Models of the formation of V1 and of its dynamics.

3.1 Models of the formation of V1

3.1.1 Introduction

Several functional properties of the brain are represented as maps: the body is represented by motor and somatosensory homunculus and the surrounding space is represented in cognitive maps of the hippocampus. The primary visual cortex is also endowed with multiple feature maps which provide a representation of the input space. Retinotopy is the mapping of spatial positions of the visual space onto the surface so that close points of the visual space are represented by neurons which are close together in the cortex. Ocular dominance of a neuron indicates how its inputs are biased toward the left or right eyes and this selectivity is organized into alternating bands. The selectivity of neurons to orientation also forms maps with special points, pinwheel singularities, around which the preferred orientation varies smoothly. Those maps are structurally related, with lines of iso-orientation orthogonal to the frontiers of ocular dominance domains. Some other features are also encoded in V1 like spatial frequency or direction of motion. In an early model proposed by Hubel and Wiesel, orientation selectivity result from the specific pattern of connections from LGN inputs but later studies highlighted the contribution of lateral connections for sharpening the orientation tuning curve.

Hebbian plasticity rule on the feedforward and lateral connections leads to the formation of orientation maps in firing rate models. Obtaining similar results with spiking neurons and spike timing dependent plasticity rules is still an open challenge and require heavy computational resources. On the other side, firing rate models are only a coarse grain version of the dynamics and don't have the fine scale complexity of spiking neurons. The column based networks of spiking we propose combine these two approaches for efficient simulation of the

visual cortex. A grid of columns of spiking neurons having the same structure as exposed in chapter 2 are used to study the fine temporal dynamics of the network after a transient stimulation and a firing rate model is used for the learning an orientation map, each unit representing a column and the connection weight between two units representing the probability of connection between two neurons of their corresponding neurons.

This chapter aims at describing the ongoing activity in models of V1, its influence on visual information processing and its structuration through learning.

Four approaches to the formation of V1 are presented and we present results on three large scale implementations of V1 models. First, in an ice cubes model of a pinwheel inspired, connections are hard wired to implement orientation selectivity and we study ongoing dynamics and the response to static or rotating bars. A second model is a large grid of columns with isotropic connection kernels for which the ongoing dynamics and the response to focalized stimulation are studied. The connections in the third network are the outcome of Hebbian learning when the connectivity of the isotropic model is used to initialize connection weights so that we can show how visual experience shapes spiking correlations in the network.

3.2 Ice cubes model.

3.2.1 Description of the model.

The simplest way to build a model of the visual cortex is to pack columns together in a "crystal-like" manner so that it achieves its function of local extraction of oriented lines. The construction of V1 in a hierarchical manner was initially proposed by Hubel and Wiesel in the 60's [82] based on their electrophysiological recordings in the cat area 17. The redundancy of coding that they discovered in the vertical direction (cortical depth) suggested that the cortical column could be the unitary building block of the cortex. The model of V1 they proposed is then obtained from an appropriate packing of these elementary units with hard wired connections. The wiring of these fixed connections is derived from the observed cortical fibers and from simplifying hypothesis when these observations are not possible.

As connections are fixed, retinotopy is imposed by simply providing a set of neighboring cortical columns with afferent connections from cells corresponding to a defined position. This group of columns defines an hypercolumn. The whole retinal space is then represented by an array of such hypercolumns, see fig 3.1. The ocular dominance domains can also be modeled by alternating an hypercolumn taking its inputs from the right eye with an hypercolumn taking its inputs from the left eye.

The orientation selectivity is implemented by considering a cylindrical parametrization (r, θ) of the hypercolumn. Columns having a similar azimuth θ code for the same orientation $\phi = \frac{\theta}{2}$ and the radial dimension is redundant. The topological charge of the pinwheel is positive if orientations are turning clockwise

and negative for anti-clockwise rotation. More detailed models consider that the r-dimension is used to code the selectivity to spatial frequencies [120]. The shape of the receptive field of simple cells could then be explained by a linear summation of aligned LGN inputs from ON and OFF channels, see fig 3.2. The resulting receptive field is a derivative of Gaussian and gives an effective description of how a cortical cell from V1 receives inputs from the eyes through the LGN. Other types of cells, like complex (phase invariant) or hypercomplex (tuned to an optimal bar length), can be built on the same principle with a pool of simple cells with same position, same orientation and all possible phases connecting to complex thus being phase invariant. This is an example where the serial and hierarchical doctrines of the with feedforward projections from the eye to the visual cortex is very efficient.

The lateral connectivity in primary visual cortex is classically modeled by a difference of Gaussians that is excitatory at short range and inhibitory at long range although, as described in the introduction of this chapter, more realistic architecture should include long range excitation.

3.2.2 Parameters of the model.

A pinwheel is composed of $N \times N$ columns placed on a regular grid with coordinates $X \in [-1, 1]$ and $Y \in [-1, 1]$. Each column receives afferent inputs from all the cells in the retinal layer and in this model all columns code for the same location, columns only differ in their preferred orientation. The retinal layer is also composed of 11×11 cells with coordinates $x \in [-1, 1]$ and $y \in [-1, 1]$. The activity of these retinal cells is modeled as inhomogeneous Poisson processes with rates corresponding to the light intensity of the visual input.

The afferent connectivity of a column c located at (X_c, Y_c) in the cortical plane with azimuth $\theta = \arctan \frac{X_c}{Y_c}$ from a cell located at a retinal position (x, y) is given by a directional second derivative of Gaussian. For $X_c = 0$, the second derivative of $G(x, y) = e^{-\frac{x^2+y^2}{\sigma^2}}$ is taken along the x-dimension:

$$\frac{\partial^2 G(x, y)}{\partial x^2} = A_0(1 - \frac{2}{\sigma^2}x^2)e^{-\frac{x^2+y^2}{\sigma^2}}$$

with A_0 a constant. For $\theta \neq 0$ and taking $A_0 = 1$, the connectivity is obtained by considering a rotation of the coordinates:

$$W^{aff}_{xy,\theta} = (1 - \frac{2}{\sigma^2}(x \cos\frac{\theta}{2} + y \sin\frac{\theta}{2})^2)e^{-\frac{(x\cos\frac{\theta}{2}+y\sin\frac{\theta}{2})^2+(-x\sin\frac{\theta}{2}+y\cos\frac{\theta}{2})^2}{\sigma^2}}$$

where σ parametrizes the sharpness of the receptive field. As density of connections, it is restricted to $[-1, 1]$, with negative weights corresponding to the inhibitory projections and positive weights to the excitatory ones. The integral of this function is zero so that a column should receive as many excitatory connections as inhibitory connections. In practice, as the retinal space is restricted to $[-1, 1] \times [-1, 1]$ and most inhibitory connections are too small to make any

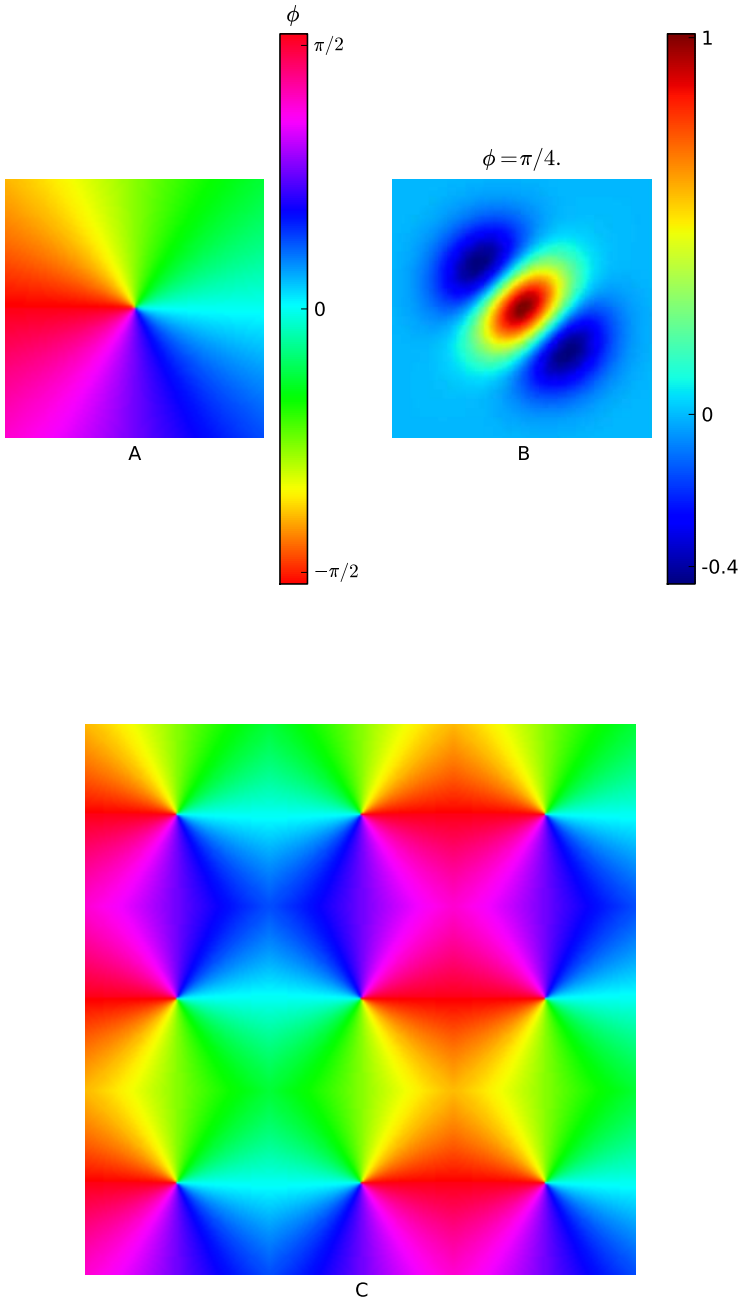

Figure 3.1: **Ice Cube model** - (A) A pinwheel with colors representing the preferred orientations of columns turning clockwise. (B) The receptive field of a column having $\frac{\pi}{4}$ as preferred orientation. (C) A "crystal-like" orientation map is built by packing pinwheels together

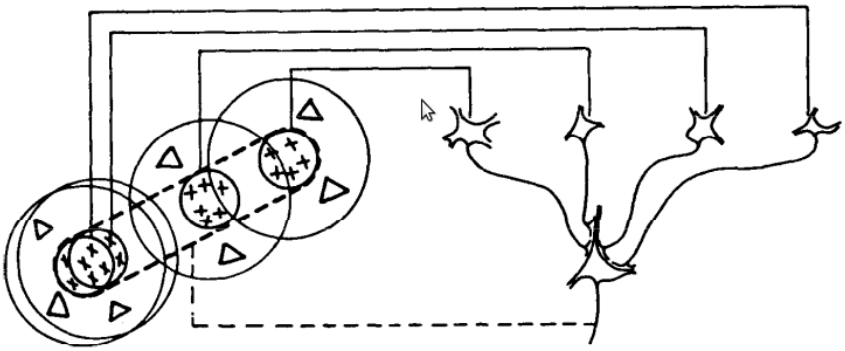

Figure 3.2: **Simple cell model.** - Cells from the LGN with ON center receptive fields feeds giving its orientation preference (Adapted from [82]).

connection [1] and the excitatory connections slightly overtake the inhibitory ones.

The lateral connectivity is restricted to the nearest neighbors, the column has a density $p = 0.2$ of self connections and the 8 neighboring column project to it with probability $p = 0.1$ both for excitation and inhibition. Parameters of the neurons are the same as described in table 2.15 of Chapter 2, a column being composed of 40 excitatory cells and 10 inhibitory cells.

During the 500 ms before the presentation of a stimulus, the cortical layer is ignited with a Poissonian stimulation at high rate, each column is connected with density 0.2 to a 5000 Hz Poisson spike train. The visual stimulation is a static or a rotating bar and the retinal activity corresponding to this stimulation is a set of Poisson spike trains with rates:

$$a_{xy} = f_0 e^{-(\frac{x\cos\alpha(t)+y\sin\alpha(t)}{\sigma_L})^2+(\frac{-x\sin\alpha(t)+ycos\alpha(t)}{\sigma_l})^2}$$

with $\alpha(t)$ the orientation of the bar,f_0 the maximum frequency $60Hz$, $\sigma_L = 0.35$ the length of the bar and $\sigma_l = 0.05$ the width of the bar.

3.2.3 Analysis of the dynamics.

On-going activity: Phase diagram The maximal conductances of excitatory (g_E) and inhibitory (g_I) is varied and the resulting mean firing rate, mean coefficient of variation of interspikes intervals (ISI) and mean pairwise correlation between spike trains are plotted in a phase diagram on fig 3.4. There are three distinct regimes:

- When $g_I = 0$, the network is highly active with regular spiking of the cells but spikes are asynchronous.

[1] For two columns A and B, each containing N neurons, for a connection probability $P_{AB} < \frac{1}{N^2}$, there will be nearly no connection from neurons of A to B.

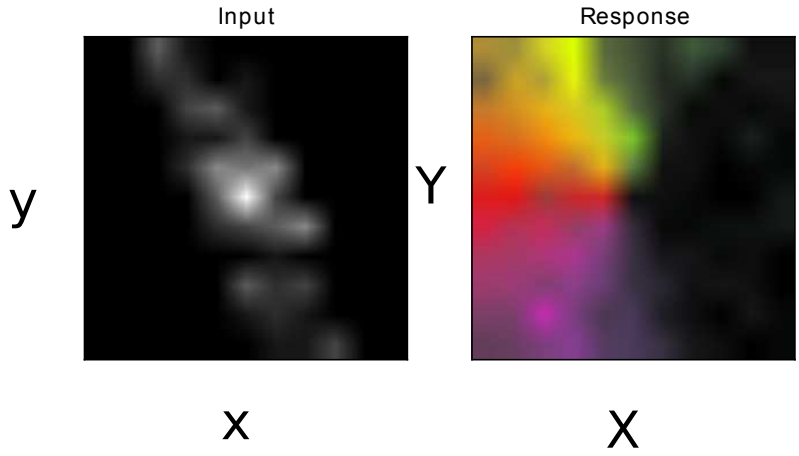

Figure 3.3: **Response of the pinwheel to an oriented bar** - (Left) Retinal activity corresponding to an oriented bar. (Right) Response of the pinwheel, the contrast indicates the firing rate of the column and the color codes for the preferred orientation of the column.

- When $g_E > g_I$, currents are dominated by excitation, the network is highly active with correlated regular spiking. This shows the synchronizing effect of inhibition.
- When $g_I > g_E$, currents are dominated by inhibition, the network activity is low and decreases when g_I is increased. Spiking occurs in an asynchronous irregular fashion.

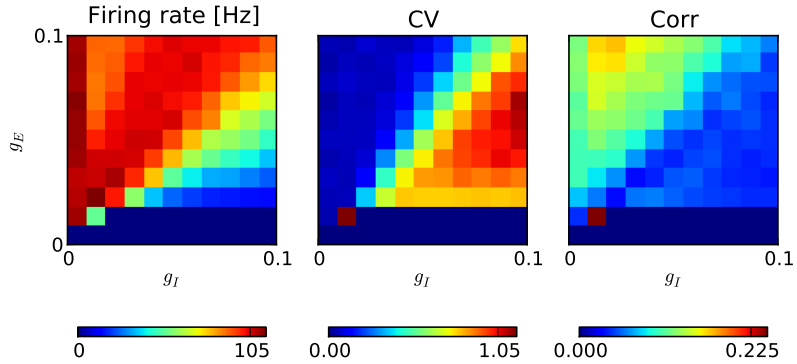

Figure 3.4: **Phase diagram for the Ice cubes model** - (Left) Average firing rate in Hertz. (Middle) Average coefficient of variation of the ISI distribution of spike trains. (Right) Average local correlation, that is correlation between spike trains of neurons belonging to the same column.

Response to a rotating bar. A rotating bar is presented to the network with a speed of 60Hz. A snapshot of the response of the network in the inhibition dominated regime is shown on fig 3.5. The activity in the inhibition dominated regime is a bump rotating at the same speed as the rotating bar, even at high rotation frequency [2], and this periodic activity is shown on fig 3.5. In the excitation dominated regime, no response specific to the stimulus can be seen.

[2] By the discretness of the grid, at very high frequency rotation, the bar is flashed periodically rather than rotating smoothly.

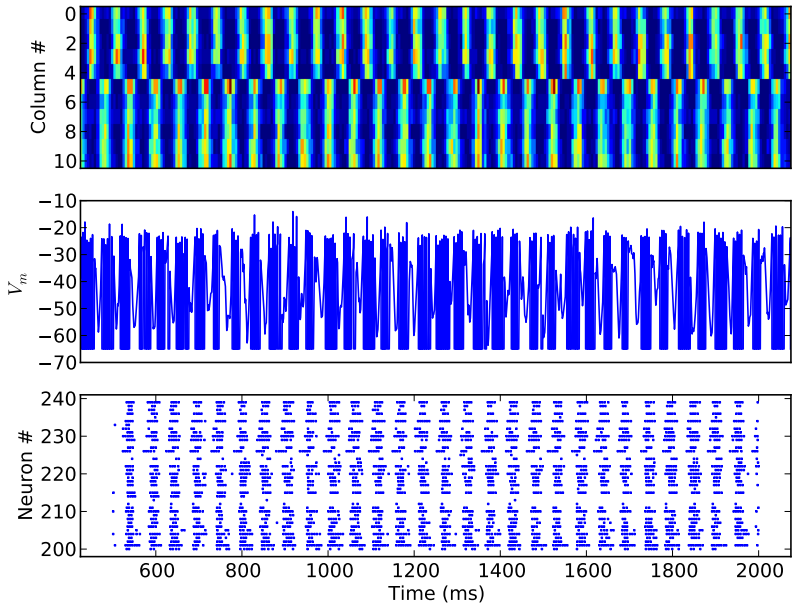

Figure 3.5: **Response to a rotating bar** - (Top) Average firing rate for a line of columns passing through the center. (Middle) Membrane potential of a neuron in column at position (0,5). (Bottom) Raster plot of the column at position (0,5).

Orientation tuning curve. As can be expected from the response to a rotating bar, the response to static bar is also very different depending on the regime of spontaneous activity. For characterizing the functional response of a cell, it is common to use the tuning curve that measures the firing rate of a cell as a function of the orientation of the presented stimulus ϕ. In our case we record the tuning curve of two columns at different positions of the pinwheel, one is close to the pinwheel singularity and the other is far (five columns away) from the pinwheel singularity. For each situation the map is shifted so that the central column of the network is the one from which the tuning curve is recorded. When the network is in the excitation dominated regime, the tuning curve is flat because the response is washed out by the spontaneous activity. In the inhibition dominated regime, the tuning curve is peaked with the maximum indicating the preferred orientation as can be seen on fig 3.6. In this regime, the tuning curve is much sharper for a column close to the pinwheel center than for a column far from the pinwheel center and the level of correlations is the same for all orientations, see fig 3.7. The level of activity in neighboring columns explains the sharpness of the orientation tuning curve close to the pinwheel center. Those columns around a column have very preferred orientation close to the pinwheel singularity and have low firing rate and thus do not excite the central column whereas far from the pinwheel, neighboring columns are excited for orientations similar to the preferred orientation of the central column.

3.3 Phenomenological models

The Ice Cubes model of $V1$ is based on the observations gathered from biological studies of the connectivity patterns, it tries to reproduce by hand these connectivity patterns and adds minimal assumptions when biology cannot help in choosing between different options. We explore here the models which forget about the biological processes but focuses on more abstract properties of $V1$. They aim to answer to simple questions like "What is the goal to achieve during the formation of $V1$?" or "What are the universal properties governing the formation of $V1$?" and can be analyzed in some simple cases. Such models are reviewed in [121].

3.3.1 The Elastic net theory

Multiple features are engrafted in V_1 through its functional maps for retinotopy, occular dominance and orientation preference. The mapping of visual stimulations parameters (X and Y positions, x and y coordinates of the preferred orientation and occularity) onto the cortical surface in layer IV is achieved during the formation of $V1$ with points close in the parameter space falling close onto the cortex. With the Elastic Net model, the formation of $V1$ is seen as an optimization process directed toward this mapping with the minimal wiring length.

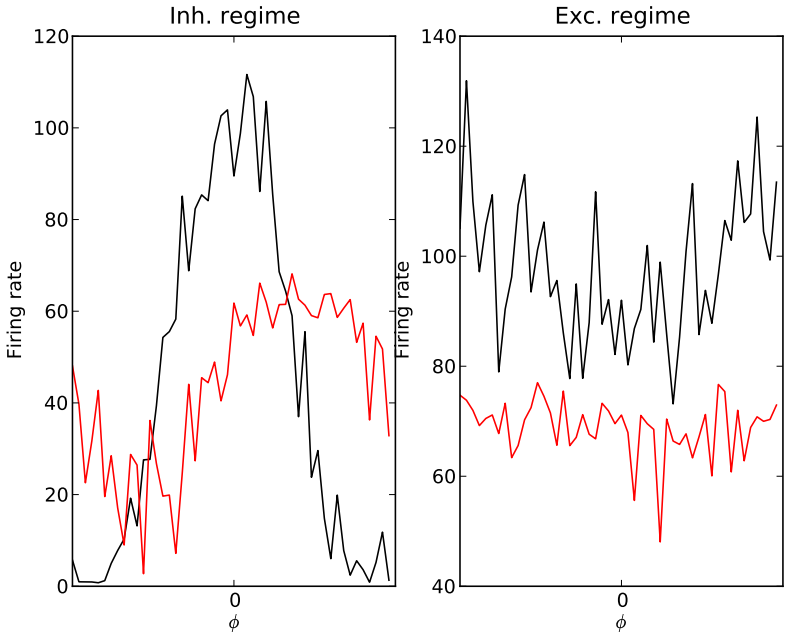

Figure 3.6: **Orientation tuning curves.** - (Left) Inhibition dominated regime, close to the pinwheel center in black and far in red. (Right) Excitation dominated regime with the same color conventions.

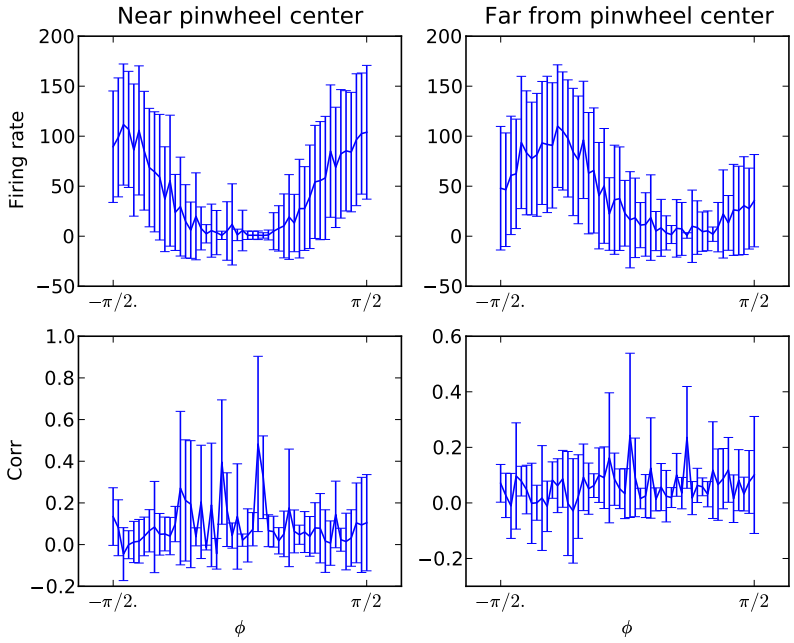

Figure 3.7: **Orientation tuning curves and correlations.** - (Left) Firing rate and correlations for a column close to the pinwheel center. (Right) Firing rate and correlations for a column far from the pinwheel center.

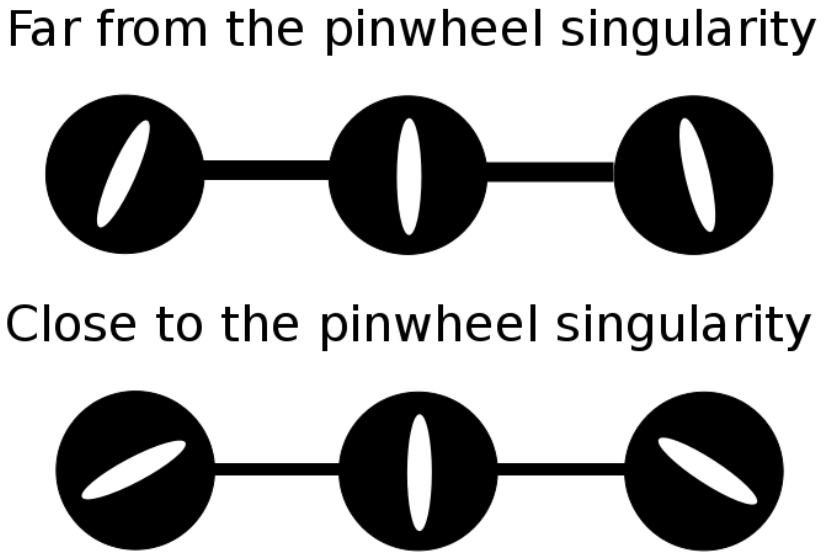

Figure 3.8: **Neighboring columns** - Preferred orientation and activity level (thickness of links) of neighboring columns.

A simple framework to apply the Elastic Net algorithm is the traveling salesman problem (TSP). A salesman has to visit N cities placed on a 2-dimensional plane and the problem is to find the shortest closed tour in which all cities are visited. This problem is NP-complete, that is it cannot be solved in a polynomial time of the number of cities to visit. An exhaustive search is possible when the number of cities is small but the computation time becomes very large as N grows. The TSP is formally equivalent to the mapping of N points regularly spaced on an elastic ring onto N points of the plane.

There are N cities with positions $(\mathbf{x_i})_{1\leq i\leq N}$ on the plane and N path points with positions $(\mathbf{y_j})_{1\leq j\leq N}$ on the elastic ring. The rule for changing the position of a path point is the following:

$$\Delta\mathbf{y_j} = \alpha\sum_i w_{ij}(\mathbf{x_i}-\mathbf{y_j}) + \beta k(\mathbf{y_{j+1}} - 2\mathbf{y_j} + \mathbf{y_{j-1}}),$$

where α scales the contribution of the cities in attracting path points and βk scales the elastic forces from the neighboring points on the path. The weight w_{ij} represents the normalized attraction that the city point i has on the path point j and is defined as

$$w_{ij} = \frac{e^{\frac{|\mathbf{x_i}-\mathbf{y_j}|^2}{2k^2}}}{\sum_p e^{\frac{|\mathbf{x_i}-\mathbf{y_p}|^2}{2k^2}}}$$

so that an energy function E can be defined:

$$E = -\alpha k\sum_i \log(\sum_j \frac{e^{\frac{|\mathbf{x_i}-\mathbf{y_j}|^2}{2k^2}}}{\sum_p e^{\frac{|\mathbf{x_i}-\mathbf{y_j}|^2}{2k^2}}}) + \frac{\beta}{2}\sum_j |\mathbf{y_{i+1}} - \mathbf{y_i}|^2$$

with $\Delta\mathbf{y_j} = -k\frac{\partial E}{\partial \mathbf{y_j}}$. The free parameter k is slowly decreased along the procedure, in a similar way as what is done in simulated annealing algorithms, so that the attraction of the cities becomes more and more specific to their closer path points. The analysis of this energy function was done in [122], for large values of k, the energy has only one minimum but the energy bifurcates as k decreases and then present several local minima. The solution reached by the Elastic Net algorithm is not necessarily the global optimum but the obtained minimum is not too far from the optimal solution [123].

After the original model of Durbin [123], Goodhill et al [124] proposed an implementation of the Elastic Net that generates stripe patterns similar to those observed for ocular dominance domains. In this setting, the cities with positions $\mathbf{x_i}$ represent LGN units scattered on two parallel planes separated by a small gap in 3 dimensional space. The cortex is then an elastic sheet with positions $\mathbf{y_j}$ having the same change rule a explained above. The width of stripes appearing in this model is controlled by the ratio between the distance separating the two LGN sheets and the distance between two neighboring LGN units. This model can test the effect of monocular deprivation by changing the value of α for one of the two eyes. The result obtained in [125] is that the stripes associated with the

eye with the smallest α get lower periodicity without changing the periodicity of the stripes associated to the other eye. Similar principles have been successfully applied to the formation of orientation selectivity.

3.3.2 Pattern formation dynamics.

In the previous model, the formation of V_1 is reduced to the solving of a mapping problem between spaces of different dimensions. The Elastic Net algorithm is successful in reproducing the stripe patterns observed for ocular dominance map and in predicting the effect of visual deprivation but it does so in a rather artificial manner and the history of this pattern formation may be very different from that observed in biology. Pattern formation is a well known phenomena in physics and biology and the possible dynamical systems used to model it are reviewed in [126] and [127]. We will see how these universal models can be used for the formation of ocular dominance stripes and orientation preference maps by integrating physical constrains and symmetries. In these kind of models some biologically relevant aspects are forgotten, for example, the fast dynamic corresponding to the neuronal activity is not taken into account to emphasize the slow dynamics of the synapses.

Ocular dominance stripes The model presented here was first implemented in [128]. It is based on simple rules for the growth of synapses:

- Locally, synapses from the same type activate their growth but inhibit the growth of the same synapses in an annular region around the growth region.
- Synapses of different eyes have opposite influences.

The influence of cells of the same type (w_{RR} or w_{LL}) can then be represented by a difference of Gaussians of the distance r between the two synapses considered:

$$w = Ae^{-\frac{r^2}{d_1}} - Be^{-\frac{r^2}{d_2}}$$

with $d_1 < d_2$ and the influence between synapses from different eyes (w_{RL} or wLR) is represented by the same functions with opposite signs. By this way, the symmetry between the left and the right eye is taken into account. The growth rate of synapses from the right eye is given by:

$$s_R = w_{RR} \star n_R + w_{LR} \star n_L$$

where $\star$ stands for the spatial convolution, with for functions f and g defined on the domain D:

$$f \star g = \int \int_D g(|\mathbf{r} - \mathbf{r}'|) f(\mathbf{r}') d\mathbf{r}'$$

and n_R, n_L are the spatial densities of synapses.

The temporal variations in the number of synapses of both types are also constrained in this model by the fact that the densities should stay in a reasonable range. As densities, it is positive and it is also supposed that at any given point $\mathbf{r}$, the local density is bounded by the maximal density N, $n_R(\mathbf{r}) < N$. The general form of temporal variations in synapses density of the right eye is thus :

$$\frac{\partial n_R(\mathbf{r})}{\partial t} = s_R f(n_R)$$

with f, a function constrained by $f(0) = 0$ and $f(N) = 0$. These constrains can be satisfied by taking for example $f(n_R) = n_R(n_R - N)$ and the dynamical system for the two spatial densities is:

$$\begin{cases} \frac{\partial n_R(\mathbf{r})}{\partial t} & = (w_{RR} \star n_R + w_{LR} n_L) n_R^2 (1 - \frac{N}{n_R}) \\ \frac{\partial n_L(\mathbf{r})}{\partial t} & = (w_{RL} \star n_R + w_{LL} n_L) n_L^2 (1 - \frac{N}{n_L}) \end{cases}$$

The total number of synapses can be considered as locally constant and uniform across the area: $n_R(\mathbf{x}) + n_L(\mathbf{x}) = N$ and the system is then reduced to a single equation:

$$\frac{\partial n_R}{\partial t} = (w_{RR} \star (2n_R - N)) f(n).$$

There are 3 fixed points to this equation:

- $n = 0$ is stable because at first order

$$\frac{\partial n}{\partial t} \approx -Nnw \star N < 0$$

.

- $n = N$ is also stable for symmetry reasons so that if all synapses becomes wired to an eye at time T, there won't be any synapses from the other eye after this time.
- $n = \frac{N}{2}$ is then unstable.

To consider how instabilities develop near the unstable state, we consider a small perturbation $x = n - \frac{N}{2}$. The evolution of such a perturbation is driven by:

$$\frac{\partial x}{\partial t} = \frac{N^2 (w \star x)}{2}.$$

To analyze the transient behavior of the perturbation, it is decomposed as follow, $x = x_0(\mathbf{r}) e^{\lambda t}$. The evolution equation becomes independent of the time variable:

$$\lambda x_0(\mathbf{r}) = \frac{N^2 (w \star x_0(\mathbf{r}))}{2}.$$

The Fourier transform of this equation gives a dispersion relation indicating the stable and unstable modes for the propagation of instabilities (Note that convolution becomes a product after Fourrier transform).

$$\lambda(\mathbf{k}) \hat{x}_0(\mathbf{k}) = \frac{N^2 W(\mathbf{k}) \hat{x}_0(\mathbf{k})}{2}.$$

Unstable modes grow when $\lambda(\mathbf{k}) > 0$ that is when

$$W(\mathbf{k}) = \int\int_D w(|\mathbf{r} - \mathbf{r}'|)e^{i\mathbf{k}.\mathbf{r}'}d\mathbf{r}' > 0.$$

The Fousrier transform of a Gaussian function is a Gaussian function so that:

$$W(\mathbf{k}) = 2\pi \int_0^\infty (Ae^{-\frac{|r-r'|^2}{d_1}} - Be^{-\frac{|r-r'|^2}{d_2}})e^{ikr'}dr'$$

$$W(\mathbf{k}) = A\sqrt{\pi d_1}e^{-\pi^2 d_1 k^2} - B\sqrt{\pi d_2}e^{-\pi^2 d_2 k^2}$$

. The periodicity of ocular dominance stripes will be given by $k^\star$ at which W is maximum, around 0.45 mm with the parameters indicated in fig 3.9.

In order to model the non-uniform distribution of biological markers, like cytochrome oxydase blobs, the upper bound on the density of synapses can be taken as spatially periodic function, $N(r) = \bar{N} + \kappa u(r)$. It was shown in [129] that the pattern formed in this modified model get aligned with the synaptic density.

Orientation preference maps. The formation of orientation columns can be treated in a similar way to that of ocular dominance stripes even if there are some differences. The ocular dominance at a point of the cortex was quantified by a single number n but the orientation preference and selectivity of a column is measured by a vector or its complex representation z. The argument of z is related to the preferred orientation θ of the column:

$$\theta = \frac{1}{2}\arctan\frac{\Re(z)}{\Im(z)}$$

with the factor $\frac{1}{2}$ restricting θ between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. The module $|z|$ quantifies the selectivity of the response of the column to a bar of orientation θ, it is related to the amount of cells in the column having θ as preferred orientation.

Orientation preference maps are organized around pinwheel singularities where all orientations collapse. Depending on whether the orientations turn clockwise or anticlockwise around the singularity, a positive or negative topological charge is attributed to the singularity. This topological charge can be calculated from the orientation preference field:

$$\mathcal{Q}_A = \frac{1}{2\pi}\oint_C \nabla\theta(\mathbf{r})d\mathbf{s}$$

where the integral is taken over a closed contour surrounding the singularity.

The orientation maps can be characterized by their spatial correlation functions between points at positions $\mathbf{r_1}$

$$C(\mathbf{r_1}, \mathbf{r_2}) =< z(\mathbf{r_1})\bar{z}(\mathbf{r_2}) >$$

$$C^\star(\mathbf{r_1}, \mathbf{r_2}) =< z(\mathbf{r_1})z(\mathbf{r_2}) >$$

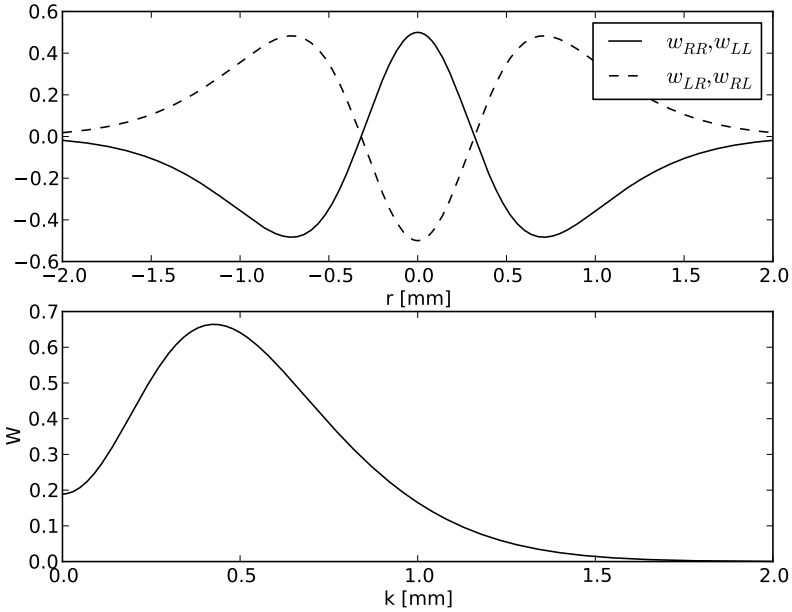

Figure 3.9: **Swindale model for the formation of ocular dominance stripes** - (Top) Kernels describing the influence of synapses from the same eye (w_{RR}, w_{LL}) or from opposite eyes (w_{RL}, w_{LR}). (Bottom) Fourrier transform of the kernel corresponding to synapses from the same eye. The wavelength, $k^\star$, for which it reaches maximum indicates the typical periodicity of the stripes.

The dynamics for the change in z can be written in a general form:

$$\frac{\partial z}{\partial t} = \mathcal{F}[z]$$

Taking a similar approach to the model for the ocular dominance stripes, the following form for $\mathcal{F}$ was proposed in [130]:

$$\mathcal{F}[z] = (w \star z) f(|z|)$$

with w the difference of Gaussians defined in the description of the model for the formation of ocular dominance stripes and $f(|z|) = Z - |z|$ providing an upper bound Z on the selectivity of a column. The only difference with the previous model for the formation of ocular domains is that the field o takes real values whereas for orientation preference, the field z takes complex values and thus similarly to what was shown for ocular dominance, the spatial frequency of the orientation map will be given by $k^\star$ for which the Fourrier transform of w_{RR} is maximal.

The following model of $V1$ was proposed in [131]:

$$\mathcal{F}[z] = \mathcal{L}[z] + \eta$$

without taking an explicit form for the deterministic kernel $\mathcal{L}$ and the noise term η but just constraining the symmetries of $\mathcal{F}$:

- $\mathcal{F}$ is invariant by translation: $\mathcal{F}[T_{\mathbf{R}} z] = T_{\mathbf{R}} \mathcal{L}[z]$ with $T_{\mathbf{R}} z(\mathbf{r}) = z(\mathbf{r} + \mathbf{R})$.
- $\mathcal{F}$ is invariant by rotation: $\mathcal{F}[R_\alpha z] = R_\alpha \mathcal{F}[z]$ with R_α the rotation matrix of angle α: $\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$
- $\mathcal{F}$ is invariant under phase shift: $\mathcal{F}[e^{i\phi} z] = e^{i\phi} \mathcal{F}[z]$.

Phase shift invariance implies that $\mathcal{F}[0] = 0$ and then the homogeneous state $z(\mathbf{r}) = 0$ is a stationary solution of the system. The phase shift invariance also implies that

$$< e^{i\phi} z(\mathbf{r_1}) e^{i\phi} z(\mathbf{r_2}) > = < z(\mathbf{r_1}) z(\mathbf{r_2}) >$$

so that $C^\star$ is null everywhere. The translation invariance implies that $C(\mathbf{r_1}, \mathbf{r_2})$ only depends on the distance $x = |\mathbf{r_1} - \mathbf{r_2}|$. The characteristic wavelength Λ of the orientation map can then be computed by considering the maximum in the Fourrier spectrum of the correlation function:

$$P(k) = \frac{1}{2\pi} \int C(x) e^{ikx} dx$$

Thanks to an analogy with the physics of defects, the analysis of this spectrum gives a lower bound for the density of pinwheels $\rho = \frac{\pi}{\Lambda^2}(1 + \alpha)$ with $\alpha > 0$. This result suggest that for animals with an abnormally low pinwheel density

compared to the wavelength of the orientation map, a phase of pinwheel anhihilation must occur before the map stabilizes. This linear model gives a lower bound on the pinwheel density but it doesn't have any saturation and if it is run without a strong randomness term, the pinwheel density will grow unbounded. A solution to this problem is to consider a non-linear part in $\mathcal{F}$.

The orientation maps are supposed to emerge through Turing instability and an explicit form for $\mathcal{F}$ up to third order can be inspired by pattern formation theory [126]:

$$\mathcal{F} = \mathcal{L} + \mathcal{N}_2 + \mathcal{N}_3$$

with

- $\mathcal{L} = a - (k_c^2 + \nabla^2)^2$, the Swift-Hohenberg operator for describing hydrodynamic instability with k_c the wavelength of the emerging instability.
- $\mathcal{N}_2 = 0$, for symmetry reasons.
- $\mathcal{N}_3[z(\mathbf{r})] = (1-g)|z(\mathbf{r})|^2 z(\mathbf{r}) + (g-2)\int K_\sigma(\mathbf{r}'-\mathbf{r}) z(\mathbf{r})(|z(\mathbf{r}')|^2 + \frac{1}{2}\bar{z}(\mathbf{r}) z(\mathbf{r}')^2) d\mathbf{r}'$ with the kernel $K_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{\mathbf{r}^2}{2\sigma^2}}$ accounting for long range interactions and g tuning the fraction of non-linearity coming from local and non-local interactions.

The detailed analysis of this model near criticality, $a << 1$, together with the phase diagram in parameters $(\frac{\sigma}{k_c}, g)$ can be found in [132]. It is shown that for this non-linear model, the pinwheel density will be close to π thus avoiding the unbounded proliferation observed with the purely linear model.

The symmetries considered above are only an approximation of what is observed in the visual cortex. The appropriate invariance is the following: $\mathcal{F}[R_\alpha z] = R_\alpha \mathcal{F}[z e^{i\alpha}]$. A model was proposed recently [133] to take this symmetry into account during development with the following non-linear dynamics for the orientation $\mathbf{s_i} = (s_{ix}, s_{ix})$ of the column i at location $\mathbf{r_i}$:

$$\frac{\partial s_i}{\partial t} = \mathbf{s_i}(1 - |\mathbf{s_i}|^2) + \sum_j [J(r_{ij})\mathbf{s_j} + K(r_{ij}(\mathbf{s_i}.\hat{\mathbf{r}}_{\mathbf{ij}})\hat{\mathbf{r}}_{\mathbf{ij}}],$$

with r_{ij} the distance between columns i and j and $\hat{\mathbf{r}}_{\mathbf{ij}}$ the unitary vector directed by $r_i - r_j$ and j running over all columns. The isotropic coupling J is positive on a disk of radius $\frac{R}{2}$ around the column i and negative on an annular region between $\frac{R}{2}$ and R. The long range term K is taken as constant over the map, it scales the anisotropic input to the column which depends on the colinearity and coaxial alignment between orientations of columns i and j, this term is invariant under joint rotation of the orientations and the grid supporting the columns. In the model with full rotation invariance ($K = 0$), pair annihilation of pinwheels leads to maps with no singularity (rainbow patterns). Adding a non-zero anisotropic term (whether positive or negative) result in stable maps with pinwheels. A similar positive anisotropic term is also included in mean field models of $V1$ dynamics and the action of the special euclidean group on $\mathbf{R}^2 \times \mathbf{S}^1$ under which it is invariant is also called the shift-twist representation of the Euclidean group [134], [135].

3.4 Learning of orientation maps.

In the models described in the two previous sections, biological details of the formation of V_1 were avoided by relying on simplifying hypothesis or universal models which behavior can be studied analytically. Models closer to biology consider the coupling between the dynamics of neuronal activity and the dynamics on the synapses. Changes in the strength of a synapse, if it lasts for few seconds to a minute, is called short-term plasticity and is responsible for synaptic depression and facilitation which are supposed to give a computational advantage for optimal detection and adaptation to changing inputs [136]. Changes can also last for hours or days and this long term plasticity implements learning. The long term potentiation (LTP), increasing the synaptic strength between two coactive neurons, was first described in the rabbit hippocampus [137]. The long term depression is necessary to avoid saturation of all synapses to their maximal strength. In the classical Hebb rule [138], summarized in "cells that fire together, get wired together", the synaptic change during learing is related to the averaged dynamics of the presynaptic and the postsynaptic neurons. In spike timing dependent plasticity [93], synaptic changes depend on the precise timing between the presynaptic and postsynaptic spike, LTP (LTD) is induced if the presynaptic neuron spikes few milliseconds before (after) the postsynaptic neuron. Hebbian learning was widely used in the 80's to implement associative memory through Hopfield network as an example of unsupervised learning. In such network, fixed point attractors of the dynamics can be used to store memories. With this unsupervised learning, after a partial presentation of the input, the network dynamics converge towards the closest memory state. In supervised learning, the network is trained on a set of examples, each example is an input and the desired output to this input and the network change synapses to improve the matching between the desired output and the effective input, for example by backpropagating an error signal [139], and the network can classify then new inputs. Computational models of the development of visual cortex were successfully developed in the 80's. These models, described below, are closer to the biological processes associated to learning than phenomenological models but are also hardly tractable analytically.

3.4.1 The Van der Malsburg model for the formation of columnar organisation.

In a seminal paper of 1973 [140], Christopher Von der Malsburg implemented the first model producing orientation maps. It includes a retinal layer with 19 cells and a cortical layer with 169 excitatory (E) and inhibitory (I) cells disposed on hexagonal grids. The activity H_k represents the firing rate of the neuron k in the cortical layer, driven by a linear differential equation:

$$\frac{dH_k(t)}{dt} = -\alpha_k H_k(t) + \sum_l p_{lk} H_l^\star(t) + \sum_i s_{ik} A_i^\star(t)$$

where $\star$ indicates a thresholded version of the signal to which it is applied.

The fixed recurrent connections are wired according to a Colonnier scheme (1966). The connectivity kernels w from which the p_{kl} are defined have truncated Gaussian shapes of extent σ with $w_{II} = 0$, $\sigma_{EE} = \sigma_{EI}$ and $\sigma_{IE} > \sigma_{EE}$. The plastic synapse s_{ik} representing the strength of connection between the afferent cell i and the excitatory cortical cell k evolves according to a version of the Hebb rule:

> *If there is a coincidence of activity in an afferent fiber i and a cortical E cell k then s_{ik} is increased to $s_{ik} + \Delta s$, Δs being proportional to the signal on the afferent fiber i and to the output signal of the E cell k. Then all the s_{jk} leading to the same cortical cell k are renormalized to keep the sum $\sum_j s_{jk}$ constant.*

A set of nine stimuli was presented to the network, each containing seven active retinal cells and representing a bar of a given orientation. After this training period, each cell fires for a specific orientation and cells are activated in clusters reaching a columnar organization close to what is observed in $V1$, see fig 3.10. The input space is thus mapped on the 2 dimensional surface of the cortex as was done in an abstract way for the elastic net.

3.4.2 Learning rules.

The classical Hebbian learning rule for the synaptic weight from cell i to cell k is the following:

$$w_{ik}(t+1) = w_{ik}(t) + \alpha a_i a_k$$

where a_i and a_k are the pre and postsynaptic activities and α is a free parameter called the learning rate. To avoid unbounded growth of the weights, a normalization scheme should be adopted, it can be substractive or multiplicative. This can be illustrated by considering two binary input neurons of activity I_R and I_L connected with positive synaptic weights w_R and w_L to a cortical neuron of activity $O = w_R I_R + w_L I_L$. With substractive normalization, the dynamics on the weights is:

$$\Delta w_R = I_R O - \frac{1}{2}(I_R O + I_L O)$$

$$\Delta w_L = I_L O - \frac{1}{2}(I_R O + I_L O)$$

replacing O with its value gives

$$\Delta w_R = \frac{1}{2}(w_R I_R^2 + (w_L - w_R) I_R I_L + w_L I_L^2)$$

$$\Delta w_L = \frac{1}{2}(w_R I_R^2 + (w_R - w_L) I_R I_L + w_L I_L^2)$$

then taking averages over input presentations with $< I_R^2 >=< I_L^2 >= 1$ and $< I_R I_L >= C$ gives the continuous time dynamics:

$$\frac{d < w_R >}{dt} = \frac{1}{2}(< w_R > +(< w_L > - < w_R >)C + < w_L >)$$

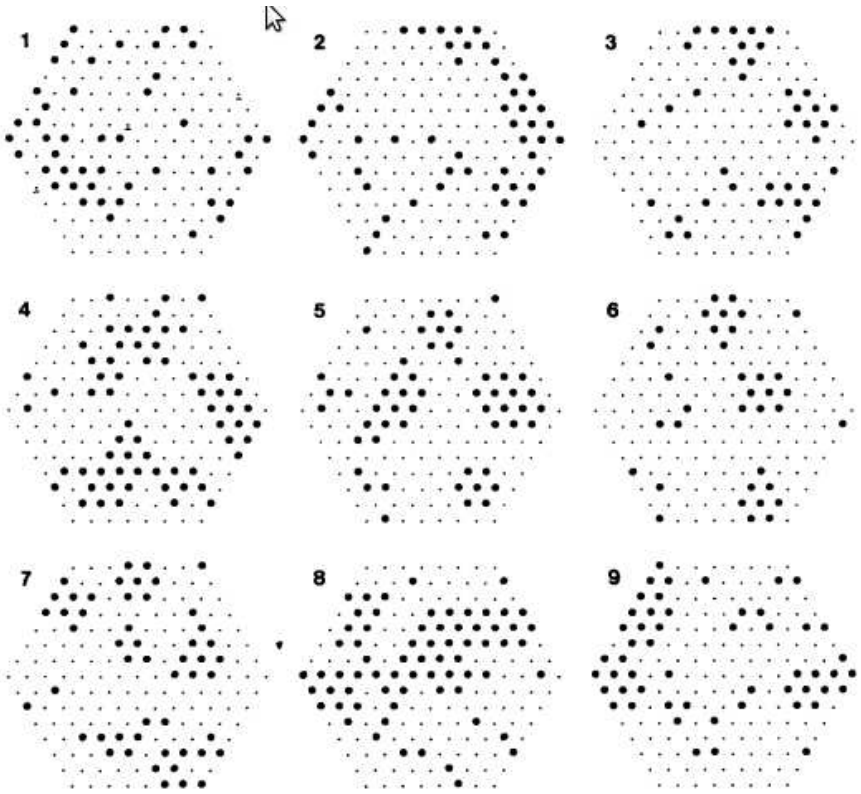

Figure 3.10: **Orientation map.** - Pattern of activity for the 9 oriented bars presented as input (Adapted from [140]).

$$\frac{d < w_L >}{dt} = \frac{1}{2}(< w_L > +(< w_R > - < w_L >)C + < w_R >)$$

which can be written:

$$\frac{d\mathbf{w}}{dt} = \mathbf{A}.\mathbf{w}$$

with $\mathbf{w} = \begin{pmatrix} < w_R > \\ < w_L > \end{pmatrix}$ and $\mathbf{A} = \frac{1}{2}\begin{pmatrix} 1-C & 1+C \\ 1+C & 1-C \end{pmatrix}$. The eigenvalue $\lambda_1 = 1$ is associated to the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and would lead to unbounded growth of the two weights. Weights are usually bounded so that this mode can be ignored. The eigenvalue $\lambda_2 = -C$ is associated to the eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ so that one weight is increasing while the other is decreasing until the system reaches $(0, 1)$ or $(1, 0)$ depending on the stimulation history. If the substractive normalization is used, all weights end up with the value 0 or one and their dynamics is driven by correlations in the inputs. The cortical neurons would then be purely monocular.

To avoid such drastic synapse elimination, the most commonly used normalization scheme is divisive, then for a synapse from cell i to cell k:

$$w_{ik}(t+1) = \frac{w_{ik}(t) + \alpha a_i a_k}{\sqrt{\sum_j (w_{jk}(t) + \alpha a_j a_k)^2}}$$

where j runs across all the input cells of the cell k can be approximated by the local Oja's rule:

$$w_{ik}(t+1) = w_{ik}(t) + \alpha a_i a_k - \alpha w_{ik} a_k^2$$

With such learning dynamics, it was demonstrated that a neuronal network can compute the principal components of the input space [141].

The Bienenstock-Cooper-Munro (BCM) learning rule [142] is another way to avoid the unbounded growth of weights. It provides a mechanism increasing the connection when the sum of weighted input to the neuron is superior to a threshold but also decreasing the weights when the post-synaptic is less than the threshold. The rule in continuous time is:

$$\dot{w}_{ik}(t) = -\epsilon w_{ik}(t) + \phi(a_k(t), \bar{a}_k(t)) a_i(t)$$

with $a_k = \sum_j w_{jk} a_j$ the sum of weighted inputs to the neuron k, $\bar{a}_k(t)$ is the average of the post-synaptic activity over a time T and the function:

$$\phi(a_k(t), \bar{a}_k(t)) < 0 \text{ if } a_k(t) < (\frac{\bar{a}_k}{a_0})^p \bar{a}_k(t)$$

and

$$\phi(a_k(t), \bar{a}_k(t)) > 0 \text{ if } a_k(t) > (\frac{\bar{a}_k}{a_0})^p \bar{a}_k(t).$$

The average post-synaptic activity thus acts as a sliding threshold determining whether the strength of the connection is increased or decreased. This plasticity rule leads to the development of orientation selectivity in a stable manner.

Another set of rules is related to the instrinsic plasticity changing the response function of a neuron rather than the strength of the connection to constrain output activity to have maximum entropy [143] or to follow an exponential distribution [144]. The output y of the neuron is related to its input x by a sigmoidal response function: $y = \frac{1}{1+e^{-(\sigma x+\theta)}}$ with $\sigma > 0$ the slope of the quasi linear part and $\theta > 0$ the threshold. The probability distribution function of y and x are related by $p_y(y) = \frac{p_x(x)}{\partial y/\partial x}$ with

$$\frac{\partial y}{\partial x} = \sigma y(1-y)$$

for a sigmoidal function. A stochastic gradient rule for maximizing a function $L(y)$ is built as follows:

$$\sigma(t+1) = \sigma(t) + \eta \frac{\partial L(y)}{\partial \sigma}$$

$$\theta(t+1) = \theta(t) + \eta \frac{\partial L(y)}{\partial \theta}$$

The differential entropy of the output is $h(y) = -\int_{-\infty}^{\infty} p_y(y) ln(p_y(y)) dy$. With the previous expressions for p_y, $h(y) = \mathbf{E}[\ln(\sigma y(1-y))] - \mathbf{E}[\ln(p_x(x))]$, but the second term doesn't contribute to the output of the neuron so that maximizing the entropy comes to maxmize $L(y) = ln\frac{\partial y}{\partial x}$. We thus obtain the Bell-Sejnowski (BS) rule:

$$\sigma(t+1) = \sigma(t) + \eta(\frac{1}{\sigma(t)} + x(t)(1-y(t))$$

$$\theta(t+1) = \theta(t) + \eta(1-2y(t))$$

with $\eta > 0$, a free parameter similar to the learning rate encountered in synaptic learning rules. The mutual information between the input and the output of the neuron is $I(x,y) = h(y) - h(y|x)$ so that this rule is also maximizing the mutual information between the input and the output. In the rule for σ, correlated activity leads to a decrease of the parameter and for that reason it is another example of anti-Hebbian learning. Constraining the output distribution to be exponential can be formalized as minimizing the Kullback-Leibler divergence between $p_y(y)$ and the target exponential distribution p_{tar} of parameter $\mu > 0$:

$$D(p_y||p_{tar}) = \int_{-\infty}^{\infty} p_y(y) ln \frac{p_y(y)}{\frac{1}{\mu} e^{-\frac{x}{\mu}}}$$

by keeping only terms depending on y, this minimization is equivalent to the maximization of $L(y) = h(y) - \frac{1}{\mu}\mathbf{E}[y]$ that is close to what was done when deriving the BS rule with an additional term for constraining the mean firing rate. The Triesch rule is then

$$\sigma(t+1) = \sigma(t) + \eta(\frac{1}{\sigma(t)} + x(t) - (2+\frac{1}{\mu})x(t)y(t) + \frac{1}{\mu}x(t)y(t)^2)$$

$$\theta(t+1) = \theta(t) + \eta(1 - (2 + \frac{1}{\mu})y(t) + \frac{1}{\mu}y(t)^2).$$

Neuromodulation mediated by acetylcholine or serotonin changes the excitability property of these neurons and could thus support such an intrinsic plasticity. A review of models of neuromodulation is provided in [145]. Hallucination patterns could originate from such changes in intrinsic excitability.

3.4.3 Kohonen network.

A simplified version of the SOM was proposed in [73]. As shortly presented in the first chapter, the algorithm is composed of two parts:

- A competitive stage where the cortical neuron r giving maximal response is selected.
- A cooperative stage where weights are updated locally around the winning neuron.

The rule for updating weights is:

$$w_{ik}(t+1) = w_{ik} + \alpha(a_i - w_{ik})h_{rk}(t)$$

with

$$h_{rk}(t) = e^{-\frac{(x_r - x_k)^2 + (y_r - y_k)^2}{\sigma(t)^2}}$$

the neighborhood function which extent σ is decreased across time in a similar way to what was presented for the elastic net. Because it doesn't include explicitly the activity of the cortical units but only the activity of the retinal units, this model is computationnally less expensive than the original model proposed by Von der Malsburg and it enables mathematical analysis. Kohonen then demonstrated some general properties of its algorithm:

- The input space is more precisely mapped along dimensions having the largest standard deviation.
- Close points in the input space are mapped onto close points in the cortical space.
- Over-represented domains of the input space are mapped onto large domains of the cortical space.

An analysis of the Kohonen network is applied to the formation of primary visual cortex was done in [146]. A cortical unit is located at position $\mathbf{r} = (r_1, r_2)$ in a 2 dimensional space with periodic boundary conditions and is represented by a 5 dimensional feature vector $\mathbf{w}(\mathbf{r})$. The feature vector is a way to code the receptive field properties of the cortical unit and contains:

- x and y, the positions in the retinal space which it is coding for.

- $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, the components of the vector representing the orientation preference and selectivity of the unit.
- $o = \frac{n_R - N/2}{N}$ characterising the ocular dominance of the unit, with n_R the density of synapses connecting to the right eye and N the density of both types of synapses (right or left). It is 0 if the unit is perfectly binocular, 1 if it receives inputs from the right eye and -1 from the left eye.

The inputs presented to the network are 2 dimensional Gaussians. Each input $\mathbf{a}$ can be represented in the feature space:

$$\mathbf{a} = \begin{pmatrix} x \\ y \\ q\cos(2\phi) \\ q\sin(2\phi) \\ o \end{pmatrix}$$

with (x, y) the position of the center of the Gaussian in the retinal space, q its ellipticity, ϕ the angle between the principal axis of the Gaussian and the horizontal axis in retinal space and o the occularity of the input. Given a probability distribution $P(\mathbf{a})$ of inputs, with $\mathbf{a} \in \mathcal{A}$, the average change is

$$\mathrm{E}(\Delta\mathbf{w}(\mathbf{r})|\mathbf{w}(\mathbf{r})) = \alpha \int_{\mathcal{A}} (\mathbf{a} - \mathbf{w}(\mathbf{r}))h_{\mathbf{s(a)r}}P(\mathbf{a})d\mathbf{a}$$

with the winner unit being selected by:

$$s(\mathbf{a}) = min_{\mathbf{r}}|\mathbf{a} - \mathbf{w}(\mathbf{r})|$$

The stationary states are the ones for which:

$$\mathrm{E}(\Delta\mathbf{w}(\mathbf{r}|\mathbf{w^0}(\mathbf{r})) = 0$$

An obvious stationary state is the purely retinotopic state :

$$\mathbf{w}^0 = \begin{pmatrix} Mr_1 \\ Mr_2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

where the cortical space is a copy of the retinal space with the magnification factor M. A whole class of stationary states can be obtained by considering translations, reflection or rotations by an angle multiple of π of $\mathbf{w}^0$. The formation of columnar organization associated to orientation and occularity ca, be understood by analyzing the stability through perturbations of these stationary states. This is done in Obermayer 1992 by writing the Fokker-Planck equation in Fourrier space for $\mathbf{u}(\mathbf{r}) = \mathbf{w}(\mathbf{r}) - \mathbf{w}^0(\mathbf{r})$. The order parameters, $(T_i)_{1 \leq i \leq 5}$,

for this equation are the standard deviation of the inputs along each dimension i of the input space:

$$T_i = \frac{\int_{\mathcal{A}}(a_i - < a_i >)^2 P(\mathbf{a})d\mathbf{a}}{\int_A P(\mathbf{a})d\mathbf{a}}$$

with $< a_i >= \int_{\mathcal{A}} a_i P(\mathbf{a})d\mathbf{a}$. The eigenvalues of the linear part of the Fokker Planck equation becomes positive, making the stationary state $\mathbf{w}^0$ unstable, when

$$T_i > T_{\text{thresh}} = \frac{1}{2}\sqrt{e}M\sigma.$$

While inputs keep being similar in their orientation and occularity dimensions, the purely retinotopic state is stable. As the variance in these dimensions crosses a threshold, columnar organization is formed. The fluctuations around the stationary state when inputs are in the suprathreshold threshold regime can be described by the correlation function calculated in [146]:

$$C_{ii}(k) = \frac{\alpha}{2}\pi T_i^2 \sigma^2 \frac{e^{-\frac{\sigma^2}{4}k^2}}{e^{-\frac{\sigma^2}{4}k^2} - \frac{T_i^2 k^2}{M^2}}$$

where $i > 2$ are the indices corresponding to the orientation and occularity dimensions. These fluctuations are finite for $k \to 0$ and grow linearly with the learning rate α.

3.4.4 Plastic lateral connections-LISSOM model

The Kohonen algorithm is based on a winner take all competition so that only one input can be presented at a time. Another pitfall is the symmetry implied by the Gaussian neighborhood function. In visual cortex, the connections between columns are not isotropic and this can be changed by imposing some dynamics on the weights associated to lateral connections.

The Lateral Interactions Synergically Self Organized Maps, LISSOM, model overcomes these problems [147]. The LISSOM can be summarized in five principles:

- The cortical layer is composed by excitatory and inhibitory units disposed regularly on a $N \times N$ grid.
- Inputs from the retina are transmitted through ON and OFF channels of the LGN ($N \times N$). Connections from the retina to the LGN are differences of Gaussian functions, mexican-hat like for the ON channel and reversed for the OFF channel.
- Lateral connections are plastic and initialized as Gaussians with inhibition wider than excitation.
- The activity of a unit is determined by summing linearly the inputs and applying a sigmoidal function or its piecewise linear approximation.

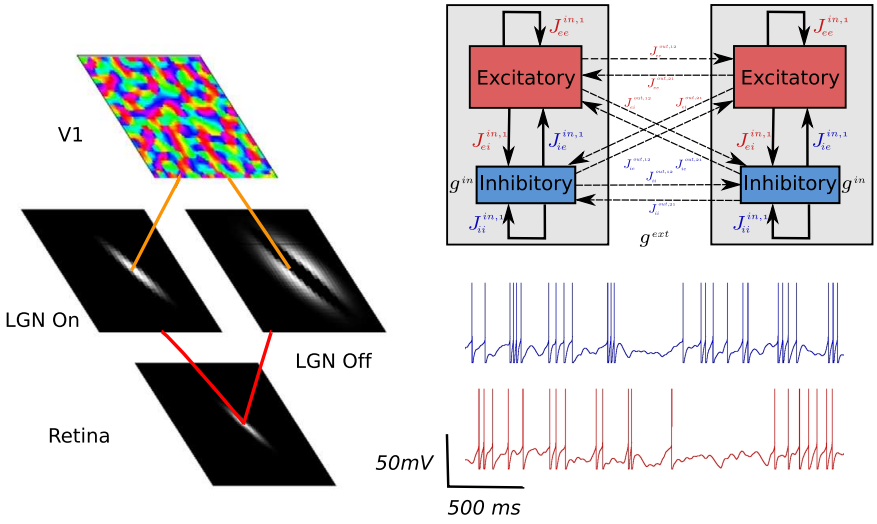

Figure 3.11: **LISSOM Model** - (Left) Description of the LISSOM architecture. Connections in red are fixed and those in orange are plastic. (Right-Top) Wiring diagram between two columns. (Right-Bottom) Typical dynamics of an excitatory (red) and inhibitory (blue) cells.

- Weight dynamics follows Hebbian rule with divisive normalization.

For each stimulus presentation a^{Ret}, the activity of a LGN unit located at position $\mathbf{y}$ is:

$$a^{LGN}(\mathbf{y}) = h(\sum_{\mathbf{x}} w^{Ret-LGN}(\mathbf{x},\mathbf{y})a^{Ret}(\mathbf{x}))$$

where h is the piecewise linear approximation of a sigmoid:

$$h(a) = \begin{cases} 0 & \text{if } a \leq \theta_1 \\ \frac{a-\theta_1}{\theta_2-\theta_1} & \text{if } \theta_1 < a < \theta_2 \\ 1 & \text{if } a \geq \theta_2 \end{cases}$$

The input from LGN to a cortical unit at location $\mathbf{y}$ is

$$s(\mathbf{y}) = \gamma_{LGN}(\sum_{\mathbf{x}\in LGN_{On}} w^{LGN_{On}}(\mathbf{x},\mathbf{y})a^{LGN}(\mathbf{x}) + \sum_{\mathbf{x}\in LGN_{Off}} w^{LGN_{Off}}(\mathbf{x},\mathbf{y})a^{LGN}(\mathbf{x}))$$

At the first step of an input presentation, the cortical activity in $\mathbf{y}$ is

$$a^{Cort}(\mathbf{y},0) = h(s(\mathbf{y})$$

and it includes lateral inputs in the next time steps

$$a^{Cort}(\mathbf{y},t) = h(s(\mathbf{y}+\gamma_E \sum_{\mathbf{x}\in Cort} w^{E}(\mathbf{x},\mathbf{y})a^{Cort}(\mathbf{x},t-1)+\gamma_I \sum_{\mathbf{x}\in Cort} w^{I}(\mathbf{x},\mathbf{y})a^{Cort}(\mathbf{x},t-1))$$

γ_{LGN}, γ_E and γ_I are the relative strengths of afferent, excitatory and inhibitory connections. The rule for updating the weights coming to a unit located in $\mathbf{y}$ is

$$w'(\mathbf{x},\mathbf{y}) = \frac{w(\mathbf{x},\mathbf{y}) + \alpha a(\mathbf{x})a(\mathbf{y})}{\sum_{\mathbf{z}} w(\mathbf{z},\mathbf{y}) + \alpha a(\mathbf{z})a(\mathbf{y})}$$

3.5 Dynamics of the spiking neurons network before and after LISSOM learning.

In the same spirit to what was done for the pinwheel in section, a network is built from a connectivty scheme obtained by LISSOM learning. The LISSOM network is run through the Topographica library [3]. The connectivity is extracted before and after learning to study the different dynamics arising from the corresponding networks.

3.5.1 LISSOM implementation and spiking neurons network.

LISSOM simulation. The parameters used in the LISSOM algorithm are listed in fig 3.12.

[3] Available at http://topographica.org/

Parameter	Value	Description
Pattern of stimulation		
n_I	2	Number of bars at each presentation.
s_I	0.088388	Size of a bar (width of the small axis of Gaussian).
ar_I	4.66667	Aspect ratio of the bar (long axis)/(small axis).
sep_I	0.595826	Minimum of separation between the centers of 2 bars.
Sheets		
d_{Ret}	28	density of retinal units.
d_{LGN}	56	density of LGN units.
d_{Cort}	56	density of cortical units.
Connections		
Retina – LGN		
σ_{center}	0.07385	Size of the center Gaussian for On and Off receptive fields.
$\sigma_{surround}$	0.29540	Size of the surround Gaussian for On and Off receptive fields.
γ_R	2.33	Strength of afferent connections to LGN.
LGN – $\mathbf{V_1}$		
σ_{LGN}	0.27083	Size of the Gaussian from which the weights of the afferent connectivity are initialized randomly.
γ_{LGN}	1.0	Strength of afferent connections to V_1.
α_{LGN}	0.47949	Learning rate for afferent connections to V_1.
$\mathbf{V_1}$ – $\mathbf{V_1}$		
σ_E	0.10417	Size of the Gaussian from which lateral excitatory connections are initialized.
σ_I	0.22917	Size of the Gaussian from which lateral inhibitory connections are initialized.
γ_E	0.9	Strength of lateral excitatory connections.
γ_I	-0.9	Strength of lateral inhibitory connections.
α_E	2.55528	Learning rate for lateral excitatory connections.
α_I	1.80873	Learning rate for lateral inhibitory connections.

Figure 3.12: Parameters used in the LISSOM model for the learning of connectivity.

During this learning phase, all connections have a delay of 0.05. After each stimulus presentation, weights are let evolving during a time $t_{settle} = 2$. A particular feature of the algorithm is a shrinkage of the excitatory kernel similar to what was described for the Elastic Net. At the 300^{th} learning time step, all excitatory connections further than 2 columns of the central unit are pruned and excitatory weights are restricted to this central region as can be visualized on the fig 3.13. The evolution in the shape of receptive fields are illustrated in fig 3.14 for afferent connections and in fig 3.15 for lateral connections. The rotational symmetry is broken by the learning and this results in the emergence of functional properties of the units for detecting orientations. An example of the resulting orientation map is shown in fig 3.16. With a reverse difference of Gaussians connectivity, only two orientations are selected and organize in a checkboard pattern. With long range excitatory connections, stable orientation maps with pinwheels can emerge only if a second $V1$ layer is added [147].

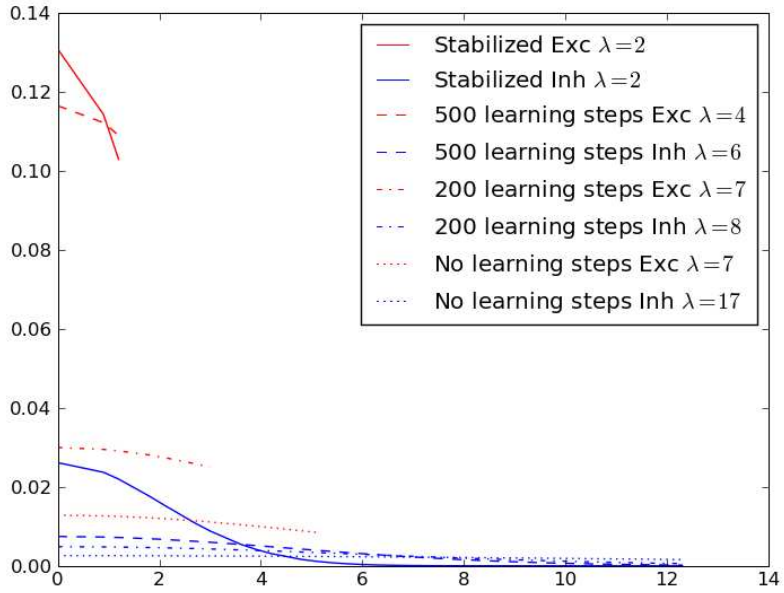

Figure 3.13: **Evolution of the connectivity profile.** - The average profile of connectivity is shown for different stages of the learning with excitatory connections in red and inhibitory connections in blue. Each profile is fitted with a Gaussian function ($e^{\frac{x^2}{\lambda^2}}$) and the corresponding λ is reported in the box.The shrinkage can be detected between the profiles at 200 and 500 learning steps.

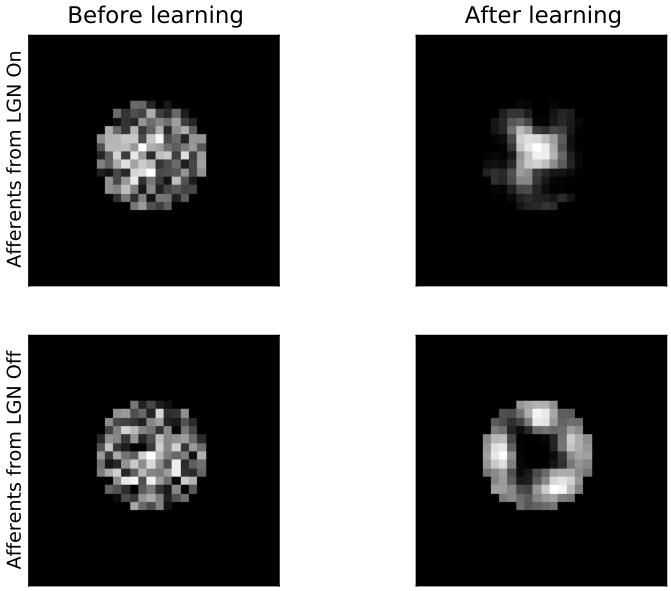

Figure 3.14: **Evolution of the afferent projective fields.** - The projective fields shows the connection weights from a central LGN afferents neuron (ON or OFF) to V_1. Before learning, it is initialized from a 2D Gaussian distribution(Bl.

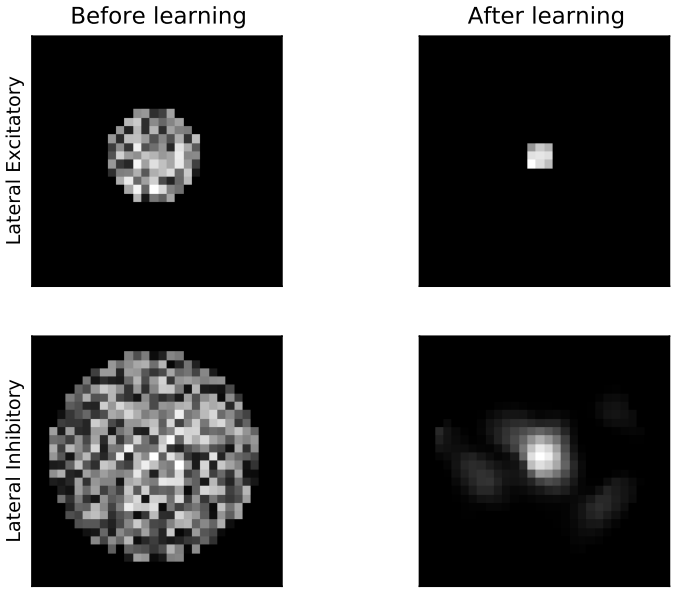

Figure 3.15: **Evolution of the lateral connections.** - The excitatory and inhibitory connections in V_1 are shown before and after learning.

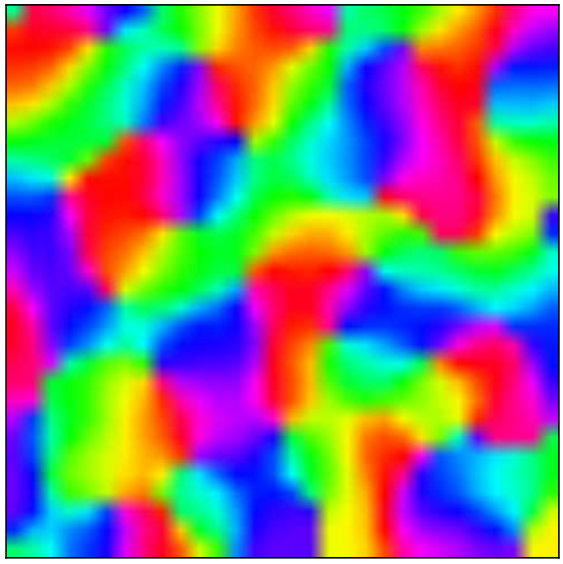

Figure 3.16: **Orientation preference map.** - Resulting map after 20000 time steps. Each color codes for an orientation between $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Exportation of the connectivity. The V_1 sheet has dimensions 56×56 but in the corresponding spiking neurons network, to avoid boundary effects on the connectivity, only the 30×30 central part is extracted. Weights obtained with LISSOM are between 0 and 1 with many small values, these values can then be used as probability of connection between two columns. This probability will be multiplied by a factor l with a saturation, so that the maximum probability of connection is 1, to scale the influence of lateral connections. The differences between the networks before and after learning can be quantified by measuring their graph properties. The graph of neurons built from the connectivity described above has directed connections with only one connection possible between two neurons. For each neuron, the in degree is the number of neurons connecting to this neuron and the out degree is the number of neurons to which it connects. Considering the same graph as undirected, if there is a connection for each pair in three neurons, there is a triangle passing through these neurons. The clustering coefficient of a node v of the graph is the fraction of possible triangles that actually exist:

$$C_v = \frac{2T(v)}{deg(v)(deg(v)-1)}$$

This is computed using the networkx Python library [4] and shown in fig 3.17. In and out degrees don't change after learning but the clustering coefficient goes from 0.01 before to 0.12 (0.05) for excitatory (inhibitory) neurons after learning. This clustering of the connections is associated to the smaller extent of connectivity profiles shown in fig 3.13 and the break in anisotropy shown in fig 3.15.

Spiking neurons network implementation Implementation of the network is done in with the PyNN implementation of the NEST simulator [148] and analysis of the result is done using Neurotools [5]. The adaptive exponential integrate and fire equations model the dynamics of each neuron with parameters described in chapter II for excitatory and inhibitory cells. Each column is composed of 80 excitatory neurons and 20 inhibitory neurons. Such network are difficult to study because many parameters are involved and insights for the analysis from coarse grained description of the network before analyzing numerical simulations.

3.5.2 Rate description of the $V1$ model before learning.

The connectivity kernels are isotropic at the beginning of the learning process and due to this symetry, the associated rate model can be analyzed easily. A important difference is that in the rate model, the spatial area is supposed to be an infinite plane whereas in the simulations it is a $30x30$ grid. In order to avoid border effects in our simulations, we looped connections to make boundary conditions periodic.

[4] Available at http://networkx.lanl.gov/.
[5] Both available at www.neuralensemble.org

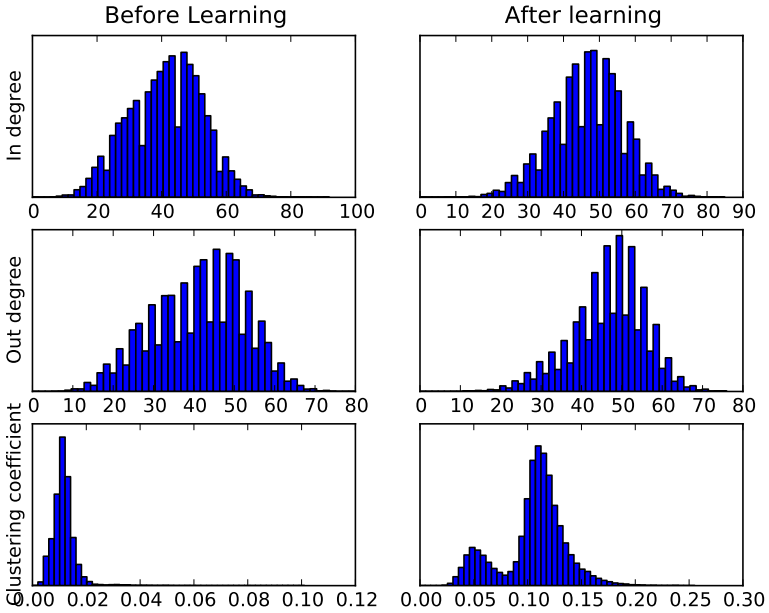

Figure 3.17: **Graph properties of the network** - (Left) Before learning. (Right) After learning.

Eigenforms of the two dimensional model.

Wilson-Cowan equations in continuous space. As we saw in chapter II, the Wilson-Cowan equations give a coarse grained description of the activity of a neuronal network by the dynamics of its average firing rate. The same model can be considered in its spatially continuous version, where the space is represented by a 2D plane with infinite or periodic boundary conditions as was done in [149] [150] [151] [152]. Note that the activity based model described in Chapter 2 can also be extended to this limit. Considering the average rate of a group of neurons located at the position $\mathbf{r}$ in the plane as $E(\mathbf{r},t)$ for the excitatory population and $I(\mathbf{r},t)$ for the inhibitory one, the dynamics evolve as:

$$\frac{\partial E}{\partial t} = -E + f_E(\alpha_{EE}\mu w_{EE} \star E - \alpha_{IE} w_{IE} \star I)$$

$$\frac{\partial I}{\partial t} = -I + f_I(\alpha_{EI}\mu w_{EI} \star E - \alpha_{II} w_{II} \star I).$$

with:

- $\star$ standing for spatial convolution, $f \star g(\mathbf{x}) = \int_{\mathbf{R}^2} f(\mathbf{x}-\mathbf{y})g(\mathbf{y})d\mathbf{y}$
- f_k, with $k \in E, I$, the same kind of response function encountered in Chapter 2, it can be taken as sigmoidal function with $f_k(0) = 0$.
- $w_{kl}(\mathbf{r},\mathbf{r}')$, with $k,l \in E,I$, is the strength of the connection from the group of neurons located in $\mathbf{r}$ to the group of neurons located in $\mathbf{r}'$. This function only depends on the relative distance $|\mathbf{r}-\mathbf{r}'|$ between the two groups. It is positive, bounded with $\int_{\mathbf{R}^2} w_{kl}(\mathbf{r})d\mathbf{r} = 1$ and its Fourrier transform $\hat{w}(\mathbf{k}) = \int_{\mathbf{R}^2} w(\mathbf{r})e^{i\mathbf{kr}}d\mathbf{r}$ is a decreasing function of $|\mathbf{k}|$. A common choice for this connectivity kernels is the Gaussian function $w(\mathbf{r} = e^{-|r|^2}$.
- μ is a parameter modulating the excitability of the network and α_{kl}, $k,l \in E,I$, scales the contribution of population k to the input of population l.

$(E_0(\mathbf{r}), I_0(\mathbf{r})) = (0,0)$ is a stationary solution of the system and we wish to track instabilities which can arise from this uniform state. The stability of the uniform solution can be deduced from the linear system equivalent to the original equations near (E_0, I_0).

$$\frac{\partial E}{\partial t} = -E + f'_E(0)\alpha_{EE}\mu w_{EE} \star E + f'_I(0)\alpha_{IE} w_{IE} \star I$$

$$\frac{\partial I}{\partial t} = -I + f'_E(0)\alpha_{EI}\mu w_{EI} \star E + f'_I(0)\alpha_{II} w_{II} \star I$$

and by the expansions $E(\mathbf{r},t) = \sum_{\mathbf{k}} \hat{E}(\mathbf{k})e^{\lambda t + i\mathbf{kx}}$ and $I(\mathbf{r},t) = \sum_{\mathbf{k}} \hat{I}(\mathbf{k})e^{\lambda t + i\mathbf{kx}}$, we reach the following system for the instability mode $\hat{\mathbf{A}} = (\hat{E}(\mathbf{k}), \hat{I}(\mathbf{k}))$ of wave vector $\mathbf{k}$:

$$\lambda \hat{\mathbf{A}}(\mathbf{k}) = B(\mu)\hat{\mathbf{A}}$$

with

$$B(\mu) = \begin{pmatrix} -1 + f'_E(0)\alpha_{EE}\mu\hat{w}_{EE}(\mathbf{k}) & f'_I(0)\alpha_{IE}\hat{w}_{IE}(\mathbf{k}) \\ f'_E(0)\alpha_{EI}\mu\hat{w}_{EI}(\mathbf{k}) & -1 + f'_I(0)\alpha_{II}\hat{w}_{II}(\mathbf{k}) \end{pmatrix}$$

with $\hat{w}_{kl}(\mathbf{k}) = \int_{\mathcal{D}} w_{kl}(\mathbf{r})e^{i\mathbf{k}.\mathbf{r}}d\mathbf{r}$, the Fourrier transform of the connectivity kernel, which for a Gaussian function is $w_{kl}(\mathbf{k}) = \sqrt{\pi}e^{-\pi^2|\mathbf{k}|^2}$. This eigenvalue problem leads to the dispersion relation $\det(B(\mu) - \lambda I) = 0$. Eigenvalues of B are

$$\lambda_{\pm} = \frac{1}{2}(tr(B) \pm \sqrt{tr(B)^2 - 4det(B)}.$$

If $\lambda_{\pm}$ both have a negative real part, the uniform state is stable. If one of them is positive, unstable modes unfolds resulting in Turing patterns and moreover if $\mathrm{tr}(B) > 0$ and $\mathrm{tr}(B)^2 < 4\det(B)$, those patterns are oscillatory which is known as the Turing-Hopf mechanism for instability [126]. When μ is increased, at least one of the eigenvalues for a given $\mathbf{k}$ becomes positive because $\hat{w}_{kl} > 0$ thus provoking the emergence of patterns of spatial frequency $\mathbf{k}$. The selected $\mathbf{k}$ is the one for which $\lambda(\mathbf{k})$ is maximum that is the one at which $\frac{d\lambda(|\mathbf{k}|)}{d|\mathbf{k}|} = 0$.

Reduced model. The previous model gives the spatial periodicity of the emerging pattern but the rotational and translational symmetries of the connectivity kernels make several doubly periodic patterns possible [6] The relative stability of patterns with different symmetries can be studied on a reduced one dimensional activity based model [7], as proposed in [153]. The activity at a position $\mathbf{r}$ follows:

$$\frac{\partial a(\mathbf{r},t)}{\partial t} = -a(\mathbf{r},t) + \int_{\mathcal{D}} w(|\mathbf{r} - \mathbf{r}'|)f(a(\mathbf{r}'))d\mathbf{r}'.$$

with f a sigmoidal function and w the difference of Gaussians connectivity kernel of extents σ_e, σ_i and amplitudes A_e, A_i as encountered in the Swindale model of part 2 in this chapter:

$$w(|\mathbf{r}|) = \frac{A_e}{\sigma_e}e^{-\frac{x^2}{\sigma_e^2}} - \frac{A_i}{\sigma_i}e^{-\frac{x^2}{\sigma_i^2}}$$

of Fourrier transform:

$$\hat{w}(|\mathbf{k}|) = \sqrt{\pi}(A_e e^{-\pi^2\sigma_e^2|\mathbf{k}|^2} - A_i e^{-\pi^2\sigma_i^2|\mathbf{k}|^2}).$$

The dispersion relation for this simplified model is :

$$\lambda(|\mathbf{k}|) = -1 + \mu\hat{w}(|\hat{k}|)$$

with $\mu = f'(0)$ as a bifurcation parameter. As μ increases there is a range of $k = |\mathbf{k}|$ for which $\lambda(k) > 0$ around the k_0 corresponding to the maximum of λ as ca be seen on fig 3.18.

[6] Periodic in the x and y dimensions.

[7] This model is obtained from the two populations model by considering that the inhibitory dynamics are faster than the excitatory one, replacing the sigmoid in the equation for inhibitory activity by its linear approximation and taking $w_{II} = 0$.

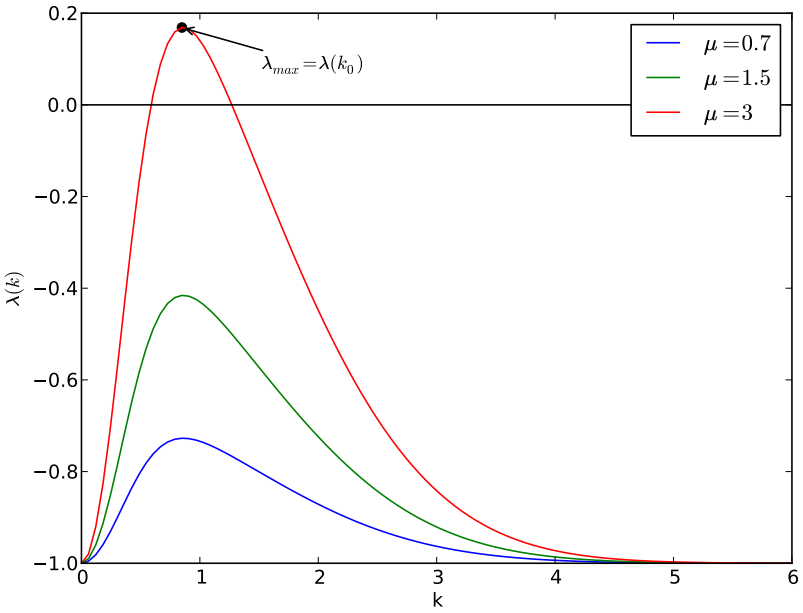

Figure 3.18: **Eigenvalue λ as a function of the norm of the wavevector k.** - This function is plotted for 3 increasing values of μ. When there exist a range over which λ is positive, the k for which λ is maximum indicates the spatial frequency of the emerging pattern.

Near the critical value of μ where a non-uniform pattern arise, the periodicity will be around k_0 but there are several patterns respecting this symmetry:

- For the square pattern, the activity can be expanded as

$$b(\mathbf{x},t) = c_1(t)e^{i\mathbf{k_1 x}} + c_1^*(t)e^{-i\mathbf{k_1 x}} + c_2(t)e^{i\mathbf{k_2 x}} + c_2^*(t)e^{-i\mathbf{k_2 x}}$$

  with $\mathbf{k}_1 = k_0(1,0)$ and $\mathbf{k}_2 = k_0(0,1)$.

- For the rhombic pattern, the activity can be expanded in a similar way but with $\mathbf{k}_2 = k_0(cos\phi, sin\phi)$.

- For the hexagonal pattern, the activity is expanded in the three terms corresponding to the three vectors generating the pattern:

$$\mathbf{k}_1 = k_0(1,0),\ \mathbf{k}_2 = k_0(\frac{1}{2}, -\frac{1}{2}) \text{ and } \mathbf{k}_3 = k_0(-\frac{1}{2}, -\frac{1}{2})$$

The equations for the dynamics of the coefficients of this expansion, c_n, can be obtained by applying perturbation methods and give the relative stability of the possible patterns. The resulting approximation is the following cubic equation for the square or rhombic pattern:

$$\frac{\partial c_n}{\partial \tau} = c_n(\mu - \mu_c - \Gamma_0|c_n|^2 - 2\Gamma_\phi|c_m|^2)$$

with $m, n = 1, 2$ and $m \neq n$. A model including anisotropic long range connections was studied in [134] [154].

Localized bump of activity and traveling wave.

Bump solution. Other types of solutions such as localized bumps and traveling waves also exist in such neural field models (see [155] for a review). An bump solution has been found in [156] when a Heaviside function [8] is used for f in the following one-dimensional layer:

$$\frac{\partial a(x,t)}{\partial t} = -a(x,t) + \int_{-\infty}^{\infty} w(|x-y|)f(a(y,t))dy$$

with w being a difference of Gaussians. A bump solution $A(x)$ of length d and centered on x_0 is defined as a stationary solution such that $A(x) = 0$ on $]-\infty, x_0 - \frac{d}{2}]$ and $A(x) > 0$ on $]x_0 - \frac{d}{2}, x_0 + \frac{d}{2}[$. For such a solution centered on $x_0 = 0$:

$$-A(x) + \int_{R[A(x)]} w(x-x')dx = 0$$

[8] The Heaviside function is defined as follows:

$$H_\theta(x) = \begin{cases} 1 & \text{if } x > \theta \\ 0 & \text{else} \end{cases}$$

with $R[A] = x|A(x) > 0$. By considering this equation at $x = -\frac{d}{2}$, we obtain $h + \int_0^d w(x)dx = h + W(d) = 0$ and this condition is also sufficient. Moreover the stability of this solution is studied by considering that the excited region have moving boundaries $x_1(t) < x_2(t)$ with the slope of the pattern at boundaries being $c_i = \frac{\partial a(x_i,t)}{\partial x}$ $(i = 1, 2)$. From the definition of the excited region, $a(x_i, t) = 0$, $a(x_i + dx_i, t + dt) = 0$ and thus

$$\frac{\partial a(x_i,t)}{\partial t}dt + \frac{\partial a(x_i,t)}{\partial x}dx = 0$$

$$\frac{dx_i}{dt} = -\frac{\partial a(x_i,t)/\partial t}{\partial a(x_i,t)/\partial x}$$

and thus

$$\frac{dx_i}{dt} = \frac{1}{c}(\pm W(x_2 - x_1) + h)$$

so that the evolution equation for the length of the excited region $l(t) = x_2(t) - x_1(t)$ is:

$$\frac{dl}{dt} = (\frac{1}{c_1} + \frac{1}{c_2})[W(l) + h]$$

. The stationary solution with length l_0 is stable if $\frac{dW(l)}{dl}|_{l=l_0} < 0$.

Travelling wave. A traveling wave for this model is a solution $a(z)$, $z = x - ct$ where c is a constant wavespeed, with:

- a monotonic.
- $0 \leq a \leq 1$.
- $a(-\infty) = 0$, $a(\infty) = 1$.

In [157], the existence and uniqueness of such a traveling wave in one dimensional rate model was demonstrated when the response function, f, checks the following properties:

- f continuously differentiable with $f' > 0$.
- $F(a) = -a + f(a)$ has three zeros at 0,$0 < a_0 < 0$ and 1.
- $f'(0) < 0$ and $f'(1) < 1$.

The solution $a(z)$ can be replaced in the model equation, with $z = x - ct$ and $z' = x' - ct$:

$$-c\frac{\partial a(z)}{\partial z} = -a(z) + \int_{-\infty}^{\infty} w(z - z')f(a(z'))dz'$$

k is such that

$$\int_{-\infty}^{\infty} k(x)dx = 1$$

and then

$$-ca'^2 f'(u) = F(a)f'(a)a' + \int_{-\infty}^{\infty} w(z-z')(f(a(z')) - f(a(z')))f'(a(z))a'(z)dz'$$

Integrating according to z gives:

$$-c\int_{-\infty}^{\infty} a'^2 f'(a)dz = \int_0^1 F(a)f'(a)da + \int_{-\infty}^{\infty} w(z-z')(f(a(z')-f(a(z))f'(a(z))a'(z)dzdz'.$$

The last integral can be rewritten:

$$\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(z-z')(f(a(z')-f(a))(f'(a(z))a'(z) - f'(a(z'))a'(z'))dzdz'$$

then by changing $z' \to z - s$, this integral is:

$$\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(s)(f(a(z-s)-f(z))(f'(a(z))a'(z) - f'(a(z-s))a'(z-s))dzds$$

which when integrated along z is 0 so that:

$$-c\int_{-\infty}^{\infty} a'^2 f'(a)dz = \int_0^1 F(a)f'(a)da.$$

Moreover

$$\int_0^1 F(a)(f'(a)-1)da = \frac{1}{2}[F(a)^2]_0^1 = 0$$

and then

$$c = -\frac{\int_0^1 F(a)da}{\int_{-\infty}^{\infty} a'^2 f'(a)da}$$

the speed of the wave thus being of the same sign as $-\int_0^1 F(a)da$. In models taking adaptation into account, dynamic patterns are found in [158] [159] [160].

3.5.3 Dynamics in the network of spiking neurons before learning

Periodic boundary conditions: Phase diagram and emergence of static patterns. As was done for the Ice Cubes model, we provide a phase diagram (g_I, g_E) on fig 3.19 indicating the mean and variance of firing rate and the mean of the coefficient of variation of the ISI. There are 3 regimes:

- When excitation dominates, neurons in the whole network fire at maximal frequency. Those regimes are located in the red area of the top panel in fig 3.19.

- When inhibition dominates, neurons of the network fire in an asynchronous irregular fashion. Those regime with high CV are located in the red area of the middle panel in fig 3.19.
- In the well-balanced regime, we see patterns with some parts of the network firing at a high frequency and other parts having a low firing frequency. Those patterns are similar to the eigenforms found in the previously described Turing mechanism. A snapshot of such a pattern is shown in fig 3.20.

The bifurcation diagram for the emergence of the pattern as g_E increases is shown in fig 3.21 for $g_I = 1.5$. At low g_E, the firing rate map is homogeneous with low firing rate. At high values of g_e, the map is also homogeneous but with high firing rate. For intermediate values of g_E, there are some non-homogeneous patterns. Note that, if we only look at the average firing over the network, this transition is not seen.

Finite square boundaries: traveling wave. In another set of simulations, we take keep finite square boundaries and we provide only a local stimulation by connecting Poissonian spike trains as input to the 4 central columns. The firing rate model suggest that the response of the network should be a static bump but due to inhomogeneities of the graph and in the input stimulation, this bump starts moving in a given direction and it is then reflected on the boundaries. For some regimes of (g_E, g_I), the initial bump splits into three bumps rotating on the map.

3.5.4 Dynamics in the network of spiking neurons after learning.

Phase diagram The first parameter varied when building the phase diagram is the balance between maximal conductance of inhibition and excitation $\frac{g_I}{g_E}$ with keeping $g_E + g_I = 0.024nS$ (verifier 0.024) and the second parameter is the lateral strength that is a factor multiplying the probabilities of connection obtained after the LISSOM learning in the rate model. As for the previous model, we detect three regimes:

- The saturated regime (S) is found for low values of $\frac{g_I}{g_E}$ and high values of the lateral strength. In this regime, all neurons fire spikes regularly at their maximal frequency such that the coefficient of variation of interspikes intervals is 0.
- An asynchronous irregular (A) regime ($CV \approx 1$) is found at high values of $\frac{g_I}{g_E}$ and low values of the lateral strength. This well-mixed regime is a good candidate for the on-going activity in awake state.
- A synchronous bursty (B) regime is found in between the two previously described regimes. In this regime, neurons fire spikes in a supra-Poissonian fashion ($CV > 1$).

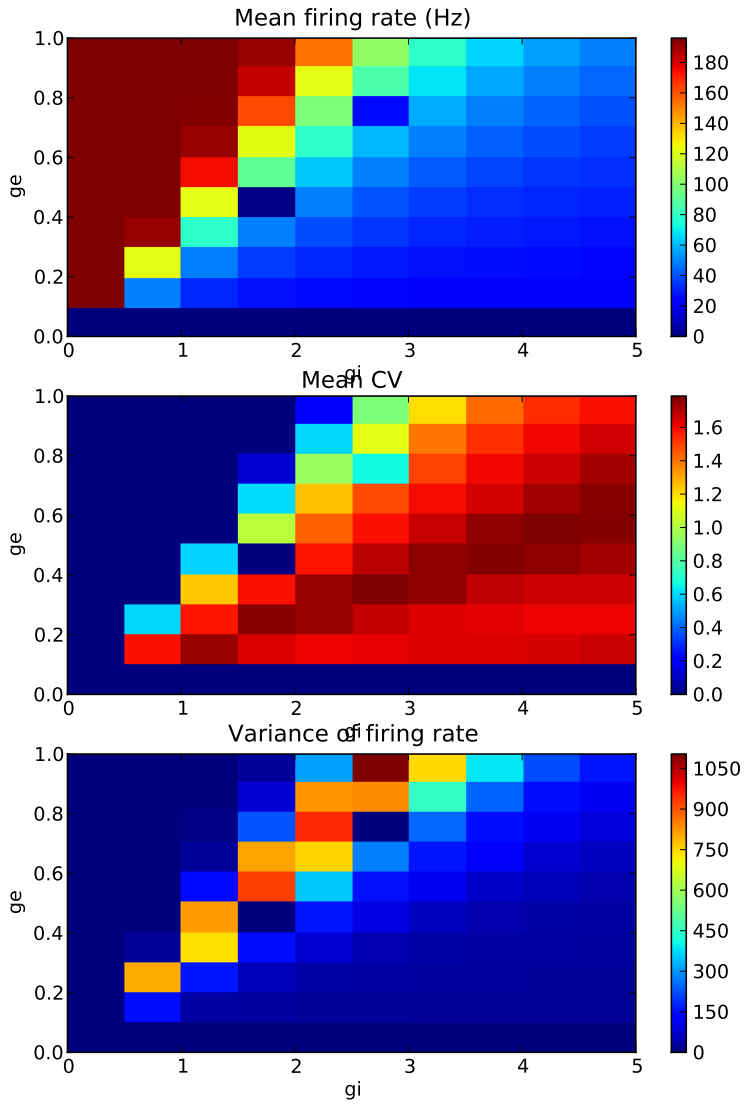

Figure 3.19: **Phase diagrams (g_E, g_I) for the network before learning** - The indicated values of conductances g_E, g_I are normalized so that it should be multiplied by $0.006nS$ to get the values used in the simulations.

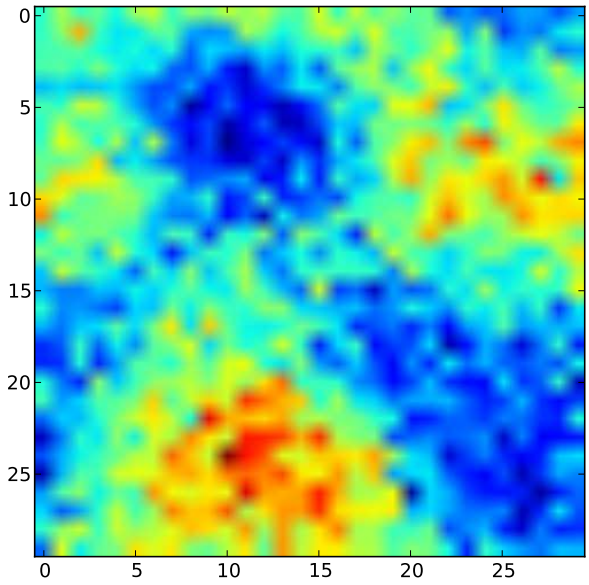

Figure 3.20: **Static pattern** - Firing rate pattern over the 30×30 obtained at the frontier between the excitation dominated regime and the inhibition dominated regime ($g_E = 0.5$,$g_I = 1.5$ in normalized units).

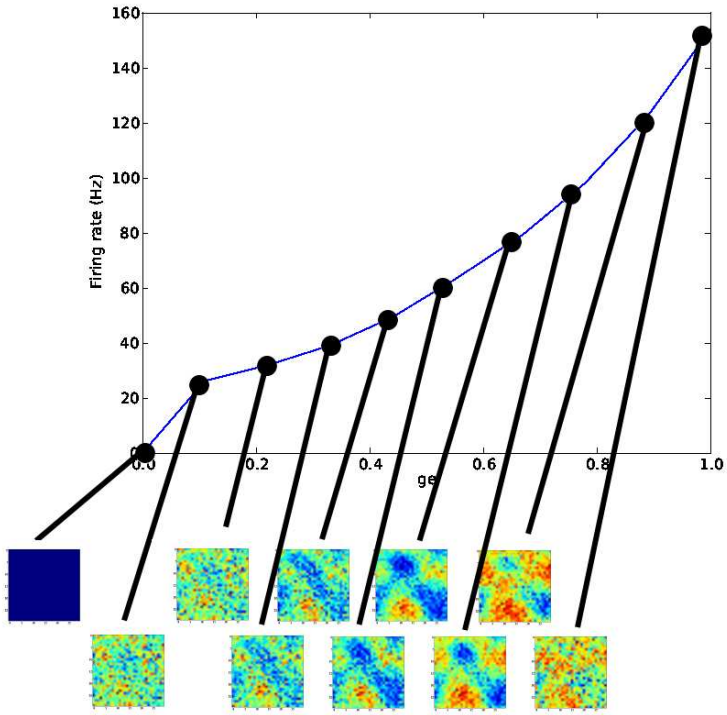

Figure 3.21: **Emergence of a static pattern** - The curve shows the average firing over the net work as a function of g_E with $g_I = 1.5$ in normalized unit and boxes show firing rate maps.

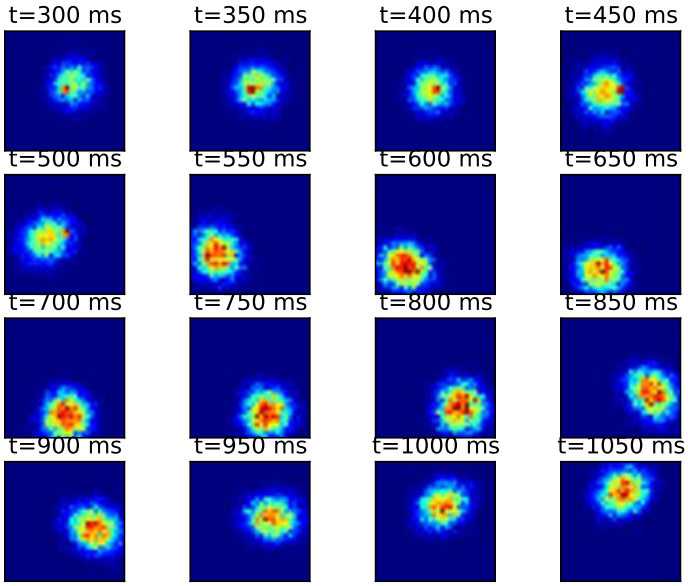

Figure 3.22: **Firing rate map after stimulation of the 4 central columns** - When the stimulation is released (500ms), the bump of activity starts moving across the map.

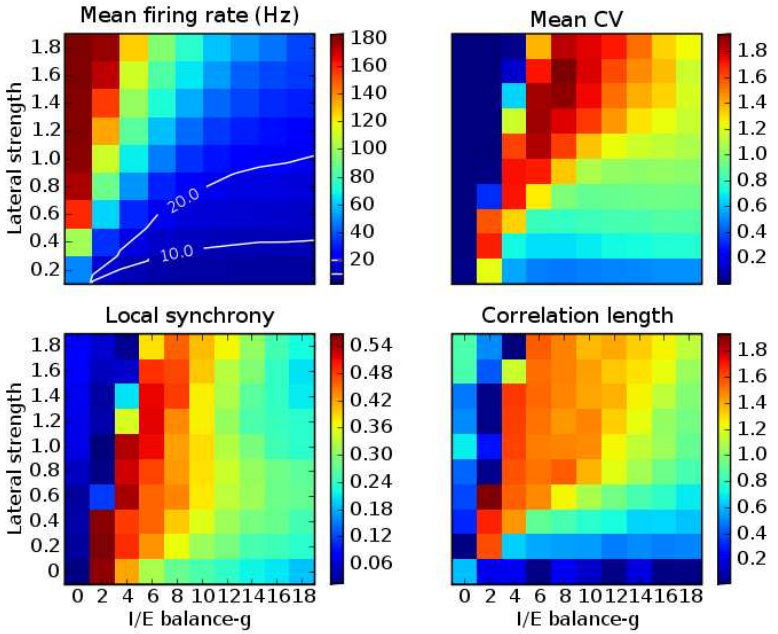

Figure 3.23: **Phase diagram** $(\frac{g_I}{g_E}, ls)$ **for the network of spiking neurons after LISSOM learning** - (Top-left) Average firing rate. (Top-right) Average coefficient of variation of the interspikes intervals. (Bottom-left) Local synchrony c_0. (Bottom-right) Correlation length λ.

Correlations. We calculate the correlation between spike trains as a function of the distance:

$$Cor(d) = \frac{1}{N} \sum_{i,j} Cor(s_i, s_j)$$

where i and j are N random pairs of neurons such that the distance between the columns from which it is extracted is d. The simplest approximation to this function is to assume an exponential decay of correlations as a function of distance:

$$Cor(d) = c_0 e^{-\frac{d}{\lambda}} + c_\infty$$

where c_0 is the local synchrony, that is the correlation averaged over spike trains of the same column. The correlation length λ gives the typical size of patterns emerging in the network. The basal synchrony c_∞ is the correlation between spike trains of columns located very far away from each other. On fig 3.23, it can be checked that the local synchrony and the correlation length have higher values in the busty regime, the correlation length is constant in all that area of the phase diagram.

Description of the activity. In fig 3.24, a phase diagram is shown for parameters (g_{ext}, g_{int}) where g_{int} is the balance $\frac{g_I}{g_E}$ for connections between neurons belonging to the same column and g_{ext} is the balance for connections between neurons belonging to different columns. The sum of conductances $g_E + g_I$ is kept constant ($0.024nS$). The 3 regimes previously described (S, A and B) can also be found in this diagram and we show the activity for a line of columns in regime A and B. In the B regime, activated up states are propagating across the network with slow velocity. The dynamics of a column is represented in fig 3.25 for an example of regime A and an example of regime B. In regime A, spikes seems to occur in an independent fashion and the distribution of the membrane potential is monomodal close to that observed in the main cluster of Chapter 1. In regime B, neurons fire spikes synchronously and the distribution of the average membrane potential is bimodal. The bimodality of cells classified could then result from the network structure when g_E and g_I are tuned in such a regime.

Correlations depend on orientation preference. In the network before learning, the network was invariant under translation and rotation so the correlations between the spike trains of two neurons depended only on the distance separating these two neurons. After learning, the symmetry is broken and the orientation map also have an effect on the correlation structure of the spiking activity in the network. As can be seen on fig 3.26, there is a decay in correlations as a function of distance but also as a function of the difference in preferred orientation. By assuming exponential decay on the distance d and the difference of preferred orientation δ_{phi}:

$$Cor(d, \delta\phi) = c_0 e^{-\frac{d}{\lambda_d} - \frac{\delta\phi}{\lambda_\phi}},$$

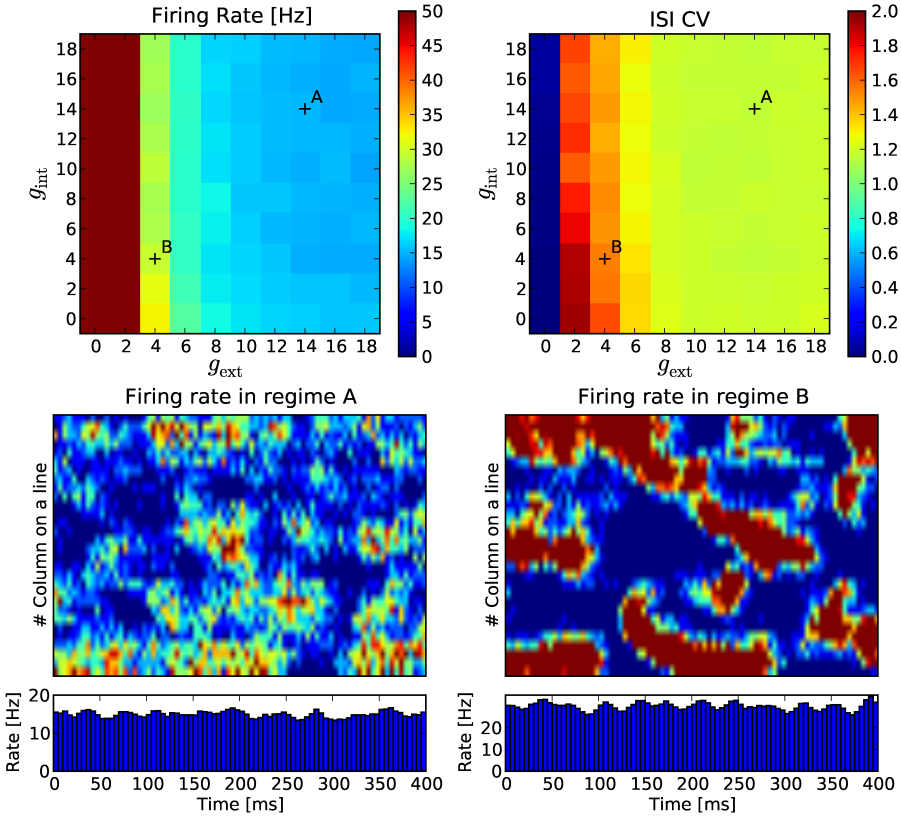

Figure 3.24: **Phase diagram (g_{ext}, g_{int}) and average firing rate on a line of columns** - (Top-left) Average firing rate as a function of (g_{ext}, g_{int}) (rescaled from $g_0 = 0.006$). (Top-right) Average coefficient of variation of the interspikes intervals as a function of (g_{ext}, g_{int}) (Bottom-left) Average firing rate in columns of the central horizontal line of the network over time in the asynchronous regime and average of the firing rate in the network over time. (Bottom-right) Same for regime B.

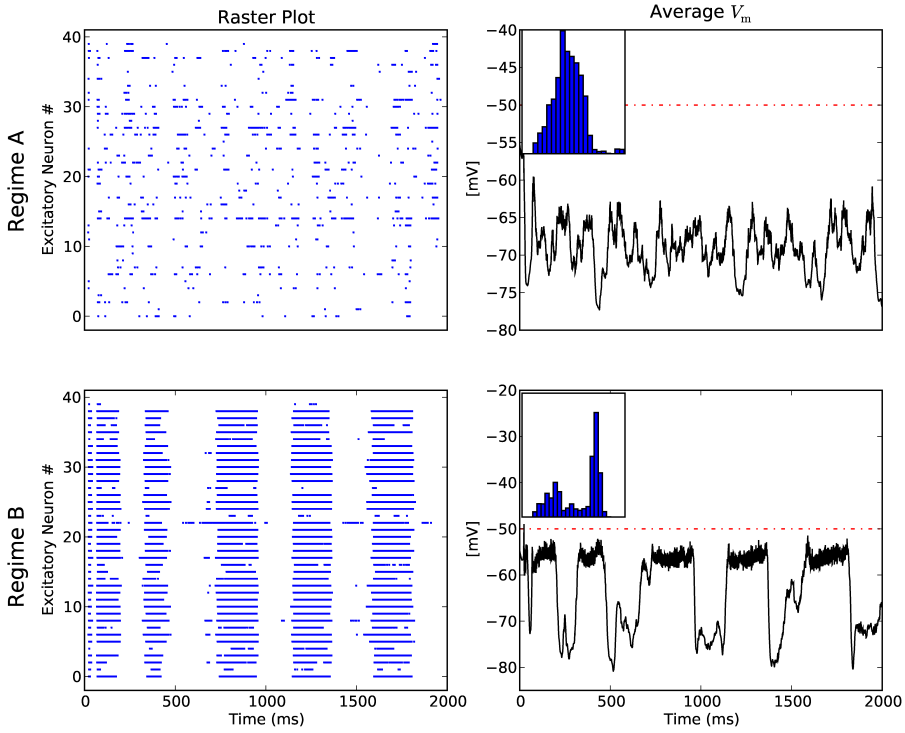

Figure 3.25: **Activity for the asynchronous regime (A) and the bursty regime (B)** - (Left) Spikes raster (each dot represent a spike) for all the neurons in a column of the network. (Right) Temporal dynamics of the average membrane potential and the corresponding distribution (inset).

there are two typical constants characterising the correlation structure of the network, λ_d the spatia correlation length and λ_ϕ the orientation correlation length. The retinotopy and the orientation preference which are reflected in the anatomy are also reflected in the dynamics of the network as was also found in recent LFP recordings [161].

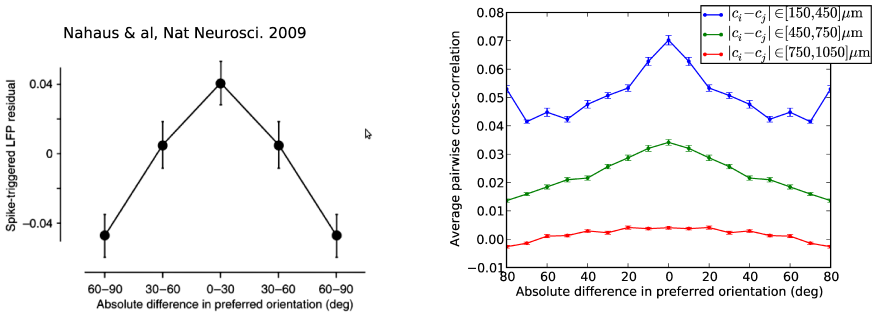

Figure 3.26: **Correlations depending on the difference of orientation preference of the columns.** - (Left) Results from biological experiments [161] showing the cross-correlations in local field potentials depending on the difference of orientation preference. (Right) Average spike trains correlation depending on the difference of preferred orientation $\delta\phi$ for columns separated by a distance d with 3 distances considered. The left part of the curve is a copy of the right part to make the picture similar to the data from [161].

3.5.5 On-going activity in a model of V_1.

The on-going activity of the V_1 model was analyzed before and after LISSOM learning. In both cases, three regimes are found:

- When excitation dominates, the network is saturated at its maximum level of activity.
- When inhibition dominates, neurons of the network fire spikes in a rather independent fashion with low firing rate, similar to the self-sustained asynchronous irregular regime (Kumar 09).
- For balanced situations, the activity in the network is not homogeneous anymore and static or dynamic instabilities occurs.

Before learning, the connectivity is isotropic and as the excitatory conductance is increased, static patterns emerge. By reducing the model to its firing

rate description, we saw that a Turing mechanism may explain the rise of these patterns. After learning, connectivity becomes anisotropic making the mathematical analysis difficult. Analytical results for the neural field model with mexican-hat connectivity and additional weakly anisotropic excitatory connectivity kernel is available in (Golubitsky) but it is very different from our situation where symmetry is broken mainly in the inhibitory kernel. So although a rigorous proof is not possible for now we suggest that the propagating up states in our model could be explained by a Turing-Hopf mechanism in the corresponding neural field model so that the typical length of patterns we observe should also be related to the characteristic length of the connectivity kernel. By this model, we were able to generate up and down states from network mechanisms which is different from other mechanisms relying on intrinsic properties of the cell [162].

3.6 Conclusion.

As described in this chapter, several approaches can be taken to explain the formation of V1 from abstract cognitive architecture to self organizing models implementing a plasticity rule, each having specific application. The ice cubes model provides a framework for natural computation, physical models have been shown to predict the structural properties of feature maps like the periodicity of ocular dominance domains or the pinwheel density of orientation maps and models including plasticity are useful to analyze the connections and dynamics of the primary visual cortex.

With column based network of spiking neurons, we found that in the excitation dominated regime, neurons fire at high rate and in a regular fashion whereas in the inhibition dominated regime, the network sets in an asynchronous irregular state similar to that described for a single column in chapter 2. In models including long range connections, we found another state for balanced excitation and inhibition where patterns of activity emerge, either static for isotropic connections or dynamic for anisotropic connections resulting from learning. Moreover, the structure of correlations in the network after learning reflects the orientation map which is related to visual experience similarly to [25]. Those patterns occurs through an neuronal analog of the Turing instability for the isotropic network. In the network after learning, the up and down states result from a combination of adaptation and inhibition and it would be interesting to study how the duration of up and down states depends on the adaptation variable. In anatomical studies, long range connections are found to be rather excitatory [163] but a simple inversion of initial connections kernels for excitation and inhibition in the rate model doesn't give rise to orientation maps because only two orientations are selected. Two layers are considered in V1 in [164] with short range excitation and inhibition of similar extent in the first layer and long range excitation in the second layer resulting in an orientation map with long range excitation. Long range excitation could also be mediated through long range inhibition targeting only inhibitory cells.

In the ice cubes model, we saw that the ongoing activity can cancel the response of the network to visual stimulation. We also found that the orientation tuning curve was sharper near the pinwheel center than far from this center consistent with the experimental observation that cells stay tuned to orientation near the pinwheel center [165]. Such properties could also be tested in the network with orientation maps resulting from learning and we suggest that the structured correlations encode priors about the statistics of the visual world and some problems in vision like the inference of 3D structure from a 2D image may be solved using such priors. It has already been shown that the level of ongoing activity before the stimulus presentation is a good predictor of the perception of an ambiguous figure [166].

Conclusion

The ongoing activity of the brain characterizes the state of consciousness. in the awake state it is irregular and asynchronous whereas in slow wave sleep there are 1 Hz collective oscillations seen at the single unit level as transitions between up and down states. Under annesthesia, we found similar dynamics to those known for the awake state and the slow wave sleep but also additional classes like cells having a membrane potential with very few fluctuations resulting from synchronous inputs or cells having micro up or micro down states in their membrane potential. The ongoing activity have a wider range than the visually evoked activity and after visual stimulation, the dynamics is close to the main cluster corresponding to the awake state. Another study from Nicolas Benech at the UNIC demonstrated an influence of the ongoing activity on the response properties, with bistable cells having longer latency. The resting state activity is described in human fMRI studies as networks having infraslow ($< 0.1Hz$) correlated fluctuations and recordings of resting state activity are usually long, so we provided a method to represent the infraslow fluctuations of the EEG in an efficient manner using wavelets. At each time, the signal is compressed into a symbol representing its local frequency content and defining a microstate. The classes of neurodynamics we found are the cellular correlates of those microstates and their definition are useful in the monitoring of anesthesia and in the understanding of patterns of ongoing activity recorded at a whole brain level [167]. We presented a collection of parameters which are used to characterize the firing, the distribution and the power spectrum of a cell and we found that those related to the bimodality of the distribution and to the power in high frequencies are useful to separate the clusters. We showed that measures relying on information theory offer a promising approach to multi channels recordings. Several approaches may be taken to classify data and we presented K means clustering, agglomerative tree building and self organized maps. We showed how the partitions resulting from these algorithms can be combined and compared.

Functional neurodynamics, like asynchronous irregular activity of the awake state, fixed point attractors associated to memory storage or limit cycles related to the binding of a coherent percept, are implemented in neuronal networks and can be studied using dynamical systems. To avoid heavy computations and to reduce the parameter space, many properties of neuronal dynamics are studied through phenomenological models which share universal properties, like

the types of bifurcation when a parameter is varied, with the detailed model. Normal form reduction thus gives the simplest polynomial system topologically equivalent to the original system and mean field equations give the evolution of average quantities in the limit of infinite size networks. We reminded how the Hodgkin Huxley model can be reduced to the 2 dimensional FitzHugh Nagumo system and we analyzed local bifurcations of codimension 1 and 2 for this system. At a codimension 3 bifurcation point, a stochastic perturbation via a Brownian motion resulted in complex dynamics mixing several timescales and we suggest that these timescales are related to limit cycles and fixed point attractors lying close to the equilibrium. Interestingly, the first harmonic in the power spectrum had a step-like evolution when the noise variance was increased with the same seed used for the generation of random numbers. A cortical column is modeled as a population of excitatory neurons and a population of inhibitory neurons and their macroscopic activity is described by Wilson-Cowan equations showing multistability and limit cycles. When Hopf bifurcation leads to periodic activity, the dynamics of coupled columns can be reduced to an equivalent phase oscillator and we showed various types of dynamics occuring in a network of phase oscillators from the transition to synchrony as the coupling is increasing to chimera states, in a model with long range connections, where a part of the network is synchronous while the other part is asychronous. In a sparsely connected network, the asynchronous irregular state results from chaotic dynamics when excitation and inhibition are balanced and we provided a Fokker Planck description of the membrane potential in a network of spiking neurons. The effective time scale is very small in the balanced network enabling fast tracking of time varying inputs and can then be used to model attentional effects. Finally, in the article in appendix, we found windows of chaotic behavior in coupled flip-flop oscillators and the ongoing dynamics in a network where cell assemblies are embedded show itinerancy among the fixed points corresponding to stored memories. There are thus wide applications of dynamical systems in neuroscience and complex dynamics modeled in neuronal networks may be used to solve computational tasks [119].

Cortical columns can be coupled together with hard wired or plastic connections to achieve visual function in a model of the primary visual cortex. Several models have been presented to explain how selectivity arises and how the features of the input space are mapped onto the surface of the cortex. In the ice cubes model, standing for a pinwheel in V1, a column have lateral connections only with its nearest neighbours and patterned inputs from the LGN result in simple cell receptive field. In this model, the ongoing dynamics depends on whether excitation or inhibition dominates. When excitation dominates, there is a synchronous regular state with neurons firing at a high rate and no specific response to an oriented bar because the response is lost in the ongoing dynamics. When inhibition dominates, we found an asynchronous irregular state similar to that described for a single column and neurons in a column have a tuning curve reflecting their prefered orientation with this curve being sharper for cells near the pinwheel singularity. Learning was also modeled in a network with long range connections and to fasten simulations, we used a coarse description of

the network with a macroscopic unit representing a column of spiking neurons and weights between two units giving the probability of connection between neurons of the 2 correspondingcolumns. In these models including long range connection, we found structured activity for balanced excitation and inhibition. Before learning, when lateral connection kernels are isotropic, these patterns are static, with some parts of the map firing at a high rate and other parts firing at low rate, and they are related to the eigenforms obtained with a mean field description of the network and considered as a model of hallucinatory perception. After learning using an Hebbian rule on both lateral and feedforward connections to V1 units, the lateral connection kernels become anisotropic, each unit gets selective to orientation and an orientation map emerges. The patterns of the balanced state in the network after learning are dynamic with neurons in each column having collective transitions between up and down states. These transitions are correlated in the network with spiking correlations between neurons of two columns decreasing exponentially as a function of the distance but also as a function of the difference between prefered orientations so that the ongoing activity reflects the visual experience during the learning. Those structured correlation may encode prior knowledge about the statistics of the visual world and used in solving problems related with visual perception. Each perception is thus an inference based on visual stimulation and priors encapsulated in ongoing dynamics.

We thus found a way to characterize several classes of neurodynamics and we found a way to switch from the asynchronous state to structured up and down states by changing the strength of excitation and inhibition in a model of the primary visual cortex. These states are usually associated with specific states of consciousness like slow wave sleep and waking but ongoing activity may also reflect attentional processes and we suggested that ongoing pattern of activity may be useful for cortical computations as reflecting internal knowledge about the world. It would be interesting to provide visual stimulation in the network after learning and to test the network in binocular rivalry, which is a typical exemple where ongoing activity interplay with visual stimulation to give rise to perception [168]. Additional plasticity mechanisms like spike timing dependent plasticity or intrinsic plasticity. The model of the visual cortex can also be used predict the effects of magnetic stimulation in brain computer interface or of a medical drug. The ice cubes model could also be extended to include several pinwheels having long range connections depending on the prefered orientations of connected columns and it would then provide a cognitive architecture for biologically inspired computing. In the column based network we assumed a decoupling of the timescales of the dynamics of neuronal activity and of the synapses but further research is needed to understand interactions between those timescales.