Macrocolumns as Decision Units

Jörg Lücke[1], Christoph von der Malsburg[1,2], and Rolf P. Würtz[1]

[1] Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany
[2] Computer Science Dept., University of Southern California, Los Angeles, USA

Abstract. We consider a cortical macrocolumn as a collection of inhibitorily coupled minicolumns of excitatory neurons and show that its dynamics is determined by a number of stationary points, which grows exponentially with the number of minicolumns. The stability of the stationary points is governed by a single parameter of the network, which determines the number of possibly active minicolumns. The dynamics symmetrizes the activity distributed among the active columns but if the parameter is increased, it forces this symmetry to break by switching off a minicolumn. If, for a state of maximal activity, the parameter is slowly increased the symmetry is successively broken until just one minicolumn remains active. During such a process minor differences between the inputs result in the activation of the minicolumn with highest input, a feature which shows that a macrocolumn can serve as decision and amplification unit for its inputs. We present a complete analysis of the dynamics along with computer simulations, which support the theoretical results.

1 Introduction

The cerebral cortex can be subdivided into *neural modules*, which are associated to different magnitudes of spatial scale ranging from *areas* of size of approximately $20cm^2$ (in humans), *maps* ($\approx 5cm^2$) to *macrocolumns* ($\approx 0.5mm^2$) and *minicolumns* of about $0.003mm^2$. The minicolumns are considered the smallest neural modules containing several tens up to a few hundred neurons, which are stacked orthogonal to the cortical surface. The grouping into minicolumns can be revealed by Nissl stains, as was first done by Cajal, by stimulus-response experiments or, more recently, by direct measurements of neural connectivity, e.g. [1]. The minicolumns themselves can be grouped into macrocolumns, neural modules, which are best studied in primary sensory areas and which are considered to process stimuli from the same source such as an area of the visual field or a patch of the body surface [2]. Although different features of columns of regions concerned with different levels of information processing can vary significantly it is widely believed that (1) the neural circuits, at least for neural modules of small scales, are of a common design which makes them universal for various computational tasks and (2) that the understanding of the interplay of the circuitries of the different modules presents the key to the understanding of information processing of the brain of vertebrates.
Several models of neural networks reflecting the modular organization of the

brain have been suggested. They range from models based on random interconnections [3] and Hopfield-like models [4] to models based on self-organizing interconnections, e.g. [5]. In this paper we study a model of a single macrocolumn that consists of a collection of coupled minicolumns and we show that it can serve as a *decision unit*, which changes its activity state by a process of symmetry breakings delicately depending on the relative inputs to the minicolumns. The model shows a dynamic behavior that differs from so far suggested ones and that makes possible the construction of networks, in which macrocolumns as principal units communicate via symmetry differences of their inputs. In such networks a macrocolumn can individually change its activity state if these differences are sufficiently non-ambiguous. Here we will be concerned with the dynamics of a single macrocolumn: In Sec. 2 we study the dynamics of a minicolumn, in Sec. 3 the macrocolumn dynamics as a coupled system of minicolumn dynamics is investigated and its properties are discussed, and in Sec. 4 we summarize the results and give a short outlook to future work.

2 Dynamics of Minicolumns

We consider a minicolumn as a network of N neurons, which are excitatorily interconnected. The neurons are modeled as threshold devices with refraction time. Thresholds and refraction times are equal for all neurons. The state of a neuron at time t, $n_i(t)$, is *one* if the neuron is active and *zero* if it is not. A non-refractory neuron is active at time $(t+1)$ if the input it receives from other neurons at time t exceeds Θ. The refraction time is chosen as one time step. The input from an active neuron j to a neuron i is given by the synaptic strength T_{ij}. The resulting dynamics is given by $(i = 1, \ldots, N)$:

$$n_i(t+1) = \mathcal{S}(\sum_{j=1}^{N} T_{ij}\, n_j(t) \; - \Theta) \cdot \underbrace{\mathcal{S}(1 - n_i(t))}_{refraction}, \quad \mathcal{S}(x) := \begin{cases} 0 \text{ if } x \leq 0 \\ 1 \text{ if } x > 0 \end{cases} . \qquad (1)$$

The connectivity is random. Each neuron has s synapses on its axon, and each synapse connects to any post-synaptic neuron with probability $\frac{1}{N}$. All synaptic weights are equal to a constant $c > 0$. Note that the resulting interconnection can include multiple connections between neurons, i.e. $T_{ij} > c$. Considering the dynamics (1) it is always possible to compensate any value of $c > 0$ by an appropriate choice of Θ. Without loss of generality, we choose c to be equal to $\frac{1}{s}$ in oder to normalize the sum over all T_{ij}, $\frac{1}{N} \sum_{i,j=1}^{N} T_{ij} = 1$. Together with the constant probability for a synapse to connect to any post-synaptic neuron we can describe (1) by a simplified dynamics for a single global observable of the network — the probability $p(t)$ of a neuron to be active at time t. The probability, $p_i(t+1)$, of a neuron i to be active at time $t+1$ depends on the probability, $P_i^A(t)$, to receive enough input and on the probability, $P_i^B(t)$, of the neuron to be non-refractory, and we assume these probabilities to be approximately independent, $p_i(t+1) = P_i^A(t)\, P_i^B(t)$. The probability $P_i^B(t)$ is simply given by the complement of the probability of the neuron i to be active at time

t, $P_i^B(t) = (1 - p_i(t))$. For the computation of $P_i^A(t)$ we have to estimate the number of excitatory post-synaptic potentials (EPSPs) received by neuron i at time t. Due to the assumptions made above the probability $P_e(x)$ of the neuron i to receive exactly x EPSPs is given by the binomial distribution,

$$P_e(x) = \binom{N s\, p(t)}{x} (\frac{1}{N})^x (1 - \frac{1}{N})^{N s\, p(t) - x}. \tag{2}$$

In order to facilitate later calculations we approximate (2) by a Gaussian distribution ($\bar{x} = s\, p(t)$, $\sigma^2 = s\, p(t)$ for $N \gg 1$). Integrating all probabilities for numbers $x > s\, \Theta$ we finally receive a compact and easy to handle description of the dynamics (1) in terms of the activation probability $p(t)$,

$$p(t+1) \;=\; \Phi_s(\frac{p(t) \,-\, \Theta}{\sqrt{p(t)}}) \;\; (1 - p(t)) \tag{3}$$

where $\Phi_s(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{s}\, x} e^{-\frac{1}{2} y^2}\, dy$. Equation (3) can be used to reproduce calculation results of [3] where (2) was approximated by a Poisson distribution. Here we will exploit (3) to introduce a special kind of inhibitory feedback.
Inhibitory neurons differ more substantially in their properties than just generating negative post synaptic potentials. On average, inhibitory neurons have thicker axons than excitatory ones and their post-synaptic targets are concentrated on the cell bodies and the proximal dendrites. This suggests to model inhibition as being generated faster than excitation. We will therefore model the inhibitory influence on the excitatory neurons to be present already in the next time-step. The dependency of the inhibition $I(t)$ on the activity of the excitatory neurons we choose to be proportional to the global activity $B(t) = \sum_{i=1}^{N} n_i(t)$ and further demand that it is equally sensed by all neurons. Such a dependency turns out to stabilize the activity in the most efficient way. Replacing Θ in (3) by $I(t) + \Theta_o$ with $I(t) \;=\; \mu\, \frac{B(t)}{N} \;=\; \mu\, p(t)$ yields:

$$p(t+1) = \Phi_s(\frac{(1-\mu)\, p(t) \,-\, \Theta_o}{\sqrt{p(t)}}) \;\; (1 - p(t))\,. \tag{4}$$

It is now possible to determine the maximal values of stationary activity $\mathcal{P}$ for parameters s, Θ_o, and μ by numerically computing the function

$$\mathcal{P}_{s,\Theta_o}(\mu) := \max\{\, p \mid p \;=\; \Phi_s(\tfrac{(1-\mu)\, p - \Theta_o}{\sqrt{p}})\; (1 - p)\}\,. \tag{5}$$

Function (5) can be compared to the values of stable stationary activity obtained by directly simulating equation (1) with $\Theta = \mu\, \frac{B(t)}{N} + \Theta_o$. For the simulations one has to consider coherence effects which are due to possibly cycling neuron activities but which can be suppressed by noise or by varying the number of possibly active synapses after each time-step (see [3]). For $s = 20$, $\Theta_o = \frac{1}{20}$, and $\mu \in [0, 2]$, e.g., it turns out that the predicted activity rates match the measured ones with absolute errors around 0.01 for values of p between 0.5 and 0.05. For lower activity rates the approximation of (2) by a Gaussian distribution gets too coarse. If the coherence effects are not suppressed, the dynamic behavior can differ significantly from the computed one.

3 Inhibitorily Coupled Minicolumns

We now consider a system of k inhibitorily coupled minicolumns. The calculations are independent of the number of minicolumns such that the results are applicable to relatively small macrocolumns ($k \approx 2, ..., 10$) as suggested, e.g., by short-ranging lateral inhibiting cells [6] [7] or to macrocolumns of about $0.5mm^2$ for k of size of several hundreds. Each minicolumn consists of M excitatory neurons with interconnection as above. In analogy to (1) the dynamics is described by $N = kM$ difference equations ($\alpha = 1, ..., k;\ i = 1, ..., M$):

$$n_i^\alpha(t+1) = \mathcal{S}(\sum_{j=1}^{M} T_{ij}^\alpha n_j^\alpha(t) - I(t) - \Theta_o) \cdot \mathcal{S}(1 - n_i^\alpha(t))\,, \tag{6}$$

where the inhibitory feedback $I(t)$ is equal for all neurons. We want to have stable stationary mean activity in the macrocolumn and therefore again choose the inhibition to be proportional to the over-all activity $B(t) = \sum_{i,\alpha} n_i^\alpha(t)$. We get in this case $I(t) = \mu \frac{B(t)}{N} = \frac{\mu}{k} \sum_{\alpha=1}^{k} p_\alpha(t)\,,$ where $p_\alpha(t) = \frac{1}{M} \sum_{i=1}^{M} n_i^\alpha(t)$ is the probability of a neuron in column α to be active. The direct simulation of dynamics (6) shows a complex behavior and for a wide range of parameters s and Θ_o we get stable ongoing activity. The dynamics favors to activate only a subset of minicolumns whereas the others are switched off. Hereby, the number of minicolumns which can be activated strongly depends on the proportionality factor of the inhibition μ. The points of stable activity and their dependency on μ can be studied again by the reformulation of (6) in terms of the activation probabilities $p_\alpha(t)$ of the different minicolumns. Calculations in analogy to above yield a system of $\alpha = 1, \ldots, k$ difference equations:

$$p_\alpha(t+1) \;=\; \Phi_s(\frac{p_\alpha(t) - \frac{\mu}{k} \sum_{\beta=1}^{k} p_\beta(t) \;-\; \Theta_o}{\sqrt{p_\alpha(t)}}) \;\; (1 - p_\alpha(t)) \;=:\; G_\alpha(\boldsymbol{p}(t)) \tag{7}$$

Equations (7) can be studied by a stability analysis and we just give the relevant results: For a macrocolumn with k minicolumns we get a family of 2^k potentially stable stationary points of the form,

$$\boldsymbol{q}_\gamma \;=\; (\underbrace{q_o,\, q_o, \ldots,\, q_o}_{\text{l-times}},\, \underbrace{0,\, 0, \ldots,\, 0}_{\text{(k-l)-times}}), \quad q_o = \mathcal{P}(\frac{l}{k}\,\mu), \tag{8}$$

and all permutations. Their stability is determined by the eigenvalues of the Jacobian of $\boldsymbol{G}(\boldsymbol{p})$ (see (7)) at these points which can be computed to be

$$\lambda_{1,2} = \frac{1 \;-\; \mathcal{P}(\frac{l}{k}\mu)}{2\sqrt{\mathcal{P}(\frac{l}{k}\mu)}} \;(1 \;\pm\; \frac{l}{k}\mu \;+\; \frac{\Theta_o}{\mathcal{P}(\frac{l}{k}\mu)})\; \Phi_s'(h(\frac{l}{k}\mu)) - \;\Phi_s(h(\frac{l}{k}\mu)), \quad \lambda_3 = 0, \tag{9}$$

where $h(\mu) \;=\; \frac{(1-\mu)\mathcal{P}(\mu) - \Theta_o}{\sqrt{\mathcal{P}(\mu)}}$. λ_1 is of multiplicity $(l-1)$, λ_2 of multiplicity 1, and λ_3 of multiplicity $(k-l)$. We get eigenvalues of magnitude greater than one

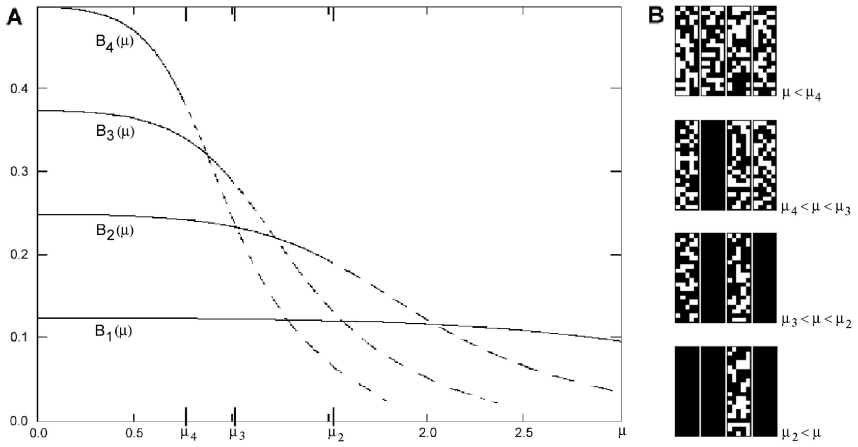

Fig. 1. A Stationary over-all activity in a macrocolumn (k=4) plotted against μ for $s = 20$ and $\Theta_o = \frac{1}{20}$. The four plots correspond to states of four to one active minicolumns (dotted parts mark unstable stationary activity). **B** Screenshots of a macrocolumn with four minicolumns. Each picture shows one activity-configuration of the respective activity probability (white pixels mark active neurons).

if and only if $\frac{l}{k}\mu$ gets greater than a critical value μ_c. Hence, the stability of a stationary point (8) with $l \geq 2$ is determined by its critical value $\mu_l := \frac{k}{l}\mu_c$. For a macrocolumn consisting, e.g., of $k = 4$ minicolumns we get a collection of 15 non-zero stationary points of type (8), whose stability is determined by the three critical points $\mu_4 = \mu_c$, $\mu_3 = \frac{4}{3}\mu_c$, and $\mu_2 = \frac{4}{2}\mu_c$. For $s = 20$ and $\Theta_o = \frac{1}{20}$ their values are $\mu_4 \approx 0.76$, $\mu_3 \approx 1.01$, and $\mu_2 \approx 1.52$. In Fig. 1A the stationary over-all activities $B_l(\mu) = lM\,\mathcal{P}(\frac{l}{k}\mu)$ are plotted for $l = 4, \ldots, 1$ active minicolumns together with the points μ_4, μ_3, and μ_2, which mark their intervals of stability. The macrocolumn's dependency on μ can be used to force the network to perform successive symmetry breakings: if we start for $\mu < \mu_4$ with the totally symmetric stable stationary point $(\mathcal{P}(\mu), \ldots, \mathcal{P}(\mu))$ and slowly increase the parameter, the macrocolumn is forced to break the symmetry by switching off one of the minicolumns as soon as $\mu > \mu_4$. The activity is then symmetrized between the minicolumns which remain active. But as soon as $\mu > \mu_3$ this symmetry is broken again. The process of symmetrizing the activity among the active columns and breaking the symmetry again continues until just one column remains active (see Fig. 1B). If the macrocolumn is exposed to input in form of externally induced EPSPs, it will keep the minicolumn with highest relative input active while successively switching off the others. After each symmetry breaking the network symmetrizes the minicolumn activities again and the next decision can be made relative to the inputs to the columns with non-zero activity. For a macrocolumn with four minicolumns of $N = 100$ neurons, with $s = 20$, $\Theta_o = \frac{1}{20}$, and μ increased from zero by 0.01 per time-step already

an average difference of three EPSPs per neuron every 10 time-steps is sufficient to select the corresponding minicolumn with a probability of more than 99%. The results were again obtained by direct simulation of (6) with coherence suppression. Simulations with a wide range of parameters show comparable results. The dynamics can further be shown to be robust against various perturbations, e.g. input or threshold noise, and against relaxation of assumptions such as strict disjointness of minicolumns.

4 Conclusion and Future Work

We have shown that a macrocolumn of inhibitorily interconnected minicolumns can have the dynamical property to symmetrize its activity among the active minicolumns and to break this symmetry spontaneously if a parameter of the inhibition is increased. This behavior was shown to be very sensitive to external input. For values of μ near to critical points small input differences are already sufficient to change the global state of the macrocolumn significantly. A macrocolumn can therefore serve to select and amplify inputs to its minicolumns. It can do so either directly with the parameter μ set near to critical points or indirectly by a succession of symmetry breakings with increasing μ. The repeated activity-suppression of the column with weakest input together with the symmetrization of the remaining activities presents a property not observed in usual winner-take-all mechanisms. Networks of interconnected macrocolumns can be expected to converge from a state of maximal activity to a state of minimal one by a process in which each macrocolumn makes a decision only if its input is sufficiently non-ambiguous. The networks can be applied to problems such as signal integration or classification and are subject of our current investigations.

Acknowledgments. Partial funding by the EU in the RTN MUHCI (HPRN-CT-2000-00111), the German BMBF in the project LOKI (01 IN 504 E9), and the Körber Prize awarded to C. von der Malsburg in 2000 is gratefully acknowledged.

References

1. A. Peters and E. Yilmaze. Neuronal organization in area-17 of cat visual-cortex. *Cereb. Cortex*, 3(1):49 – 68, 1993.
2. O.V. Favorov and D.G. Kelly. Minicolumnar organization within somatosensory cortical segregates I. *Cereb. Cortex*, 4:408–427, 1994.
3. P.A. Anninos, B. Beek, T.J. Csermely, E.M. Harth, and G. Pertile. Dynamics of neural structures. *J. Theo. Biol.*, 26:121 – 148, 1970.
4. E. Fransen and A. Lansneer. A model of cortical associative memory based on a horiz. netw. of connected columns. *Netw.-Comp.Neur.Sys.*, 9(2):235 – 264, 1998.
5. O.V. Favorov and D.G. Kelly. Stimulus-response diversity in local neuronal populations of the cerebral cortex. *Neuroreport*, 7(14):2293 – 2301, 1996.
6. J. DeFelipe, M.C. Hendry, and E.G. Jones. Synapses of double bouquet cells in monkey cerebral cortex. *Brain Res.*, 503:49 – 54, 1989.
7. J.M.L. Budd and Z.F. Kisvarday. Local lateral connectivity of inhibitory clutch cells in layer 4 of cat visual cortex. *Exp. Brain Res.*, 140(2):245 – 250, 2001.