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

Antonino Casile
antonino.casile@uni-tuebingen.de
Laboratory for Action Representation and Learning, Department of Cognitive Neurology, Hertie Institute for Clinical Brain Research, University Clinic, 72072 Tübingen, Germany

Michele Rucci
rucci@cns.bu.edu
Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215, U.S.A.

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 to the Hebbian maturation of the receptive fields of V1 simple cells (Rucci, Edelman, & Wray, 2000; Rucci & Casile, 2004). This letter examines the origins of such a contribution. Using quasilinear models of lateral geniculate nucleus 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. This input signal produces a regime of thalamocortical activity similar to that present before eye opening and compatible with the Hebbian maturation of cortical receptive fields.

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 postsynaptic 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. At the level of the retina, spontaneous activity appears to be correlated on a narrow spatial scale in the order of tens of arcmins (Mastronarde, 1983), whereas natural visual stimulation is known to be characterized by broad spatial correlations in the order of degrees of visual angle (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 (Ratliff & Riggs, 1950; Yarbus, 1967; Steinman, Haddad, Skavenski, & Wyman, 1973; Ditchburn, 1980). 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 et al., 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 & Raushecker, 1982).

In previous studies, we simulated the responses of lateral geniculate nucleus (LGN) and V1 neurons to analyze the second-order statistics of thalamic (Rucci, Edelman, & Wray, 2000) and thalamocortical activity (Rucci & Casile, 2004) before and after eye opening. 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 letter, to better understand the possible influence of fixational instability on visual development, we used quasilinear 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, Hebbian synapses change their strengths proportionally to the levels of correlation between the responses of pre- and postsynaptic elements. In this article, instead of explicitly modeling synaptic changes, we estimated the emerging structure of thalamocortical connectivity by directly analyzing levels of correlation between geniculate and cortical cells. Since V1 neurons exhibit orientation-selective responses at the time of eye opening, the goal of this analysis was to examine the compatibility between a strictly Hebbian mechanism of synaptic plasticity and the preservation of preexisting patterns of connectivity.

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

$$R_\eta(\boldsymbol{x}) = \left\langle \eta_{x_\eta}(t)\left[\alpha^{\mathrm{ON}}_{x_\alpha}(t) - \alpha^{\mathrm{OFF}}_{x_\alpha}(t)\right]\right\rangle_{\mathcal{I},t},$$

where $\eta_{x_\eta}(t)$ is the activity of the cortical neuron and $\alpha^{\mathrm{ON}}_{x_\alpha}(t)$ and $\alpha^{\mathrm{OFF}}_{x_\alpha}(t)$ are the responses of an ON- and an OFF-center geniculate cell with receptive fields centered at $\boldsymbol{x}_\alpha$. The average is evaluated over time t and an ensemble of input images $\mathcal{I}$.

A positive value of $R_\eta(\boldsymbol{x})$ 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 $\boldsymbol{x} = \boldsymbol{x}_\eta - \boldsymbol{x}_\alpha$. The opposite holds for a negative value of $R_\eta(\boldsymbol{x})$. R_η can be seen as η's receptive field predicted by purely Hebbian synapses. To preserve and refine the spatial organization of η's receptive field, R_η needs to be positive at locations that correspond to ON subregions and negative in correspondence of the OFF subregions.

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

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

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

where A_η 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, ω_η and ϕ are the angular velocity and phase of the plane wave, and $\mathbf{R}$ is a rotation matrix that introduces the angle ρ 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, Table 1).

Spatial kernels of geniculate units were modeled as differences of gaussians,

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

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 degrees and 15 degrees 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. $s_\alpha^{\mathrm{ON}} = -s_\alpha^{\mathrm{OFF}}$.

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 $t_1 = t_2 = 0$, $n_1 = n_2 = 2$, $k_1 = 1$, $k_2 = 0.6$, $c_1 = 60\mathrm{s}^{-1}$, $c_2 = 40\mathrm{s}^{-1}$.

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 responses,

$$R_\eta(\boldsymbol{x}) = \left\langle \eta_{x_\eta}(t)\left[\alpha^{\mathrm{ON}}_{x_\alpha}(t) - \alpha^{\mathrm{OFF}}_{x_\alpha}(t)\right]\right\rangle_{\mathcal{I},t} = \langle \eta_{x_\eta}(t)\alpha_{x_\alpha}(t)\rangle_{\mathcal{I},t}, \tag{2.1}$$

where $\alpha(t) = \alpha^{\mathrm{ON}}(t) - \alpha^{\mathrm{OFF}}(t) = k^{\mathrm{ON}}_\alpha(\boldsymbol{x}, t) \star I(\boldsymbol{x}, t)$. While this choice of rectification parameters simplified the mathematical analysis of this letter, our previous simulation data ensure that results remain valid for a wide range of thresholds. In this letter, correlation difference maps were estimated on the basis of equation 2.1, without explicitly simulating cell responses. Examples of traces of neuronal activity can be found in Figure 6 of our previous study (Rucci & Casile, 2004).

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 (Papoulis, 1984),

$$R_\eta(x) = \mathcal{F}^{-1}\{K_\eta(\omega_x, \omega_t)\overline{K_\alpha}(\omega_x, \omega_t)C_{\mathrm{SA}}(\omega_x, \omega_t)\}|_{t=0}, \tag{3.1}$$

where $\mathcal{F}^{-1}$ indicates the operation of inverse Fourier transform, $C_{\mathrm{SA}}(\omega_x, \omega_t)$ is the power spectrum of spontaneous activity in the retina, 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, equation 3.1 gives

$$R_\eta(x) = T\mathcal{F}^{-1}\{S_\eta(\omega_x)\overline{S_\alpha}(\omega_x)S_{\mathrm{SA}}(\omega_x)\}, \tag{3.2}$$

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 Mastronarde (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 σ_{SA} that ranged from 0.18 degree at eccentricity 5 degrees to 0.35 degree at 25 degrees. Use in equation 3.2 of a gaussian approximation for retinal spontaneous activity gives, after some algebraic manipulations, an analytical expression for the structure of correlated activity,

$$R_\eta(x) \propto \widehat{A}e^{-\frac{x^2}{2\hat{\sigma}^2}}\cos(\widehat{\omega}x+\phi) + \widetilde{A}e^{-\frac{x^2}{2\tilde{\sigma}^2}}\cos(\widetilde{\omega}x+\phi) = \widehat{R}_{\eta\alpha}(x) + \widetilde{R}_{\eta\alpha}(x), \quad (3.3)$$

where the parameters are given by:

$$\begin{cases} \widehat{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{cnt}^2 + \sigma_{SA}^2} \\ \widehat{\omega} = \frac{\sigma_\eta^2}{\hat{\sigma}^2}\omega_\eta \\ \widehat{A} = \frac{A_\eta A_{cnt} A_{SA}\sigma_{cnt}\sigma_\eta\sigma_{SA}2\pi}{\hat{\sigma}}e^{-\frac{\sigma_\eta^2\omega_\eta(\hat{\omega}-\omega_\eta)}{2}} \end{cases} \text{and} \begin{cases} \widetilde{\sigma} = \sqrt{\sigma_\eta^2 + \sigma_{srn}^2 + \sigma_{SA}^2} \\ \widetilde{\omega} = \frac{\sigma_\eta^2}{\tilde{\sigma}^2}\omega_\eta \\ \widetilde{A} = \frac{A_\eta A_{srn} A_{SA}\sigma_{srn}\sigma_\eta\sigma_{SA}2\pi}{\tilde{\sigma}}e^{-\frac{\sigma_\eta^2\omega_\eta(\tilde{\omega}-\omega_\eta)}{2}} \end{cases}. \quad (3.4)$$

Substitution of cell-receptive field parameters in equation 3.4 yields $\widehat{A} \gg \widetilde{A}$ at all considered angles of visual eccentricity. Thus, the second term of equation 3.3 can be neglected, and correlation difference maps are described by Gabor functions:

$$R_\eta(x) \approx \widehat{R}_\eta(x) = \widehat{A}e^{-\frac{x^2}{2\hat{\sigma}^2}}\cos(\widehat{\omega}x+\phi). \quad (3.5)$$

Since the spatial receptive fields of modeled V1 units are also represented by Gabor functions, the similarity between correlation difference maps and cortical receptive fields can be quantified by directly comparing the two parameters of the Gabor maps: σ, the width of the gaussian, and ω, the spatial frequency of the plane wave.

Figure 1 compares the correlation difference maps given by equation 3.5 to the receptive fields of modeled V1 units. Since the precise locations of the receptive fields of recorded cells were not reported by Jones and Palmer

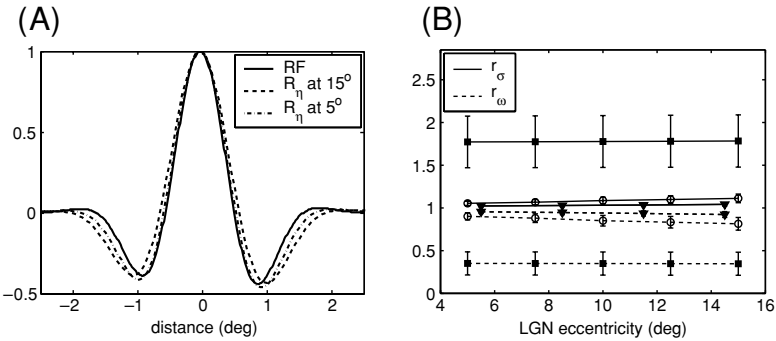

Figure 1: Comparison between the spatial organization of simple cell receptive fields and the structure of thalamocortical activity for retinal inputs with various levels of spatial correlation. (a) Results for one of the 10 modeled simple cells in the case of spontaneous activity. The correlation difference maps (R_η) measured between the considered simple cell and arrays of geniculate units located around 5 and 15 degrees 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. Dashed and solid curves show, respectively, the ratios $r_\omega = \widehat{\omega}/\omega_\eta$ and $r_\sigma = \widehat{\sigma}/\sigma_\eta$ evaluated in the presence of retinal spontaneous activity (○), white noise (▼), and broad spatial correlations ($\sigma_{SA} = 1^\circ$) at the level of the retina (■). 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. Each curve represents average values over 10 modeled simple cells. Error bars represent standard deviations. The x-axis marks the angle of visual eccentricity of the receptive fields of geniculate units.

(1987a, 1987b), we estimated the patterns of correlation that each modeled V1 unit would establish with LGN cells located at various angles of visual eccentricity. Figure 1a shows an example for one of the modeled V1 units. The patterns of correlated activity measured at both 5 and 15 degrees of visual eccentricity closely resembled the receptive field profile of the cortical cell. The curves marked by solid triangles in Figure 1b show, respectively, the mean values of the two ratios $r_\omega = \widehat{\omega}/\omega_\eta$ and $r_\sigma = \widehat{\sigma}/\sigma_\eta$ evaluated over all 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 $\bar{r}_\sigma = 1.08 \pm 0.05$ and $\bar{r}_\omega = 0.86 \pm 0.07$, respectively. Thus, in the model, the structure of thalamocortical activity present immediately before eye opening matched the spatial organization of simple cell receptive fields.

It is important to notice that the similarity between receptive fields and correlation difference maps shown in Figure 1 originated from the narrow

spatial correlations of spontaneous activity. Indeed, when no spatial correlation was present at the level of the retina, that is, when spontaneous activity was modeled as white noise ($C_{SA}(\omega_x, \omega_t) = 1$), correlation difference maps calculated from equation 3.1 were virtually identical to the simulated receptive fields. Mean ratios over all simulated cells and angles of visual eccentricity were $\bar{r}_\sigma = 1.03 \pm 0.02$ and $\bar{r}_\omega = 0.94 \pm 0.03$, indicating that correlation difference maps and cortical receptive fields were highly similar. This similarity did not occur in the presence of large input spatial correlations. For example, in the case of $\sigma_{\mathrm{SA}} = 1^\circ$, the mean matching ratios were $\bar{r}_\sigma = 1.77 \pm 0.31$ and $\bar{r}_\omega = 0.35 \pm 0.14$. This analysis shows that the narrow correlations of spontaneous retinal activity were responsible for the compatibility between the structure of thalamocortical activity and the Hebbian maturation of cortical receptive fields observed in our previous modeling studies (Rucci et al., 2000; Rucci & Casile, 2004).

4 Thalamocortical Activity After Eye Opening

After eye opening, the assumption of narrow spatial correlations 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). Figure 2 examines the impact of these broad input correlations on the structure of thalamocortical activity. Following the approach of section 3, correlation difference maps were given by

$$R_\eta(\boldsymbol{x}) = C\mathcal{F}^{-1}\{S_\eta(\boldsymbol{\omega})\overline{S_\alpha}(\boldsymbol{\omega})\mathcal{N}(\boldsymbol{\omega})\}, \tag{4.1}$$

where $\mathcal{N}(\boldsymbol{\omega})$ is the power spectrum of natural images and C is a multiplicative factor equal to $H_\alpha(0)H_\eta(0)$. The power spectrum $\mathcal{N}(\boldsymbol{\omega})$ was estimated from a set of 15 natural images (van Hateren & van der Schaaf, 1998). Its radial mean was best interpolated by $\bar{\mathcal{N}}(\omega) \propto \omega^{-2.02}$, which is in agreement with previous measurements (Field, 1987; Ruderman, 1994).

Similar to the results of our previous study (Rucci & Casile, 2004), patterns of correlated activity did not match the receptive fields of simple cells during static presentation of natural scenes. An example for one of the 10 modeled simple cells is shown in Figure 2a, which compares the profile of the cell receptive field to sections of the correlation difference maps measured at 5 and 15 degrees of visual eccentricity. The mismatch is particularly evident in correspondence of the side lobes of the receptive field, where levels of correlation predicted stabilization of afferents from geniculate cells with the wrong polarity (ON- instead of OFF-center).

Figure 2b shows average results obtained over the entire population of simulated simple cells. Since, in this case, an analytical expression of $R_\eta(\boldsymbol{x})$ was not available, correlation difference maps obtained by solving

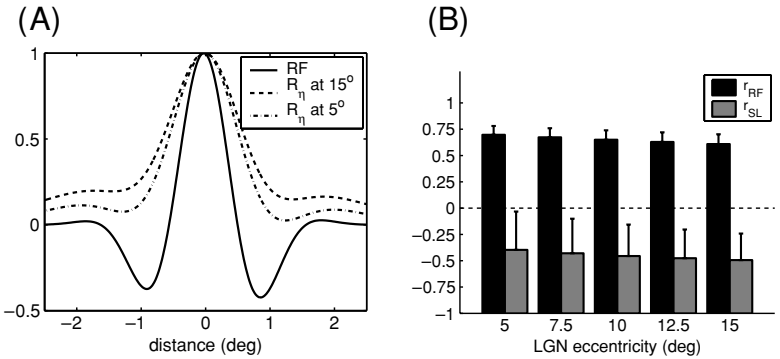

Figure 2: Comparison between the spatial organization of simple cell receptive fields and the structure of thalamocortical activity in the case of static presentation of natural images. (a) Results for one of the 10 modeled simple cells. The two correlation difference maps (R_η) measured between the considered simple cell and arrays of geniculate units located around 5 and 15 degrees 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 the text for details). The x-axis represents the angle of visual eccentricity of simulated geniculate units. Vertical lines represent the standard deviation.

numerically equation 4.1 were compared to receptive fields by means of the 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. In addition to the mean correlation 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. At all considered eccentricities, a clear mismatch was present between correlation maps and receptive fields. Average correlation coefficients were $\bar{r}_{RF} = 0.65 \pm 0.09$ over the entire receptive fields and $\bar{r}_{SL} = -0.45 \pm 0.3$ 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 the broad correlations of natural images was not compatible with a Hebbian refinement of the receptive fields of simple cells.

The results of Figure 2 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 article we model fixational instability by means of a two-dimensional ergodic process $\boldsymbol{\xi}(t) = [\xi_x(t),\ \xi_y(t)]^{\mathrm{T}}$. 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 $I(\widetilde{\boldsymbol{x}})$ of a natural image in the neighborhood of a generic point $\boldsymbol{x}$ can be approximated as

$$I(\widetilde{\boldsymbol{x}}) \approx I(\boldsymbol{x}) + \nabla I(\boldsymbol{x})^{\mathrm{T}} \cdot [\widetilde{\boldsymbol{x}} - \boldsymbol{x}].$$

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

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

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

$$\begin{aligned} \eta_{x_0}(t) &= k_\eta(\boldsymbol{\mu}, \tau) \star I_r(\boldsymbol{\mu}, \tau)|_{(x_0,t)} \approx k_\eta(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu})|_{(x_0,t)} \\ &\quad + k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau)\bigg|_{(x_0,t)} + k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau)\bigg|_{(x_0,t)} \\ &= \eta^{\mathrm{S}}_{x_0}(t) + \eta^{\mathrm{D}}_{x_0}(t) \qquad (4.2) \\ \alpha_{x_1}(t) &= k_\alpha(\boldsymbol{\mu}, \tau) \star I_r(\boldsymbol{\mu}, \tau)|_{(x_1,t)} \approx k_\alpha(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu})|_{(x_1,t)} \\ &\quad + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau)\bigg|_{(x_1,t)} + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau)\bigg|_{(x_1,t)} \\ &= \alpha^{\mathrm{S}}_{x_1}(t) + \alpha^{\mathrm{D}}_{x_1}(t), \end{aligned}$$

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

$$\begin{cases} \eta^{\mathrm{S}}_{x_0}(t) = k_\eta(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu})|_{(x_0,t)} \\ \eta^{\mathrm{D}}_{x_0}(t) = k_\eta(\boldsymbol{\mu}, \tau) \star \dfrac{\partial I(\boldsymbol{\mu})}{\partial \mu_x} \xi_x(\tau)\bigg|_{(x_0,t)} + k_\eta(\boldsymbol{\mu}, \tau) \star \dfrac{\partial I(\boldsymbol{\mu})}{\partial \mu_y} \xi_y(\tau)\bigg|_{(x_0,t)} \\ \qquad = \eta^{\mathrm{D}_x}_{x_0}(t) + \eta^{\mathrm{D}_y}_{x_0}(t) \end{cases}$$

$$\begin{cases} \alpha^{\mathrm{S}}_{\boldsymbol{x}_1}(t) = k_\alpha(\boldsymbol{\mu}, \tau) \star I(\boldsymbol{\mu})\big|_{(\boldsymbol{x}_1,t)} \\ \alpha^{\mathrm{D}}_{\boldsymbol{x}_1}(t) = k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x}\xi_x(\tau)\Big|_{(\boldsymbol{x}_1,t)} + k_\alpha(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_y}\xi_y(\tau)\Big|_{(\boldsymbol{x}_1,t)} \\ \qquad\quad = \alpha^{\mathrm{D}_x}_{\boldsymbol{x}_1}(t) + \alpha^{\mathrm{D}_y}_{\boldsymbol{x}_1}(t). \end{cases}$$

That is, cell responses can be decomposed into a static component with nonzero mean (η^S and α^S) and a zero-mean dynamic component introduced by fixational instability (η^D and α^D). η^{D_x}, α^{D_x}, η^{D_y}, and α^{D_y} are the contributions to cell responses generated by the instability of visual fixation along the x- and y-axes.

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

$$R_\eta(\boldsymbol{x}) = R^{\mathrm{S}}_\eta(\boldsymbol{x}) + R^{\mathrm{D}}_\eta(\boldsymbol{x}). \tag{4.3}$$

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 $\eta_{\boldsymbol{x}_0}(t)$ and $\alpha_{\boldsymbol{x}_1}(t)$ have nonzero means. The first of these terms is given by

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

where $N(\boldsymbol{x})$ is the autocorrelation function of natural images. Since this term depends on only the static components of cell responses, it represents the correlation difference map that would be obtained in the absence of fixational instability (see equation 4.1).

The second term is given by

$$\begin{aligned} &\left\langle \eta^{\mathrm{D}_x}_{\boldsymbol{x}_0}(t)\alpha^{\mathrm{D}_x}_{\boldsymbol{x}_1}(t)\right\rangle_{\xi,\mathcal{I},t} \\ &\quad = \left\langle \left(k_\eta(\boldsymbol{\mu}, \tau) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x}\xi_x(\tau)\Big|_{\boldsymbol{x}_1,t} \right) \left(k_\alpha(\boldsymbol{\mu}, t) \star \frac{\partial I(\boldsymbol{\mu})}{\partial \mu_x}\xi_x(\tau)\Big|_{\boldsymbol{x}_0,t} \right) \right\rangle_{\xi,\mathcal{I},t} \\ &\quad = (s_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N_x(\boldsymbol{\mu})|_{\boldsymbol{x}_1-\boldsymbol{x}_0})\langle h_\eta(\tau) \star h_\alpha(-\tau) \star R_{\xi_x\xi_x}(\tau)|_{\tau=0}\rangle_t \\ &\quad = Ds_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'_x(\boldsymbol{\mu})|_{\boldsymbol{x}}, \end{aligned} \tag{4.4}$$

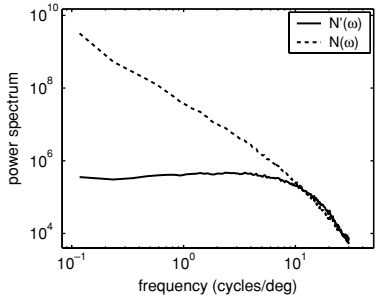

Figure 3: Comparison between the power spectra of natural images $\mathcal{N}(\boldsymbol{\omega})$ and the dynamic power spectrum $\mathcal{N}'(\boldsymbol{\omega}) = \mathcal{N}'_x(\boldsymbol{\omega}) + \mathcal{N}'_y(\boldsymbol{\omega})$ given by the sum of the power spectra of the x and y components of the gradient of natural images. The two curves are radial averages over a set of 15 natural images.

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,

$$\left\langle \eta_{x_0}^{\mathrm{D}_y}(t)\alpha_{\boldsymbol{x}_1}^{\mathrm{D}_y}(t)\right\rangle_{\xi,\mathcal{I},t} = Ds_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'_y(\boldsymbol{\mu})\big|_{\boldsymbol{x}}$$

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

$$\begin{aligned} R_\eta(\boldsymbol{x}) = {} & Cs_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N(\boldsymbol{\mu})|_x \\ & + Ds_\eta(\boldsymbol{\mu}) \star s_\alpha(-\boldsymbol{\mu}) \star N'(\boldsymbol{\mu})|_x = R_\eta^{\mathrm{S}}(\boldsymbol{x}) + R_\eta^{\mathrm{D}}(\boldsymbol{x}), \end{aligned} \tag{4.5}$$

which proves equation 4.3.

Equation 4.5 shows that fixational instability adds a contribution $R_\eta^{\mathrm{D}}(\boldsymbol{x})$ to the correlation map $R_\eta^{\mathrm{S}}(\boldsymbol{x})$ 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 $N(\boldsymbol{x})$ (or, equivalently, its power spectrum $\mathcal{N}(\boldsymbol{\omega})$), the term $R_\eta^{\mathrm{D}}(\boldsymbol{x})$ introduced by the jittering of the eye depends on the autocorrelation function of the gradient of the stimulus, $N'(\boldsymbol{x})$ (or, equivalently, its power spectrum $\mathcal{N}'(\boldsymbol{\omega})$, the dynamic power spectrum).

Figure 3a 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

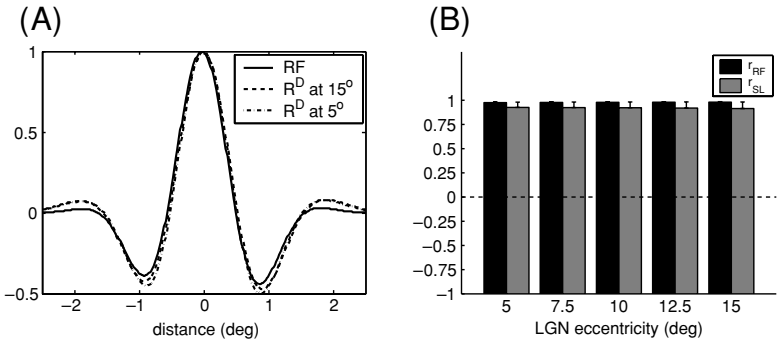

Figure 4: Comparison between the spatial organization of simple cell receptive fields and patterns of correlated activity measured when images of natural scenes were examined in the presence of fixational instability (the term $R_\eta^{\mathrm{D}}(\boldsymbol{x})$ in equation 4.5). The layout of the panels and the graphic notation are the same as in Figure 2.

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, it was uncorrelated. Thus, in the presence of natural images, fixational instability adds an input signal that discards spatial correlations.

It should be observed that the whitening of the dynamic power spectrum is a direct consequence of the scale invariance of natural images and has a simple explanation in the frequency domain. Since the Fourier transforms of the two partial derivatives $\frac{\partial I(\boldsymbol{x})}{\partial x}$ and $\frac{\partial I(\boldsymbol{x})}{\partial y}$ 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) and the product $|\boldsymbol{\omega}|^2\mathcal{N}(\boldsymbol{\omega})$ produce a dynamic power spectrum $\mathcal{N}'(\boldsymbol{\omega})$ with uniform spectral density. In other words, our analysis shows that whereas the intensity values of natural images tend to be correlated over large distances, local changes in intensity around pairs of pixels are uncorrelated. Therefore, fixational instability represents an optimal decorrelation strategy for visual input with power spectrum that declines as $|\boldsymbol{\omega}|^{-2}$.

We have already shown in Figure 2 that the patterns of correlated activity $R_\eta^{\mathrm{S}}(\boldsymbol{x})$ measured with static presentation of images of natural scenes did not match the receptive fields of modeled simple cells. Figure 4 analyzes the contribution of fixational instability, the term $R_\eta^{\mathrm{D}}(\boldsymbol{x})$ in equation 4.5, to the structure of correlated activity. In this case, correlation difference maps closely resembled the spatial organization of cortical receptive fields irrespective of the eccentricity of simulated geniculate units. The mean

matching index was $\bar{r}_{RF} = 0.98 \pm 0.006$ over the entire receptive fields and $\bar{r}_{SL} = 0.92 \pm 0.06$ 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 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.

To summarize, equation 4.5 shows that in the presence of the self-motion of the retinal image that occurs during natural viewing, the second-order statistics of thalamocortical activity depend on both the spatial configuration of the stimulus and how its retinal projection changes during visual fixation. The first component is represented in equation 4.5 by the term R_η^{S}, which depends on the power spectrum of the stimulus $\mathcal{N}(\boldsymbol{\omega})$. The latter component is represented by R_η^{D}, which is determined by the dynamic power spectrum $\mathcal{N}'(\boldsymbol{\omega})$, a spectrum that discards the broad spatial correlations of natural images. Of these two terms, only R_η^{D} matches the spatial organization of simple cell receptive fields (compare Figure 4 with Figure 2). The overall structure of correlated activity is given by the weighted sum of the results of Figures 2 and 4.

During fixational instability, the relative influence of R_η^{S} and R_η^{D} depends on two elements: (1) the powers of the two inputs $\mathcal{N}(\boldsymbol{\omega})$ and $\mathcal{N}'(\boldsymbol{\omega})$ and (2) neuronal sensitivity to both input signals. In natural images, most energy is concentrated at low spatial harmonics. Since $\mathcal{N}'(\boldsymbol{\omega})$ attenuates the low spatial frequencies of the stimulus, it tends to possess less power than $\mathcal{N}(\boldsymbol{\omega})$. For example, for the two power spectra in Figure 3, the ratio of power dynamic/static in the range 0 to 10 cycles/deg was only 0.08. That is, $\mathcal{N}(\boldsymbol{\omega})$ provided over 10 times more power than $\mathcal{N}'(\boldsymbol{\omega})$ within the main spatial range of sensitivity of geniculate cells. However, in equation 4.5, $\mathcal{N}(\boldsymbol{\omega})$ and $\mathcal{N}'(\boldsymbol{\omega})$ are modulated by the multiplicative terms C and D, which depend on the temporal characteristics of cell responses (both C and D) and fixational instability (D only). Since geniculate neurons respond more strongly to changing stimuli than stationary ones, D tends to be higher than C. For example, a retinal image motion with gaussian temporal correlation (the term $R_{\xi_x\xi_x}$ in equation 4.4) characterized by a standard deviation of 30 ms and a mean amplitude of 10, values that are consistent with the instability of fixation of several species, produced a ratio $D/C \approx 950$. Thus, although $\mathcal{N}'(\boldsymbol{\omega})$ provided less power than $\mathcal{N}(\boldsymbol{\omega})$, the weighted ratio of the total power $(D\mathcal{N}'(\boldsymbol{\omega}))/(C\mathcal{N}(\boldsymbol{\omega}))$ in the range 0 to 10 cycles/deg was approximately 76. Since the term R_η^{D} dominated the sum of equation 4.5, the matching between correlation difference maps and receptive fields increased from $\bar{r}_{RF} = 0.65 \pm 0.09$ and $\bar{r}_{SL} = -0.45 \pm 0.3$ (the values obtained with static presentation of natural images) to $\bar{r}_{RF} = 0.90 \pm 0.05$ and $\bar{r}_{SL} = 0.12 \pm 0.55$. That is, 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 postsynaptic elements (Sejnowski, 1977):

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

In the case in which mean activity levels are estimated over periods of fixation, $\overline{\eta} = \eta^{\mathrm{S}}$ and $\overline{\alpha} = \alpha^{\mathrm{S}}$, yielding $R_\eta(\boldsymbol{x}) = R_\eta^{\mathrm{D}}(\boldsymbol{x})$. Thus, the term $R_\eta^{\mathrm{S}}(\boldsymbol{x})$ 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 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

$$\eta(t) = (1 - \beta) \cdot \eta^{\mathrm{S}}(t) + \eta^{\mathrm{D}}(t),$$

where the constant $\beta \in [0, 1]$ defines the degree of attenuation. With this modification, correlation difference maps are given by

$$R_\eta(\boldsymbol{x}) = (1 - \beta) R_\eta^{\mathrm{S}}(\boldsymbol{x}) + R_\eta^{\mathrm{D}}(\boldsymbol{x}). \tag{4.6}$$

Figure 5 compares the receptive fields of simulated simple cells with the correlation difference maps estimated with various degrees of attenuation. It is clear by comparing these data to those of Figure 2 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. A 60% attenuation 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), in the presence of fixational instability, a nonlinear attenuation of simple cell responses leads to a regime of correlated activity that is compatible with a Hebbian refinement of the spatial organization of simple cell receptive fields.

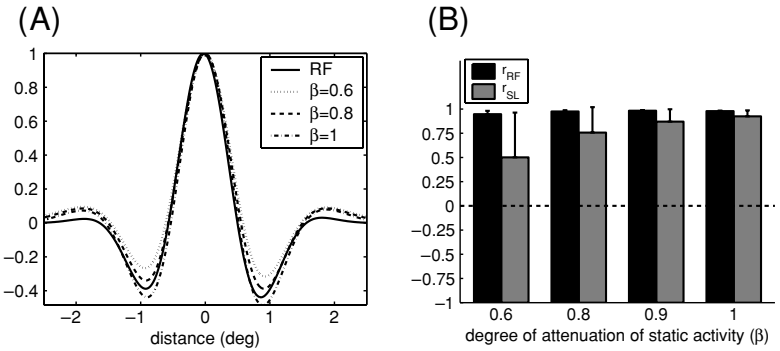

Figure 5: Effect of nonlinear attenuation of simple cells responses to unchanging stimuli. (a) Results for one of the 10 modeled simple cells. The correlation difference maps (R_η) estimated from equation 4.6 for three values of the attenuation factor β are compared to the profile of the simple cell 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 degrees.

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 article, 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 in agreement 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 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). This study shows that in the presence of images of natural scenes, these input fluctuations lack spatial correlations. 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. Thus, our analysis shows that a direct scheme of Hebbian plasticity can be added to the category of activity-dependent mechanisms compatible with the maturation of cortical receptive fields in the presence of decorrelated natural visual input (Law & Cooper, 1994; Olshausen & Field, 1996).

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 disrupt 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 & Raushecker, 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, Macknik, & Hubel, 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).

This study relied on two important assumptions. A first assumption was the use of linear models to predict cell responses to visual stimuli. Linear spatiotemporal models enabled the derivation of analytical expressions of 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 (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 settle on lower steady-state levels 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 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 equation 4.5. However, the results presented in this article 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). Furthermore, although a statistical analysis of the instability of visual fixation under natural viewing conditions has not been performed, the motion of the retinal image as subjectively experienced by a jitter after-effect (Murakami & Cavanagh, 1998) appears to be compatible, at least qualitatively, with our modeling assumptions.

It is worth emphasizing that during natural viewing, other elements, in addition to eye movements, contribute to the instability of visual fixation. In particular, small movements of the head and body and imperfections in the vestibulo-ocular reflex (Skavenski, Hansen, Steinman, & Winterson, 1979) are known to amplify the self-motion of the retinal image. Our analysis aims to address the joint effect of all these movements. It can be shown analytically that the factor D, which in equation 4.5 modulates the impact of a moving retinal stimulus (the term $R^{D}_{\eta}(\boldsymbol{x})$), depends in a quadratic manner on the spatial extent of fixational instability. Therefore, within the limits of validity of the Taylor approximation of equation 4.2, the larger the amplitude of fixational instability, the stronger its influence on the structure of correlated activity. It should also be noted that while this article focuses on the examination of static images of natural scenes, our analysis applies to any jittering stimulus on the retina, regardless of the origin of motion, whether self-generated or external. For example, the trembling of leaves on a tree exposed to the wind might produce a decorrelation of neural activity similar to that of fixational instability.

Our results appear to contrast with a previous proposal according to which the spatial response characteristics of retinal and geniculate neurons are sufficient to decorrelate the spatial signals provided by images of natural

scenes (Atick & Redlich, 1992). According to this hypothesis, a neuronal sensitivity function that increases linearly with the spatial frequency would counterbalance the power spectrum of natural images and produce a decorrelated pattern of neural activity. However, neurophysiological recordings have shown that 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). Such deviation is not compatible with Atick and Redlich's proposal and, in the absence of fixational instability, would lead to a regime of thalamocortical activity strongly influenced by the broad spatial correlations of natural images (see Figure 2). In contrast to this static decorrelation mechanism, the decorrelation of visual input produced 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. A number of recent studies have suggested an important role for fixational instability in the neural encoding of visual stimuli (Ahissar & Arieli, 2001; Greschner, Bongard, Rujan, & Ammermüller, 2002; Snodderly et al., 2001). The results presented here suggest that fixational instability, by decreasing statistical dependencies between neural responses, might contribute to discarding broad input correlations, thus establishing efficient visual representations of natural visual scenes (Barlow, 1961; Attneave, 1954). Further theoretical and experimental studies are needed to characterize and test this hypothesis.

Acknowledgments

We thank Matthias Franz, Alessandro Treves, and Martin Giese for many helpful comments on a preliminary version of this article. This work was supported by the Volkswagen Stiftung, the National Institutes of Health grant EY015732-01, and the National Science Foundation grants CCF-0432104 and CCF-0130851. Correspondence and requests for materials should be addressed to A.C.

References

Ahissar, E., & Arieli, A. (2001). Figuring space by time. *Neuron*, *32*, 185–201.

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, MA: MIT Press.

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.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Res.*, *35*(1), 7–24.

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), 1118–1135.

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.

Greschner, M., Bongard, M., Rujan, P., & Ammermüller, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature*, *5*(4), 341–347.

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.

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.

Mastronarde, D. N. (1983). Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X and Y cells. *J. Neurophysiol.*, *49*, 303–323.

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.

Papoulis, A. (1984). *Probability, random variables and stochastic processes*. New York: McGraw-Hill.

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.*, *21*(5), 725–738.

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., & Raushecker, 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 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.

Tolhurst, 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.

Received January 5, 2005; accepted July 19, 2005.