correlations with either ON- or OFF-center geniculate units only when the receptive fields of these units

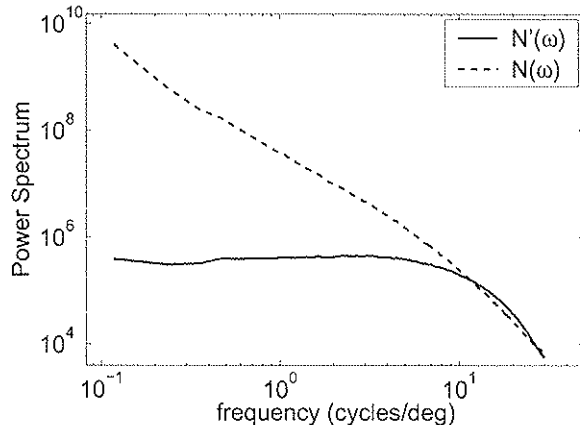


Figure 5: Comparison between the power spectrum of natural images  $\mathcal{N}(\omega)$  and the dynamic power spectrum  $\mathcal{N}'(\omega) = \mathcal{N}'_x(\omega) + \mathcal{N}'_y(\omega)$  given by the sum of the power spectra of the  $x$  and  $y$  components of the gradient of natural images. The two curves are radial averages over a set of 15 natural images.

overlapped an ON or an OFF subregion. Similar to the case of spontaneous retinal activity, this pattern of correlated activity is compatible with a Hebbian refinement of simple cell receptive fields.

In natural images most energy is concentrated at low spatial harmonics. Since  $\mathcal{N}'(\omega)$  attenuates the low spatial frequencies of the stimulus, it has less spatial power than  $\mathcal{N}(\omega)$  (see Fig. 5). However, the impact of  $\mathcal{N}'(\omega)$  on the global structure of correlated activity depends on the spatiotemporal characteristics of fixational instability, which are represented in Eq. 11 by the amplification factor  $D$ . That is, the relative contribution of  $R^D$  and  $R^S$  in Eq. 11 depends on how the retinal image moves during the acquisition of visual information. For example, a retinal image motion with Gaussian temporal correlation (the term  $R_{\xi_x, \xi_x}$  in Eq. 10) characterized by a standard deviation of 30ms and a mean amplitude of  $10^\circ$ , values that are consistent with the instability of fixation of several species, increased matching from  $\bar{r}_{RF} = 0.65 \pm 0.03$  and  $\bar{r}_{SL} = -0.45 \pm 0.04$  (the values obtained with static presentation of natural images, see Fig. 4) to  $\bar{r}_{RF} = 0.90 \pm 0.02$  and  $\bar{r}_{SL} = 0.12 \pm 0.1$ . Thus, in the presence of fixational instability, the responses of simulated cortical units tended to be correlated with those of geniculate units with correct polarity.

It is important to observe that several mechanisms might further enhance the impact of fixational instability on the refinement of thalamocortical connectivity. A first possibility is a rule of synaptic plasticity that depends on the covariance (and not the correlation) between the responses of pre- and post-synaptic elements (Sejnowski,

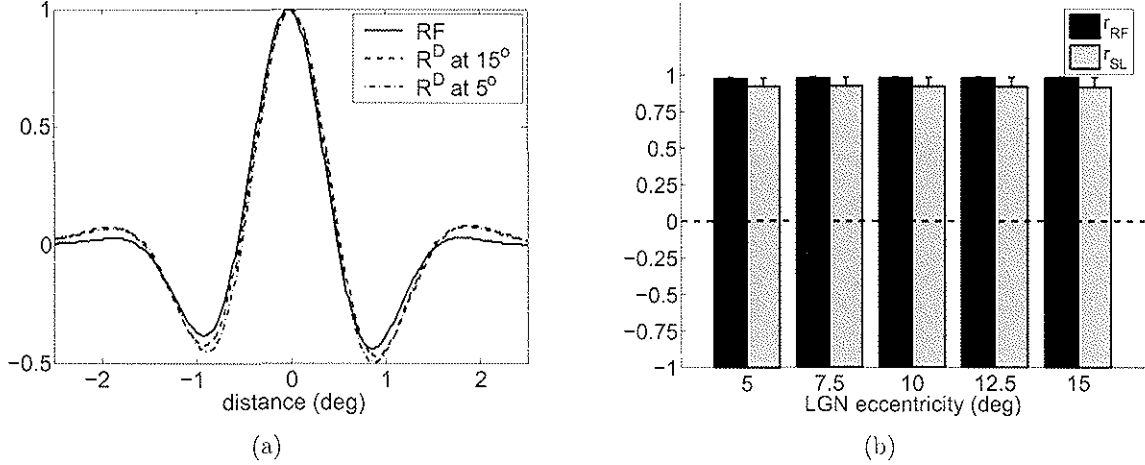


Figure 6: Comparison between the spatial organization of simple cell receptive fields and patterns of correlated activity measured when images of natural scenes were examined in the presence of fixational instability (the term  $R_\eta^D(\mathbf{x})$  in Eq. 11). The layout of the panels and the graphic notation are the same as in Fig. 4.

1977):

$$R_\eta(\mathbf{x}) = \langle (\eta(t) - \bar{\eta})(\alpha(t) - \bar{\alpha}) \rangle.$$

In the case in which mean activity levels are estimated over fixational periods,  $\bar{\eta} = \eta^S$  and  $\bar{\alpha} = \alpha^S$  yielding  $R_\eta(\mathbf{x}) = R_\eta^S(\mathbf{x})$ . Thus, the term  $R_\eta^D(\mathbf{x})$  does not affect synaptic plasticity, and the structure of thalamocortical activity is compatible with the spatial organization of the receptive fields of simple cells. This is consistent with the results of our previous simulations in which we analyzed the statistics of geniculate activity during natural viewing (Rucci et al., 2000).

A second mechanism that might enhance the influence of fixational instability is a nonlinear attenuation of the responses of simple cells to unchanging stimuli. Systematic deviations from linearity have been observed in the responses of simple cells. In particular, it has been shown that responses to stationary stimuli tend to decline faster and give lower steady-state levels of activity than it would be expected from linear predictions (Tolhurst, Walker, Thompson, & Dean, 1980; DeAngelis et al., 1993a). This attenuation can be incorporated into our model by assuming that, after an initial transitory period following the onset of visual fixation, a simple cell responds as:

$$\eta(t) = (1 - \beta) \cdot \eta^S(t) + \eta^D(t)$$

where the constant  $\beta \in [0, 1]$  defines the degree of attenuation. With this modification,

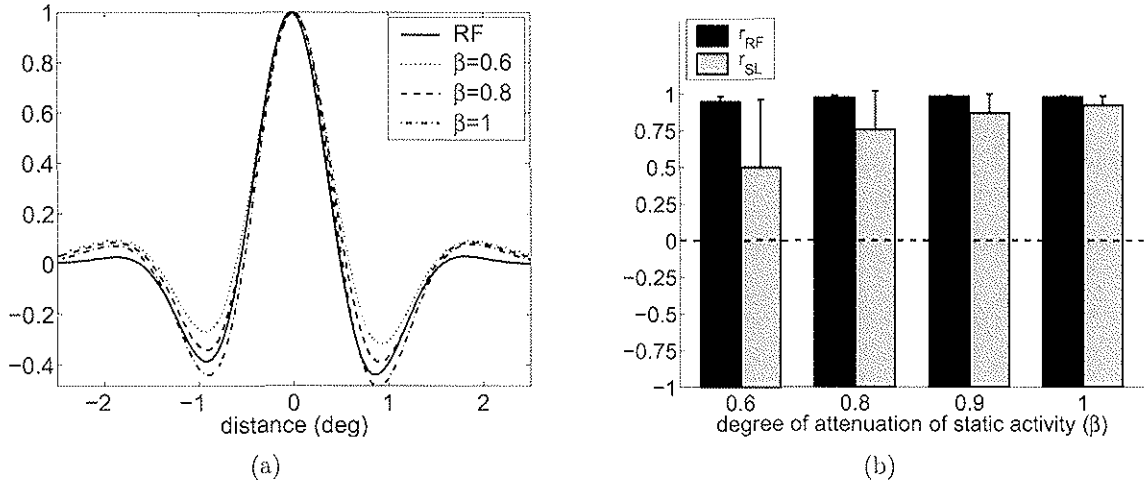


Figure 7: Effect of nonlinear attenuation of simple cells responses to unchanging stimuli. (a) Results for one of the 10 modeled simple cells. The correlation difference maps ( $R_\eta$ ) measured for three values of the attenuation factor  $\beta$ , are compared to the profile of the simple cell receptive field (RF). (b) Mean matching indices over all modeled V1 units as a function of the attenuation factor. Both correlation coefficients evaluated over the entire receptive field ( $r_{RF}$ ) and the secondary subregions ( $r_{SL}$ ) are shown. Parameters of LGN units simulated an eccentricity of  $10^\circ$ .

correlation difference maps are given by:

$$R_\eta(\mathbf{x}) = (1 - \beta)R_\eta^S(\mathbf{x}) + R_\eta^D(\mathbf{x})$$

Fig. 7 compares the receptive fields of the simulated simple cells with the correlation difference maps estimated with various degrees of attenuation. It is clear by comparing these data to those of Fig. 4 that even a partial attenuation of cortical responses to unchanging stimuli resulted in a substantial improvement in the degree of similarity between patterns of correlation and receptive fields. An attenuation of 60% was sufficient to produce an almost perfect matching ( $\bar{r}_{RF} = 0.97 \pm 0.02$  and  $\bar{r}_{SL} = 0.50 \pm 0.44$ ). Thus, consistent with our previous simulations of thalamocortical activity (Rucci & Casile, 2004), a nonlinear attenuation of simple cell responses leads to regime of correlated activity that is compatible with a Hebbian refinement of the spatial organization of simple cell receptive fields.

## 5 Conclusions

Many of the response characteristics of V1 neurons develop before eye opening and refine with exposure to pattern vision (Hubel & Wiesel, 1963; Blakemore & van Sluyters, 1975; Buisseret & Imbert, 1976; Pettigrew, 1974). After eye opening, small eye and body movements keep the retinal image in constant motion. The statistical analysis of this paper, together with the results of our previous simulations (Rucci et al., 2000; Rucci & Casile, 2004), indicate that the physiological instability of visual fixation contributes to decorrelating cell responses to natural stimuli and establishing a regime of neural activity similar to that present before eye opening. Thus, at the time of eye opening, no sudden change occurs in the second-order statistics of thalamocortical activity, and the same correlation-based mechanism of synaptic plasticity can account for both the initial emergence and the later refinement of simple cell receptive fields.

In this study, we have used independent models of LGN and V1 neurons to examine whether the structure of thalamocortical activity is compatible with a Hebbian maturation of the spatial organization of simple cell receptive fields. The results of our analysis are consistent with a substantial body of previous modeling work. Before eye opening, in the presence of spontaneous retinal activity, a modeled simple cell established strong correlations with ON- and OFF-center geniculate units only when the receptive fields of these units overlapped, respectively, the ON and OFF subregions within its receptive fields. This pattern of correlated activity is consistent with the results of previous studies that modeled the activity-dependent development of cortical orientation selectivity (Linsker, 1986; Miller, 1994; Miyashita & Tanaka, 1992).

After eye opening, the visual system is exposed to the broad spatial correlations of natural scenes. In the absence of retinal image motion, these input correlations would coactivate geniculate units with the same polarity (ON- or OFF-center) and with receptive fields at relatively large separations, a pattern of neural activity that is not compatible with a Hebbian refinement and maintenance of the organization of simple cell receptive fields. During natural fixation, however, neurons receive input signals that vary in time as their receptive fields move with the eye (Gur & Snodderly, 1997). We have shown that, in the presence of images of natural scenes, these input fluctuations lack spatial correlations. This result is a direct consequence of the statistical structure of natural images. Although in a natural image the intensity values of pairs of pixels tend to be correlated over large distances, the local changes in intensity around each pixel are on average uncorrelated. In the model, this spatially uncorrelated input signal strongly influenced neuronal responses and produced patterns of thalamocortical activity that were similar to those measured immediately before eye opening. Our analysis shows that a direct scheme of Hebbian plasticity can be added to the category of activity-dependent mechanisms capable of accounting for the emergence of cortical receptive fields in the presence of decorrelated natural visual input (Law & Cooper, 1994; Lee, Blais, Shouval, & Cooper, 2000; Olshausen & Field, 1996; Mayer, Herrmann, & Geisel, 2003).

The fact that fixational instability might have such a strong effect on the development of cortical receptive fields should not come as a surprise. Consistent with the results of our analysis, several experimental studies have shown that prevention and manipulation of eye movements during the critical period disrupts the maturation of the response properties of cortical neurons (for a review, see Buisseret, 1995). For example, no restoration of cortical orientation selectivity (Gary-Bobo et al., 1986; Buisseret et al., 1978) and ocular dominance (Freeman & Bonds, 1979; Singer & Rauschecker, 1982) is observed in dark-reared kittens exposed to visual stimulation with their eye movements prevented. In addition, neurophysiological results have shown that fixational eye movements strongly influence the responses of geniculate and cortical neurons (Gur et al., 1997; Leopold & Logothetis, 1998; Martinez-Conde et al., 2002). In the primary visual cortex of the monkey, bursts of spikes have been recorded following fixational saccades (Martinez-Conde et al., 2000), and distinct neuronal populations have been found that selectively respond to the two main components of fixational eye movements, saccades and drifts (Snodderly et al., 2001).

An important assumption of this study was the use of linear models to predict the responses to visual stimuli. Linear spatio-temporal models enabled the derivation of analytical expressions of the levels of correlation in thalamocortical activity. A substantial body of evidence shows that LGN X cells act predominantly as linear filters. Responses to drifting gratings contain most power in the first harmonic (So & Shapley, 1981), and responses to both flashed and complex naturalistic stimuli are well captured by linear predictors (Dan, Atick, & Reid, 1996; Stanley, Li, & Dan, 1999; Cai et al., 1997). Also the responses of V1 simple cells contain a strong linear component (Jones & Palmer, 1987b; DeAngelis, Ohzawa, & Freeman, 1993b). However, for these neurons important deviations from linearity have been reported. In particular, it has been observed that responses to stationary stimuli decline faster and give lower steady-state levels of activity than would be expected from linear predictions (Tolhurst et al., 1980; DeAngelis et al., 1993a). We have shown that a nonlinear attenuation of cortical responses to unchanging stimuli enhances the influence of fixational instability on the structure of correlated activity. In the model, the broad correlations of natural scenes had little impact on the second-order statistics of thalamocortical activity in the presence of strong nonlinear attenuation.

A second important assumption concerned the way we modeled the self motion of the retinal image. In this study, the physiological instability of visual fixation was modeled as a zero-mean stochastic process with uncorrelated components along the two Cartesian axes. These assumptions simplified our analysis and led to the elimination of several terms in Eq. 11. However, the results presented in this paper do not critically depend on them. Simulations in which retinal image motion replicated the cat's oculomotor behavior have produced patterns of correlated activity that are very similar to the theoretical predictions of this study (Rucci et al., 2000; Rucci & Casile, 2004). During natural viewing, several elements, including movements of the eye (Ditchburn, 1980; Steinman, Haddad, Skavenski, & Wyman, 1973) and body as

well as imperfections in the vestibulo-ocular reflex (Skavenski, Hansen, Steinman, & Winterson, 1979), contribute to fixational instability. Although a statistical analysis of fixational instability under natural viewing conditions has not been performed, the motion of the retinal image as subjectively experienced by a jitter after-effect appears to qualitatively follow our assumptions (Murakami & Cavanagh, 1998).

Our results appear to contrast with a previous proposal according to which the contrast sensitivity functions of retinal and geniculate neurons decorrelate the input signals provided by images of natural scenes (Atick & Redlich, 1992). In the low frequency range, human contrast sensitivity increases proportionally with the spatial frequency in a way that counterbalances the scaling invariance of the power spectrum of natural images. Psychophysical measurements of contrast sensitivity were taken by Atick and Redlich (1992) to represent a cumulative envelope of the frequency responses of large neuronal ensembles. However, in addition to neuronal characteristics, these measurements are also likely to include the impact of fixational instability. In our analysis, all model parameters were based on neurophysiological data. In both the cat and the monkey, the frequency responses of cells in the retina and the LGN deviate significantly from linearity in the low spatial frequency range (So & Shapley, 1981; Linsenmeier et al., 1982; Derrington & Lennie, 1984; Croner & Kaplan, 1995; Cheng, Chino, Smith III, Hamamoto, & Yoshida, 1995). Since most of the power of natural images is in this frequency range, such deviation is not compatible with the decorrelation mechanism proposed by Atick and Redlich (1992). It is important to observe that the decorrelation of visual input operated by fixational instability does not depend on the spatial response properties of geniculate and cortical units. Thus, the proposed mechanism is highly robust with respect to individual neuronal differences in spatial contrast sensitivity functions.

While in this study we have focused on the developmental consequences of a chronic exposure to fixational instability, our results also have important implications concerning the way visual information is represented in the early visual system. It is a long-standing proposal that an important function of early visual processing is the elimination of redundant information (Barlow, 1961; Attneave, 1954). The results of our analysis suggest that fixational instability, by decreasing statistical dependencies between neural responses, might contribute to discarding broad input correlations and establishing efficient visual representations of natural visual scenes. Further theoretical and experimental studies are needed to characterize and test this hypothesis.

**Acknowledgments** We would like to thank Matthias Franz, Alessandro Treves and Martin Giese for many helpful comments on a preliminary version of this manuscript. This work was supported by the Deutsche Volkswagen Stiftung, by the National Institute of Health grant EY15732-01 and the National Science Foundation grants BIC-0432104 and CCF-0130851

## References

- Alonso, J., Usrey, M., & Reid, R. C. (2001). Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. *J Neurosci*, *21*(11), 4002-4015.
- Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? *Neural Comput*, *4*, 196-210.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychol Rev*, *61*(3), 183-193.
- Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication*. Cambridge : MIT.
- Blakemore, C., & van Sluyters, R. C. (1975). Innate and environmental factors in the development of the kitten's visual cortex. *J Physiol*, *248*, 663-716.
- Buisseret, P. (1995). Influence of extraocular muscle proprioception on vision. *Physiol Rev*, *75*(2), 323-338.
- Buisseret, P., Gary-Bobo, E., & Imbert, M. (1978). Ocular motility and recovery of orientational properties of visual cortical neurons in dark-reared kittens. *Nature*, *272*, 816-817.
- Buisseret, P., & Imbert, M. (1976). Visual cortical cells: their developmental properties in normal and dark-reared kittens. *J Physiol*, *255*, 511-525.
- Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Appl Opt*, *26*, 157-170.
- Cai, D., DeAngelis, G. C., & Freeman, R. D. (1997). Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J Neurophysiol*, *78*, 1045-1061.
- Changeux, J. P., & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, *264*, 705-712.
- Cheng, H., Chino, Y. M., Smith III, E. L., Hamamoto, J., & Yoshida, K. (1995). Transfer characteristics of lateral geniculate nucleus X neurons in the cat: Effects of spatial frequency and contrast. *J Neurophysiol*, *74*(6), 2548-2557.

- Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Res*, *35*(1), 7-24.
- Dan, Y., Atick, J. J., & Reid, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus : Experimental test of a computational theory. *J Neurosci*, *16*(10), 3351-3362.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993a). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J Neurophysiol*, *69*(4), 1091-1117.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993b). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J Neurophysiol*, *69*(4), 1118-35.
- Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J Physiol*, *357*, 219-240.
- Ditchburn, R. (1980). The function of small saccades. *Vision Res*, *20*, 271-272.
- Ferster, D., Chung, S., & Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, *380*, 249-252.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A*, *4*, 2379-2394.
- Fiorentini, A., Maffei, L., & Bisti, S. (1979). Change of binocular properties of cortical cells in the central and paracentral visual field projections of monocularly paralyzed cats. *Brain Res*, *171*, 541-544.
- Freeman, R. D., & Bonds, A. B. (1979). Cortical plasticity in monocularly deprived immobilized kittens depends on eye movement. *Science*, *206*, 1093-1095.
- Gary-Bobo, E., Milleret, C., & Buisseret, P. (1986). Role of eye movements in developmental processes of orientation selectivity in the kitten visual cortex. *Vision Res*, *26*(4), 557-567.
- Gur, M., Beylin, A., & Snodderly, D. M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J Neurosci*, *17*(8), 2914-2920.
- Gur, M., & Snodderly, D. M. (1997). Visual receptive fields of neurons in primary visual cortex (V1) move in space with the eye movements of fixation. *Vision Res*, *37*(3), 257-265.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture of the cat's visual cortex. *J Physiol*, *160*, 106-154.
- Hubel, D. H., & Wiesel, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *J Neurophysiol*, *26*, 994-1002.
- Jones, J. P., & Palmer, L. A. (1987a). An evaluation of the two-dimensional

- Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol*, 58(6), 1233-1258.
- Jones, J. P., & Palmer, L. A. (1987b). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol*, 58(6), 1187-1211.
- Law, C. C., & Cooper, L. N. (1994). Formation of receptive fields in realistic visual environments according to the Bienenstock, Cooper and Munro (BCM) theory. *Proc Natl Acad Sci USA*, 91, 7797-7801.
- Lee, A. B., Blais, B., Shouval, H. Z., & Cooper, L. N. (2000). Statistics of lateral geniculate nucleus (LGN) activity determines the segregation of ON/OFF subfields for simple cells in visual cortex. *Proc Natl Acad Sci USA*, 97(23), 12875-12879.
- Leopold, D. A., & Logothetis, N. K. (1998). Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp Brain Res*, 123, 341-345.
- Linsenmeier, R. A., Frishman, L. J., Jakiela, H. G., & Enroth-Cugell, C. (1982). Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. *Vision Res*, 22, 1173-1183.
- Linsker, R. (1986). From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proc Natl Acad Sci USA*, 83, 8390-8394.
- Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2000). Microsaccadic eye movements and firing of single cells in the macaque striate cortex. *Nat Neurosci*, 3(3), 251-258.
- Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2002). The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proc Natl Acad Sci USA*, 99(21), 13920-13925.
- Mastrorarde, D. N. (1983). Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X and Y cells. *J Neurophysiol*, 49, 303-323.
- Mayer, N., Herrmann, M., & Geisel, T. (2003). Shaping of receptive fields in the visual cortex during retinal maturation. *J Comput Neurosci*, 15, 307-320.
- Miller, K. D. (1994). A model of the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF- center inputs. *J Neurosci*, 14(1), 409-441.
- Miller, K. D., Erwin, E., & Kayser, A. (1999). Is the development of orientation selectivity instructed by activity? *J Neurobiol*, 41, 55-57.
- Miyashita, M., & Tanaka, S. (1992). A mathematical model for the self-organization of orientation columns in visual cortex. *Neuroreport*, 3, 69-72.
- Murakami, I., & Cavanagh, P. (1998). A jitter after-effect reveals motion-based

- stabilization of vision. *Nature*, 395, 798-801.
- Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607-609.
- Pettigrew, J. D. (1974). The effect of visual experience on the development of stimulus specificity by kitten cortical neurons. *J. Physiol*, 237, 49-74.
- Ratliff, F., & Riggs, L. A. (1950). Involuntary motions of the eye during monocular fixation. *J Exp Psychol*, 40, 687-701.
- Reid, R. C., & Alonso, J. M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, 378, 281-284.
- Rucci, M., & Casile, A. (2004). Decorrelation of neural activity during fixational instability: Possible implications for the refinement of V1 receptive fields. *Vis Neurosci*. (in press)
- Rucci, M., Edelman, G. M., & Wray, J. (2000). Modeling LGN responses during free-viewing: A possible role of microscopic eye movements in the refinement of cortical orientation selectivity. *J Neurosci*, 20(12), 4708-4720.
- Ruderman, D. L. (1994). The statistics of natural images. *Network*, 5, 517-548.
- Sejnowski, T. J. (1977). Storing covariance with nonlinearly interacting neurons. *J Math Biol*, 4(4), 303-321.
- Singer, W., & Rauschecker, J. (1982). Central-core control of developmental plasticity in the kitten visual cortex II. electrical activation of mesencephalic and diencephalic projections. *Exp Brain Res*, 47, 223-233.
- Skavenski, A. A., Hansen, R., Steinman, R. M., & Winterson, B. J. (1979). Quality of retinal image stabilization during small natural and artificial body rotations in man. *Vision Res*, 19, 365-375.
- Snodderly, D. M., Kagan, I., & Gur, M. (2001). Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Vis Neurosci*, 18, 259-277.
- So, Y. T., & Shapley, R. (1981). Spatial tuning of cells in and around the lateral geniculate nucleus of cat: X and Y relay cells and perigeniculate interneurons. *J Neurophysiol*, 45(1), 107-120.
- Stanley, G. B., Li, F. F., & Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J Neurosci*, 19(18), 8036-8042.
- Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature eye movement. *Science*, 181(102), 810-819.
- Stent, G. S. (1973). A physiological mechanism for Hebb's postulate of learning. *Proc Natl Acad Sci USA*, 70, 997-1001.
- Tollhurst, D., Walker, N., Thompson, I., & Dean, A. F. (1980). Non-linearities

- of temporal summation in neurones in area 17 of the cat. *Exp Brain Res*, 38, 431-435.
- van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc R Soc Lond B Biol Sci*, 265, 359-366.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.