



**HAL**  
open science

# Retinal processing: insights from mathematical modelling

Bruno Cessac

► **To cite this version:**

Bruno Cessac. Retinal processing: insights from mathematical modelling. J. Imaging, In press. hal-03454859v1

**HAL Id: hal-03454859**

**<https://inria.hal.science/hal-03454859v1>**

Submitted on 29 Nov 2021 (v1), last revised 17 Jan 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Retinal processing: some mathematical insights

Bruno Cessac <sup>1</sup>  0000-0003-1523-4187

<sup>1</sup> Université Côte d'Azur, INRIA, France, Biovision team and Neuromod Institute; bruno.cessac@inria.fr

\* Correspondence: bruno.cessac@inria.fr.

**Abstract:** The retina is the entrance of the visual system. Although based on common biophysical principles the dynamics of retinal neurons is quite different from their cortical counterparts, raising interesting problems for modellers. In this paper I address some mathematically stated questions in this spirit, discussing, in particular: (1) How could lateral (amacrine cells) connectivity shape the spatio-temporal spike response of retinal ganglion cells? (2) How could spatio-temporal stimuli correlations and retinal network dynamics shape the spike train correlations at the output of the retina. I also briefly discuss some potential consequences of these results at the cortical level.

**Keywords:** Retinal network; visual system; spatio-temporal spike correlations; linear response; non stationarity

## 0. Introduction

Let us start with a very simple experiment. Look around you... That's it, the experiment is over. A very ordinary experience, isn't it? Is it really though? Let us first point out that looking around you to see, that is, having the sense of sight, is indeed ordinary — except for those who have partially or totally lost their ability to see. We will come back to this point at the end of the paper. Now, excluding visual impairments, vision is everything but ordinary.

Think of it. A flux of photons, with frequencies in the visible spectrum range, emitted by the external world around us enters into our eyes, then "something" happens, and we see. Thanks to constant progress in experimental and theoretical neuroscience, we understand better and better this "something", the mechanisms of vision, although our view of it is far from being complete. Especially, in these times of artificial intelligence, bio-inspired computing, computer vision and ... global warming, it might be helpful to understand how our brain is able to handle the complex visual information coming from the external world so rapidly and efficiently ... with an energy consumption of the order of a few Watts.

Certainly, the retina plays a central role in this process. It is known for long that this is definitely not a camera. The retina is smart [1] and it has to be. Think especially of the difference of scale between the retina and the visual cortex, in terms of size but also numbers of neurons and synapses. As everything that goes to the visual cortex comes from the retina, this little membrane, at the back of the eye, half a millimetre thick, with an area of order a  $\text{cm}^2$  (for humans), has somewhat to filter the visual information, leaving out "irrelevant detail" and stick at crucial events, and then, signal them in the proper way to the brain via spike trains. As a matter of fact, the question(s) of "efficiently" encoding information by spikes has been the subject of many fascinating papers [2–4], especially in the seminal paper from Barlow [5] with concepts such as reducing redundancy, information compression and efficient coding. These concepts are regularly updated with recent experimental and theoretical investigations [6–15]. We come back to this point, at the end of the paper too.

The retina has, roughly, the following structure (Fig. 1, left). For more detail see e.g. [16] or <https://webvision.med.utah.edu/book/part-i-foundations/simple-anatomy->

**Citation:** Cessac B. Retinal processing: mathematical insights. *J. Imaging* 2021, 1, 0. <https://doi.org/>

Received:

Accepted:

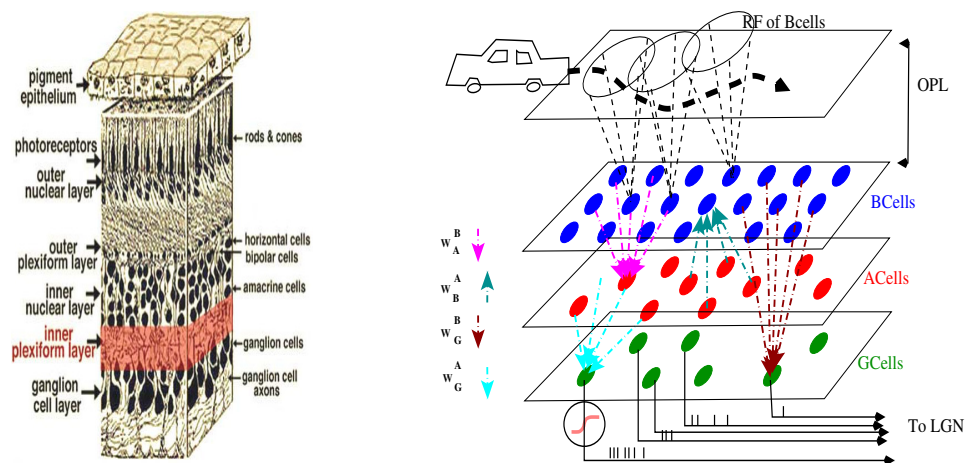
Published:

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2021 by the authors. Submitted to *J. Imaging* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

of-the-retina/)). It is organized in five neuronal types : Photo-receptors, rod and cones (P), horizontal cells (HCells), bipolar cells (BCells), amacrine cells (ACells), ganglion cells (GCells), to which are added glial cells (Mueller's cells). These neurons types are connected by chemical and electric synapses, in specific functional circuits or "pathways" (like the rod-cone pathway [17,18]) which are a key in the retinal capacity to convert the light coming from a visual scene into spike patterns sent to the visual cortex, through the Lateral Geniculate Nucleus (LGN), via the optic nerve made of GCells axons. In particular, there are in the retina very specific synapses like the ribbon synapse enabling neurons to transmit light signals from photoreceptors to BCells over a dynamic range of several orders of magnitude in intensity [19]. Roughly, two main connectivity structure can be distinguished: feed-forward, the P-B-G path which leads from the photo transduction to the spike trains emitted by the GCells towards the cortex. There is also a lateral connectivity due to Hcells, at the origin of the Center-Surround structure of the receptive fields, and the Acells whose role is still poorly understood and which are one of the main object of study of this paper.

57



**Figure 1. Left. Structure of the biological retina.** From <http://webvision.org/es/gross-anatomy-of-the-eye/1-2-simple-anatomy-of-the-retina/> **Right. Structure of the model introduced in section 1.1.** A moving object (here, presumably, a car) moves along a trajectory (dashed black line). Its image is projected by the eye optics to the upper retina layers (Photoreceptors and HCells) and stimulates them. In the model, this corresponds to the convolution of the stimulus with the Receptive Field (RF) of BCells. This provides to BCells what we call the "OPL" input to BCells. BCells (blue points) are connected to ACells (red points) via excitatory synapses (pink arrows, denoted  $W_A^B$ ) and to GCells (green points) via excitatory synapses (brown arrows, denoted  $W_G^B$ ). ACells are connected to BCells via inhibitory synapses (green arrows, denoted  $W_B^A$ ) and to GCells via inhibitory synapses (cyan arrows, denoted  $W_G^A$ ). The voltage of GCells is sent through a non linearity (pink curve in the black circle) so as to produce spike trains conveyed to the LGN.

58 The structure of the retina and its behaviour are thus well studied on the experi-  
 59 mental side. There are comparatively less modelling studies and quite less mathematical  
 60 results on the way how retinal structure, especially, lateral ACells connectivity shapes  
 61 the spike response to spatio-temporal stimuli. One of the goals of this paper is to elicit  
 62 reflections in this direction, grounded on recent mathematical developments fed by the  
 63 recent progresses in the knowledge of retina physiology and structure. This is a humble  
 64 and partial point of view, issued from my collaboration with neuro-biologists experts  
 65 in the retina. More precisely, I am addressing the following problems on mathematical  
 66 grounds. In the main text I focus on neuroscience modeling perspective, whereas, in the  
 67 conclusion section, I discuss potential consequences of these results out of the field of

68 neuroscience.

69

70 **Problem 1. How does the structure of the retina, in particular, amacrine lateral**  
 71 **connectivity condition the retinal response to dynamic stimuli?**

72 The problem can be addressed at several levels.

73

**Level 1. Single cell response to stimuli.** The individual response of ganglion cells is usually expressed in terms of their receptive field. This notion is on the one hand phenomenological: it is observed that each ganglion cell responds preferentially to stimuli, localized in space, with a characteristic spatio-temporal structure. For example, a ON-Center cell preferentially responds to an increase in luminance in a circular area corresponding to the central part of the receiving field. This notion is also expressed mathematically by a kernel  $\mathcal{K}_G$ , i.e. a function of space and time, so that the response of a GCell to a spatio-temporal stimulus  $S(x, y, t)$ , takes the form:

$$\left[ \mathcal{K}_G \begin{smallmatrix} x,y,t \\ * \end{smallmatrix} S \right] (t) = \int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} \int_{s=-\infty}^t \mathcal{K}_G(x - x_C, y - y_C, t - s) S(x, y, s) dx dy ds, \quad (1)$$

74 where  $\begin{smallmatrix} x,y,t \\ * \end{smallmatrix}$  means space  $(x, y)$ -time  $(t)$  convolution.  $x_C, y_C$  are the coordinates of the RF  
 75 center. The integrals are well defined since the kernel decreases fast enough to infinity,  
 76 in space and time, to guarantee convergence. The upper bound in time,  $t$ , expresses  
 77 causality, whereas the lower bound,  $-\infty$ , implicitly assumes that the stimulus has  
 78 been applied in a distant past compared to  $t$ , quite longer than the characteristic times  
 79 involved in GCell response.

80 Equation (1) corresponds to a linear response. It is therefore only valid for stimuli  
 81 of low amplitude in voltage. More generally, the voltage response to the stimulus is  
 82 a functional of the stimulus that one can, under well posed mathematical conditions,  
 83 write as a Volterra expansion [20], (1) being the lowest order (linear) term. Unfortunately,  
 84 higher-order terms are essentially inaccessible experimentally and one usually constrains  
 85 instead the non-linearity of the response under other modalities. Especially, taking into  
 86 account that the response of a ganglion cell to a stimulus is, ultimately, a sequence of  
 87 spikes, one writes the probability density of emitting a spike between  $t$  and  $t + dt$  in  
 88 the form  $f\left(\left[\mathcal{K}_G \begin{smallmatrix} x,y,t \\ * \end{smallmatrix} S\right](t) + b\right)$  where  $f$  is a nonlinear positive increasing function  
 89 (typically, a sigmoid),  $b$  is a threshold constraining the level of activity of the GCell in  
 90 the absence of stimulation. This procedure defines an inhomogeneous Poisson process  
 91 called the linear-non-linear Poisson (LNP) model [21–23]. Experimentally the kernel  
 92  $\mathcal{K}_G$  is determined by Spike-Triggered Average or Spike-Triggered Correlation technique,  
 93 studying the response at a white noise [21]. Nonlinearity is then determined, typically  
 94 by the Levenberg-Marquart method [24]. This modelling asks however the following  
 95 questions:

- 96 (i) How is the kernel  $\mathcal{K}_G$  of the GCells constrained by the structure/dynamics of  
 97 the upper layers of retinal cells ?
- 98 (ii) The forms (1) implicitly assumes that  $\mathcal{K}_G$  does not depend on the stimulus. Can  
 99 one write mathematical conditions that guarantee such an independence?
- 100 (iii) To which extent is the notion of Ganglion Cells Receptive Field compatible with  
 101 non linear effects reported in retinal neurons and synapses, such as voltage  
 102 rectification or gain control ?

103 **Level 2. Collective response to stimuli and spike statistics.** GCells do not inter-  
 104 act directly, but amacrine connectivity induces an effective, causal interaction between  
 105 them. What is therefore the structure of the spatio-temporal correlations induced by the  
 106 conjunction of the spatio-temporal stimulus and the response of the retinal network, in  
 107 particular, the amacrine lateral connectivity ? A classical paradigm in neural coding is  
 108 to assume that the retina decorrelates GCells outputs to maximize information transfer

109 [6–11,13–15]. It is in particular believed that ACells play a central role in this decorrela-  
110 tion process (see [15] and references therein). What can be, at the mathematical level, the  
111 conditions, on the stimulus and dynamics, that allow a network of neurons interacting  
112 with each other to produce vanishing, or at least, *weak* correlations ? When does weak  
113 mean negligible ? These questions are actually closely related to the second problem.

114

### 115 **Problem 2. Spike statistics.**

116 More generally, considering the retina as a dynamical system forced by non-  
117 stationary, spatially inhomogeneous stimuli, what could be a general form for the  
118 (non-stationary) statistic of spike trains emitted by ganglion cells, taking into account  
119 that spike trains emitted by the retina are all what the LGN and cortex see ? One can  
120 attempt to construct a canonical form of probability distributions of the retinal spike  
121 trains taking into account that:

- 122 (i) Stimuli, thus statistics, are not stationary;
- 123 (ii) The cortex (and before, the LGN) only receive spikes, thus have no information  
124 about the biophysical processes which have generated those spikes and no  
125 information on the underlying dynamics of the retina (voltages, activation  
126 variables, conductances). All the information is contained in the spatio-temporal  
127 structure of spikes;
- 128 (iii) Spike trains distribution may have a long memory.

129 In this paper I address these problems with the help of two models. The first,  
130 presented in section 1.1, grounded on biology and e.g. the papers [25–28] mimics the  
131 Bipolar-Amacrine-Ganglion cells network and is used, in sections 2.1, 2.2, to make pro-  
132 gresses in elucidating problem 1. I first show how one can obtain an explicit form for  
133 the kernel (1) featuring the ACells lateral connectivity. This RF explicitly depends on  
134 the BCs-ACs network through the eigenvalues and eigenvectors of a operator I call  
135 "transport operator". I discuss some consequences of this result, especially in terms of  
136 response to propagating stimulus. This result is valid when cells act as linear integrators.  
137 However, cells are in general rectified by non linearities. I propose piecewise-linear  
138 rectifications (as used in several retina model) and I discuss how rectification acts on  
139 the RF of eq. (1). A striking conclusion is that, if the convolution form (1) is preserved,  
140 this is to the price of having a RF depending on the stimulus. A consequence of this  
141 analysis is that spike correlations may depend on the stimulus and are expected to be  
142 quite different when considering e.g. objects moving along trajectories in comparison to  
143 static images.

144

145 The second model, introduced in section 1.2 and analysed in section 2.3 attempts to  
146 propose a canonical form of probability distributions of the retinal spike trains based  
147 on the constraints (i), (ii), (iii) above. These sections essentially presents the conclusions  
148 of works published elsewhere [29–39]. As I argue, these constraints leads to a natural  
149 notion of spike probabilities, somewhat extending the statistical physics notion of Gibbs  
150 distribution to the non stationary case. In this setting, one establishes a linear response for  
151 a network of interacting spiking cells, that can mimic a set of GCells coupled via effective  
152 interactions corresponding to the ACells network influence. This linear response theory  
153 not only gives the effect of a non stationary stimulus to first order spike statistics (firing  
154 rates) but also its effect on higher order correlations. Indeed, spike correlations are  
155 modified by a spatio-temporal stimulus and can be computed thanks to the knowledge  
156 of spontaneous correlations. The linear response formula is expressed as a convolution  
157 where the kernel can be explicitly computed for an Integrate and Fire conductance  
158 based model. Moreover, as I argue, these spike trains distributions have close links with  
159 information geometry. Especially, they induce a natural metric in an abstract space of  
160 probabilities, with close potential links with the neuro-geometry introduced by Sarti,  
161 Citti, Petitot et al [40–43]. This is discussed in the conclusion section.

162 More generally, the application and discussion sections shortly proposes possible  
 163 extension of this work to several domains: Retinal prostheses, section 3.1; Convolutional  
 164 networks, section 3.2; Implications for cortical models, section 4.1; Neuro-geometry,  
 165 section 4.2.

## 166 1. Materials and Methods

### 167 1.1. Modelling the retinal network

#### 168 1.1.1. Specificities of the retina

169 Neurons in the retina have the same biophysics as their cortical counterparts.  
 170 However, they operate under different modalities. Remarkably, with the exception of the  
 171 GCells, the retinal neurons do not emit action potentials. Their activity and interactions  
 172 therefore take place through graded (continuous) membrane potentials as opposed to the  
 173 sharp peak of an action potential. Furthermore, there is no long-term synaptic plasticity  
 174 in the retina. Finally, the main "computational" elements in the retina are functional  
 175 circuits [18] made of a few neurons and synapses, in large contrast with "computational"  
 176 units in the visual cortex, such as cortical columns, involving thousands of neurons.  
 177 A modelling consequence is that mean-field or neural masses description used in the  
 178 cortex might not be relevant to study the retina.

179 The goal of this paper is to address mathematical questions about the dynamics and  
 180 behaviour of the retina embedded in the visual system. To instantiate these questions on  
 181 a firm mathematical ground we are going to consider a model of the retinal network,  
 182 based on a few fundamental facts briefly exposed in the previous section:

- 183 1. The retina is a high dimensional, non autonomous and noisy dynamical system,  
 184 layered and structured, with non stationary and spatially inhomogeneous entries  
 185 (visual scenes).
- 186 2. Most retinal neurons are not spiking, except GCells. Thus, retina performs analogic  
 187 computing.
- 188 3. Local retinal circuits efficiently process the local visual information. These local  
 189 circuits are connected together, spanning the whole retina in a regular tiling. In  
 190 this perspective, it is important to consider individual neurons and synapses, in  
 191 contrast, e.g., to cortical modelling where it is relevant to consider mean-field  
 192 approaches averaging over populations.

193 Thus, the model presented below and in Fig. 1, right, is non stationary, with a  
 194 layered retina's like structure, where dynamics ruling BCells, ACells, and GCells voltage  
 195 is piecewise linear. As we discuss, the model affords additional non linearities like gain  
 196 control although we will not delve into the maths of the gain controlled case (see [28] for  
 197 a study of this case). For GCells, the spiking process is mimic by a non linear firing rate  
 198 so that our model enters in the class of LNP models.

#### 199 1.1.2. Structure of the retina model

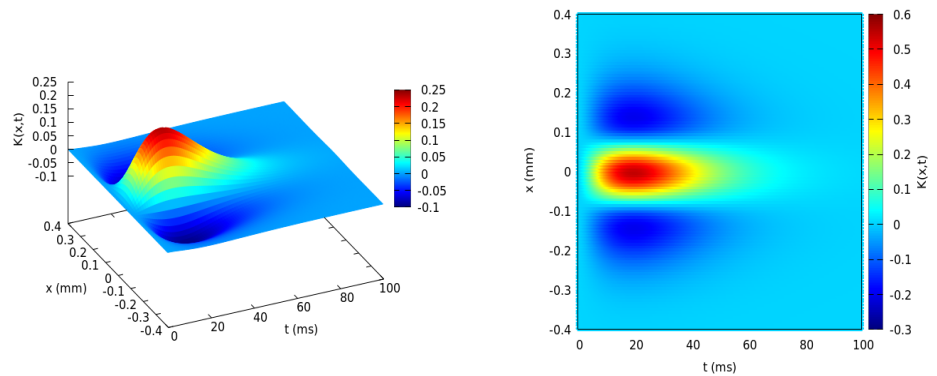
200 We assimilate the retina to a superimposition of 3 layers, each one being a flat, two  
 201 dimensional square of edge length  $L$  mm where spatial coordinates are noted  $x, y$  (Fig.  
 202 1, right). Each layer corresponds to a cell population (BCells, ACells, GCells) where  
 203 the density of cells is taken uniform. We note  $\delta_p$  the lattice spacing in mm, and  $N_p$  the  
 204 total number of cells in the layer  $p$ . Without loss of generality we assume that  $L$ , the  
 205 retina's edge size, is a multiple of  $\delta_p$ . We note  $L_p = \frac{L}{\delta_p}$ , the number of cells  $p$  per row or  
 206 column so that  $N_p = L_p^2$ . Each cell in the population  $p$  has thus Cartesian coordinates  
 207  $(x, y) = (i_x \delta_p, i_y \delta_p)$ ,  $(i_x, i_y) \in \{1, \dots, L_p\}^2$ . To avoid multiples indices, we associate to  
 208 each pair  $(i_x, i_y)$  a unique index  $i = i_x + (i_y - 1) L_p$ . The cell of population  $p$ , located at  
 209 coordinates  $(i_x \delta_p, i_y \delta_p)$  is then denoted by  $p_i$ .

210 One can roughly subdivide the real retina into two blocks (Fig. 1). The first, that we  
 211 name in short, for modelling purposes<sup>1</sup> OPL, (Outer Plexiform Layer), includes the P,  
 212 HCells, BCells and the related synapses. As an "input" of this block is the flow of photons  
 213 emitted by the outside world and picked up by the photo-receptors. In our model, this  
 214 corresponds to a "stimulus", i.e. a function  $\mathcal{S}(x, y, t)$  where  $t$  is the time coordinate. As  
 215 we don't consider color sensitivity here  $\mathcal{S}$  characterizes a black and white scene, with  
 216 a control on the level of contrast  $\in [0, 1]$ . The "output" of the OPL is sent to Bcells in  
 217 the form of a "drive" voltage, defined in eq. (2) below. In the real retina, the voltage  
 218 of each BCell integrates, spatially and temporally, the local visual information of the  
 219 photo-receptors which are connected to it, with a lateral modulation due to the HCells.  
 220 Each BCells is thus sensitive to specific local characteristics of the visual scene, defining  
 221 its Receptive Field (RF). Thus, BCells, like GCells, have a receptive field. But, as they are  
 222 earlier in the vertical pathway they integrate less features.

We label BCells (layer 1) with the index  $i = 1, \dots, N_B$  and we model the RF of BCells  
 by a convolution kernel,  $\mathcal{K}_{B_i}$ , such that the voltage of BCell  $i$  is stimulus-driven by the  
 term:

$$V_{i_{drive}}(t) = \left[ \mathcal{K}_{B_i} \overset{x,y,t}{*} \mathcal{S} \right](t). \quad (2)$$

223 The center of the RF, located at  $x_i, y_i$ , also corresponds to the coordinates of the BCell  $i$ .  
 224 A typical shape for the RF of BCells is illustrated in Fig. 2, although the explicit form  
 225 does not play a role in the subsequent developments.



**Figure 2. Receptive Field of a ON BCell.** **Left.** Example of a spatio-temporal RF of BCells (ON center cell) represented in 3D (one dimension of space,  $x$  and time  $t$ ). There is inhibition at the surround, physiologically due to HCells. **Right.** Spatio-temporal RF representation with a color map.

226 The second block, that we name in short IPL (Inner Plexiform Layer), comprises the  
 227 ACells and GCells and the afferent synapses. Its "input" is the output of the OPL, and its  
 228 output, the trains of action potentials emitted by the GCells. ACells are difficult to study  
 229 experimentally because hardly accessible from electrophysiology measurements. There  
 230 are also a large number of cell subtypes in the ACells class (around 40), of which only a  
 231 small number have duly identified functions. It is however recognized that they play  
 232 an essential role in the treatment of motion [17,44,45]. Here we address mathematically  
 233 the question of the RGCell receptive field form, resulting from the pooling of BCells, as  
 234 illustrated in Fig. 1, each with a specific RF as exemplified in Fig. 2, and modulated by  
 235 ACells lateral connectivity.

<sup>1</sup> Note that the terminology OPL and IPL refers actually to synaptic layers. "The outer plexiform layer has a wide external band composed of inner fibers of rods and cones and a narrower inner band consisting of synapses between photoreceptor cells and cells from the inner nuclear layer." "The inner plexiform layer consists of synaptic connections between the axons of bipolar cells and dendrites of ganglion cells" (ref <https://www.sciencedirect.com/topics/medicine-and-dentistry/>). In our model, these naming are short cuts to distinguish the network input (OPL) and the network processing (IPL).

### 236 1.1.3. BCells-ACells interactions

237 We label ACells (second layer) with the index  $j = 1 \dots N_A$ . We note  $W_{B_i}^{A_j}$  the  
 238 synaptic weight from ACell  $j$  to BCell  $i$  and  $W_{A_j}^{B_i}$  the synaptic weight from BCell  $i$  to  
 239 ACell  $j$ . We set  $W_{B_i}^{A_j} \leq 0$ , (ACells are in general inhibitory although some excitatory  
 240 ACells exist, not considered here) whereas  $W_{A_j}^{B_i} \geq 0$ . The synaptic weight matrices BCells  
 241 to ACells and ACells to BCells are noted  $W_{A_j}^B, W_B^A$ . They are not squared in general.  
 242 There also exist electric synapses (gap junctions) between BCells and ACells (e.g. in the  
 243 Rod Cone pathway [17]) but we will not consider them, for simplicity. Note however  
 244 that adding gap junctions would simply result in adding linear terms to equations (3),  
 245 (6), (9) (when considering passive gap junctions) and modify characteristic time scales,  
 246 without changing the global analysis.

The voltage of BCell  $i$ ,  $V_{B_i}$ , evolves according to:

$$\frac{dV_{B_i}}{dt} = -\frac{1}{\tau_{B_i}} V_{B_i} + \sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathcal{N}_A(V_{A_j}) + F_{B_i}(t). \quad (3)$$

Here,  $\tau_{B_i}$  is the characteristic time scale of BCell  $i$  response (in ms). The function:

$$\mathcal{N}_A(V) = \begin{cases} V - \theta_A, & \text{if } V_A > \theta_A; \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

247 is a linear rectifier ensuring that the synapse  $j \rightarrow i$  becomes silent when the voltage  
 248 of the pre-synaptic ACell  $j$ ,  $V_{A_j}$ , is lower than a threshold  $\theta_A$ . This corresponds to a  
 249 biophysical fact : a synapse cannot change its sign. For simplicity we consider  $\theta_A$  to  
 250 be the same for all ACells, although the present formalism can be extended, e.g., to  
 251 several families of ACells having different thresholds. Note that linear rectifiers of type  
 252 (4) rectify cell's voltage "from below". Rectification "from above" also exist, ensuring that  
 253 the cell's voltage does not increase without bounds. A typical mechanism is gain control,  
 254 where an additional variable, called the activity, increasing as voltage increases, triggers  
 255 a gain function non linearly dropping down the voltage when it exceeds an upper  
 256 threshold [25,27]. Under some mild assumptions gain control can also be implemented  
 257 as a piecewise linear function of the activity (see [28]).

Finally,  $F_{B_i}(t)$  is the OPL input term. To match classical retina models as developed  
 e.g. in [25,27] it reads:

$$F_{B_i}(t) = \frac{V_{i_{drive}}}{\tau_B} + \frac{dV_{i_{drive}}}{dt} = \left[ \mathcal{K}_{B_i}^{x,y,t} * \left( \frac{\mathcal{S}}{\tau_B} + \frac{d\mathcal{S}}{dt} \right) \right](t), \quad (5)$$

258 (where  $\mathcal{K}_{B_i}(x, y, 0) = 0$ ). In short,  $F_{B_i}(t)$  is chosen so that, in the absence of ACells  
 259 interaction,  $V_{B_i}(t) = V_{i_{drive}}(t)$ . Note that  $F_{B_i}(t)$  implements therefore a time derivative of  
 260 the drive, which makes, e.g. BCells response to moving objects sensitive to changes in  
 261 directions or speed.

262

ACells are connected to BCells with chemical synapses. The differential equation  
 obeyed by the voltage of ACell  $j$  is:

$$\frac{dV_{A_j}}{dt} = -\frac{1}{\tau_{A_j}} V_{A_j} + \sum_{i=1}^{N_B} W_{A_j}^{B_i} \mathcal{N}_B(V_{B_i}), \quad (6)$$

263 where  $\tau_{A_j}$  is the characteristic time scale of ACell  $j$  response, and  $\mathcal{N}_B$  has the same form  
 264 as (4), with a threshold  $\theta_B$ . Note that, in contrast to BCells, ACells do not receive an OPL  
 265 input.



## 266 1.1.4. GCells.

We label GCells (third layer) with the index  $k = 1 \dots N_G$ . They are connected to BCells with excitatory synaptic weights,  $W_{G_k}^{B_i} \geq 0$  (e.g. glutamatergic synapses) and to ACells with inhibitory synaptic weights,  $W_{G_k}^{A_j} \leq 0$  (e.g. glycinergic or GABA-ergic synapses). Their voltage,  $V_{G_k}$ , evolves according to:

$$\frac{dV_{G_k}}{dt} = -\frac{1}{\tau_G} V_{G_k} + \sum_{i=1}^{N_B} W_{G_k}^{B_i} \mathcal{N}_B(V_{B_i}) + \sum_{j=1}^{N_A} W_{G_k}^{A_j} \mathcal{N}_A(V_{A_j}). \quad (7)$$

GCells are spiking. In the model their spiking activity (firing rate) is defined by a LNP model [21–23]. It depends on the voltage via a non linear function  $\mathcal{N}_G(V_G) \equiv f\left(\frac{V_G(t) - \theta_G}{\sigma_G}\right)$ , where  $f$  is typically a sigmoid. Although the detailed form of  $f$  does not matter here, it will be convenient, in the sequel, to consider:

$$\mathcal{N}_G(V_G) = \frac{A}{\sqrt{2\pi}} \int_{-\infty}^{\frac{V_G - \theta_G}{\sigma_G}} e^{-\frac{x^2}{2}} dx. \quad (8)$$

267 The parameters  $\theta_G$  (spiking threshold) and  $\sigma_G$  (controlling the slope of the sigmoid at  
268  $V_G = \theta_G$ ) corresponds, in the case where  $\mathcal{N}_G$  has the form (8), to the probability that a  
269 Gaussian centered Ornstein-Uhlenbeck processes with mean-square deviation  $\sigma_G$  crosses  
270 the threshold  $\theta_G$ .

## 271 1.1.5. Joint dynamics

The joint dynamics of all cells voltage is given by the dynamical system:

$$\begin{cases} \frac{dV_{B_i}}{dt} = -\frac{1}{\tau_B} V_{B_i} + \sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathcal{N}_A(V_{A_j}) + F_{B_i}(t), & i = 1 \dots N_B; \\ \frac{dV_{A_j}}{dt} = -\frac{1}{\tau_A} V_{A_j} + \sum_{i=1}^{N_B} W_{A_j}^{B_i} \mathcal{N}_B(V_{B_i}), & j = 1 \dots N_A; \\ \frac{dV_{G_k}}{dt} = -\frac{1}{\tau_G} V_{G_k} + \sum_{i=1}^{N_B} W_{G_k}^{B_i} \mathcal{N}_B(V_{B_i}) + \sum_{j=1}^{N_A} W_{G_k}^{A_j} \mathcal{N}_A(V_{A_j}), & k = 1 \dots N_G; \end{cases} \quad (9)$$

272 whereas GCells spikes are produced by the LNP mechanism described above.

273 The system of eq. (9) can be summarized as follows (Fig. 1, right). BCells re-  
274 ceive the visual input via the term  $F_{B_i}(t)$  which depends on the stimulus and on the  
275 Bcell's receptive field. They are inhibited by ACells via the synaptic weights  $W_{B_i}^{A_j} < 0$ .  
276 ACells are excited by BCells via the synaptic weights  $W_{A_j}^{B_i} > 0$ . BCells are connected to  
277 GCells via the synaptic weights  $W_{G_k}^{B_i} > 0$ . ACells are connected to GCells via the synap-  
278 tic weights  $W_{G_k}^{A_j} < 0$ . Note that we do not impose any constraint on the connectivity here.  
279

To study mathematically the dynamical system (9) we write it in a more convenient form. We use Greek indices  $\alpha, \beta, \gamma = 1 \dots N \equiv N_A + N_B + N_G$ , and define the state vector  $\vec{\mathcal{X}}$ , with entries:

$$\mathcal{X}_\alpha = \begin{cases} V_{B_i}, & \alpha = i, & i = 1 \dots N_B; \\ V_{A_j}, & \alpha = N_B + j, & j = 1 \dots N_A; \\ V_{G_k}, & \alpha = N_B + N_A + k, & k = 1 \dots N_G. \end{cases} \quad (10)$$

We introduce  $\vec{\mathcal{F}}(t)$ , the non stationary input, with entries:

$$\mathcal{F}_\alpha(t) = \begin{cases} F_{B_i}(t), & \alpha = i, & i = 1 \dots N_B; \\ 0, & \alpha > N_B; \end{cases}$$

and  $\vec{\mathcal{R}}(\vec{\mathcal{X}})$ , the rectification term, with entries:

$$\mathcal{R}_\alpha(\vec{\mathcal{X}}) = \begin{cases} \mathcal{N}_B(V_{B_i}), & \alpha = i, & i = 1 \dots N_B; \\ \mathcal{N}_A(V_{A_j}), & \alpha = N_B + j, & j = 1 \dots N_A; \\ 0, & \alpha = N_B + N_A + k, & k = 1 \dots N_G. \end{cases}$$

280  $\mathcal{R}_\alpha$  can be extended to include gain control; it becomes then a function of voltage and  
281 cell's activity [25,28].

We use the notation  $0_{n_1 n_2}$  for the  $n_1 \times n_2$  matrix with zero entries. We introduce the  $N \times N$  matrices:

$$\mathcal{T} = \begin{pmatrix} -\text{diag}[\tau_{B_i}]_{i=1 \dots N_B} & 0_{N_B N_A} & 0_{N_B N_G} \\ 0_{N_A N_B} & -\text{diag}[\tau_{A_j}]_{j=1 \dots N_A} & 0_{N_A N_G} \\ 0_{N_G N_B} & 0_{N_G N_A} & -\text{diag}[\tau_{G_k}]_{k=1 \dots N_G} \end{pmatrix}, \quad (11)$$

characterizing the characteristic integration times of cells,

$$\mathcal{W} = \begin{pmatrix} 0_{N_B N_B} & W_B^A & 0_{N_B N_G} \\ W_A^B & 0_{N_A N_A} & 0_{N_A N_G} \\ W_G^B & W_G^A & 0_{N_G N_G} \end{pmatrix}, \quad (12)$$

282 summarizing chemical synapses interactions. Note that, to our best knowledge, there  
283 are no synapse from GCells to GCells, but they could be added in this formalism.

Then, the dynamical system (9) reads, in vector form:

$$\frac{d\vec{\mathcal{X}}}{dt} = \mathcal{T}^{-1} \cdot \vec{\mathcal{X}} + \mathcal{W} \cdot \vec{\mathcal{R}}(\vec{\mathcal{X}}) + \vec{\mathcal{F}}(t). \quad (13)$$

284 We remark that (13) has a specific skew-product structure: the dynamics of GCells is  
285 driven by BCells and ACells with no feedback. This means that one can study first the  
286 coupled dynamics of BCells and ACells and then the effect on GCells. This corresponds  
287 to a biological reality as, to our best knowledge, there is no feedback from GCells to  
288 BCells or to ACells.

### 289 1.1.6. Piecewise linear evolution

290 We assume here that  $\mathcal{F}_\alpha(t)$  is bounded, as well as synaptic weights. Thus, the phase  
291 space  $\Omega$  of (13) can be taken compact. Indeed, trajectories cannot escape to infinity thanks  
292 to the rectification terms  $\mathcal{N}_B, \mathcal{N}_A$ , (eq. (4)) and thanks to the sign of synaptic weights  
293  $W_{A_j}^{B_i}, W_{B_i}^{A_j}$ . More precisely,  $V_{B_i}$  cannot become arbitrary large and positive because the  
294 input term  $\mathcal{F}_\alpha(t) \equiv F_{B_i}(t)$  is bounded and because  $\sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathcal{N}_A(V_{A_j}) \leq 0$ . Assume  
295 indeed that  $V_{B_i}$  increases (due to a large enough  $\mathcal{F}_\alpha > 0$  making the r.h.s. of eq. (3)  
296 positive). This leads to an increase of connected ACells voltages  $V_{A_j}$  (eq. (6)), thus to  
297 a decrease of the term  $\sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathcal{N}_A(V_{A_j}) \leq 0$  until the point where the r.h.s. of (3)  
298 becomes negative, thereby decreasing  $V_{B_i}$  and preventing it from becoming arbitrary  
299 large. This, implies as well that  $V_{A_j}$ s cannot become arbitrary large. On the opposite, if  
300  $V_{B_i}$  (resp.  $V_{A_j}$ ) becomes smaller than  $\theta_B$  (resp.  $\theta_A$ ) it does not play any more role in the  
301 dynamics because of rectification.

302 Due to the specific form (4) of the rectification terms, the dynamical system (13) is  
303 piecewise linear. More precisely, we can partition the phase space  $\Omega$  into sub domains  
304  $\Omega^{(n)}$ ,  $n = 1 \dots 2^{N_B + N_A}$  defined as follow. To each cell  $\alpha = 1 \dots N_B + N_A$  (BCell or ACell)  
305 we associate a "rectification label"  $\eta_\alpha = 1$  if the cell  $\alpha$  is rectified and  $\eta_\alpha = 0$  otherwise.  
306 Because of the form (4) of the rectification, the label  $\eta_\alpha$  corresponds to a partition of  
307 the voltage  $X_\alpha$ 's domain of variation into two sub domains (e.g., for a BCell,  $\eta_\alpha = 1$  if

308  $V_{B_i} < \theta_B$  and  $\eta_\alpha = 0$  if  $V_{B_i} \geq \theta_B$ ). One can actually generalize this definition to gain  
 309 control by extending the phase space (adding activity variables) and approximating the  
 310 gain control function by a step function [28]. Now, the set  $\{0, 1\}^{N_B+N_A}$  is made of chains  
 311  $\eta = (\eta_1 \dots \eta_{N_B+N_A})$  composed of the rectification labels  $\eta_\alpha$  of all BCells and ACells. To  
 312 each such sequence is therefore associated a convex domain  $\Gamma^{(n)}$  of  $\mathbb{R}^{N_B+N_A}$  where all  
 313 cells  $\alpha$  such that  $\eta_\alpha = 0$  have their voltage  $X_\alpha$  larger than the rectification threshold, thus,  
 314 are not rectified, and all cells such that  $\eta_\alpha = 1$  are rectified. To each such  $\eta$  is associated  
 315 a unique integer (e.g.  $n = \sum_{\alpha=1}^{N_B+N_A} \eta_\alpha 2^{\alpha-1}$ ,  $\eta$  is then the binary coding of  $n$ ). Finally,  
 316 we set  $\Omega^{(n)} = \Gamma^{(n)} \times \mathbb{R}^{N_G}$ , where the product with the subspace  $\mathbb{R}^{N_G}$  integrates the  
 317 states space of GCells dynamics. They are slaved by BCells and ACells dynamics, but  
 318 they are not rectified. In this setting,  $\Omega^{(0)}$  is the subset of  $\Omega$  such that neither BCells nor  
 319 ACells are rectified;  $\Omega^{(1)}$  the subset of the phase space where only BCell 1 is rectified;  
 320  $\Omega^{(3)}$  the subset where only BCells 1, 2 are rectified;  $\Omega^{(2^{N_B})}$  the subset where only ACell 1  
 321 is rectified and so on.

It is easy to check that the sets  $\Omega^{(n)}$  are disjoint and cover  $\mathbb{R}^N$ , thus, make a partition  
 of the phase space. The vector  $\vec{\mathcal{R}}(\vec{\mathcal{X}})$  has now the form:

$$\mathcal{R}_\alpha(\vec{\mathcal{X}}) = \begin{cases} (1 - \eta_\alpha) (X_\alpha - \theta_B), & \alpha = 1 \dots N_B; \\ (1 - \eta_\alpha) (X_\alpha - \theta_A), & \alpha = N_B \dots N_B + N_A; \\ 0, & \alpha = N_B + N_A + k, \quad k = 1 \dots N_G, \end{cases}$$

and is piecewise-linear in  $\vec{\mathcal{X}}$ . For  $\vec{\mathcal{X}} \in \Omega^{(n)}$ , the transformation  $\mathcal{T}^{-1} \cdot \vec{\mathcal{X}} + \mathcal{W} \cdot \vec{\mathcal{R}}(\vec{\mathcal{X}})$  can  
 therefore be written  $\mathcal{L}^{(n)} \cdot \vec{\mathcal{X}} + \vec{\mathcal{C}}^{(n)}$ , where  $\vec{\mathcal{C}}^{(n)}$  is the vector with entries:

$$\vec{\mathcal{C}}^{(n)} = \begin{cases} -\theta_B (1 - \eta_\alpha), & \alpha = 1 \dots N_B; \\ -\theta_A (1 - \eta_\alpha), & \alpha = N_B \dots N_B + N_A; \\ 0, & \alpha = N_B + N_A + 1, \dots, N. \end{cases} \quad (14)$$

The matrix:

$$\mathcal{L}^{(n)} = \begin{pmatrix} -\text{diag} \left[ \frac{1}{\tau_{B_i}} \right]_{i=1 \dots N_B} & W_B^A \cdot D_A^{(n)} & 0_{N_B N_G} \\ W_A^B \cdot D_B^{(n)} & -\text{diag} \left[ \frac{1}{\tau_{A_j}} \right]_{j=1 \dots N_A} & 0_{N_A N_G} \\ W_G^B \cdot D_B^{(n)} & W_G^A \cdot D_A^{(n)} & -\text{diag} \left[ \frac{1}{\tau_{G_k}} \right]_{k=1 \dots N_G} \end{pmatrix}, \quad (15)$$

322 is called the *transport operator in the domain*  $\Omega^{(n)}$ . This terminology is further explained  
 323 in section 1.1.8, but, in short,  $\mathcal{L}^{(n)}$  acts as a flow (or a propagator) characterizing the evo-  
 324 lution of a trajectory within  $\Omega^{(n)}$ . In eq. (15), the matrices  $D_B^{(n)} = \text{diag}[1 - \eta_\alpha]_{\alpha=1 \dots N_B}$ ,  
 325  $D_A^{(n)} = \text{diag}[1 - \eta_\alpha]_{\alpha=N_B+1 \dots N_B+N_A}$  are projecting onto the subspace of non rectified  
 326 cells in the domain  $\Omega^{(n)}$ . In other words, when the state  $\vec{\mathcal{X}}$  is  $\Omega^{(n)}$ , a rectified cell  $\alpha$  gives  
 327 a zero contribution to the dynamics of other cells, which corresponds to have a row and  
 328 column  $\alpha$  made of zeros in  $D_A^{(n)}, D_B^{(n)}$ .

The dynamical system (13) reads now:

$$\frac{d\vec{\mathcal{X}}}{dt} = \mathcal{L}^{(n)} \cdot \vec{\mathcal{X}} + \vec{\mathcal{F}}^{(n)}(t), \quad \vec{\mathcal{X}} \in \Omega^{(n)}, \quad (16)$$

329 where we wrote  $\vec{\mathcal{F}}^{(n)}(t) = \vec{\mathcal{C}}^{(n)} + \vec{\mathcal{F}}(t)$ . Thanks to the decomposition of the phase  
 330 space into convex sub-domains  $\Omega^{(n)}$ , (16) is now linear. This technique of phase space  
 331 decomposition is classical and has been used in domains such as ergodic theory and  
 332 billiards, self-organized criticality [46,47] or neurosciences [29,30,32,48]. See especially  
 333 the recent paper by A. Rajakumar et al [49] very much in the spirit of the present model.

### 334 1.1.7. Spectra and fixed points

335 It is important to consider in detail the spectrum<sup>2</sup> of  $\mathcal{L}^{(n)}$  for further studies. We  
 336 note  $\lambda_\beta^{(n)}, \beta = 1 \dots N$ , the eigenvalues of  $\mathcal{L}^{(n)}$  and its right eigenvectors are noted,  $\mathcal{P}_\beta^{(n)}$ .  
 337 These vectors are the columns of the matrix  $\mathcal{P}^{(n)}$  transforming  $\mathcal{L}^{(n)}$  in diagonal form  
 338 (assuming it is diagonalizable).  $(\mathcal{P}^{(n)})^{-1}$  is the inverse matrix. Its rows are the left  
 339 eigenvectors of  $\mathcal{L}^{(n)}$ .

340 As  $D_A^{(n)}, D_B^{(n)}$  are projections matrices, it is easy to see, from the form (15), that a  
 341 rectified cell generates an eigenvalue  $-\frac{1}{\tau_\alpha}$  and an eigenvector  $\vec{e}_\alpha$ , the canonical basis  
 342 vector of  $\mathbb{R}^N$  in the direction  $\alpha$ . The non rectified cells span a subspace of  $\mathbb{R}^N$  and  
 343 the projection of  $\mathcal{L}^{(n)}$  on this subspace has a spectrum depending on the connectivity  
 344 matrices  $W_A^B, W_B^A$  and on other parameters like characteristic times.

345 The corresponding eigenvalues  $\lambda_\beta^{(n)}, \beta = 1 \dots N$  can be real or complex, with a  
 346 positive or a negative real part. In the case where  $W_A^B$  and  $W_B^A$  commute it is actually  
 347 possible to explicitly compute the eigenvalues and the eigenvectors and obtain conditions  
 348 for stability (all eigenvalues with real negative part) and real/complex eigenvalues [28].  
 349 If we further assume that  $W_A^B$  and  $W_B^A$  have no zero eigenvalues, the sign constraints  
 350 on these matrices implies that  $\mathcal{L}^{(n)}$  is invertible for all  $n$ . This is what we are going to  
 351 assume from now.

352 It follows that, in the absence of external stimulus ( $\vec{\mathcal{F}}(t) = \vec{0}$ ), equation (16) has,  
 353 for each  $n$ , a unique fixed point  $\vec{\mathcal{X}}^{(n)} = -(\mathcal{L}^{(n)})^{-1} \cdot \vec{\mathcal{C}}^{(n)}$ . Note, however, that this point  
 354 *may not be* in  $\Omega^{(n)}$ . This is a typical situation for piecewise linear dynamical systems  
 355 (like Iterated Function Systems [52–54]) where dynamics can have complex attractors  
 356 even if maps are linear into sub-domains of the phase space. The simplest non trivial  
 357 case is when dynamics generates a periodic orbit, but more complex attractors (fractal  
 358 sets) can be obtained. Here, it is reasonable to assume, at least, that cells at rest are not  
 359 rectified. Mathematically, this means that the fixed point of  $\mathcal{L}^{(0)}$ ,  $\vec{\mathcal{X}}^* = -(\mathcal{L}^{(0)})^{-1} \cdot \vec{\mathcal{C}}^{(0)}$ ,  
 360 belongs to  $\Omega^{(0)}$  and this is what we are going to assume for now. This imposes a set of  
 361 constraints linking synaptic weights and thresholds. A simple assumption consists of  
 362 having vanishing thresholds  $\theta_A = \theta_B = 0$ , in which case the rest state is  $\vec{0}$ . We will also  
 363 assume that  $\vec{\mathcal{X}}^*$  is stable (eigenvalues of  $\mathcal{L}^{(0)}$  have a negative real part), which imposes  
 364 additional assumptions on synaptic weights and cells integration times. On biophysical  
 365 grounds it means that the rest state is stable to small perturbations, like noise. Because  
 366 rectified cells produce stable eigenvalues the following holds. Taking an initial condition  
 367 in any domain  $\Omega^{(n)}$  spontaneous dynamics (without stimulus) eventually drives the  
 368 trajectory back to  $\Omega^{(0)}$  and, then, to the rest state. This is further commented below  
 369 (section 1.1.8, remark 2).

### 370 1.1.8. Solutions

We now consider the general situation where dynamics is in the rest state at times  
 $t < 0$ , and, from time  $t = 0$  on, the stimulus  $\mathcal{S}(t)$  is applied, resulting in a non stationary  
 drive  $\vec{\mathcal{F}}(t)$ . In general, the stimulus is applied over a finite time. After this the system  
 eventually returns to rest. Under this stimulation the trajectory  $\left\{ \vec{\mathcal{X}}(t) \right\}_{t \geq 0}$  is going to  
 cross a sequence of domains  $\Omega^{(n_k)}, k = 1, \dots$ , with  $n_1 = 0$ , entirely determined by the  
 stimulus and the network characteristics. Call  $t_-^{(n_{k+1})}$  the time where the trajectory enters

<sup>2</sup> Another approach consists of considering the Schur decomposition instead of the diagonalisation [49–51].

the domain  $\Omega^{(n_{k+1})}$  and  $t_+^{(n_{k+1})}$  the time where it gets out. Note that  $t_-^{(n_{k+1})} = t_+^{(n_k)}$ . By direct integration of eq. (16), we have:

$$\vec{\mathcal{X}}(t) = e^{\mathcal{L}^{(n_{k+1})}(t-t_-^{(n_{k+1})})} \cdot \vec{\mathcal{X}}(t_-^{(n_{k+1})}) + \int_{t_-^{(n_{k+1})}}^t e^{\mathcal{L}^{(n_{k+1})}(t-s)} \cdot \vec{\mathcal{F}}^{(n_{k+1})}(s) ds, \quad t \in [t_-^{(n_{k+1})}, t_+^{(n_{k+1})}], \quad (17)$$

where  $\vec{\mathcal{X}}(t_-^{(n_{k+1})})$ , corresponding to the state of  $\vec{\mathcal{X}}$  when entering  $\Omega^{(n_{k+1})}$ , is given by the integration of the past trajectory and can be computed explicitly. This is:

$$\vec{\mathcal{X}}(t_-^{(n_{k+1})}) = \vec{\mathcal{X}}(t_+^{(n_k)}) = \sum_{m=0}^k \mathcal{H}_m^k \vec{\Phi}_m, \quad (18)$$

where  $\mathcal{H}_m^k$  is a sequence of matrices satisfying:

$$\mathcal{H}_k^k = \mathcal{I}_N; \quad \mathcal{H}_m^k = \mathcal{H}_{k-1}^k \mathcal{H}_m^{k-1}; \quad \mathcal{H}_{k-1}^k = e^{\mathcal{L}^{(n_k)}(t_+^{(n_k)} - t_+^{(n_{k-1})})}, \quad (19)$$

where  $\mathcal{I}_N$  is the identity matrix of dimension  $N$ . The matrix  $\mathcal{H}_m^k$  transports the flow from the exit point of  $\Omega^{(n_m)}$  to the exit point of  $\Omega^{(n_k)}$ . The vectors  $\vec{\Phi}_m$  are defined by:

$$\vec{\Phi}_0 = \vec{\mathcal{X}}(0); \quad \vec{\Phi}_m = \int_{t_-^{(n_m)}}^{t_+^{(n_m)}} e^{\mathcal{L}^{(m)}(t_+^{(n_m)} - s)} \cdot \vec{\mathcal{F}}^{(m)}(s) ds. \quad (20)$$

371 The proof of (18) is easily done by recurrence.

### 372 1.1.9. Remarks

373 Let us now make some remarks on the structure of these solutions.

- 374 1. In the definition of  $\mathcal{H}_m^k$  the operators  $\mathcal{L}_{n_k}$  do not commute in general.
- 375 2. Eigenvalues of some  $\mathcal{H}_m^k$  can have a positive real part leading to exponential  
376 increase along the corresponding eigen-direction. This means that some cells  
377 voltage increases exponentially in absolute value. However, when voltages become  
378 too large, voltage rectification takes place, corresponding to the trajectory entering  
379 a new continuity domain. Here, unstable cells do not contribute any more to  
380 dynamics which is projected on the subspace of non rectified cells. This has  
381 the effect of transforming unstable eigenvalues into stable ones preventing the  
382 trajectories  $\vec{\mathcal{X}}(t)$  to diverge. Actually, the spectrum of  $\mathcal{H}_m^k$ , controlling stability,  
383 resembles the Lyapunov spectrum in ergodic theory [55], with two main differences.  
384 First, we are simply considering product of matrices without multiplying by the  
385 adjoint so that eigenvalues can be complex. Second, we are not assuming stationarity  
386 and the existence of an invariant measure. Instead, the product  $\mathcal{H}_m^k$  is constrained  
387 by the non stationary stimulus and dynamical system parameters which fixes the  
388 sequence of times  $n_k$ s.
- 389 3. Rectification induces a weak form of non linearity where e.g. the contraction/expansion  
390 in the phase space depends on the domain  $\Omega^{(n_k)}$  (whereas in a differentiable non  
391 linear system it would depend on the point in the phase space). This has deep  
392 consequences on cells response, as commented in the results sections.

### 393 1.2. Spike statistics

394 As pointed out in the introduction, it might be helpful to propose a mathematical  
395 setting taking into account non stationarity and potentially long memory in spike trains  
396 probabilities. Such a setting exists since long but has not been applied to spike train  
397 statistics until recently. It is inherited from statistical physics on one hand [56] and  
398 on extensions of Markov chains to unbounded memory on the other hand [57]. The  
399 material briefly sketched here has been published in [29–39].

### 400 1.2.1. Mathematical setting for spike trains

401 Neurons variables such as membrane potential or ionic currents are described  
 402 by continuous-time equations. In contrast, spikes resulting from the experimental  
 403 observation are discrete events, binned with a certain time resolution  $\delta$ , say of order  
 404 a millisecond. We consider a network of  $N$  spiking neurons, labelled with an index  
 405  $k = 1 \dots N$ . We define a spike variable  $\omega_k(n) = 1$  if neuron  $k$  has emitted a spike in the  
 406 time interval  $[n\delta, (n+1)\delta[$ , and  $\omega_k(n) = 0$  otherwise. We denote by  $\omega(n) = [\omega_k(n)]_{k=1}^N$   
 407 the spike-state of the entire network at time  $n$ , which we call a *spiking pattern*. A *spike*  
 408 *block* denoted by  $\omega_m^n$ ,  $n \geq m$ , is the sequence of spike patterns  $\omega(m), \omega(m+1) \dots \omega(n)$ .  
 409 The range of a block  $\omega_m^n$  is  $n - m + 1$ , the number of time steps from  $m$  to  $n$ . We call  
 410 a spike train an infinite sequence of spikes both in the past and in the future, and, to  
 411 simplify notations we note a spike train  $\omega$  (instead of  $\omega_{-\infty}^{+\infty}$ ). Of course, on operational  
 412 grounds spike trains are finite, but it is mathematically more convenient to work on a  
 413 space of bi-infinite spike sequences.

### 414 1.2.2. Mathematical setting for spiking probabilities

415 We now consider a family of transition probabilities of the form  $\mathbb{P}_n \left[ \omega(n) \mid \omega_{-\infty}^{n-1} \right]$ ,  
 416 which represent the probability, that at time  $n$ , one observes the spiking pattern  $\omega(n)$   
 417 given the network spike history, extending to an infinite past. This is an extension of  
 418 Markov chains where probabilities have the form  $\mathbb{P}_n \left[ \omega(n) \mid \omega_{n-D}^n \right]$ , where  $D$  is the  
 419 memory depth of the Markov chain. Letting the memory be possibly infinite corresponds  
 420 to situation where one cannot precisely fix the memory depth necessary to characterize  
 421 the probability of a spike pattern given the past spike history. An example of a model  
 422 requiring this context is presented in section 1.2.3 below. Having infinite memory  
 423 imposes mathematical constraints on the memory decay that has to be sufficiently fast  
 424 (typically, exponential) so that the situation is close to Markov chains. In addition to the  
 425 model presented below, neural models with infinite memories have been considered by  
 426 several authors such as E. Loecherbach and A. Galves [58]. A few remarks about this  
 427 form of probability:

- 428 1. We do not assume stationarity.  $\mathbb{P}_n$  may depend explicitly on time. This is actually  
 429 the reason why we have an index  $n$ . A time translation invariant probability will  
 430 simply be written  $\mathbb{P}$ .
- 431 2. For such probabilities to be well defined and useful, one need to make assumptions  
 432 on their structure. Beyond technical assumptions such as measurability, summabil-  
 433 ity, non nullness and continuity [59,60], the most important assumption here is that  
 434 the dependence in the past (memory) decays fast enough, typically, exponentially,  
 435 so that, even if this chain has infinite memory it is very close to Markov.
- 436 3. As one can associate to Markov chains an equilibrium probability (under condi-  
 437 tions actually quite more general than detailed-balance) the system of transition  
 438 probabilities  $\{\mathbb{P}_n\}_{n \in \mathbb{Z}}$  also admits, under the mathematical conditions sketched in  
 439 the item 2 above, an equivalent notion called “chains with complete connections”  
 440 or a “chain swith unbounded memory” [57].
- 441 4. These distributions are formally (left-sided) Gibbs distributions where the Gibbs po-  
 442 tential is  $\Phi(n, \omega) \stackrel{\text{def}}{=} \log \mathbb{P}_n \left[ \omega(n) \mid \omega_{-\infty}^{n-1} \right]$  (the non-nullness assumption imposes  
 443 that  $\mathbb{P}_n \left[ \omega(n) \mid \omega_{-\infty}^{n-1} \right] > 0$ ). This establishes a formal link to statistical physics.  
 444 In particular, when the chain is stationary, expanding the potential in product  
 445 of spikes events up to second order one recovers the maximum entropy models  
 446 used in the literature of spike trains analysis, including the so-called Ising model  
 447 [35,61–63]. However, the chains we consider are not necessarily stationary.

### 448 1.2.3. A model of effective interactions between GCells

449 The visual cortex has no clue on which biophysical processes are taking place in the  
 450 retina. All the visual information it receives is encoded in spike trains. This leads to the  
 451 idea of proposing models of spiking GCells network where dynamics of GCells voltage  
 452 is only constrained by GCells spikes history. Here, one assumes that GCells dynamics is  
 453 controlled by the interactions with hidden layers, for example, the BCells-ACells layers  
 454 in the model (13), in a situation where an observer is just recording the spikes emitted by  
 455 GCells, while having no clue of the dynamics in the upper layers. These hidden layers  
 456 result in providing, *causal, effective interactions* between GCells that one can interpolate  
 457 by fitting the statistics. The idea is then to construct a dynamical model where the  
 458 spiking of a GCell depends on the spike history emitted by the network, with virtual  
 459 interactions that mimic hidden causal effects [64]. This strategy lead us to propose the  
 460 model presented in the next paragraph. The advantage of this approach is that one can  
 461 explicitly write the transition probabilities  $\mathbb{P}_n \left[ \omega(n) \mid \omega_{-\infty}^{n-1} \right] > 0$  and infer, from this,  
 462 a linear response formula telling us how statistical quantities such as firing rates, but  
 463 also spike correlations are modified by a time dependent stimulus. These results are  
 464 presented in the "Results" subsection 1.2.3.

465 The model is inspired from the generalized Integrate and Fire model (gIF) proposed  
 466 by Rudolph and Destexhe [65] and generalizes the Leaky-Integrate and Fire model  
 467 [66,67]. We have  $N$  neurons (say GCells) characterized by their voltage  $V_k, k = 1 \dots N$ .  
 468 One fixes a voltage threshold  $\theta$  such that, whenever  $V_k(t) = \theta$  a spike is emitted by  
 469 neuron  $k$  at time  $t$ , and is reset to a reset value (typically,  $V_{reset} = 0$ ). Below  $\theta$ , the  
 470 dynamics of voltage (sub-threshold dynamics) is governed by eq. (23) below.

In the LIF model, synaptic conductances are constant. In the gIF model, in contrast,  
 the synaptic conductance  $g_{kj}$  between the pre-synaptic neuron  $j$  and the post-synaptic  
 neuron  $k$  depends on spike history as:

$$g_{kj}(t, \omega) = G_{kj} \alpha_{kj}(t, \omega), \quad (21)$$

where:

$$\alpha_{kj}(t, \omega) = \sum_{n=-\infty}^t \alpha_{kj}(t-n) \omega_j(n). \quad (22)$$

471 The notation  $g_{kj}(t, \omega)$  means that function  $g_{kj}$  depends on spikes occurring before time  
 472  $t$ .  $G_{kj} \geq 0$  is the maximal conductance between  $j$  and  $k$ . It is zero when there is  
 473 no synaptic connection between neurons  $j$  and  $k$ . In (22), the function  $\alpha_{kj}(t)$ , called  
 474  $\alpha$ -kernel, summarizes the complex dynamical process underlying the generation of  
 475 a post-synaptic potential after the emission of a pre-synaptic spike [68]. It has the  
 476 typical form  $\alpha_{kj}(t) = P(t) e^{-\frac{t}{\tau_{kj}}} H(t)$  where  $P(t)$  is a polynomial in time and  $H(t)$  is the  
 477 Heaviside function. What matters on mathematical grounds is the exponential tail of  
 478  $\alpha_{kj}(t)$  [33]. The function  $\alpha_{kj}(t, \omega)$  depends on the spike history preceding  $t$ . It records  
 479 the spikes emitted by the pre-synaptic neuron  $j$  before  $t$ , corresponding to  $\omega_j(n) = 1$  and  
 480 adds up a contribution  $\alpha_{kj}(t-n)$  to the post synaptic conductance from pre-synaptic  
 481 neuron  $j$  to post-synaptic neuron  $k$ .

Now, the gIF dynamics reads [30,33,34]:

$$C_k \frac{dV_k}{dt} + g_L (V_k - E_L) + \sum_j g_{kj}(t, \omega) (V_k - E_j) = S_k(t) + \sigma_B \tilde{\zeta}_k(t), \quad \text{if } V_k(t) < \theta,$$

482 where  $g_L, E_L$  are respectively the leak conductance and the leak reversal potential,  $E_j$  the  
 483 reversal potential characterizing the synaptic transmission between  $j$  and  $k$ . Finally,  $\tilde{\zeta}_k(t)$   
 484 is a white noise, introducing stochasticity in dynamics. Its intensity is  $\sigma_B$ .

Setting:  $W_{kj} = G_{kj}E_j$ ,  $i_k(t, \omega) = g_L E_L + \sum_j W_{kj} \alpha_{kj}(t, \omega) + S_k(t) + \sigma_B \zeta_k(t)$ ,  $g_k(t, \omega) = g_L + \sum_{j=1}^N g_{kj}(t, \omega)$ , one can finally write the gIF dynamics in the form:

$$C_k \frac{dV_k}{dt} + g_k(t, \omega) V_k = i_k(t, \omega), \quad \text{if } V_k(t) < \theta, \quad (23)$$

where  $i_k(t, \omega)$  depends, on the network spike history via  $\alpha_{kj}(t, \omega)$ , on the stimulus, and contains a stochastic term. As the reversal potential  $E_j$  can be positive or negative, the synaptic weights  $W_{kj}$  define an oriented and signed graph, whose vertices are the neurons. These weights are what we call effective interactions.

What makes the gIF model very rich is that it proposes a biophysically grounded way to construct a dynamical system where the variables (here, voltages) are constrained by the only information of spike train history. The price to pay is that dynamics actually depends on the *whole spike history*, which is potentially infinite. Actually, the gIF model has an infinite memory. This is essentially because the conductance depends on the whole history, and, contrarily to voltages is not reset when the neuron fires. Nevertheless, the exponential decay in the alpha profile actually ensures the existence (and uniqueness) of transition probabilities of the form  $\mathbb{P}_n \left[ \omega(n) \mid \omega_{-\infty}^{n-1} \right]$  [34–38].

Note that the integration of (23) does not only requires the knowledge of voltages  $V_k$ , stimulus and noise at time  $t$ . It requires, in addition, the knowledge of the spike train  $\omega$  emitted by the network before  $t$ . In this sense, this is not a classical dynamical system. Nevertheless, eq. (23) can be explicitly integrated [34,39].

## 2. Results

### 2.1. How could lateral ACells connectivity shape the receptive field of a ganglion cell ?

The response of a GCell to visual stimuli is shaped by the retina structure depicted in Fig. 1. Here, with the model introduced in section 1.1, we would like to characterize the respective effects of the stimulus and of the network connectivity, especially ACells, and understand under which condition can the conjugated effect of network dynamics and stimulus be represented by a convolution of the form (1) where the kernel  $\mathcal{K}_{G_\alpha}$  is *intrinsic* to the cell, i.e. does not depend on the stimulus ?

#### 2.1.1. Non rectified case

The answer is relatively easy when no rectification takes place, i.e. when the trajectory of (13) stays in the domain  $\Omega^{(0)}$  (see section 1.1.6 for the definition). Indeed, in this case evolution is ruled by equation (17) which holds from the initial time  $t = t_0$  where the stimulus starts to be applied, to the current time  $t$ . Actually, we can consider that  $t_0$  starts far in the past and let it tend to  $-\infty$ . This corresponds to considering that the stimulus is applied on a time scale quite longer than the characteristic times in the problem (i.e. the inverse of the real part of eigenvalues). Then eq. (17) reads  $\vec{\mathcal{X}}(t) = \int_{-\infty}^t e^{\mathcal{L}^{(0)}(t-s)} \vec{\mathcal{F}}^{(0)}(s) ds$ , which is  $\vec{\mathcal{X}}(t) = \left[ e^{\mathcal{L}^{(0)} t} * \vec{\mathcal{F}}_0 \right](t)$ . This equation actually makes sense only if all eigenvalues of  $\mathcal{L}^{(0)}$  are stable, as we assumed above. Note also that  $\vec{\mathcal{F}}^{(0)} = \vec{\mathcal{C}}^{(0)} + \vec{\mathcal{F}}$  where  $\vec{\mathcal{C}}^{(0)}$  is a constant, depending on thresholds (eq. (14)) and whose integration in the convolution product gives  $-\left(\mathcal{L}^{(0)}\right)^{-1} \vec{\mathcal{C}}^{(0)} = \vec{\mathcal{X}}^*$ , the base line activity of  $\vec{\mathcal{X}}(t)$  without stimulus. We may ignore this constant in the sequel and focus on the time varying part of the response,  $\left[ e^{\mathcal{L}^{(0)} t} * \vec{\mathcal{F}} \right](t)$ . As  $\vec{\mathcal{F}}$  is itself defined in terms of a convolution (eq. (5)) with the stimulus and its derivative,  $\vec{\mathcal{X}}(t)$  is a convolution with the stimulus and its derivative. Here, it is useful to express  $\vec{\mathcal{X}}(t)$  in components.



One can then show that [28]:

$$\mathcal{X}_\alpha(t) = V_{\alpha_{drive}}(t) + \mathcal{E}_{\alpha_{net}}^{(0)}(t), \quad \alpha = 1 \dots N, \quad (24)$$

where:

$$\mathcal{E}_{\alpha_{net}}^{(0)}(t) = \sum_{\beta=1}^N \sum_{\gamma=1}^{N_B} \mathcal{P}_{\alpha\beta}^{(0)} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \omega_{\beta\gamma}^{(0)} \int_{-\infty}^t e^{\lambda_{\beta}^{(0)}(t-s)} V_{\gamma_{drive}}(s) ds, \quad (25)$$

532 where  $\omega_{\beta\gamma}^{(0)} = \lambda_{\beta}^{(0)} + \frac{1}{\tau_{B\gamma}}$ . The term  $V_{\alpha_{drive}}(t)$  in eq. (24) is the stimulus drive and acts  
 537 only on BCells (it vanishes for  $\alpha > N_B$ ). The term (25) contains the network effects.  
 538 The drive imposed on BCells impacts ACells via the connectivity and, thereby, have a  
 539 feedback effect on BCells. In addition, the join activity of BCells and ACells drive the  
 540 GCells response ( $\alpha > N_B + N_A$ ). In particular, this equation allows to compute explicitly  
 541 the RF of a GCell.

For this, we introduce the function  $e_{\beta}^{(0)}(t) \equiv e^{\lambda_{\beta}^{(0)} t} H(t)$  so that  $\int_{-\infty}^t e^{\lambda_{\beta}^{(0)}(t-s)} V_{\gamma_{drive}}(s) ds \equiv$   
 $\left[ e_{\beta}^{(0)} \overset{t}{*} V_{\gamma_{drive}} \right](t)$ , which according to (2) is  $\left[ e_{\beta}^{(0)} \overset{t}{*} \mathcal{K}_{B\gamma} \overset{x,y,t}{*} \mathcal{S} \right](t)$ . Thus, by identifica-  
 tion with (1), the kernel of GCell  $\alpha = N_B + N_A + 1 \dots N_G$  is:

$$\mathcal{K}_{G_\alpha}(x, y, t) = \sum_{\beta=1}^N \sum_{\gamma=1}^{N_B} \mathcal{P}_{\alpha\beta}^{(0)} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \omega_{\beta\gamma}^{(0)} \left[ e_{\beta}^{(0)} \overset{t}{*} \mathcal{K}_{B\gamma} \right]. \quad (26)$$

532 This provides an explicit equation for the kernel of a GCell, embedded in a network of  
 533 BCells, ACells, GCells with dynamics (13), when no rectification take place.

### 534 2.1.2. Consequences

535 **Interpretation.** The kernel obtained in (26) is the response of the GCell to a Dirac  
 536 pulse corresponding, in experiments, to a brief light (or dark) full-field flash. It can also  
 537 obtained from a white noise stimulus, corresponding, in experiments, to the so-called  
 538 Spike Triggered Average (STA) [21–23]. It corresponds therefore to the functional def-  
 539 inition of the receptive field of GCells used in experiments. In addition, eq. (24), (25)  
 540 give us the voltage of *all* cells in the network at time  $t$  under the influence of a stimulus.  
 541 Interestingly, thus, these equations allow us to visualize the join evolution of BCells  
 542 and ACells as well as their action of GCells. Note that BCells and ACells are difficult to  
 543 access experimentally. Given a prescribed connectivity (matrices  $W_A^B, W_B^A, W_G^B, W_G^A$ ), eq.  
 544 (24) provides us, therefore, a mathematical insight on the potential, hidden, dynamics  
 545 of BCells and ACells leading to the experimentally observed response of GCells. Thus,  
 546 this give us possible scenarios characterizing the potential effects of ACells networks on  
 547 GCells response. In addition, eq. (26) also provides the RF for BCells ( $\alpha = 1 \dots N_B$ ) and  
 548 ACells ( $\alpha = N_B + 1 \dots N_B + N_A$ ). We observe in particular that, in a network, the RF of  
 549 a BCell is therefore not only what comes from the OPL - the term  $V_{\alpha_{drive}}(t)$  - it integrates  
 550 as well lateral ACells connectivity. This is similar to the center-surround shaping of  
 551 OPL output due to HCells, but here, we might have different effects, due to the different  
 552 physiology of ACells.

553 **Space-time separability.** The GCell kernel, in general, does not factorise into a  
 product of a function of space and a function of time (separability). Even in the case  
 where the BCells RF is separable, i.e.  $\mathcal{K}_{B\gamma}(x, y, t) = \mathcal{K}_{B_{S\gamma}}(x, y) \mathcal{K}_{B_{T\gamma}}(t)$  where  $\mathcal{K}_{B_{S\gamma}}$  is the  
 spatial part, centred at  $x_\gamma, y_\gamma$  and  $\mathcal{K}_{B_{T\gamma}}$  the temporal part, the GCell kernel reads:

$$\mathcal{K}_{G_\alpha}(x, y, t) = \sum_{\beta=1}^N \mathcal{P}_{\alpha\beta}^{(0)} \left( \sum_{\gamma=1}^{N_B} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \omega_{\beta\gamma}^{(0)} \left[ e_{\beta}^{(0)} \overset{t}{*} \mathcal{K}_{B_{T\gamma}} \right] \times \mathcal{K}_{B_{S\gamma}}(x, y) \right), \quad (27)$$

and is not separable either. Now, if BCells have the same temporal kernel  $\mathcal{K}_{B_T}$ , independent of  $\gamma$  and the same characteristic time  $\tau_B$ , such that  $\omega_{\beta\gamma}^{(0)} = \lambda_{\beta}^{(0)} + \frac{1}{\tau_B}$  is independent of  $\gamma$ , we can write :

$$\mathcal{K}_{G_{\alpha}}(x, y, t) = \sum_{\beta=1}^N \mathcal{P}_{\alpha\beta}^{(0)} \omega_{\beta}^{(0)} \left[ e_{\beta}^{(0)} \overset{t}{*} \mathcal{K}_{B_T} \right] \left( \sum_{\gamma=1}^{N_B} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \mathcal{K}_{B_{S_{\gamma}}}(x, y) \right). \quad (28)$$

554 This kernel is not yet strictly separable as the term  $\sum_{\gamma=1}^{N_B} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \mathcal{K}_{B_{S_{\gamma}}}(x, y)$  still de-  
 555 pends on  $\beta$ , the eigenmode index, via  $\left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1}$ . Now, the eigenmodes depends  
 556 on connectivity. Especially, the BCell to GCell connectivity corresponds to a pooling  
 557 of BCells located in the vicinity of GCell  $\alpha$ . The simplest case is when there is  
 558 no lateral connectivity and where each GCell  $\alpha$  is contacted by only BCell with in-  
 559 dex  $\gamma_{\alpha}$  (this implies  $N_B = N_G$ ). In this case:  $\mathcal{P}_{\alpha\beta}^{(0)} = \delta_{\alpha\beta}$ ,  $\left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} = \delta_{\beta\gamma}$  so that

560  $\mathcal{K}_{G_{\alpha}}(x, y, t) = \omega_{\alpha}^{(0)} \left[ e_{\alpha}^{(0)} \overset{t}{*} \mathcal{K}_{B_T} \right] \mathcal{K}_{B_{S_{\alpha}}}(x, y)$  is separable. More generally, pooling im-

561 plies that  $\mathcal{P}_{\alpha\beta}^{(0)}$  and  $\left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1}$  are locally spread around  $\alpha$  resulting in a spatial part

562  $\sum_{\gamma=1}^{N_B} \left( \mathcal{P}_{\beta\gamma}^{(0)} \right)^{-1} \mathcal{K}_{B_{S_{\gamma}}}(x, y)$  depending only on  $\alpha$ .

563

564 **Resonances.** The eigenvalues of  $\mathcal{L}^{(0)}$  can be complex, going by conjugated pairs. It  
 565 is actually quite easy to obtain such a situation mathematically, even considering nearest  
 566 neighbours interactions [28]. A straightforward consequence is the existence of preferred  
 567 time frequencies (resonance) for a GCell. In other words, applying periodic sequences  
 568 of brief flashes with a varying frequency, one might observe a peak in the amplitude of  
 569 the GCell response, for specific frequencies. This remark could, e.g., explain the "bump"  
 570 observed in experiments when the retina is submitted to the so-called "Chirp" stimulus  
 571 [69], a stimulus composed of different phases of flashes stimulation where one varies  
 572 duration, frequency, and amplitude. In the phase where the amplitude is constant but  
 573 frequency is varying, some GCells exhibit a resonance like peak (see e.g. Fig. 1 b in [69]).  
 574 Of course, such resonances could also be explained by intrinsic cells properties, like  
 575 ion channels response. The potential effect of lateral ACells connectivity would have  
 576 to be tested experimentally by, e.g. inhibiting ACells synaptic transmission for GCells  
 577 exhibiting resonance peaks.

578

579 **Stimulus induced waves.** This is a general fact that networks of coupled units can  
 580 produce waves. Spontaneous waves are actually reported in the developmental retina,  
 581 induced, in the so-called stage II and stage III by ACells [70]. They are generated by non  
 582 linear mechanisms and closeness to bifurcations [71]. This is not the type of wave we  
 583 are dealing with here, though. Instead, we are referring to waves triggered by a moving  
 584 stimulus, say a moving bar. The idea is that such a stimulus can induce, via ACells con-  
 585 nectivity, a wave of connectivity which can be *ahead* of the stimulus, for a certain range  
 586 of parameters (e.g. synaptic coupling intensity) compatible with physiology. Stimulus  
 587 induced waves, in advance with respect to the stimulus, have been reported in the visual  
 588 cortex [72]. They are due to lateral cortical activity and induce cortical anticipation. The  
 589 mathematical analysis made in [28] suggests that such anticipatory waves could as well  
 590 exist in the retina thanks to ACells lateral connectivity, conjugated with non linear gain  
 591 control already known to induce a form of retinal anticipation [25,27].

592

593 **Stimulus adaptation.** Short term plasticity has been reported in the retina at the  
 594 synapses between BCells-ACells and ACells-GCells[26,73]. Note actually that, although  
 595 most models of plasticity, referring to cortical neurons, are considering spiking neurons  
 596 [74], the physiology of short term synapse adaptation does not necessarily require spikes

and is compatible with inner retinal networks dynamics. The effect of synaptic plasticity can be integrated in the model (13). It will result in a variations of eigenvalues and eigenvectors of the transport operator  $\mathcal{L}$  with potential changes in dynamics. Although, potential and highly relevant phenomena such as bifurcations induced by plasticity would require considering a non linear version of (13) (at least, rectification to avoid exponential instability), we can ask about simple linear effect of plasticity on the GCells response. A straightforward potential effect could be frequency adaptation to periodic flashes.

### 2.1.3. Rectification.

Let us now investigate the role of rectification. In the general case, a trajectory crosses several domains, and is characterized by eq. (17). Starting from the domain  $\Omega^{(0)}$  (rest state) the state of the network submitted to a stimulus, enters a new domain  $\Omega^{(n_1)}$  at time  $t_+^{(n_1)}$  where some cells are rectified and so on. Can one still define a response formula of type (1) ? This raises several technical difficulties, first because some eigenvalues can be unstable. As we have seen above, this does not lead to an exponential explosion though precisely rectification prevents cells voltage to diverge. Mathematically, this is expressed by the exit of the trajectory from the domain with positive eigenvalue and a projection on the subspace spanned by non-rectified cells. Another difficulty also comes from the constants  $\vec{c}^{(n)}$  defined in (14) coming from the threshold in the rectification function. They can be removed by assuming that all thresholds are equal to 0. This is what we are going to do now for the sake of simplicity. One can then define domain-dependent flows  $\Phi^{(n)}(\vec{\mathcal{X}}, t) \equiv e^{\mathcal{L}^{(n)} t} \Theta(\vec{\mathcal{X}}(t) \in \Omega^{(n)})$ , where  $\Theta$  is the indicator function so that  $\left[ \Phi^{(n)}(\vec{\mathcal{X}}, \cdot) \overset{t}{*} \vec{\mathcal{F}} \right](t) = \sum_{n_m=n} \int_{t_-^{(n_m)}}^{t_+^{(n_m)}} e^{\mathcal{L}^{(m)}(t_+^{(n_m)}-s)} \cdot \vec{\mathcal{F}}(s) ds$  where the sum holds on indices  $n_m$  in the trajectory such that  $n_m = n$ . This allows to express the recurrence formula (18) in terms of a convolution and thereby to express the whole trajectory in terms of a convolution with a transport operator.

However, there are several important differences with the non rectified case. First, the kernel defined this way *depends on the trajectory*. As the sequence of domains met by the trajectory, and the time where the trajectory enters in these domains, depend on the stimulus, the RF of rectifiable cells *depends now on the stimulus*. Note that the situation would actually be even worse for non linear cells. Indeed, the question hidden behind these remarks is: "to what extent the *linear response* assumption defining a RF via a convolution equation such as (2) is valid". We will actually come back to a similar question in section 2.3 for a network of spiking neurons. Linear response essentially requires the perturbation to be "weak enough", which in our case, means that cells are not rectified. The formulation in terms of a piecewise linear system allows to extend the notion of RF to rectified cells, but the price to pay is that RF now depends on the stimulus. With respect to biology, this effect would for example mean that cells identified e.g. to be ON with a STA approach, responds differently (e.g. ON-OFF) to a more sophisticated stimulus like the "chirp" stimulus [69].

In the rectified cases, the eigenvalues  $\lambda_\beta^{(n)}$ ,  $\beta = 1 \dots N$  and eigenvectors  $\mathcal{P}_\beta^{(n)}$  depend on the domain, i.e. on the list of rectified cells and are different from the domain  $\Omega^{(0)}$  of the rest state. They actually differ in two ways. First, rectified cells provide eigenvalues  $-\frac{1}{\tau_\beta}$  and eigenvectors  $\vec{e}_\beta$  so that  $\mathcal{P}_{\alpha\beta}^{(n)} = \delta_{\alpha\beta}$  for these cells so that they do not contribute any more to the network response. The second effect is more intricate. Indeed, the mere fact of rectifying one cell, has, in general, the effect of *modifying the whole spectrum and eigenvectors*, with strong effects on the cells response. This can be easily understood. Consider the (not really retinal-realistic) situation where a cell is a hub in a network. Silencing it have in general dramatic effects on the global dynamics of this network.

#### 646 2.1.4. Conclusion of section 2.1

647 In this section, we have given a mathematical answer to the problem 1, level 1,  
 648 posed in the introduction. On the basis of a simplified model of BCells - ACells -  
 649 GCells interactions, we have produced a formalism allowing us to compute this network  
 650 response to spatio-temporal stimuli. We have been able to write explicitly the RF of  
 651 individual GCells appearing in eq. (1) where the kernel depends explicitly on lateral  
 652 connectivity. As we showed, however, the linear response formula (1), where the kernel  
 653 is independent of the stimulus, holds when the stimuli has a weak enough amplitude so  
 654 that cells are not rectified. As soon as rectification takes place the convolution form (1)  
 655 implies, in general, that the kernel can change with the stimulus. This effect could be  
 656 observed in experiments if the cell type, characterized via STA, provides a different type  
 657 of response to other stimuli.

#### 658 2.2. How could spatio-temporal stimuli correlations and retinal network dynamics shape the 659 spike train correlations at the output of the retina ?

660 In the section, we extrapolate the previous analysis of the model (13) to analyse  
 661 how spike trains emitted by GCells can be correlated via the network and especially  
 662 ACells connectivity. We especially want to make mathematical statements on how could  
 663 ACells decorrelate GCells, as claimed on the basis of experiments [15] ? We consider first  
 664 the non rectified case and then analyse how rectification can modify correlations.

#### 665 2.2.1. Voltages correlations

666 We now compute the GCells voltage correlations induced by a non stationary  
 667 spatio-temporal stimulus in the model (13). Note that correlations requires some notion  
 668 of probability, thus, of randomness. Moreover, it is more convenient when such a  
 669 probability is stationary, while we want here to consider a non-stationary problem. This  
 670 is not contradictory though. There are two simple (not incompatible) ways to address  
 671 this point. First, one may consider that the dynamical system (13) has random initial  
 672 conditions, drawn with respect to a stationary probability measure. Second, one can add  
 673 to the dynamics (13) noise, which always present in biological systems. We can make the  
 674 assumption that noise is stationary and that it is Brownian (which is a pure mathematical  
 675 convenience). In biology, spike correlations are usually obtained by averaging over  
 676 repeats of the same experiment where a stimuli is presented to the retinal network. This  
 677 corresponds therefore to averaging over initial conditions in the presence of noise.

We consider therefore a noisy version of (13), first in the non rectified case. For  
 simplicity we set  $\vec{C}^{(0)} = \vec{0}$ . Dynamics of cells voltage is now given by:

$$\frac{d\vec{\mathcal{X}}}{dt} = \mathcal{L}^{(0)}.\vec{\mathcal{X}} + \vec{\mathcal{F}}(t) + \sigma_B\vec{\zeta}(t), \quad (29)$$

678 where  $\vec{\zeta}(t)$  is a white noise (we wrote the stochastic dynamics the physicists way). The  
 679 noise intensity  $\sigma_B$  is here a constant. We could consider more elaborated versions where  
 680 noise depends on neurons indices, but this would unnecessarily make the analysis more  
 681 complex. Note however that we add noise to GCells too, but this was already implicitly  
 682 done in the LNP formulation (8). As a consequence, these two noise will somehow add  
 683 up, as made clear in eq. (30). This is harmless however to our analysis on correlations.  
 684 As in the previous section, we consider that the stimulus is applied from an initial time  
 685 far in the past, so that we integrate the trajectory from  $-\infty$  to  $t$ .

The solution of (29) is a Gaussian process with probability law  $\mathbb{P}$ , mean:

$$\vec{m}^{(0)}(t) = \int_{-\infty}^t e^{\mathcal{L}^{(0)}(t-s)}.\vec{\mathcal{F}}(s) ds \equiv \left[ e^{\mathcal{L}^{(0)} \cdot t} * \vec{\mathcal{F}} \right](t), \quad (30)$$

and correlation matrix:

$$\mathcal{C}(t_1, t_2) = \sigma_B^2 e^{\mathcal{L}^{(0)}(t_2-t_1)} \cdot \int_{-\infty}^{t_1} e^{\tilde{\mathcal{L}}^{(0)}(t_1-s)} \cdot e^{\mathcal{L}^{(0)}(t_1-s)} ds, \quad (31)$$

686 for  $t_2 \geq t_1$ , where  $\tilde{\mathcal{L}}^{(0)}$  is the transpose of  $\mathcal{L}^{(0)}$ . The integral is a function of  $t_1$  and  $\mathcal{L}^{(0)}$ .  
 687 From these equations we see that, in general, the probability law of  $\vec{X}$  depends on the  
 688 spectrum of  $\mathcal{L}^{(0)}$ . Let us make this dependence more explicit.

689

In the general case  $\mathcal{L}^{(0)}$  is not symmetric and does not commute with  $\tilde{\mathcal{L}}^{(0)}$ . One can then compute  $\mathcal{C}(t_1, t_2)$  in terms of the (common) spectrum of  $\mathcal{L}^{(0)}$ ,  $\tilde{\mathcal{L}}^{(0)}$  using the spectral decomposition theorem  $\mathcal{L}^{(0)} = \sum_{\alpha=1}^N \lambda_{\alpha}^{(0)} v_{\alpha}^{(0)} \cdot \tilde{w}_{\alpha}^{(0)}$  where  $v_{\alpha}^{(0)}$  is the right eigenvector  $\alpha$  of  $\mathcal{L}^{(0)}$  (the  $\alpha$ -th column of  $\mathcal{P}^{(0)}$ ) and  $\tilde{w}_{\alpha}^{(0)}$  is the left eigenvector  $\alpha$  of  $\mathcal{L}^{(0)}$  (the  $\alpha$ -th row of  $(\mathcal{P}^{(0)})^{-1}$ ). In general, right (left) eigenvectors are not mutually orthogonal but  $\tilde{w}_{\alpha}^{(0)} \cdot v_{\beta}^{(0)} = \delta_{\alpha\beta}$  so that  $v_{\alpha}^{(0)} \cdot \tilde{w}_{\alpha}^{(0)}$  is the projector on eigendirection  $\alpha$ . From this, one obtains the correlation matrix:

$$\mathcal{C}(t_1, t_2) = -\sigma_B^2 \sum_{\alpha=1}^N e^{\lambda_{\alpha}^{(0)}(t_2-t_1)} v_{\alpha}^{(0)} \cdot \tilde{w}_{\alpha}^{(0)} \sum_{\beta=1}^N \frac{v_{\beta}^{(0)} \cdot \tilde{w}_{\beta}^{(0)}}{\lambda_{\alpha}^{(0)} + \lambda_{\beta}^{(0)}}, \quad (32)$$

690 where eigenvalues are real or complex conjugate and are assumed to be stable (negative  
 691 real part). Note that eigenvalues and projectors combine so that, in fine, the correlation  
 692 matrix is real. It is stationary and *stimulus independent*.

693

We will keep this general form for further discussions on the rectified case, but here, it is insightful to consider the case where  $\mathcal{L}^{(0)}$  is symmetric. Here, it is diagonalizable in a orthogonal basis, with  $(\mathcal{P}^{(0)})^{-1} = \tilde{\mathcal{P}}^{(0)}$  and with real eigenvalues  $\lambda_{\beta} \equiv -s_{\beta}$ ,  $\beta = 1 \dots N$ , where  $s_{\beta}$  is real, positive. Then, (32) reduces, in form of components, to:

$$C_{\alpha_2, \alpha_1}(t_2 - t_1) = \frac{\sigma_B^2}{2} \sum_{\beta=1}^N \frac{P_{\alpha_2\beta} P_{\alpha_1\beta}}{s_{\beta}} e^{-s_{\beta}(t_2-t_1)}. \quad (33)$$

This is the noise-induced correlation between cell  $\alpha_2$  at time  $t_2$  and cell  $\alpha_1$  at time  $t_1$ . It is useful to express, from (33), the variance of cell  $\alpha_1$ 's voltage (independent of time due to stationarity):

$$\sigma_{\alpha_1}^2 = \frac{\sigma_B^2}{2} \sum_{\beta=1}^N \frac{P_{\alpha_2\beta} P_{\alpha_1\beta}}{s_{\beta}}. \quad (34)$$

694

These computations provide the so-called "noise correlations" between cells voltage.  
 695 Note that this concerns *all* cell types (not only GCells). The wording "noise correlations"  
 696 is actually a bit misleading. Indeed, a simple glance at (33) (or even (31)), shows that  
 697 cells correlations are essentially constrained by the network structure, so that, a change  
 698 of parameters, e.g. plasticity or pharmacology, induces dramatic variations in this corre-  
 699 lation.

700

We now compute spike correlations of GCells. We assume a spiking probability of the form (8). The probability that GCell  $\alpha_1 (> N_B + N_A)$  spikes at time  $t_1$  is induced

by the voltage probability  $\mathbb{P}$  and is given by  $v_{\alpha_1}(t_1) \equiv \mathbb{E}\left[f\left(\frac{V_G(t)-\theta_G}{\sigma_G}\right)\right]$  where the expectation is taken with respect to  $\mathbb{P}$ . Taking the form (8) for  $f$  this is:

$$v_{\alpha_1}(t_1) = f\left(\frac{m_{\alpha_1}(t_1) - \theta_G}{\sqrt{\sigma_G^2 + \sigma_{\alpha_1}^2}}\right). \quad (35)$$

701 As pointed out above, two sources of noise add up here: the implicit noise, with variance  
 702  $\sigma_G^2$  appearing in the LNP formulation (8), which is intrinsic to the cell, and the network  
 703 induced noise, explicit in the term  $\sigma_{\alpha_1}^2$ .

704

Likewise, the probability that GCell  $\alpha_1 (> N_B + N_A)$  spikes at time  $t_1$  and GCell  $\alpha_2 (> N_B + N_A)$  spikes at time  $t_2$  is:

$$v_{\alpha_1\alpha_2}(t_1, t_2) = \int f\left(\frac{\sqrt{\mu_1} \cos(\phi) y_1 - \sqrt{\mu_2} \sin(\phi) y_2 + m_{\alpha_1}(t_1) - \theta_G}{\sigma_G}\right) f\left(\frac{\sqrt{\mu_1} \sin(\phi) y_1 + \sqrt{\mu_2} \cos(\phi) y_2 + m_{\alpha_2}(t_2) - \theta_G}{\sigma_G}\right) DY, \quad (36)$$

705 where the integral holds on  $\mathbb{R}^2$  and where  $DY = \frac{1}{2\pi} e^{-\frac{y_1^2 + y_2^2}{2}} dy_1 dy_2$ . Here,  $\mu_1, \mu_2$  are the  
 706 eigenvalues of the pairwise correlation matrix  $\mathcal{C} = \begin{pmatrix} \sigma_{\alpha_1}^2 & \mathcal{C}_{\alpha_1\alpha_2}(t_1 - t_2) \\ \mathcal{C}_{\alpha_2\alpha_1}(t_2 - t_1) & \sigma_{\alpha_2}^2 \end{pmatrix}$   
 707 which is diagonalizable in an orthogonal basis with an orthogonal transformation, a  
 708 rotation with angle  $\phi$  determined by the coefficients of  $\mathcal{C}$ .

## 709 2.2.2. Consequences

710 **Decorrelation.** It is evident that the double integral (36) factorizes only in the case where  
 711  $\mathcal{C}$  is diagonal ( $\phi = 0, \mu_1 = \sigma_{\alpha_1}, \mu_2 = \sigma_{\alpha_2}$ ), and it reduces to  $v_{\alpha_1\alpha_2}(t_1, t_2) = v_{\alpha_1}(t_1)v_{\alpha_2}(t_2)$ .  
 712 Thus, spikes of GCell  $\alpha_1$  at time  $t_1$  and of GCell  $\alpha_2$  at time  $t_2$  are decorrelated if and only  
 713 if the correlation matrix (33) is diagonal. In particular, as we saw, simple networks with  
 714 ACells have, in general, the effect to correlate voltages and thereby spikes. This is in  
 715 contrast with the claim, found in deep experimental papers stating that "the inhibition"  
 716 (mediated by ACells) "decorrelates visual feature representations in the inner retina [15]".  
 717 What could be the origin of this discrepancy ?

718 First, even if the correlations we compute are non vanishing they can nevertheless  
 719 be weak. The weakness of pairwise correlations in the retina has actually been reported  
 720 by many authors [61,62]. It is known since Lancaster, 1957 [75] that the passage of  
 721 two correlated Gaussian variables through a subsequent non linearity always reduces  
 722 the correlation of the two signals, regardless of the shape of the nonlinearity. Thus,  
 723 in our case, the non linear function of the LNP model reduces the decorrelation. This  
 724 fact was reported by Pitkow and Meister in their paper "Decorrelation and efficient  
 725 coding by retinal ganglion cells" [12] where they say: "The classical theory of retinal  
 726 decorrelation attributed that phenomenon to filtering by center-surround receptive fields  
 727 and explained its purpose as serving the efficient transmission of visual information  
 728 through the optic nerve". Although their experiments conclude "that the spike trains  
 729 of retinal ganglion cells were indeed decorrelated compared with the visual input"  
 730 they insist on the prominent role of nonlinearities: "most of the decorrelation was  
 731 accomplished not by the receptive fields, but by nonlinear processing in the retina"  
 732 and "in the LN model, the nonlinearity decorrelates if it has a high threshold, ensuring  
 733 that each neuron spends much of the time silent except for sharp and sparse firing  
 734 events." From these remarks they conclude about information transmission by the retinal  
 735 network: "At very high thresholds, the information transmission is poor. Notably,  
 736 transmission also drops at low thresholds. Thus, the choice of threshold involves a  
 737 trade-off between rarely using reliable symbols, such as high spike counts, or frequently  
 738 using unreliable symbols, such as low spike counts".

739 Thus, non linearities play a role in retinal coding making the spike rate of RGCs  
 740 as sparse as possible, so that these cells are silent most of the time and fire at high rate

741 only when salient features of the stimulus make it necessary. This effect should be even  
 742 more prominent for moving objects which is clearly an example of a stimulus with salient  
 743 features and strong spatio-temporal correlations induced by its trajectory, especially if  
 744 this trajectory shows sharp changes.

745

746 **Rectification.** Now, the LNP non linearity is not the only source of decorrelation. Rec-  
 747 tification also plays a crucial role. What happens, indeed, in the rectified case? Math-  
 748 ematically, one can extend equations (17) providing the general solution of (29) in the  
 749 presence of noise, but the main, quite intricate problem is now that the entrance and  
 750 exit time of domains  $t_{-}^{(n_k)}, t_{+}^{(n_k)}$  are *themselves* random. This is again a consequence of  
 751 the stimulus dependence of these times. The computation of the voltage correlations in  
 752 this case being, for the moment, out of reach, I am going to give some straightforward  
 753 although insightful remarks.

754

The non rectified case correspond to a trajectory staying in the domain  $\Omega^{(0)}$  (forget-  
 755 ting about conditions on noise ensuring that this holds for an infinite time). Now, the  
 756 computation of voltage correlation is essentially the same if the trajectory stays in the  
 757 domain  $\Omega^{(n)}$ . The only difference is that eigenvalues and projectors have a superscript  
 758 ( $n$ ) instead of ( $0$ ). This difference is essential though, because rectification induces a  
 759 projection on the space of non rectified cells. The contribution to rectified cells to voltage  
 760 correlations with other cells vanishes thereby transforming the voltage correlations  
 761 matrix. By permutation of rows and columns, one can convert this matrix in a form  
 762 containing a diagonal block (correlations rectified cells  $\leftrightarrow$  rectified cells) and a block  
 763 characterizing the correlations non rectified cells  $\leftrightarrow$  all cells. This reduces the model  
 764 dimensionality and the global correlations. This effect, composed with the LN non  
 765 linearity can even reduce correlation.

766

Now, the most important remark here is that rectification implies that GCells corre-  
 767 lations are *stimulus dependent*. This effect might not be noticeable with full field stimuli  
 768 or white noise, which weakly solicit the lateral ACell connectivity, but it could be more  
 769 prominent when studying spatio-temporal stimuli, in particular moving trajectories.  
 770 From this perspective, the standard decorrelation argument, where centre and surround  
 771 components of Bcells receptive fields interact via ACells lateral inhibition to decorrelate  
 772 bipolar cell output in the spatial and temporal domains, may not apply for non stationary  
 773 stimuli, which constitute most of real visual scenes.

### 774 2.2.3. Conclusion of section 2.2

775

In this section, we have mathematically investigated the structure of correlations  
 776 induced by the model (13), Fig. 1 right. Our conclusion is essentially that the stimulus  
 777 generates GCells spike correlations modulated by the BCells-ACells networks, and  
 778 more precisely, by the eigenvalues-eigenmodes of the transport operator. While, for  
 779 non rectified cells, the spike correlations are independent of the stimulus, introducing  
 780 rectification dramatically changes the picture. As well as the RF, correlations are stimulus  
 781 dependent. However, non linearities (LNP transfer function and rectification) have the  
 782 general effect of reducing pairwise correlations, confirming a result emphasized by  
 783 Pitkow and Meister [12]. Let us also remark that rectification makes the stochastic  
 784 process of voltages non Gaussian, because the times of entering and exiting domains  
 785 are now random variables too. As a consequence, spike statistics involves higher order  
 786 correlations. Although it has to be further investigated on experimental grounds, this  
 787 would lead to important consequences in terms of coding. As pointed out, again, by  
 788 Pitkow and Meister [12], "for highly non-Gaussian signals, such as neural spike trains  
 789 and natural images, correlation may be only weakly related to redundancy."

790

Sticking at the model we may ask the following questions. Assume that we submit  
 791 the model to different type of stimuli: the "classical" ones such as white noise, "Chirp"  
 792 stimulus, natural images; but also more elaborated ones such as moving objects with  
 793 different type of trajectories, or "natural movies" including motion and "surprise". For

794 example, a bird crossing the visual scene, with, on the background, a forest of trees in the  
795 wind. It is known that the retina is able to filter the "noisy" motion of tree leaves while  
796 signalling the bird, thanks to dedicated circuits involving ACells [1,44]. Such circuits  
797 can be easily implemented in the model (13) [28]. What will be the structure of its spike  
798 trains, depending on the different type of stimuli ? How can one "efficiently" decode the  
799 stimulus from the mere knowledge of those spike trains ? How efficient is a decoding  
800 scheme based on independent, decorrelated Gcells ? In contrast, would cooperative  
801 network effects make the code more precise affording faster responses to motion [14] ?

802 Although we are not going to answer these questions here (there is still a long  
803 way to it), we give, in the next sections, several insightful mathematical results in this  
804 direction.

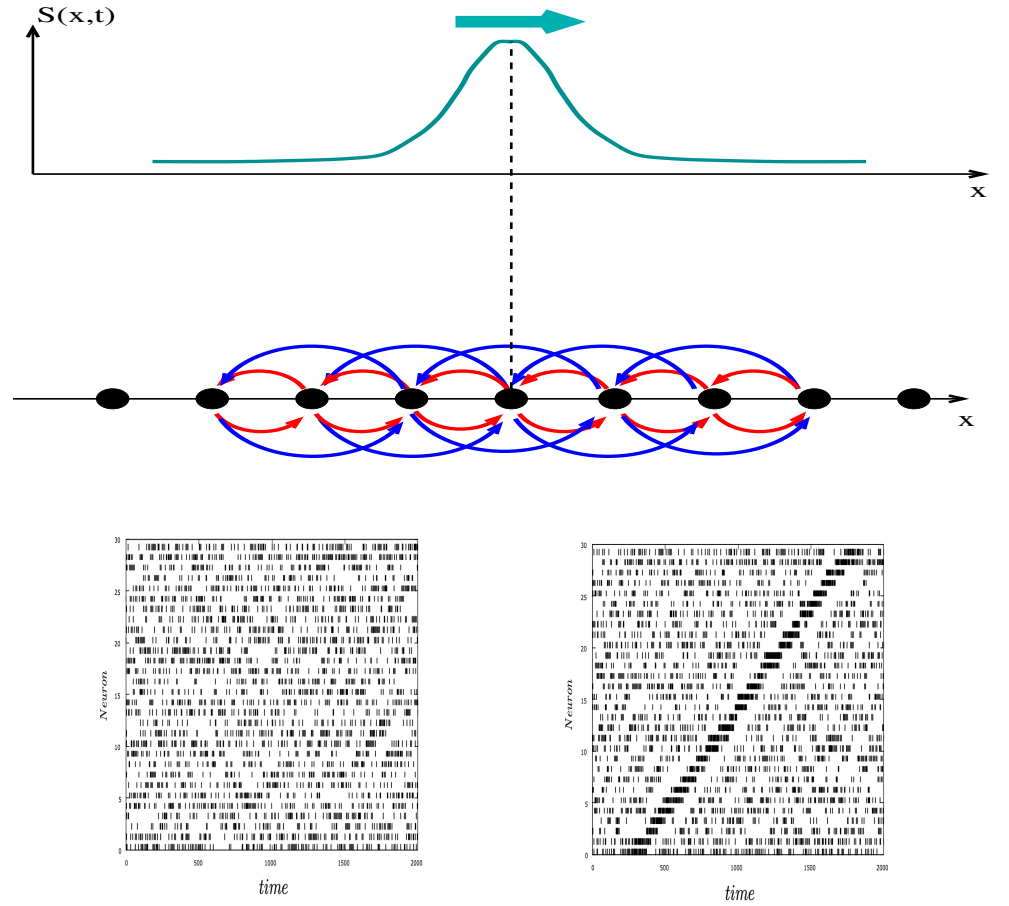
### 805 2.3. Computing the mixed effect of network and stimulus on spike correlations

806 Let us now consider the retina from the point of view of its output. We sit on the  
807 optic nerve and measure the spikes sent to the LGN and cortex via the optic nerve. We  
808 have no access to the biophysical machinery taking place in the retina and generating  
809 those spikes, but we know that the spike trains contains information about the external  
810 world stimuli that we want to extract. We can measure as many quantities as we want  
811 such as firing rate, or higher correlations. More generally, we are seeking the (time  
812 dependent) joint probability of spikes adopting the approach described in section 1.2,  
813 Methods.

814 In this context, assume a retina "at rest" i.e. receiving no stimulus or stationary  
815 stimuli like noise. We can describe the spike trains emitted by this retina by a stationary  
816 transition probability  $\mathbb{P}$ , associated with a stationary probability  $\mu^{(sp)}$  (for "spontaneous").  
817 In general, this probability has spike correlations of order 2 and higher. Assume now  
818 that, from time  $t_0$  on a stimulus (say a moving object) is getting through the visual field  
819 of this retina. As exposed in section 2.2 one expects the spike correlations (at any order)  
820 to be modified by this stimulation. Typically, a moving object carries spatio-temporal  
821 correlations in its trajectory which will superimposed upon the network correlations,  
822 resulting in a mixed effect where non linearities can also play a role. Can we predict, for  
823 a given stimulus, how correlations will be modified ?

824 Let us give an example. Consider a linear chain of neurons, as depicted in Fig. 3 top.  
825 Each neuron (black points), is connected to its neighbours with an excitatory connection  
826 (red arrows) and to its second nearest neighbours with an inhibitory connection (blue  
827 arrows). In spontaneous activity, we assume an asynchronous dynamics, such as Fig.  
828 3 bottom, left. In addition, each neuron is able to sense external stimuli. Consider a  
829 moving stimulus propagating from left to right (cyan, bell shaped curve) Fig. 3 top. This  
830 stimulus is going to modify the spike patterns, as seen in Fig. 3 bottom, left, where one  
831 sees clearly the effect of the stimulus, nearest neighbours excitation and second nearest  
832 neighbours inhibition. The remarkable fact is that the stimulus not only modifies the  
833 firing rates of neurons, but also *their correlations*. The question is: can we compute this  
834 effect ?





**Figure 3. Top. Network of spiking neurons sensing a stimulus.** Each neuron, represented as a black point, is connected to its neighbours with an excitatory connection (red arrows) and to its second nearest neighbours with a inhibitory connection (blue arrows). In addition, each neuron is able to sense external stimuli  $S(x, t)$  (cyan, bell shaped curve). **Bottom. Left.** Spontaneous spiking activity. **Bottom. Right.** Spiking activity in the presence of the moving stimulus.

835 This question has been solved in the paper [39] for the gIF model (23). Here, we  
 836 briefly state the main result (see the paper for technical details). Consider a function  
 837  $f(t, \omega)$  (observable) depending on time and spike history up to time  $t$ . Let  $\mu^{(sp)}$  be  
 838 the joint probability distribution of spikes in spontaneous activity (no stimulus), and  $\mu$   
 839 the joint probability distribution of spikes in the presence of a spatio-temporal stimulus  
 840  $S(x, t)$ . We note  $\delta\mu[f](t) = \mu[f](t) - \mu^{(sp)}[f]$ , where  $\mu[f](t)$  is the average of  $f$ , at  
 841 time  $t$ , in the presence of the stimulus and  $\mu^{(sp)}[f]$  the average of  $f$  in spontaneous  
 842 activity (which does not depend of time because spontaneous dynamics is stationary).  
 843  $\delta\mu[f(t)]$  characterizes how much the time dependent mean of  $f(t, \omega)$  under stimulation  
 844 departs from the spontaneous mean at time  $t$ . In the simplest case  $\delta\mu[f(t)]$  characterizes  
 845 the variation in the firing rate of neuron  $k$ , if  $f(t, \omega) = \omega_k(t)$ , or the variation in the  
 846 correlation between neuron  $k_1$  at time  $t_1$  and neuron  $k_2$  at time  $t_1 + t$  if  $f(t, \omega) =$   
 847  $(\omega_{k_1}(t_1) - \mu^{(sp)}[\omega_{k_1}]) (\omega_{k_2}(t_1 + t) - \mu^{(sp)}[\omega_{k_2}])$ , and so on.

One can show that, when the stimulus amplitude is weak enough,  $\delta\mu[f(t)]$  is given by a linear response formula of the form:

$$\delta\mu[f(t)] = [\kappa_f * S](t) \quad (37)$$

848 That is, by the convolution of the stimulus with a specific kernel,  $K_f$ , depending on the  
849 observable  $f$  and on the spontaneous distribution  $\mu^{(sp)}$ . We do not give the expression of this  
850 kernel here, for simplicity, but the reader can refer to the paper [39].

### 851 2.3.1. Consequences

852 Let us comment this result.

853  
854 **Convolution.** Similarly to (1) (GCells response to stimuli) or (2) (BCells response  
855 to stimuli), we have here again a linear response where the effect of a stimulus on a  
856 system is expressed by a convolution. We are however in a completely different per-  
857 spective. Indeed, while we were considering formerly voltage response of individuals  
858 cells (shaped by network effects), we are now working on a more abstract level, where  
859 we attempt to measure the effect of a stimulus on *statistics*. This is of course due to the  
860 difference in what is accessible by experiments, what the observer is able to deal with  
861 in his observations, here spikes. Thus, the mathematical machinery allowing to extract  
862 the response requires to define spike statistics in a non stationary setting, where the  
863 influence of the stimulus can be inferred.

864  
865 **Kernel.** The kernel  $K_f$  can be explicitly computed in the gIF model. It depends on  
866 several features. First, on network characteristics (especially the effective interaction  
867  $W_{kj}$ , and more generally, the parameters shaping the model dynamics). It also depends  
868 on the observable  $f$ . However, the main content of this result is that the kernel  $K_f$  is  
869 actually determined by spike correlations in *spontaneous activity*. In other word, it is  
870 possible to anticipate the response to a non stationary stimulus from the knowledge  
871 of the spontaneous activity. Although this result is expected from Kubo theory in non  
872 equilibrium statistical physics [76,77] or from Volterra-Wiener expansions [20], it has  
873 interesting consequences when dealing with neural dynamics, and more specifically here,  
874 with retina outputs. First, it provides a consistent treatment of the expected perturbation  
875 of higher-order correlations, beyond the known linear perturbation of firing rates and  
876 instantaneous pairwise correlations; in particular, it extends to time-dependent corre-  
877 lations. In addition it reveals how the stimulus-response and dynamics are entangled  
878 in a complex manner. For example, the response of a neuron  $k$  to a stimulus applied  
879 on neuron  $i$  does not only depends on the synaptic weight  $W_{ki}$  but, in general, on all  
880 synaptic weights, because the dynamics create complex causality loops which build up  
881 the response of neuron  $k$  [36,78,79]. The linear response formula is written in terms of the  
882 parameters of a spiking neuronal network model and the spike history of the network.  
883 Although a linear treatment may seem a strong simplification, our results suggest that al-  
884 ready, in this case, the connectivity architecture should not be neglected. In the presence  
885 of stimuli, the whole architecture of synaptic connectivity, history and the dynamical  
886 properties of the networks are playing a role in the spatio-temporal correlations structure.

887  
888 **Linear response and higher order corrections.** The derived formula provides  
889 a good agreement with simulations in the gIF model under time dependent stimuli  
890 (typically, a moving object). It requires however that the stimulus amplitude is weak  
891 enough. That is, higher corrections are weaker than the leading order. For larger  
892 amplitude stimuli one would to compute higher order correlations. This can be done  
893 using the same formalism [80] although it might not be the best approach. Indeed,  
894 this method requires to measure spontaneous correlations which are difficult to obtain  
895 experimentally for orders higher than 2. This is actually one of the reason why LNP-like  
896 models exist. The expected non linearity in the response is handled by a static non linear  
897 function. Exploring what could be the best non linear correction to the linear response  
898 in such models is definitely an interesting mathematical challenge.

## 899 2.4. Conclusion

900 **Beyond naive RF description.** This linear response theory actually shows how the  
901 neuronal network substrate and stimulus response are entangled. Indeed, in contrast to  
902 naive RF representation where the convolution kernel is assumed to depend only on the  
903 *cell*, here, mathematics show that it depends as well on the *observable*. The explicit form of  
904 the kernel is also tightly constrained by the neurons connections. Finally, a convolution  
905 implies an integration over histories, requiring thereby to consider spikes probabilities  
906 with memory, instead of "instantaneous" spikes probabilities (not or weakly depending  
907 on the past). Of course, one may always argue that, on experimental grounds, long tail  
908 memory are just impossible to measure so "instantaneous" [61] or first order Markov  
909 [81] models are largely sufficient. But what means "sufficient" ? This is a difficult ques-  
910 tion which requires sophisticated methods to determine the "best performing" memory  
911 depth from data [82–84]. Actually, numerical computations of the kernel use, of course  
912 Markovian approximations [39], although, with a memory depth that can be controlled.

913  
914 **Link with the retina model.** Can we relate the formalism developed here with  
915 the retinal model presented in the section 1.1 ? As GCells voltage are Gaussian it is  
916 in principle possible to compute transition probabilities using the transport operator  
917 formalism. However, even in the non rectified case, the computation promises to be  
918 a formidable task, unless one adds some additional constraints. For example, a big  
919 advantage of Integrate and Fire models is that a spiking neuron loses memory after  
920 spiking, a property which is not implemented in LNP like models.

921  
922 **Information geometry.** There is a close link between Gibbs distributions and in-  
923 formation geometry. This theory, developed by Shun'ichi Amari and his collaborators  
924 (see [85] and references therein) on the basis of early work from Rao [86] establishes a  
925 geometric theory of information where probabilities are considered as points on Riemannian  
926 manifolds. A prominent family of probability measures is called the exponential  
927 family. It contains the Gibbs distributions in the standard statistical physics sense, i.e.  
928 probabilities having the form  $\frac{e^{-\beta H}}{Z}$  where the energy  $H$  does not depend on time. In this  
929 case, the metric is given by the Hessian of  $\log(Z)$ , the free energy, and is tightly linked  
930 to Fisher information on one hand and to linear response on the other hand. The linear  
931 response is actually a correlation function, from the fluctuation dissipation theorem.  
932 Thus, correlation functions induce a natural geometry for Gibbs distributions providing  
933 strong insights on how these distribution are modified by smooth, local, transformations  
934 of their parameters (like learning [87]) or under a stimulation of weak amplitude. In this  
935 last case, the stimulus action corresponds to a perturbation in the tangent space of the  
936 manifold [88,89]. Although information geometry has not been extended, to our best  
937 knowledge, to the type of Gibbs distribution we study here (they are non stationary)  
938 the mathematical formalism is similar when considering finite memory and stationary  
939 distributions. This essentially tells us that the structure of spatio-temporal correlations  
940 observed in spike trains reveals an hidden geometrical structure which, somewhat,  
941 shapes the response of the retina, and, henceforth of cortex, to stimuli. We come back to  
942 this point in the conclusion section.

## 943 3. Applications

944 The OPL-BCells-ACells processing is based on graded potentials departing from  
945 the classical paradigm of binary spike processing. Mathematically, this has strong conse-  
946 quences in terms of response to a spatio-temporal stimulus: existence of eigenmodes,  
947 potentially modulated by non linear effects, inducing properties such as activity waves  
948 ahead of the stimulus (anticipation), resonances, correlations modified by the stimulus.  
949 In this section, and although this paper is essentially theoretical I would like to shortly  
950 propose possible applications of these results, outside the field of neuroscience.

### 951 3.1. Retinal prostheses

952 Retinal pathologies such as Age Macular Degeneration or Retinitis Pigmentosa, are  
953 due to the degeneration of photo-receptors [90]. In addition, they induce morphological  
954 and structural changes in the retina with significant pathogenic effects: inflammation,  
955 change in connectivity, the appearance of large-scale spontaneous electrical oscillations,  
956 and, of course, attenuation of response to visual stimuli [91–94]. In this process of  
957 degeneration, however, the Gcells are the last to be deficient, maintaining, therefore, a  
958 link between the retina and the brain, provided they are suitably stimulated. The strategy  
959 of retinal prostheses is to stimulate the retina electrically by an array of electrodes.  
960 Stimulation of an electrode generates, in the visual cortex, a phosphene, the perception of  
961 a light spot. By stimulating the electrodes one induces in the cortex an image "pixelised"  
962 by the phosphenes, with resolution limited, on the one hand, by the number of electrodes,  
963 and, on the other hand, by the size of the phosphenes, which can be enlarged by diffusion  
964 and non linear effects [95]. Technological solutions, taking into account the physiological  
965 limitation on the electrical power that can be injected in an electrode, improve resolution  
966 [96]. However, there are still obstacles which cannot be resolved by purely technological  
967 solutions (hardware). In addition, a valid stimulation strategy at a given period of the  
968 pathology may not be later because the retina degeneration evolves in time.

969 Stimulation strategies use processor pre-processing to calculate, from a given image  
970 (captured by a camera) the pattern of stimulation of the prosthesis, by mimicking the  
971 calculation that a healthy retina would make, or by incorporating corrections taking  
972 into account the pathology [97]. These algorithms might be improved using what  
973 we know about the retinal structure, especially ACells lateral connectivity, where a  
974 model like (13) can be easily implemented with a relatively low energy consumption  
975 cost. The idea would be to improve electrodes stimulation sequences so as to allow an  
976 implanted patient to perceive in real time a moving object. The model (13) with ACells  
977 lateral connectivity and gain control is known to produce a wave of activity ahead of a  
978 stimulus, performing a form of anticipation [28]. This could be used to compensate for  
979 the processing times imposed by the equipment, in the same way that the visual system  
980 knows how to compensate for the delays induced by photo-transduction [25]. The ideal  
981 would also be to have adaptive algorithms, i.e., depending on parameters adjustable  
982 according to the patient and the course of his pathology.

### 983 3.2. Convolutional networks

984 Several recent studies attempt to understand how retinal response to stimuli is  
985 related to circuit processes using convolutional neural network models [98] to under-  
986 stand the structure of retinal prediction [99]. Reciprocally, these networks can be used to  
987 encode deep-learning models to encode dynamic visual scenes with important potential  
988 outcomes in the domain of computer vision. Especially, a recent work by Zheng et  
989 al, [100], shows the important role played by recurrence in encoding complex natural  
990 scenes. To my best knowledge (which is quite scarce in this field) there is no mathemati-  
991 cal analysis of the dynamics of these models, especially the dependence in parameters  
992 and robustness of the training schemes. The present study could bring some insights in  
993 this perspective. Even if the model (13) is different from what these researchers were  
994 using, the techniques of piecewise linear phase decomposition and eigenmodes study  
995 could be insightful to better understand the dynamical evolution of these convolutional  
996 networks and the role played by rectification.

## 997 4. Discussion

998 In this paper we have addressed mathematically the potential effect of ACells  
999 lateral connectivity on retinal response to spatio-temporal stimuli. We have seen how,  
1000 mathematically, the retina structure and the collective dynamics of retinal cells organized  
1001 in local circuits spanning the whole retina might constrain this response. In particular,

1002 the structure of correlations is expected to depend on the stimulus, as soon as non linear  
1003 effects are involved.

1004 These properties are established on the basis of theoretical results though, based on  
1005 incomplete modelling of the retina, and specific assumptions. Their validation would  
1006 require experiments, some of which may require time and others are not yet accessible,  
1007 for example, simultaneously measuring retina and cortex. As a matter of fact, one may  
1008 argue that the models presented here are far too simplistic compared to the real retina(s)  
1009 having a large number of BCells, ACells, GCells type, making complex circuits [18]  
1010 and whose characteristics depend, in addition, on species, age, pathologies. However,  
1011 the idea behind mathematical modelling is precisely to try and infer some generic  
1012 mechanism underlying the real object under study, here the retina. This is the simplicity  
1013 of the structure which makes it generic. The question is: "Would the addition of more  
1014 elaborated retinal features make the response to stimuli simpler?"

1015 In the next section I further discuss some further implications of this work leading  
1016 to some new questions.

#### 1017 4.1. Cortical response

1018 If a dynamical stimulus, combined with the retinal network and non linearities  
1019 produces non negligible dynamical spatio-temporal correlations, what could be the  
1020 consequences at the cortical level<sup>3</sup>? There is a physiological transformation, called  
1021 retinotopy, which maps smoothly the retina topology to the cortical V1 topology. In  
1022 models, it is usually considered to be the identity map, although it is not. This is a non  
1023 linear transformation, depending, in addition on the species [101–103]. Nevertheless,  
1024 what matters here is that this mapping is smooth and invertible. Therefore, retinotopy  
1025 transports, in a smooth and invertible way the spatio-temporal retinal correlations to the  
1026 visual cortex. This leads to a first question.

1027  
1028 **How to define a cortical model taking into account spatio-temporal spike corre-**  
1029 **lations.** Cortical models are usually based on mean-field approximations where one  
1030 features firing rates evolution, but not spike correlations. This is the case of the Wilson-  
1031 Cowan model [104–106] or neural field models [107–109]. I know about 2 mean-field  
1032 approaches taking care of spike correlations.

1033 The first approach is the one initiated by S. El Boustani et A. Destexhe [110] using a  
1034 Markovian approach to write down mean field equations of second order (i.e. includ-  
1035 ing pairwise spatial correlations) and a non static thalamic entry, that can feature the  
1036 retinal-LGN input. This model can be used to construct a retino-cortical model [111],  
1037 although the mathematical consequences of having correlated retinal entries have not  
1038 been explored yet.

1039 The second approach is based on the so-called Ott-Antonsen Ansatz [112] and  
1040 has been used by Montbrio, Pazo, Roxin to propose an exact mean field approach  
1041 with second order statistics [113]. Since their paper there has been a lot of activity in  
1042 developing this model, especially in connection with cortical imaging, with impressive  
1043 results [114–117]. It is a promising track.

1044 All these approaches could certainly provide powerful numerical and mathematical  
1045 tools to better understand how spatio-temporal retinal correlations could be processed.  
1046 In particular, having a retino-(LGN)-cortical model allows to do a task which is currently  
1047 impossible experimentally: measure simultaneously retina and cortex.

#### 1048 4.2. Retinal correlations and neurogeometry.

1049 We have also seen that retinal correlations and Gibbs distributions naturally define  
1050 a metric on a Riemannian manifold where probabilities are points on this manifold. In  
1051 particular, the application of a weak amplitude stimulus corresponds to a perturbation

<sup>3</sup> For simplicity, I am going to consider the LGN as a simple relay

1052 along the tangent space of this manifold. What is the image of this metric under the  
 1053 retinotopic transformation ? Let us make this question a bit more precise.

1054 The visual system has evolved to map as efficiently as possible retinal output to  
 1055 cortical structures. The shaping of the visual system during development is actually a  
 1056 highly dynamical process involving retinal waves and synaptic plasticity [70]. These  
 1057 processes provides the visual system a structure allowing to respond in a fast and efficient  
 1058 way to the stimuli coming from the external world, via the retina. In particular, the  
 1059 capacity of the visual cortex to respond to spike trains with spatio-temporal correlations  
 1060 induced by natural stimuli should be somewhat imprinted in the cortical connectivity.

1061 The visual perception is actually highly geometrically structured and shaped by the  
 1062 structure of the cortical connectivity. This leads researchers to introduce a link between  
 1063 the geometry of cortex and the geometry of vision in the concept of neurogeometry  
 1064 (or neuromathematics) where the functional architecture of V1 is considered as a Lie  
 1065 group of symmetry with a Riemannian geometry (see [40–43] and reference therein). In  
 1066 this approach cortical columns are point-like processors detecting visual features where  
 1067 functional connectivity is represented in terms of geodesics. To my best knowledge,  
 1068 neurogeometry essentially deals with V1 and static percepts, although extensions to  
 1069 motor cortex [118] and motion areas [119] have been done. Now, a natural question is:  
 1070 "Is there a relation between the cortical metric of neurogeometry and the metric induced  
 1071 by spatio-temporal spike correlations observed by the retina ?".

1072 Let us address the problem the other way round: Project the cortical metric back  
 1073 to the retina via the inverse retinotopy map, what do we find ? Is there a physiological  
 1074 correspondence with the retina structure and especially lateral connectivity ? What  
 1075 could be the consequences on spike trains statistics and on the way retina processes  
 1076 visual stimuli ? What does cortical metric tell us about retinal spikes correlations ?

1077 Dealing with neural coding of vision, the simplest assumption consists of assuming  
 1078 that GCells are independent encoders and that cortex makes the job of restoring the  
 1079 spatio-temporal correlations existing in the visual scene (e.g. in the trajectory of a moving  
 1080 object). The alternative proposition, where spatio-temporal correlations imprinted in the  
 1081 RGCs spike trains are deciphered by the cortex makes the question of stimuli decoding  
 1082 by the cortex more challenging, but opens up far more possibilities.

1083 **Acknowledgments:** I would like to warmly acknowledge all the neuroscientist collaborators from  
 1084 which I learned about this beautiful object, the retina, and, more generally, about vision: Frédéric  
 1085 Chavane, Gerrit Hilgen, Olivier Marre, Adrian Palacios, Serge Picaud and Evelyne Sernagor.

1086 **Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gollisch, T.; Meister, M. Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron* **2010**, *65*, 150–164. doi:doi:10.1016/j.neuron.2009.12.009.
2. Attneave, F. Some informational aspects of visual perception. *Psychological Review* **1954**, *61*, 183–193.
3. Selfridge, O.G. Pandaemonium: A paradigm for learning. Proceedings of the Symposium on the mechanisation of thought processes; , 1959.
4. Lee, C. Representation of Switching Functions by Binary Decision Programs. *Bell Systems Technical Journal* **1959**, pp. 985–999.
5. Barlow, H. Possible principles underlying the transformation of sensory messages. *Sensory communication* **1961**, pp. 217–234.
6. Atick, J.J.; Redlich, A.N. Towards a Theory of Early Visual Processing. *Neural Computation* **1990**, *2*, 308–320. doi:10.1162/neco.1990.2.3.308.
7. Olshausen, B.A.; Field, D.J. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research* **1997**, *37*, 3311–3325. doi:https://doi.org/10.1016/S0042-6989(97)00169-7.
8. Vinje, W.E.; Gallant, J.L. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science* **2000**, *287*, 1273–1276. doi:10.1126/science.287.5456.1273.
9. Simoncelli, E.; Olshausen, B. Natural IMAGE STATISTICS AND NEURAL REPRESENTATION. *Annual Review of Neuroscience* **2001**, *24*, 1193–1216.
10. Simoncelli, E.P. Vision and the Statistics of the Visual Environment. *Current Opinion in Neurobiology*, 2003, pp. 144–149.
11. Zhaoping, L. Theoretical understanding of the early visual processes by data compression and data selection. *Network: Computation in Neural Systems* **2006**, *17*, 301–334. PMID: 17283516, doi:10.1080/09548980600931995.
12. Pitkow, X.; Meister, M. Decorrelation and efficient coding by retinal ganglion cells. *Nature neuroscience* **2012**, *15*, 628–635.

13. Zhaoping, L., *Understanding Vision: Theory, Models, and Data*; Oxford University Press, 2014; chapter The efficient coding principle.
14. Denève, S.; Machens, C.K. Efficient codes and balanced networks. *Nature neuroscience* **2016**, *19*, 375–382.
15. Franke, K.; Berens, P.; Schubert, T.; Bethge, M.; Euler, T.; Baden, T. Inhibition decorrelates visual feature representations in the inner retina. *Nature* **2017**, *542*, 439–444.
16. Besharse, J.; Bok, D. *The Retina and its Disorders*; Elsevier Science, 2011.
17. Nelson, R.; Kolb, H. ON and OFF pathways in the vertebrate retina and visual system. *The Visual Neurosciences* **2004**, *1*, 260–278.
18. Azeredo da Silveira, R.; Roska, B. Cell Types, Circuits, Computation. *Current Opinion in Neurobiology* **2011**, *21*, 664–671. Networks, circuits and computation, doi:<https://doi.org/10.1016/j.conb.2011.05.007>.
19. Oesch, N.; Diamond, J. Ribbon synapses compute temporal contrast and encode luminance in retinal rod bipolar cells. *Nat Neurosci.* **2011**, *23*, 1555–61.
20. Rieke, F.; Warland, D.; de Ruyter van Steveninck, R.; Bialek, W. *Spikes: Exploring the Neural Code*; Bradford Books, 1997.
21. Chichilnisky, E.J. A simple white noise analysis of neuronal light responses. *Network: Comput. Neural Syst.* **2001**, *12*, 199–213.
22. Simoncelli, E.P.; Paninski, L.; Pillow, J.; Schwartz, O. Characterization of neural responses with stochastic stimuli. In *The Cognitive Neurosciences, III*; Gazzaniga, M., Ed.; MIT Press, 2004; chapter 23, pp. 327–338.
23. Schwartz, O.; Pillow, J.W.; Rust, N.C.; Simoncelli, E.P. Spike-triggered neural characterization. *Journal of Vision* **2006**, *6*, 13–13. doi:10.1167/6.4.13.
24. Motulsky, H.; Christopoulos, A. *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*; Oxford University Press: Oxford, 2004.
25. Berry, M.; Brivanlou, I.; Jordan, T.; Meister, M. Anticipation of moving stimuli by the retina. *Nature* **1999**, *398*, 334–338.
26. Hosoya, T.; Baccus, S.A.; Meister, M. Dynamic predictive coding by the retina. *Nature* **2005**, *436*, 71–77.
27. Chen, E.Y.; Marre, O.; Fisher, C.; Schwartz, G.; Levy, J.; da Silveira, R.A.; Berry, M.J.I. Alert response to motion onset in the retina. *Journal of Neuroscience* **2013**, *33*, 120–132.
28. Souihel, S.; Cessac, B. On the potential role of lateral connectivity in retinal anticipation. *J. Math. Neuro. to appear* **2020**.
29. Cessac, B. A discrete time neural network model with spiking neurons. Rigorous results on the spontaneous dynamics. *J. Math. Biol.* **2008**, *56*, 311–345.
30. Cessac, B.; Viéville, T. On Dynamics of Integrate-and-Fire Neural Networks with Adaptive Conductances. *Frontiers in neuroscience* **2008**, *2*.
31. Vasquez, J.C.; Cessac, B.; Rostro-Gonzalez, H.; Viéville, T. How Gibbs Distributions may naturally arise from synaptic adaptation mechanisms. *BMC Neuroscience* 2009; Central, B., Ed., 2009, Vol. 10, (Suppl 10), p. 213. doi:doi:10.1186/1471-2202-10-S1-P213.
32. Cessac, B. A discrete time neural network model with spiking neurons II. Dynamics with noise. *Journal of Mathematical Biology* **2011**, *62*, 863–900. doi:10.1007/s00285-010-0358-4.
33. Cessac, B. Statistics of spike trains in conductance-based neural networks: Rigorous results. *The Journal of Mathematical Neuroscience* **2011**, *1*, 1–42. doi:10.1186/2190-8567-1-8.
34. Cofré, R.; Cessac, B. Dynamics and spike trains statistics in conductance-based integrate-and-fire neural networks with chemical and electric synapses. *Chaos, Solitons & Fractals* **2013**, *50*, 3.
35. Cessac, B.; Cofré, R. Spike train statistics and Gibbs distributions. *Journal of Physiology-Paris* **2013**, *107*, 360–368. Special issue: Neural Coding and Natural Image Statistics.
36. Cofré, R.; Cessac, B. Exact computation of the maximum-entropy potential of spiking neural-network models. *Phys. Rev. E* **2014**, *89*.
37. Cessac, B.; Cofré, R.; Ampuero, I. Linear response for spiking neuronal networks with unbounded memory. *Dynamics Days*; , 2020.
38. Cofré, R.; Maldonado, C.; Cessac, B. Thermodynamic Formalism in Neuronal Dynamics and Spike Train Statistics. *Entropy* **2020**, *22*. doi:10.3390/e22111330.
39. Cessac, B.; Ampuero, I.; Cofré, R. Linear Response of General Observables in Spiking Neuronal Network Models. *Entropy* **2021**, *23*. doi:10.3390/e23020155.
40. Sarti, A.; Citti, G. The constitution of visual perceptual units in the functional architecture of V1. *Journal of computational neuroscience* **2014**, pp. 1–16.
41. Citti, G.; Sarti, A., Eds. *Neuromathematics of Vision*; Springer, 2014.
42. Petitot, J. *Elements of neurogeometry*; Springer, 2017.
43. Citti, G.; Sarti, A. Neurogeometry of Perception: Isotropic and Anisotropic Aspects. *Axiomathes* **2019**. doi:10.1007/s10516-019-09426-1.
44. Baccus, S.; Meister, M. Fast and Slow Contrast Adaptation in Retinal Circuitry. *Neuron* **2002**, *36*, 909–919.
45. Johnston, J.; Lagnado, L. General features of the retinal connectome determine the computation of motion anticipation. *Elife* **2015**.
46. Blanchard, P.; Cessac, B.; Krueger, T. A dynamical system approach to SOC models of Zhang’s type. *J. Stat. Phys.* **1997**, *88*, 307–318.
47. Blanchard, P.; Cessac, B.; Krueger, T. What can one learn about Self-Organized Criticality from Dynamical System theory ? *Journal of Statistical Physics* **2000**, *98*, 375–404.

48. Coombes, S.; Lai, Y.M.; Şayli, M.; Thul, R. Networks of piecewise linear neural mass models. *European Journal of Applied Mathematics* **2018**, *29*, 869–890. doi:10.1017/S0956792518000050.
49. Rajakumar, A.; Rinzel, J.; Chen, Z.S. Stimulus-Driven and Spontaneous Dynamics in Excitatory-Inhibitory Recurrent Neural Networks for Sequence Representation. *Neural Computation* **2021**, *33*, 2603–2645. doi:10.1162/neco\_a\_01418.
50. Goldman, M.S. Memory without Feedback in a Neural Network. *Neuron* **2009**, *61*, 621–634. doi:https://doi.org/10.1016/j.neuron.2008.12.011.
51. Murphy, B.; Miller, K. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **2009**, *61*, 635–648.
52. Falconer, K.J. *The Geometry of Fractal Sets*; Cambridge University Press: USA, 1985.
53. Barnsley, M.; Rising, H. *Fractals Everywhere*; Elsevier Science, 1993.
54. Falconer, K. *Techniques in Fractal Geometry*, 1997. John Wiley & Sons, Ltd., Chichester.
55. Young, L.S. Mathematical theory of Lyapunov exponents. *Journal of Physics A: Mathematical and Theoretical* **2013**, *46*, 254001. doi:10.1088/1751-8113/46/25/254001.
56. Georgii, H.O. *Gibbs Measures and Phase Transitions*; De Gruyter, 1988.
57. Onicescu, O.; Mihoc, G. Sur les chaînes statistiques. *C. R. Acad. Sci. Paris* **1935**, *200*, 511–512.
58. Galves, A.; Löcherbach, E. Infinite Systems of Interacting Chains with Memory of Variable Length—A Stochastic Model for Biological Neural Nets. *J Stat Phys* **2013**, *151*, 896–921.
59. Fernandez, R.; Maillard, G. Chains with complete connections : General theory, uniqueness, loss of memory and mixing properties. *J. Stat. Phys.* **2005**, *118*, 555–588.
60. LeNy, A. Introduction to (Generalized) Gibbs Measures. *Ensaïos Matematicos* **2008**, *15*, 1–126.
61. Schneidman, E.; Berry, M.; Segev, R.; Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **2006**, *440*, 1007–1012.
62. Shlens, J.; Field, G.; Gauthier, J.; Grivich, M.; Petrusca, D.; Sher, A.; Litke, A.; Chichilnisky, E. The Structure of Multi-Neuron Firing Patterns in Primate Retina. *Journal of Neuroscience* **2006**, *26*, 8254.
63. Nghiem, T.A.; Telenczuk, B.; Marre, O.; Destexhe, A.; Ferrari, U. Maximum-entropy models reveal the excitatory and inhibitory correlation structures in cortical neuronal activity. *Phys. Rev. E* **2018**, *98*, 012402. doi:10.1103/PhysRevE.98.012402.
64. Cocco, S.; Leibler, S.; Monasson, R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *PNAS* **2009**, *106*, 14058–14062. doi:10.1073/pnas.0906705106.
65. Rudolph, M.; Destexhe, A. Analytical Integrate and Fire Neuron models with conductance-based dynamics for event driven simulation strategies. *Neural Computation* **2006**, *18*, 2146–2210.
66. Gerstner, W.; Kistler, W.M. Mathematical formulations of Hebbian learning. *Biological Cybernetics* **2002**, *87*, 404–415.
67. Dayan, P.; Abbott, L. *Theoretical Neuroscience : Computational and Mathematical Modeling of Neural Systems*; MIT Press, 2001.
68. Destexhe, A.; Mainen, Z.F.; Sejnowski, T.J., *Methods in Neuronal Modeling*; The MIT Press, 1998; chapter Kinetic models of synaptic transmission, pp. 1–25.
69. Baden, T.; Berens, P.; Franke, K.; Rosón, M.R.; Bethge, M.; Euler, T. The functional diversity of retinal ganglion cells in the mouse. *Nature* **2016**.
70. Sernagor, E.; Hennig, M. Chapter 49 - Retinal Waves: Underlying Cellular Mechanisms and Theoretical Considerations. In *Cellular Migration and Formation of Neuronal Connections*; Rubenstein, J.L.; Rakic, P., Eds.; Academic Press: Oxford, 2013; pp. 909 – 920. doi:https://doi.org/10.1016/B978-0-12-397266-8.00151-4.
71. Karvouniari, D.; Gil, L.; Marre, O.; Picaud, S.; Cessac, B. A biophysical model explains the oscillatory behaviour of immature starburst amacrine cells. *Scientific Reports* **2019**, *9*, 1859. doi:10.1038/s41598-018-38299-4.
72. Benvenuti, G.; Chemla, S.; Arjan Boonman, G.M.; Chavane, F. Anticipation of an approaching bar by neuronal populations in awake monkey V1. *Journal of Vision* **2015**.
73. Kastner, D.; Ozuysal, Y.; Panagiotakos, G.; Baccus, S. Adaptation of Inhibition Mediates Retinal Sensitization. *Current Biology* **2019**, *29*. doi:10.1016/j.cub.2019.06.081.
74. Hennig, M. Theoretical models of synaptic short term plasticity. *Frontiers in Computational Neuroscience* **2013**, *7*, 154. doi:10.3389/fncom.2013.00154.
75. LANCASTER, H.O. Some properties of the bivariate normal distribution considered in the form of a contingency table. *Biometrika* **1957**, *44*, 289–292. doi:10.1093/biomet/44.1-2.289.
76. Kubo, R. Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems. *Journal of the Physical Society of Japan* **1957**, *12*, 570–586. doi:10.1143/JPSJ.12.570.
77. Kubo, R. The fluctuation-dissipation theorem. *Reports on Progress in Physics* **1966**, *29*, 255–284. doi:10.1088/0034-4885/29/1/306.
78. Cessac, B.; Sepulchre, J. Stable resonances and signal propagation in a chaotic network of coupled units. *Phys. Rev. E* **2004**, *70*.
79. Cessac, B.; Sepulchre, J. Transmitting a signal by amplitude modulation in a chaotic network. *Chaos* **2006**, *16*.
80. Ruelle, D. Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms. *Nonlinearity* **1998**, *11*, 5–18. doi:10.1088/0951-7715/11/1/002.
81. Marre, O.; Boustani, S.E.; Fr'egnac, Y.; Destexhe, A. Prediction of Spatiotemporal Patterns of Neural Activity from Pairwise Correlations. *Phys. rev. Let.* **2009**, *102*, 138101.
82. Nasser, H.; Marre, O.; Il, M.J.B.; Cessac, B. Spatio temporal Gibbs distribution analysis of spike trains using Monte Carlo method. AREADNE 2012 Research in Encoding And Decoding of Neural Ensembles, 2012.



83. Nasser, H.; Cessac, B. Parameters estimation for spatio-temporal maximum entropy distributions: application to neural spike trains. *Entropy* **2014**, *16*, 2244–2277. doi:10.3390/e16042244.
84. Cessac, B.; Kornprobst, P.; Kraria, S.; Nasser, H.; Pamplona, D.; Portelli, G.; Viéville, T. PRANAS: A New Platform for Retinal Analysis and Simulation. *Frontiers in Neuroinformatics* **2017**, *11*, 49. doi:10.3389/fninf.2017.00049.
85. Amari, S.i.; Nagaoka, H. *Methods of Information Geometry*; Vol. 191, Oxford, 2000.
86. Rao, C. Information and Accuracy Attainable in the Estimation of Statistical Parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81–91.
87. Amari, S.i. Information geometry in optimization, machine learning and statistical inference. *Frontiers of Electrical and Electronic Engineering in China* **2010**, *5*, 241–260. doi:10.1007/s11460-010-0101-3.
88. Ruelle, D. Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *J. Statist. Phys.* **1999**, *95*, 393–468.
89. Cessac, B., Recent Trends in Chaotic, Nonlinear and Complex Dynamics, in honour of Prof. Miguel A.F. Sanjuán on his 60th Birthday; World Scientific, 2020; chapter The retina as a dynamical system.
90. Lok, C. Curing blindness: Vision quest. *Nature* **2014**, *513*, 160.
91. Jones, B.; Kondo, M.; Terasaki, H.; Lin, Y.; McCall, M.; Marc, R. Retinal remodeling. *Jpn J Ophthalmol* **2012**, *56*, 289–306.
92. Marc, R.E.; Jones, B.W.; Watt, C.B.; Strettoi, E. Neural remodeling in retinal degeneration. *Progress in retinal and eye research* **2003**, *22*, 607–655.
93. Barrett, J.; Degenaar, P.; Sernagor, E. Blockade of pathological retinal ganglion cell hyperactivity improves optogenetically evoked light responses in rd1 mice. *Frontiers in cellular neuroscience* **2015**, *9*, 330.
94. Barrett, J.; Hilgen, G.; Sernagor, E. Dampening Spontaneous Activity Improves the Light Sensitivity and Spatial Acuity of Optogenetic Retinal Prosthetic Responses. *Sci Rep* **2016**, *6*, 33565.
95. Roux, S.; Matonti, F.; Dupont, F.; Hoffart, L.; Takerkart, S.; Picaud, S.; Pham, P.; Chavane, F. Probing the functional impact of sub-retinal prosthesis. *eLife* **2016**, *5*.
96. Pham, P.; Roux, S.; Matonti, F.; Dupont, F.; Agache, V.; Chavane, F. Post-implantation impedance spectroscopy of subretinal micro-electrode arrays, OCT imaging and numerical simulation: towards a more precise neuroprosthesis monitoring tool. *Journal of Neural Engineering* **2013**, *10*, 046002. doi:10.1088/1741-2560/10/4/046002.
97. Al-Atabany, W.; McGovern, B.; Mehran, K.; Berlinguer-Palmini, R.; Degenaar, P. A Processing Platform for Optoelectronic/Optogenetic Retinal Prosthesis. *IEEE Transactions on Biomedical Engineering* **2013**, *60*, 781–791.
98. Maheswaranathan, N.; McIntosh, L.; Kastner, D.B.; Melander, J.B.; Brezovec, L.; Nayebi, A.; Wang, J.; Ganguli, S.; Baccus, S.A. Deep learning models reveal internal structure and diverse computations in the retina under natural scenes. *bioRxiv* **2018**, p. 340943.
99. Tanaka, H.; Nayebi, A.; Maheswaranathan, N.; McIntosh, L.; Baccus, S.; Ganguli, S. From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction. *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada; Wallach, H.M.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E.B.; Garnett, R., Eds., 2019, pp. 8535–8545.*
100. Zheng, Y.; Jia, S.; Yu, Z.; Liu, J.K.; Huang, T. Unraveling neural coding of dynamic natural visual scenes via convolutional recurrent neural networks. *Patterns* **2021**, *2*, 100350. doi:https://doi.org/10.1016/j.patter.2021.100350.
101. Gias, C.; Hewson-Stoate, N.; Jones, M.; Johnston, D.; Mayhew, J.; Coffey, P. Retinotopy within rat primary visual cortex using optical imaging. *NeuroImage* **2005**, *24*, 200–206. doi:https://doi.org/10.1016/j.neuroimage.2004.08.015.
102. Schira, M.M.; Tyler, C.W.; Spehar, B.; Breakspear, M. Modeling Magnification and Anisotropy in the Primate Foveal Confluence. *PLOS Computational Biology* **2010**, *6*, 1–10. doi:10.1371/journal.pcbi.1000651.
103. Ayzenshtat, I.; Gilad, A.; Zurawel, G.; Slovlin, H. Population Response to Natural Images in the Primary Visual Cortex Encodes Local Stimulus Attributes and Perceptual Processing. *The Journal of Neuroscience* **2012**, *32*, 13971 – 13986.
104. Amari, S. Characteristics of randomly connected threshold element networks and neural systems. *Proc. IEEE* **1971**, *59*, 35–47.
105. Wilson, H.; Cowan, J. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal* **1972**, *12*, 1–24.
106. Wilson, H.; Cowan, J. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics* **1973**, *13*, 55–80.
107. Bressloff, P. Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D: Nonlinear Phenomena* **2001**, *155*, 83–100.
108. Bressloff, P.; Cowan, J.; Golubitsky, M.; Thomas, P.; Wiener, M. Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Phil. Trans. R. Soc. Lond. B* **2001**, *306*, 299–330.
109. Bressloff, P. Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math* **2009**, *70*, 1488–1521.
110. ElBoustani, S.; Destexhe, A. A master equation formalism for macroscopic modeling of asynchronous irregular activity states. *Neural computation* **2009**, *21*, 46–100.
111. Cessac, B.; Souihel, S.; Di Volo, M.; Chavane, F.; Destexhe, A.; Chemla, S.; Marre, O. Anticipation in the retina and the primary visual cortex : towards an integrated retino-cortical model for motion processing. *Workshop on visuo motor integration*, 2019.
112. Ott, E.; Antonsen, T.M. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2008**, *18*, 037113. doi:10.1063/1.2930766.
113. Montbrió, E.; Pazó, D.; Roxin, A. Macroscopic Description for Networks of Spiking Neurons. *Physical Review X* **2015**, *5*, 021028.

114. Bick, C.; Goodfellow, M.; Laing, C.; Martens, E.A. Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions: a review. *J. Math. Neurosc.* **2020**, *10*. doi:10.1186/s13408-020-00086-9.
115. Bi, H.; Segneri, M.; di Volo, M.; Torcini, A. Coexistence of fast and slow gamma oscillations in one population of inhibitory spiking neurons. *Phys. Rev. Research* **2020**, *2*, 013042. doi:10.1103/PhysRevResearch.2.013042.
116. Buendía, V.; Villegas, P.; Burioni, R.; Muñoz, M.A. Hybrid-type synchronization transitions: Where incipient oscillations, scale-free avalanches, and bistability live together. *Phys. Rev. Research* **2021**, *3*, 023224. doi:10.1103/PhysRevResearch.3.023224.
117. Volo, M.D.; Segneri, M.; Goldobin, D.; Politi, A.; Torcini, A. Coherent oscillations in balanced neural networks driven by endogenous fluctuations, 2021.
118. Mazzetti, C. A mathematical model of the motor cortex. PhD thesis, Bologna, 2017.
119. Barbieri, D.; Citti, G.; Cocci, G.; Sarti, A. A Cortical-Inspired Geometry for Contour Perception and Motion Integration. *Journal of Mathematical Imaging and Vision* **2014**, *49*, 511–529. doi:10.1007/s10851-013-0482-z.