

A Dynamical Model for Receptive Field Self-organization in V1 Cortical Columns

Jörg Lücke

Gatsby Computational Neuroscience Unit, UCL, London WC1N 3AR, UK

Abstract. We present a dynamical model of processing and learning in the visual cortex, which reflects the anatomy of V1 cortical columns and properties of their neuronal receptive fields (RFs). The model is described by a set of coupled differential equations and learns by self-organizing the RFs of its computational units – sub-populations of excitatory neurons. If natural image patches are presented as input, self-organization results in Gabor-like RFs. In quantitative comparison with *in vivo* measurements, we find that these RFs capture statistical properties of V1 simple-cells that learning algorithms such as ICA and sparse coding fail to reproduce.

1 Introduction

Self-organizing systems are commonly used to study learning in biological networks and/or to learn from a set of presented inputs. Classical examples are systems that learn input categories [1] or systems that are learning the neighborhood relationship of the input data [2]. Based on recent results on the anatomical fine-structure of cortical columns [3], we show how self-organization can be used to extract basic constituents of the input. The presented bottom-up approach is based on earlier work on this subject [4, 5] and shows the applicability of the approach to natural images. We find that the components extracted by the model have a higher degree of similarity with *in vivo* measurements of simple-cells than the classical algorithmic approaches of ICA [6, 7] and sparse coding [8, 9].

2 System Dynamics

Our model column consists of k neuron populations or *hidden units* p_1, \dots, p_k and N input units y_1, \dots, y_N . The inputs $\tilde{I}_1, \dots, \tilde{I}_k$ to the hidden units originate from external neural units y_1, \dots, y_N and influences the hidden units via afferent fibers $R_{\alpha j}$ with $I_\alpha = \sum_j R_{\alpha j} y_j$. We implement a feed-forward inhibition that ensures that the inputs $\tilde{I}_1, \dots, \tilde{I}_k$ sum to zero: $\tilde{I}_\alpha = I_\alpha - \frac{1}{k} \sum_\beta I_\beta$. It follows that $\tilde{I}_\alpha = \sum_j R_{\alpha j}^{\text{eff}} y_j$ with $R_{\alpha j}^{\text{eff}} = R_{\alpha j} - \frac{1}{k} \sum_\beta R_{\beta j}$. $\mathbf{R}_1^{\text{eff}}, \dots, \mathbf{R}_k^{\text{eff}}$ will be referred to as *effective RFs* or just *RFs* of the column. Fig. 1 visualizes the afferents and the internal connectivity of the system. Note that the dynamics is formulated on the population level and that there are different alternatives of its implementation. Fig. 1 is a visualization of entities that are relevant for the dynamics.

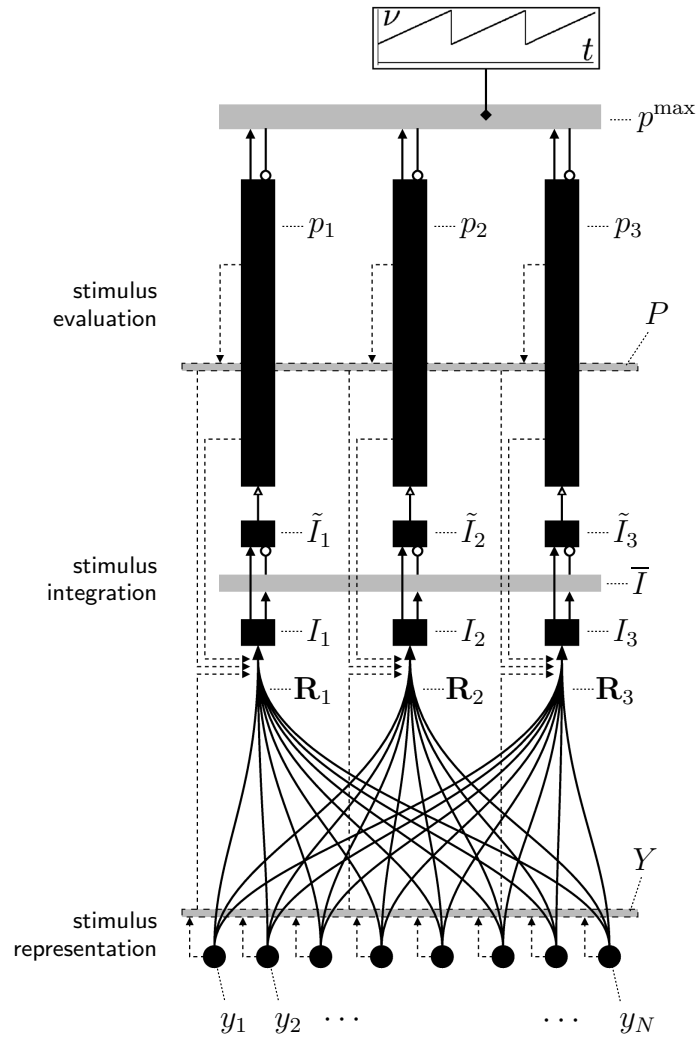


Fig. 1. Sketch of a cortical column with $k = 3$ sub-populations of excitatory neurons (visualized as black vertical bars). Input to the column originates from N external units y_1 to y_N . In a first processing stage the input is integrated. The inputs I_1 to I_3 are via feed-forward inhibition transformed to mean-free inputs \tilde{I}_1 to \tilde{I}_3 . These inputs drive the self-excitatory sub-populations with activities p_1 to p_3 . \bar{I} is the mean input. Triangular arrow hats denote excitatory (solid) and mixed (hollow) influences, empty circles inhibitory influences. Lateral inhibition between the self-excitatory populations is modulated by the bifurcation parameter ν . Dashed lines indicate the influence of dynamic variables on the modification of the populations' receptive fields \mathbf{R}_1 to \mathbf{R}_3 . P and Y are the sums of input unit activities and population activities, respectively. p^{\max} is equal to the greatest population activity.

The hidden units of the system model sub-networks of excitatory neurons as found in cortical columns [3]. We use a polynomial approximation of such self-excitatory units as suggested in [10]. Together with a particular inhibitory coupling between the self-excitatory units the dynamics is given by:

$$\nu(t) = (\nu_{\max} - \nu_{\min})\tilde{t} + \nu_{\min}, \quad \tilde{t} = \frac{1}{T}\text{mod}(t, T), \quad (1)$$

$$\frac{d}{dt}p_\alpha = a \left(\underbrace{p_\alpha^2}_{\text{self-excitation}} - \underbrace{\nu(t)p_\alpha \max_{\beta=1, \dots, k} \{p_\beta\}}_{\text{lat. inhibition}} - \underbrace{p_\alpha^3}_{\text{self-inhibition}} \right) + \underbrace{\kappa \tilde{I}_\alpha}_{\text{input}} + \underbrace{\sigma p_\alpha \eta_t}_{\text{noise}}, \quad (2)$$

where κ is the coupling to input \tilde{I}_α and where σ parameterizes multiplicative Gaussian white noise. Eqn. 1 represents a linear increase of the dynamics' bifurcation parameter ν . The time \tilde{t} runs from 0 to 1 within the time interval $[0, T)$. After time $t = T$ a new input is presented and the increase of ν starts anew. We refer to such a cycle as ν -cycle. Dynamics (1) and (2) implement a particular kind of lateral competition between the hidden units that has proven to be advantageous for learning distributed input encodings. The parameter ν increases competition between the hidden units, and initially active units are deactivated during a ν -cycle. The dynamics of neural activity, (1) and (2), has been studied earlier [10] and represents an abstraction of a model that was based on sub-populations of explicitly modeled excitatory neurons [4].

The non-linear evaluation of an input pattern according to (1) and (2) couples into a dynamics of Hebbian-type synaptic plasticity given by:

$$\frac{d}{dt}R_{\alpha j} = \frac{\epsilon}{N} ([p_\alpha]^+ y_j - [p_\alpha]^+ Y R_{\alpha j}) \quad \text{iff } P(t) < \chi, \quad (3)$$

where $P = \sum_{\alpha=1}^k p_\alpha$ and $Y = \sum_{j=1}^N y_j$ are overall-activities of hidden units and input units, respectively, and where $[p_\alpha]^+ = p_\alpha$ if $p_\alpha \geq 0$ and $[p_\alpha]^+ = 0$ otherwise. To learn with slowly increasing competition between the RFs¹ we change, after each ν -cycle, the threshold for learning χ and the maximal level of lateral inhibitory coupling ν_{\max} :

$$\Delta\chi = -\lambda_\chi (\chi - a_\chi P) \quad \text{and} \quad \Delta\nu_{\max} = -\lambda_\nu (a_\nu - P), \quad (4)$$

where λ_χ and λ_ν are modification rates. The second equation increases ν_{\max} to counteract the effect of RFs that are increasingly specialized to the input.

Self-organization of an initially unstructured system is commonly characterized by: (A) random fluctuations that result in structural seeds, (B) a positive feed-back loop amplifying certain structures or modes, and (C) a negative feed-back loop that counteracts amplification and finally keeps the system in an active equilibrium. For our system we start in a state with all afferents initialized to the same value $R_{\alpha j} = \frac{1}{N}$. (A) After an input is presented, ν is increased and the noise in (2) breaks the symmetry among the activities p_α (compare [10]). Hidden

¹ which is related but not identical to the competition between the hidden units

units are deactivated until $P(t) < \chi$. During the remainder of the ν -cycle the RFs are modified according to their activities (3). This implies that just some or only one RF is significantly changed to become more similar to the presented input. (B) Hidden units with RFs similar to a certain type of input are likely to remain active if such an input is presented. Their RFs will therefore further specialize to this input type. (C) Lateral inhibition in (2) forces the system to specialize to different input patterns and the negative term in (3) prevents the afferents from growing infinitely.

For our dynamics the control of lateral inhibition represents the crucial part. If we learn *after* inhibition selects a single unit, i.e. as in winner-take-all (WTA) networks, a distributed encoding of presented input is not observed even if it consists of easily to identify combinations of basic components. The same applies if we learn *after* inhibition has selected, e.g., K units (K -WTA). In contrast, dynamics (3) modifies RFs *during* the process of deactivation of hidden units. Furthermore, learning favors input that results in a relatively sparse activation of the hidden layer (small $P(t)$). If we learn according to dynamics (1) to (4) and if the presented input consists of combinations of basic constituents, the system self-organizes its RFs to represent these constituents. Continuously increasing competition between the RFs (4) crucially helps in guiding the self-organization process. With increasing competition groups of initially similar RFs (we initialize with a relatively large χ) decay into ever smaller sub-groups. For a similar system, such a type of self-organization was therefore termed *hierarchical self-organization* in [11].

Equations (1) to (4) represent a dynamical model of a cortical column. Before applying the dynamics to input, we have to find a set of parameters that lets the system operate in the interesting non-linear regime between no competition and competition with WTA characteristics. Choosing parameters is straightforward and good results can be obtained for a large range of different parameters. To determine the particular set of parameters used in this paper² we have tuned the parameters using the so-called bars test [12] as a benchmark. Once the set of parameters is chosen, the system is extraordinarily robust with respect to different types of input.

3 Simulation Results

To model input to our dynamics that resembles input received by cortical columns in V1, we will use gray-level images patches as input. We expect the column model to self-organize its RFs such that the activities of hidden units can represent the input distributedly. Input is taken from 20 randomly selected images of the van Hateren database [7] that do not show man-made structures. We use difference of Gaussians (DoG) transformed versions of these images to emulate the preprocessing of visual input by the retina and the lateral geniculate nucleus (LGN). We used a standard deviation of $\sigma_+ = 1.0$ pixels for the positive part

² Eqn. 1: $\nu_{\min} = 0.4$, $T = 25\text{ms}$; Eqn. 2: $a = 200\text{ms}^{-1}$, $\kappa = 1.0\text{ms}^{-1}$, $\sigma = 0.01\text{ms}^{-1}$; Eqn. 3: $\epsilon = 0.02$; Eqn. 4: $\lambda_\chi = 5 \times 10^{-5}$, $a_\chi = 1.2$, $\lambda_\nu = 10^{-3}$, $a_\nu = 0.7$.

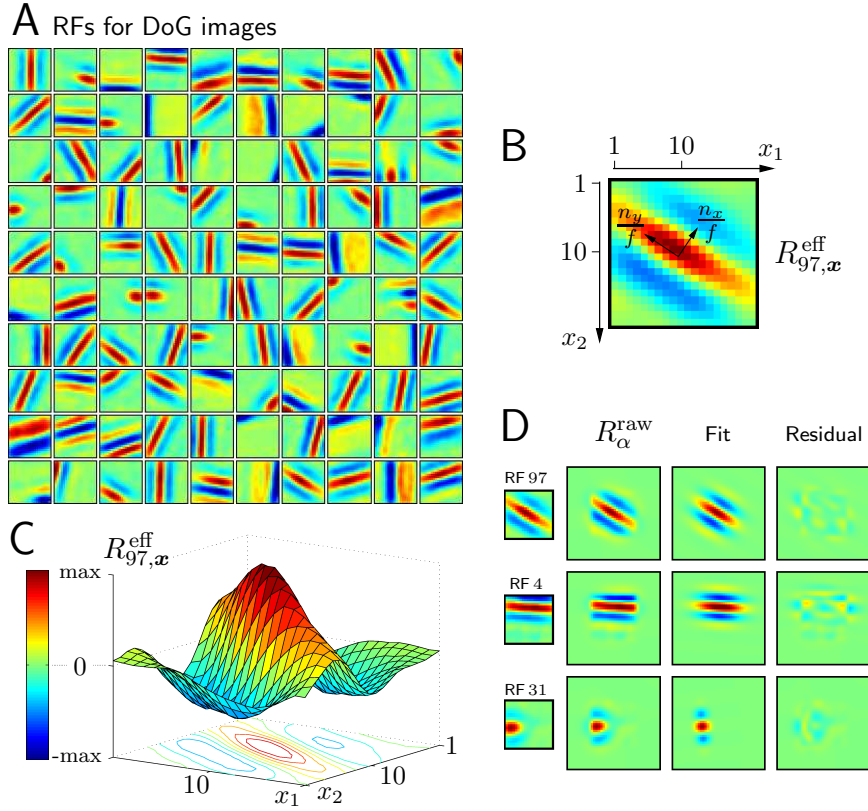


Fig. 2. RFs of the model column if natural images are used as input. **(A)** Effective RFs $\mathbf{R}_{\alpha}^{\text{eff}}$ after 2×10^6 ν -cycles if difference of Gaussians (DoG) filtered images are used as input. For each of $k = 100$ RFs, values are color coded to lie between dark blue (-max) and dark red (+max) where max is the maximal absolute value of the RF, $\text{max} = \max_j |R_{\alpha_j}^{\text{eff}}|$, which ensures that $\mathbf{R}_{\alpha}^{\text{eff}} = 0$ is assigned to the same color (green) for all RFs. See **(C)** for the color coding scheme. **(B)** Effective RF $\alpha = 97$ of the simulation displayed in **(A)**. The index j is replaced by the two-dimensional vector \mathbf{x} . The same coding scheme as in **(A)** is used with $\text{max} = 1.67 \times 10^{-3}$ in this case. The small arrows in the center represent the principal axes of the Gaussian envelope after the RF was matched with a Gabor wavelet. The lengths of the arrows are the wavelet's standard deviations $\sigma_x = \frac{n_x}{f}$ and $\sigma_y = \frac{n_y}{f}$. The dimensionless entities n_x and n_y (f is the wavelet frequency) are used for further analysis (see Fig. 4). **(C)** The same RF as in **(B)** plotted in three dimensions to illustrate the color coding. **(D)** Three examples of matching the RFs with Gabors. RFs 4 and 31 illustrate artifacts of rectangular sampling and Gabor matching, respectively. Columns show original RFs (1st column), corresponding filters that act on the raw pixel images (2nd column), Gabor wavelet matches of the raw filters (3rd column), and differences (residuals) of raw filters and Gabor fits, which shows the error made by the wavelet approximation (4th column).

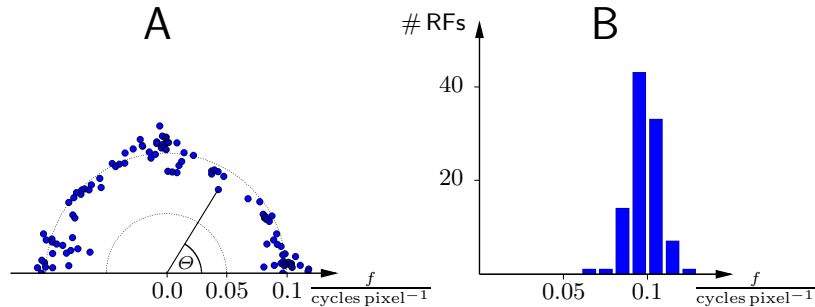


Fig. 3. Orientation and frequency distribution of RFs. **(A)** Distribution of frequency vs. angle for the RFs displayed in Fig. 2A after being transformed to filters on the raw pixel values and matched with Gabor wavelets. RFs are relatively homogeneously distributed in orientation space apart from a preference of horizontal and vertical RFs. **(B)** Distribution of RFs in frequency space. Receptive fields cluster around $f \approx 0.1 \frac{\text{cycles}}{\text{pixel}}$ and cover approximately one octave.

and $\sigma_- = 3.0$ pixels for the negative part, consistent with the biologically measured ratio [13] of $\frac{\sigma_+}{\sigma_-} \approx \frac{1}{3}$. From the 20 images we randomly selected patches of $N = 20 \times 20$ pixels whose values were scaled linearly to lie in the interval $[0, 1]$. In Fig. 2A the effective RFs of a system with $k = 100$ hidden units are shown for DoG preprocessed input after being trained for 2×10^6 ν -cycles³. The RFs have the familiar Gabor-like shape (see Fig. 2C) and represent filters acting on the already DoG transformed images. For comparison with physiological data as obtained in simple-cell recordings, we first have to compute the corresponding filters that act on the raw images. For DoG preprocessed input this amounts to a convolution of the RFs in Fig. 2A using the same DoG-kernel as for the preprocessing of the image. In Fig. 2D the resulting filters are shown for three examples. The filters are, in theory, infinitely large but are virtually zero for all pixels outside a central region of 40×40 pixels. For further analysis, and as is customary in the literature, we match the real filters using Gabor wavelet functions (see e.g. [14, 15]). Matching works well in most cases but the artificial rectangular sampling can result in notable artifacts. The effect of these artifacts on the later analysis can be well understood, however, and they will be discussed below using RFs 4 and 31 in Fig. 2D.

An analysis of the parameters of the matched Gabors shows a relatively even distribution of RF positions (data not shown). Plotting orientation vs. frequency (Fig. 3A) shows a distribution similar to the ones obtained by using independent

³ To reduce undesirable boundary effects, image patches (20×20 pixels) are large compared to patch sizes used by other methods (e.g. 12×12 in ICA [6] and sparse coding [8, 9]). Larger patch sizes exceed the already extensive computational resources required to simulate dynamics (1) to (4). Note that the range of preferred spatial frequencies is determined by the DoG preprocessing (Fig. 3B). Increasing this frequency with a smaller DoG kernel is not possible because the standard deviation of its positive part is already as small as one pixel.

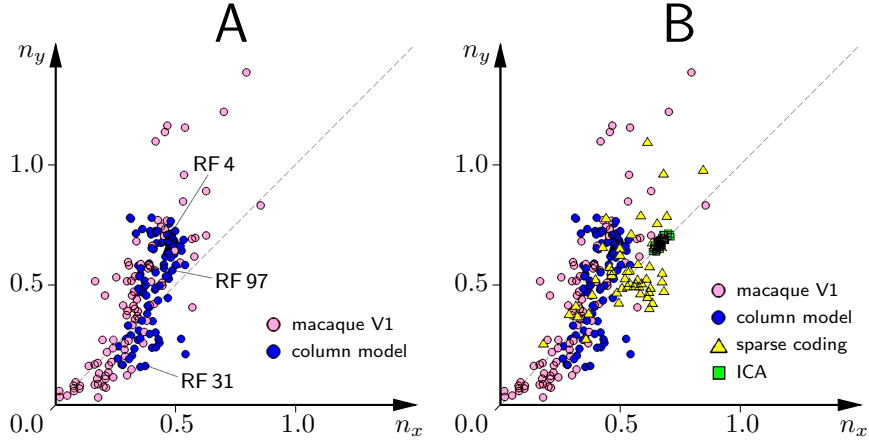


Fig. 4. Distributions of RF extensions parallel and orthogonal to the RFs' wave vector (compare Fig. 2B). n_x and n_y are parameters of wavelets that were matched to the RFs acting on the raw input, \mathbf{R}^{raw} (compare Fig. 2D). n_x is the size of the Gaussian envelope in wave vector direction expressed in terms of the wave length of the wavelet and n_y is the size of the Gaussian envelope orthogonal to the wave vector. **(A)** Distribution in the n_x/n_y -plane of the $k = 100$ RFs displayed in Fig. 2A (blue) together with the distribution of macaque simple-cells (bright magenta) as measured *in vivo* [15]. The data points of the three RFs displayed in Fig. 2D are explicitly labeled. **(B)** The distributions displayed in **(A)** overlaid with the corresponding distributions of RFs obtained using sparse coding (yellow) as described in [9] and ICA (green) as described in [7]. Data are taken from [15]. The dashed diagonal line is the bisection line.

component analysis (ICA) [6, 7] and sparse coding [8, 9]. The orientation preferences are relatively evenly distributed apart from a stronger preference for vertical and horizontal orientations. As for sparse coding and ICA, and unlike RFs measured *in vivo*, the RFs obtained in our model are clustered around a preferred frequency which is given by $f \approx 0.1 \frac{\text{cycles}}{\text{pixel}}$ (see Fig. 3). In our case, the frequency distribution is largely explained by the bandpass properties of the DoG preprocessing, which prefers frequencies in the range of the obtained filters.

The most notable feature of the RFs in Fig. 2A is the variation in the widths and lengths of their Gaussian envelopes. In [15] an analysis of properties of the envelopes relative to the wavelets' frequency was suggested as a means for comparing the RFs of computational models with data. For quantitative comparison the RF parameters are plotted in the n_x/n_y -plane with $n_x = \sigma_x f$ and $n_y = \sigma_y f$ where f is the frequency of the matched wavelet and σ_x and σ_y are, respectively, the standard deviations of the Gaussian envelope in the direction of, and perpendicular to, the wave vector (compare Fig. 2B). Fig. 4A shows the distribution predicted by our model together with *in vivo* measurements of simple-cell RFs in macaque primary visual cortex [15]. Both distributions have similar variance and show the same correlation between n_x and n_y , with a preference for RFs to

be elongated in the n_y -direction if they are distant from the origin. Note that the RF distribution in cat primary visual cortex also shows this property [14, 15]. Two notable differences between the measured distribution and the RFs of the model are the absence of model RFs with values near the origin and with values far away from it.

The absence of RFs distant from the origin can be explained by the chosen data representation which restricts RFs to a rectangular patch size. As can be seen in Fig. 2D (2nd row), the Gabor match results in a Gaussian envelope that is artificially restricted in n_y -direction. Because of the range of preferred frequencies, Fig. 3, this prevents values of n_y from being larger than $n_y \approx 0.8$. The cluster of data points near (0.5, 0.7) in Fig. 4A can therefore be interpreted as a consequence of rectangular patch sizes (compare RF 4). Less artificial or larger patches would move some of these points, e.g. RF 4, further away from the origin and presumably closer to the measured data.

RFs with values near to the origin are associated with what was called ‘globular’ RFs in [15], i.e., RFs with no or weak orientation preference. Although RFs with weak orientation preference are actually developed by our system (see, e.g., RF 31 in Fig. 2D), the plot in Fig. 4 does not show any points near the origin. The reason for this is that we use Gabor wavelets to match localized RFs whose positive and negative parts sum to zero. Even in the case of perfectly radial symmetric RFs, Gabor matching would break the symmetry to a preferred orientation (compare Fig. 2D, 3rd row). RFs measured *in vivo* have values near the origin in Fig. 4 because they are not subject to this effect. Many of the simple-cell RFs in [15] are therefore essentially matched by a Gaussian, which represents a degenerated wavelet with zero frequency.

Comparison of the distribution predicted by our model and distributions predicted by sparse coding [8, 9] and ICA [6, 7] are shown in Fig. 4B. Sparse coding seems to partly predict the measured n_y/n_x -correlation but shows an incorrect preference for RFs elongated in n_x -direction near (0.5, 0.5). The distribution of RFs obtained using ICA is concentrated in a small region on the bisecting line near (0.7, 0.7). ICA neither predicts the variety of Gaussian envelopes nor any preference for RF elongation in any direction.

In contrast to sparse coding and ICA, the distribution predicted by our model is in good agreement with *in vivo* measurements (see Fig. 4). It has to be clearly stated, however, that the results depend on various choices made in the experiment and the analysis. In particular these are: the DoG preprocessing, rectangular patches that restrict RFs, and Gabor matching which can artificially remove globular RFs. Likewise, different preprocessing and analysis techniques can affect the RF properties of other computational systems (see [15] for a discussion). Bearing in mind these various influences, it can nevertheless be stated that the dynamics described in this paper captures a property of simple-cell RFs that has not been reproduced by earlier systems. That is, the broad distribution of RFs in the n_y/n_x -space and their n_y/n_x -correlation with elongation of RFs in n_y -direction if they are distant from the origin (Fig. 4). For our RFs this property is already recognizable by considering the raw RF data as displayed in Fig. 2A.

4 Discussion

Motivated by cortical interconnectivity and earlier modeling work we have developed and analyzed a dynamical bottom-up model of a cortical column. The system models the first stages of columnar processing and self-organization of afferent fibers. Its dynamics is based on self-excitatory populations, feed-forward inhibition, and modulation of lateral inhibitory coupling (compare [3] and [16]). Anatomically the model predicts that two stages of processing (see Fig. 1) are required to enable the emergence of a distributed stimulus encoding. The integration stage has linear response properties and projects to an evaluation stage with non-linear responses of neurons. For V1 our model predicts that these two stages are a pre-requisite for the emergence of Gabor-like RFs. For natural images, no simple-cell-like RFs were obtained without feed-forward inhibition. With the use of feed-forward inhibition, self-organization generated a rich diversity of RFs if randomly selected and DoG filtered image patches were presented. Similar to properties of simple-cell RFs and classical models thereof the obtained RFs show sensitivity to different spatial orientations, frequencies, and locations. Furthermore, as analyzed by matching Gabor wavelets, the RFs show a specific variety in the extents of their Gaussian envelopes relative to their frequencies. This feature is consistent with *in vivo* measurements of simple-cell like RFs [14, 15] (Fig. 4A) and has earlier not been reproduced to such an extent as reported here. Only very recently, in two functionally motivated approaches developed in parallel to this work, distributions of Gaussian envelopes comparable to the one presented here were reported. The resulting distributions in [17] are broader than *in vivo* measurements, however. In the model suggested in [18], which implements a form of sparse coding, RF distributions are obtained that contain globular RFs. These RFs can be matched by degenerated Gabor wavelets and correspond to points near the origin of an n_x/n_y -plot. Otherwise the RFs are reminiscent of those obtained by sparse coding (compare Fig. 4B). That is, they partly match the measured distributions well but show, distant from the origin, numerous RFs that are incorrectly elongate in n_x - instead of n_y -direction. Also note that the model in [18] was tuned to fit the data. The classical models for the emergence of simple-cell RFs [8, 7, 6], in particular ICA, do not accurately reproduce the experimental data (see Fig. 4B). Bottom-up models for the emergence of Gabor-like RFs include BCM [19] and CBA [20], which model learning based on single neurons. Quantitative comparison of the RFs obtained with these systems is difficult. To the knowledge of the author, no data about the variability of Gaussian envelopes (as used in Fig. 4) is available for BCM; and RFs obtained with CBA do not seem localized enough for an analysis using Gabor matching.

To conclude, we have defined and simulated a dynamical model of a cortical column. The dynamics evaluates input using a balance between excitation and two forms of inhibitory interaction. If coupled to Hebbian synaptic plasticity, the dynamics induces a self-organization process of afferent fibers. If natural images are used as input, self-organization results in Gabor-like RFs of columnar sub-populations. These RFs match the variability of simple-cell RFs better than classical methods. The dynamics models important functions associated with

cortical columns and represents, by directly combining cortical anatomy and the emergence of neuronal RFs, a coherent model of the first stages in columnar processing.

Acknowledgments. I gratefully acknowledge funding by the Gatsby Charitable Foundation and by the EU project FP6-2005-015803.

References

1. G. A. Carpenter and S. Grossberg. Adaptive resonance theory. In Arbib, ed., *The handbook of brain theory and neural networks*, 87 – 90. MIT Press, 2003.
2. T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, 1995.
3. Y. Yoshimura, J. L. Dantzker, and E. M. Callaway. Excitatory cortical neurons form fine-scale functional networks. *Nature*, 433:868 – 873, 2005.
4. J. Lücke and C. von der Malsburg. Rapid processing and unsupervised learning in a model of the cortical macrocolumn. *Neural Computation*, 16:501 – 533, 2004.
5. J. Lücke and J. D. Bouecke. Dynamics of cortical columns – self-organization of receptive fields. In *Proc. ICANN*, LNCS 3696, pages 31 – 37. Springer, 2005.
6. A. J. Bell and T. J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision Research*, 37(23):3327 – 38, 1997.
7. J. H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B*, 265:359 – 366, 1998.
8. B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607 – 609, 1996.
9. B. A. Olshausen. *Probabilistic Models of the Brain: Perception and Neural Function*, chapter Sparse Codes and Spikes, pages 257 – 272. MIT Press, 2002.
10. J. Lücke. Dynamics of cortical columns – sensitive decision making. In *Proc. ICANN*, LNCS 3696, pages 25 – 30. Springer, 2005.
11. J. Lücke. Hierarchical self-organization of minicolumnar receptive fields. *Neural Networks*, 17/8–9:1377 – 1389, 2004.
12. P. Földiák. Forming sparse representations by local anti-Hebbian learning. *Biological Cybernetics*, 64:165 – 170, 1990.
13. D. Somers, S. Nelson, and M. Sur. An emergent model of orientation selectivity in cat visual cortical simple cells. *The Journal of Neuroscience*, 15:5448 – 5465, 1995.
14. J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1233 – 1258, 1987.
15. D. L. Ringach. Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*, 88:455 – 463, 2002.
16. J. L. Dantzker and E. M. Callaway. Laminar sources of synaptic input to cortical inhib. interneurons and pyram. neurons. *Nature Neuroscience*, 3:701 – 707, 2000.
17. S. Osindero, M. Welling, and G. E. Hinton. Topographic product models applied to natural scene statistics. *Neural Computation*, 18:381 – 414, 2006.
18. M. Rehn and F. T. Sommer. A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. *Journal of Computational Neuroscience*, 2006.
19. E. L. Bienenstock, L. N. Cooper, and P. W. Munro. Theory for the development of neuron selectivity. *The Journal of Neuroscience*, 2(1):32 – 48, 1982.
20. J. Jerez, M. Atencia, and F. J. Vico. A Learning Rule to Model the Development of Orientation Selectivity in Visual Cortex. *Neural Processing Letters*, 21:1-20, 2005.