How does the inner retinal network shape the ganglion cells receptive field : a computational study

Kartsaki Evgenia, Hilgen Gerrit, Sernagor Evelyne, Bruno Cessac

How does the inner retinal network shape the ganglion cells receptive field : a computational study

Evgenia Kartsaki[1,2], Gerrit Hilgen[2,3], Evelyne Sernagor[2], and Bruno Cessac[1]

[1]Université Côte d'Azur, Inria, Biovision team and Neuromod Institute, Sophia Antipolis, France

[2]Biosciences Institute, Newcastle University, Newcastle upon Tyne NE2 4HH, UK

[3]Health and Life Sciences, Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK

Abstract

We consider a model of basic inner retinal connectivity where bipolar and amacrine cells interconnect, and both cell types project onto ganglion cells, modulating their response output to the brain visual areas. We derive an analytical formula for the spatio-temporal response of retinal ganglion cells to stimuli taking into account the effects of amacrine cells inhibition. This analysis reveals two important functional parameters of the network: (i) the intensity of the interactions between bipolar and amacrine cells, and, (ii) the characteristic time scale of these responses. Both parameters have a profound combined impact on the spatiotemporal features of RGC responses to light. The validity of the model is confirmed by

faithfully reproducing pharmacogenetic experimental results obtained by stimulating excitatory DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) expressed on ganglion cells and amacrine cells subclasses, thereby modifying the inner retinal network activity to visual stimuli in a complex, entangled manner. Our mathematical model allows us to explore and decipher these complex effects in a manner that would not be feasible experimentally and provides novel insights in retinal dynamics.

1 Introduction

Vision is the most fundamental of our senses and visual processing involves some of the most complex neural networks in the vertebrate central nervous system (Marr, 1982; Chalupa and Werner, 2004; Besharse and Bok, 2011; Daw, 2012). The retina is the entry point to our visual system. Located at the back of the eye, this thin neural tissue receives the light that the cornea and lens have captured from different parts of the visual scene, converts it into electrical signals and finally, transmits these signals to the brain visual areas. In particular, light follows a vertical excitatory pathway in the retina, from photoreceptors to bipolar cells (BCs) and onwards to retinal ganglion cells (RGCs), modulated laterally by inhibitory interneurons; horizontal (HCs) and amacrine cells (ACs) (Figure 1, left). RGCs serve as a bridge between the retina and the brain, conveying highly processed and integrated signals from the upstream retinal neurons to downstream visual processing cortical areas. Amazingly, the human brain can recreate images from interpreting parallel streams of information of about one million RGCs, the sole retinal output neurons. This ability is partially due to the astonishing functional, anatomical and molecular diversity across the retinal layers. However, how the different cell classes interact to ultimately encode, via RGCs, a visual scene into spike trains (action potentials) deciphered by the brain remains largely a mystery.

RGCs are indeed embedded in a complex network and their response is roughly driven by two "controllers": 1) The output of BCs that includes both the intrinsic response properties of these cells and the actions of ACs upon them (lateral connectivity); 2) The direct input from ACs via chemical synapses or gap junctions, helping spike synchrony between neighbor RGCs (Masland, 2012; Demb and Singer, 2015). As a consequence, the response of RGCs to visual stimuli does not only depend on local or

physiological characteristics (Sanes and Masland, 2015), but also on the network they are embedded in, and, on the stimulus itself (Cessac, 2022). Previous studies have thus attempted to investigate how the inner retinal neurons are organised into parallel circuits across different cell types and converge onto RGCs (Wässle, 2004; Gollisch and Meister, 2010). This has been studied extensively at the level of bipolar cells, leading to a fairly good understanding of their function (Euler et al., 2014). Other studies have investigated the functional role of amacrine cell (ACs) types in retinal processing (Asari and Meister, 2012; Franke and Baden, 2017; Diamond, 2017; Schröder et al., 2020), suggesting either specific functions such as direction selectivity (starbust ACs) or more general computations, like motion anticipation (Berry et al., 1999; Souihel and Cessac, 2021). Nevertheless, the potential role of the BCs-ACs network on the RGCs response, both from a theoretical and experimental perspective, hasn't been sufficiently explored yet. The scope of the present study is to make one step further in this direction.

On a theoretical ground, we consider a network of BCs connected via ACs, both cell types being connected with chemical synapses to RGCs. Based on a mathematical study, we derive an analytical formula of the RGCs receptive field (RF) that takes into account the lateral ACs connectivity and shows how the response of RGCs to spatio-temporal stimuli is shaped. Especially, we focus on two important parameters: (i), the average intensities of the interactions BCs - ACs and ACs-BCs, and, (ii), the characteristic time scales of these cells response. Varying these parameters acts on the shape of the response to light with potential prominent effects such as a switch from monophasic to biphasic in the temporal RF.

We illustrate our predictions by analysing experimental data obtained from the pharmacological action of excitatory DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) in genetically modified mice. DREADDs are activated by

"designer drugs" such as clozapine-n-oxide (CNO), resulting in an increase in free cytoplasmic calcium and profound increase in excitability in the cells that express these DREADDs (Roth, 2016). In the mouse lines we used, we found DREADD expression both in subsets of RGCs and in ACs in the inner nuclear layer (INL) (Hilgen et al., 2022). In these conditions, CNO would act both on ACs and RGCs providing a way to unravel the entangled effects of: (i) direct excitation of DREADD-expressing RGCs and increase inhibitory input onto RGCs originating from DREADD-expressing ACs; (ii) change in cells response time scale (via a change in the membrane conductance), thereby providing an experimental set up to validate our theoretical predictions.

In the following, we propose a model for the BCs - ACs - RGCs network. The concerted activity of BCs - ACs -RGCs in response to a visual stimulus is described by a large dimensional dynamical system whose mathematical study allows an explicit computation of the RF of BCs, ACs, RGCs and, more generally, to anticipate the effects resulting from light stimulation conjugated with the network activity due to ACs lateral inhibition, in control conditions and in the presence of CNO. Computing the receptive fields of of all cell types (especially RGCs) allows us to to disentangle the concerted effect of ACs lateral connectivity and CNO on the RF of RGCs and provides an excellent agreement with experimental data. We argue, on the basis of the model and analytical computations, that the BCs-ACs network shapes the RF of RGCs via two main parameters, one characterizing the ratio in the time scales of ACs and BCs response, and the other, characterizing the ratio between the synaptic weight BCs $\to$ ACs and ACs $\to$ BCs. This defines a two dimensional map, called the "RFs map". We show how the experimental data settle in this map and we link the observed changes in the shape of the RF, when applying CNO, to a displacement in this map. Finally, we extrapolate our model predictions to situations where the RGCs response could not be

explained by purely local mechanisms but would involve the ACs network.

2 Methods

2.1 The retina model

2.1.1 Structure

We assimilate the retina to a superimposition of 3 layers, each one corresponding to a cell type (BCs, ACs, RGCs), and being a flat, two-dimensional square of edge length L mm where spatial coordinates are noted x, y (Figure 1).

We consider a simplified form of connectivity, inspired from the real connectivity in the retina, illustrated in Figure 1. First, there are as many BCs as ACs and RGCs (N cells per type so that the total number of cells is $3N$). BCs are labelled with an index $i = 1 \dots N$, ACs with an index $j = 1 \dots N$, RGCs with an index $k = 1 \dots N$. BCs are connected to ACs. We note $W_{B_i}^{A_j} \leq 0$ the weight of the inhibitory synapse from AC j to BC i. We use the convention that $W_{B_i}^{A_j} = 0$ when AC j is not connected to BC i. Likewise, we note $W_{A_j}^{B_i} \geq 0$ the weight of the excitatory synapse from BC i to AC j, $W_{G_k}^{B_i} \geq 0$ the weight of the excitatory synapse from BC i to RGC k and $W_{G_k}^{A_j} \leq 0$ the weight of the inhibitory synapse from AC j to RGC k. We note W_B^A the (square) matrix of connections from ACs to BCs and so on.

2.1.2 Visual input

Each neuron has a receptive field (RF), a specific region of the visual field where light stimulus will modify the activity of this neuron. The term RF is not limited only to the spatial region but it is often extended to include the temporal structure within this

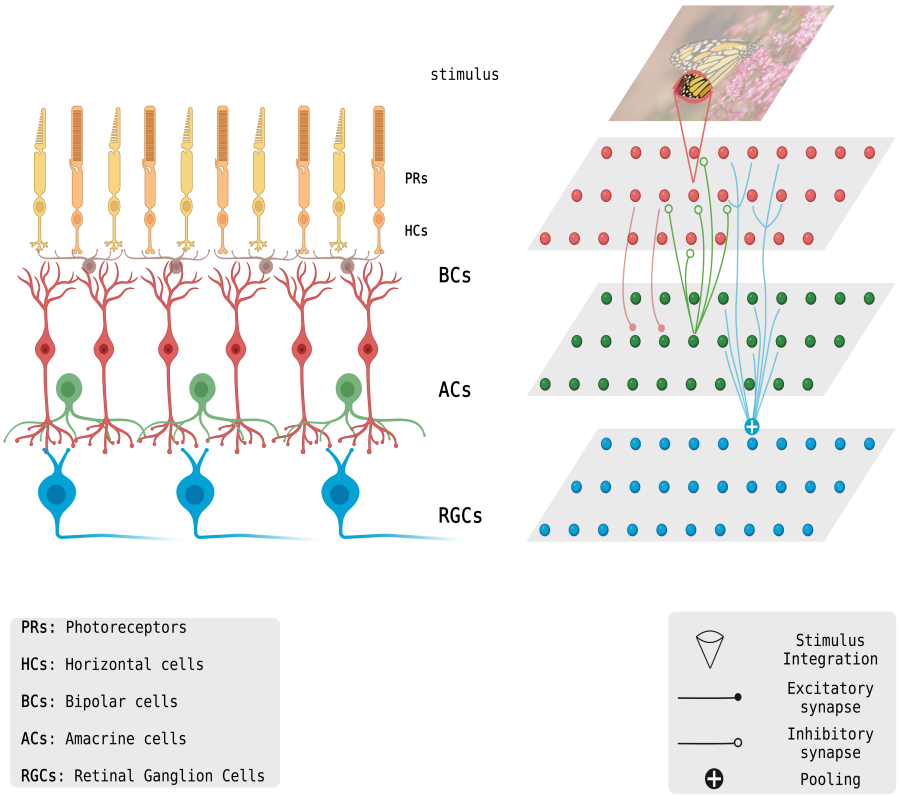

Figure 1: **Translation of the retinal circuit to a computational network model. Left.** Schematic of the retina. Light activates the photoreceptor cells (PRs), that transduce the input into a cascade of biochemical and electrical events that can stimulate BCs and onwards RGCs. This vertical excitatory pathway is modulated by inhibitory interneurons comprising two groups; horizontal (HCs) and amacrine cells (ACs). All these neural signals are integrated by RGCs and finally converted into action potentials going to the brain. **Right**. Schematic view of the model. The joint integration of PRs and HCs in a limited region of space is modelled by a spatio-temporal kernel mathematically defining the OPL input to a BC, corresponding to eq. (1). The convolution of this kernel with the stimulus drives the BC evolution, modulated by inhibitory connections with ACs. BCs indeed make excitatory synaptic connections with ACs and ACs inhibit BCs. Finally, RGCs pool over many BCs and ACs in their neighbourhood.

region. The RF usually exhibits a centre-surround organisation and it is assumed that BCs are the first neurons along the visual pathway to follow this principle (Werblin and Dowling, 1969). Each BC receives synaptic inputs from its upstream circuitry, a combination of dendritic excitatory inputs from photoreceptors (rods and cones) and inhibitory inputs from horizontal cells (Franke and Baden, 2017) occurring at the level of the Outer Plexiform Layer (OPL) and this interaction emerges on their RF.

A popular approach to simplify the complex process involved is to model the RF as a single spatio-temporal linear filter that essentially represents the opposition between the centre of the receptive field, driven by photoreceptors, and the surround signal transmitted by horizontal cells. The membrane potential of the BC's soma can then be linearly approximated by the convolution of the spatio-temporal kernel $\mathscr{K}_{B_i}$, featuring the biophysical processes at the OPL, with the visual stimulus $\mathscr{S}(x,y,t)$. As we do not consider color sensitivity here, $\mathscr{S}$ characterizes a black and white scene, with a control on the level of contrast $\in [0,1]$. The voltage of BC i is stimulus-driven by the term ("OPL input"):

$$\begin{aligned} \mathscr{V}_i^{(drive)}(t) &= \left[\mathscr{K}_{B_i} \overset{x,y,t}{*} \mathscr{S}\right](t) \\ &= \int_{x=-\infty}^{+\infty}\int_{y=-\infty}^{+\infty}\int_{s=-\infty}^{t} \mathscr{K}(x-x_i, y-y_i, t-s)\,\mathscr{S}(x,y,s)\,dx\,dy\,ds, \end{aligned} \tag{1}$$

where $\overset{x,y,t}{*}$ means space-time convolution. We consider only one family of BCs so that the kernel $\mathscr{K}$ is the same for all BCs. What changes is the center of the RF, located at x_i, y_i, which also corresponds to the coordinates of the BC i . We consider in the paper separable kernel $\mathscr{K}(x,y,t) = \mathscr{K}_S(x,y)\,\mathscr{K}_T(t)$ where $\mathscr{K}_S$ is the spatial part and $\mathscr{K}_T$ the temporal part. We restrict to ON or OFF BCs with a *monophasic* spatio temporal kernel

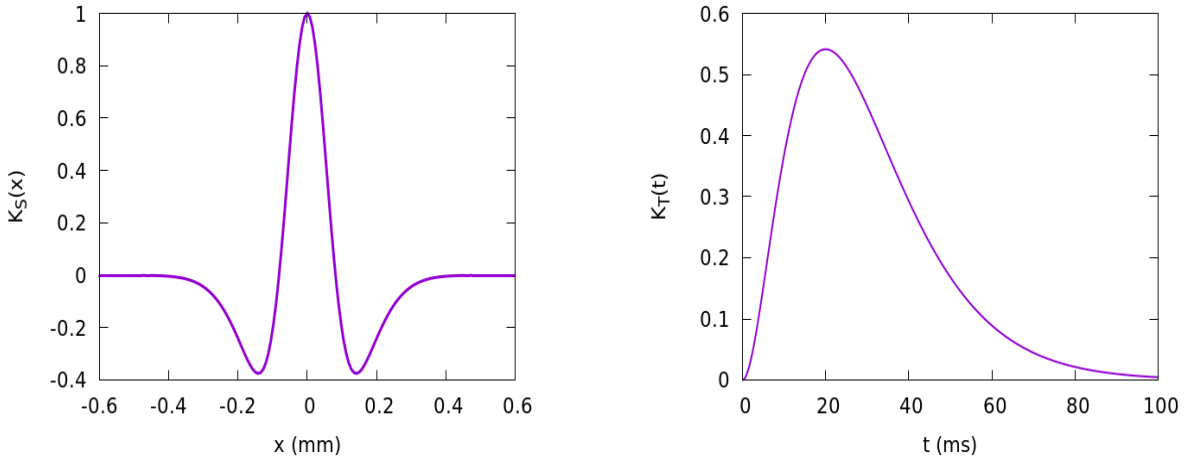

Figure 2: **OPL kernel. Left.** Spatial part in one spatial dimension. **Right.** Temporal part.

of the form:

$$\mathscr{K}_T = \left(A_0 \frac{t^2}{2\,\tau_{RF}^3} e^{-\frac{t}{\tau_{RF}}} + b_0 \right) H(t). \tag{2}$$

where A_0 controls the amplitude of the OPL input and can be positive (ON BC) or negative (OFF BC), while b_0 controls the level of residual polarisation observed in experiments, when light stimulation has stopped. The spatial part, $\mathscr{K}_S$, is a classical difference of Gaussians. This spatio temporal kernel is illustrated in Figure 2. Although the OPL input to BCs is monophasic, the ACs lateral inhibition can make the RF of a BC biphasic, depending on parameters.

2.1.3 Voltage dynamics

In the model neurons are characterized by their voltage. Although RGCs are spiking, we will be indeed interested only in their voltage variations, as a function of the stimulus and network effects. Spiking could be obtained from voltage e.g. using a Linear Nonlinear Poisson process (LNP) for spiking neurons (Berry et al., 1999; Chen et al.,

2013; Souihel and Cessac, 2021). We note V_{B_i} the voltage of BC i, V_{A_j} the voltage of AC j, V_{G_k} the voltage of RGC k. The joint dynamics of voltages is given by the dynamical system (Souihel and Cessac, 2021):

$$\begin{cases} \frac{dV_{B_i}}{dt} &= -\frac{1}{\tau_{B_i}} V_{B_i} + \sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathscr{N}^{(A)}\left(V_{A_j}\right) + F_{B_i}(t), \quad i = 1 \dots N \\ \\ \frac{dV_{A_j}}{dt} &= -\frac{1}{\tau_{A_j}} V_{A_j} + \sum_{i=1}^{N_B} W_{A_j}^{B_i} \mathscr{N}^{(B)}\left(V_{B_i}\right) + \zeta_A, \quad j = 1 \dots N, \\ \\ \frac{dV_{G_k}}{dt} &= -\frac{1}{\tau_{G_k}} V_{G_k} + \sum_{i=1}^{N_B} W_{G_k}^{B_i} \mathscr{N}_B(V_{B_i}) + \sum_{j=1}^{N_A} W_{G_k}^{A_j} \mathscr{N}_A(V_{A_j}) + \zeta_G, \quad k = 1 \dots N. \end{cases} \tag{3}$$

$\tau_{B_i}, \tau_{A_j}, \tau_{G_k}$ are respectively the characteristic integration times of the BC i, AC j and RGC k.

The rectification term:

$$\mathscr{N}_A(V_{A_j}) = \begin{cases} V_{A_j} - \theta_A, & \text{if } V_{A_j} > \theta_A; \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

ensures that the synapse $j \to i$ is inactive when the pre-synaptic voltage V_{A_j} is smaller than the threshold θ_A. The same holds for the rectification term $\mathscr{N}_B(V_{B_i})$ with a threshold θ_B.

The term:

$$F_{B_i}(t) = \frac{\mathscr{V}_i^{(drive)}}{\tau_{B_i}} + \frac{d}{dt}\mathscr{V}_i^{(drive)}, \tag{5}$$

is the stimulus driven input to BCs. It takes this form to ensure that, in the absence of ACs coupling, $V_{B_i}(t) = \mathscr{V}_i^{(drive)}(t)$.

CNO bind very specific to designed receptors leading either to an excitatory or inhibitory response, depending on the receptors (G_i or G_q) (Roth, 2016; Urban and Roth,

2015). The CNO effect can be modelled by a current of the form $-g_{CNO_T}\,(V-\mathscr{E}_{CNO_T})$, where V is the voltage of a cell sensitive to CNO, T is the cell type (i.e. only ACs or RGCs according to the experimental set-up); g_{CNO_T} is the conductance of channels sensitive to CNO that is zero in the absence of CNO while it increases with CNO concentration and $\mathscr{E}_{CNO_T}$ is the corresponding reversal potential. As CNO changes the membrane conductance it also induces a change of polarization, which is characterized by the parameter ζ_A for ACs and ζ_G for RGCs, with a general form:

$$\zeta_T = \frac{\mathscr{E}_{CNO_T}}{C}\, g_{CNO_T} \tag{6}$$

where $T = A, G$, as well as a change in the characteristic time scale of the membrane, that are the parameters τ_{A_j}, τ_{G_k}. We also assume an indirect effect: CNO modifies the synaptic weight from a CNO sensitive cell to the cells it is connected to, that is the synaptic weights W_B^A, W_G^A. We note, however, that this simplification of the actual biological process is for mathematical convenience.

2.1.4 Parameters dependence

The model depends on many parameters that constrain the dynamical evolution of the system (3). Although some of our results, like the explicit form for RFs (12), are quite general, it is easier, for analytic derivations and for simulations to further simplify these parameters. These simplifications, detailed below, essentially converge to considering that our network is composed of identical cells for each type.

First, we consider that the synaptic weights intensity from a cell type to another, e.g. the synapse from ACs to BCs, are controlled by a unique parameter, that is $W_B^A = -w^-\,\Gamma_B^A$ where $w^- > 0$ (inhibitory synapse) and where Γ_B^A is a connectivity matrix

($\Gamma_{B_i}^{A_j} = 1$ if the j-th AC connects to the i-th BC, and $\Gamma_{B_i}^{A_j} = 0$ otherwise). Likewise, $W_A^B = w^+ \Gamma_A^B$ with $w^+ > 0$, $W_G^B = w_G^B \Gamma_G^B$ with $w_G^B > 0$, and $W_G^A = w_G^A \Gamma_G^A$ with $w_G^A < 0$. Second, we consider in the mathematical derivation, that the characteristic integration times of cells only depend on the cell type, that is $\tau_{B_i} \equiv \tau_B$, $\tau_{A_j} \equiv \tau_A$, $\tau_{G_k} \equiv \tau_G$.

In addition, we consider a simple form of connectivity, where $\Gamma_A^B = \Gamma_B^A$ are nearest neighbors connectivity matrices. The retina is regularly tiled by different cell types so this approximation is reasonable, although here one cell connects to more than 4 neighbors. Our "cells" must actually be considered as effective cells with effective interactions. In particular, our parameters w^-, w^+ correspond to many synaptic contacts. Γ_G^B and Γ_G^A are pooling matrices: cell i connects to cell j with a Gaussian probability depending on the distance between the two cells (Berry et al., 1999; Chen et al., 2013).

We fix the thresholds θ_A, θ_B to zero so that the voltages in the rest state of eq. (3) are vanishing. Most of the analysis below will be done considering that no rectification takes place so that we essentially consider a linear model. For a more general analysis please check the supplementary, as well as the discussion section.

2.1.5 Experimental set-up

In our experiments, excitatory DREADDs (hM3Dq) were activated using CNO on RGCs and ACs co-expressing a certain gene (*Scnn1a* or *Grik4*), triggering a calcium release from organelles and thus, leading to increase of intracellular concentration of free calcium. This resulted in membrane depolarisation and higher neuronal excitability. Our experiments suggested that subclasses of ACs and RGCs could be simultaneously sensitive to CNO but we did not observe any evidence of an effect on BCs.

Detailed experimental details can be found in our recent publication (Hilgen et al., 2022). All experimental procedures were approved by the ethics committee at New-

castle University and carried out in accordance with the guidelines of the UK Home Office, under the control of the Animals (Scientific Procedures) Act 1986. Recordings were performed on the BioCamX platform with high-density-multielectrode array (HD-MEA) Arena chips (3Brain GmbH, Lanquart, Switzerland), integrating 4096 square microelectrodes in a 2.67×2.67 mm^2 area and aligned in a square grid with 42 µm spacing. Light stimuli were projected onto the retina using a LED projector. Briefly, the projector irradiance was attenuated using neutral density filters to mesopic light levels.

3 Results

In this section we present the theoretical and numerical results based on our retina model. We provide only the main conclusions of the mathematical derivations, which are presented in detail in the supplementary section.

3.1 Model fitting of ganglion cells receptive fields characterized from experimental data

RGC responses emanate from a dynamic balance of synaptic excitation and inhibition, originating from the interactions of BCs and ACs. We believe that such network connectivity gives rise to various response patterns and we show that our model can capture these joint effects, by providing an analytic form of the RF of the cells. As we demonstrate, this computation provides us an algorithmic way to fit the model parameters to the light responses recorded from mouse RGCs. One can then infer the possible behaviour of ACs and BCs leading to this RGC response, *even if we do not measure them experimentally.*

3.1.1 Mathematical form of the RF of retinal cells

The results presented below hold for all cell types. Thus, we label cells with a generic index $\alpha = 1 \dots 3N$. BCs have an index $\alpha = 1 \dots N$, ACs have an index $\alpha = N+1 \dots 2N$, RGCs have an index $\alpha = 2N+1 \dots 3N$ and we write $\mathscr{X}_\alpha$ the voltage of cell α.

The time evolution of the dynamical system in eq. (3) is controlled by a matrix, $\mathscr{L}$, called "transport operator" and explicitly written in the supplementary section 4.1. $\mathscr{L}$ depends on the connectivity matrices $W_B^A, W_A^B, W_G^A, W_G^B$ and on all the parameters controlling the dynamics. The form of $\mathscr{L}$ also depends on the set of rectified cells. In the following, we assume that cells are not rectified i.e. hyperpolarised BCs do not reach the rectification threshold (the rectified case is discussed in the conclusion section and in the supplementary material). Consequently, the dynamical system (3) is linear.

In this case, the eigenvalues λ_β, $\beta = 1 \dots 3N$ and the eigenvectors $\mathscr{P}_\beta$ of $\mathscr{L}$ characterize the evolution of cells' voltages. We note $\mathscr{P}$ the matrix transforming $\mathscr{L}$ in diagonal form (the columns of $\mathscr{P}$ are the eigenvectors $\mathscr{P}_\beta$) and $\mathscr{P}^{-1}$ its inverse.

In this context, we show in the supplementary section 4.1.3 that the voltage of a cell with index α is the sum of 4 terms:

$$\mathscr{X}_\alpha(t) = \mathscr{V}_\alpha^{(drive)}(t) + \mathscr{E}_\alpha^{(drive)}(t) + \mathscr{E}_\alpha^{(CNO_A)} + \mathscr{E}_\alpha^{(CNO_G)}, \quad \alpha = 1 \dots 3N. \tag{7}$$

Stimulus drive. The first term, $\mathscr{V}_\alpha^{(drive)}(t)$ corresponds to (1), and is non zero for BCs only. It corresponds to the BCs response in the absence of the ACs network.

CNO effects. The terms:

$$\mathscr{E}_\alpha^{(CNO_A)} = \zeta_A \sum_{\beta=1}^{3N} \sum_{\gamma=N+1}^{2N} \frac{\mathscr{P}_{\alpha\beta}\, \mathscr{P}^{-1}_{\beta\gamma}}{\lambda_\beta}, \quad \alpha = N+1 \dots 2N; \tag{8}$$

and

$$\mathscr{E}_{\alpha}^{(CNO_G)} = \zeta_G \, \tau_{G_\alpha}, \quad \alpha = 2N+1 \ldots 3N. \tag{9}$$

correspond, respectively, to the impact of CNO on the voltages of ACs and RGCs. There are important nonlinear effects hidden in the terms $\frac{\mathscr{P}_{\alpha\beta}\,\mathscr{P}^{-1}_{\beta\gamma}}{\lambda_\beta}$ (eq. (8)). Thus, the polarization level of ACs and RGCs is not only fixed by the direct effect of CNO on the cell, but is also tuned by entangled network effects.

Stimulus-network interaction term. In eq. (7), the term :

$$\mathscr{E}_{\alpha}^{(drive)}(t) = \sum_{\beta=1}^{3N} \sum_{\gamma=1}^{N} \mathscr{P}_{\alpha\beta}\, \mathscr{P}^{-1}_{\beta\gamma} \varpi_{\beta\gamma} \int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \mathscr{V}_{\gamma}^{(drive)}(s)\, ds, \quad \alpha = 1 \ldots 3N, \tag{10}$$

where $\varpi_{\beta\gamma} = \frac{1}{\tau_{B_\gamma}} + \lambda_\beta$, corresponds to the indirect effect, via the network connectivity, of the stimulus drive on (i) BCs, for $\alpha = 1 \ldots N$; (ii) ACs for $\alpha = N+1 \ldots 2N$; (iii) RGCs $\alpha = 2N+1 \ldots 3N$. Thus, this equation describes the response of *all* cells to the stimulus. Especially, it tells us how the direct input (1) to BCs is modulated by the concerted activity of BCs and ACs.

Mathematically, the term (10) can be interpreted as follows. The drive (index $\gamma = 1 \ldots N$) triggers the eigenmodes $\beta = 1 \ldots 3N$ of $\mathscr{L}$, with a weight proportional to $\mathscr{P}^{-1}_{\beta\gamma}$. The mode β, in turn, acts on the voltage of cell $\alpha = 1 \ldots 3N$ with a weight proportional to $\mathscr{P}_{\alpha\beta}$. The time dependence and the effect of the drive are controlled by the integral $\int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \mathscr{V}_{\gamma}^{(drive)}(s)\, ds$.

The Receptive Field of all cell types. Introducing the function $e_\beta(t) \equiv e^{\lambda_\beta t} H(t)$ so that $\int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \mathscr{V}_{\gamma}^{(drive)}(s)\, ds \equiv \left[e_\beta \overset{t}{*} \mathscr{K}_{B_T} \overset{t}{*} \left(\mathscr{K}_{B_{S_\gamma}} \overset{x,y}{*} \mathscr{S} \right) \right](t)$, and using the sepa-

rated kernel form (2), the response (10) reads:

$$\mathscr{E}_{\alpha}^{(drive)}(t) = \left[\mathscr{K}_{\alpha} \overset{x,y,t}{*} S \right](t), \tag{11}$$

with:

$$\mathscr{K}_{\alpha}(x,y,t) = \sum_{\beta=1}^{3N} \left(\mathscr{P}_{\alpha\beta} U_{\beta}(t) \times \sum_{\gamma=1}^{N} \mathscr{P}_{\beta\gamma}^{-1} \varpi_{\beta\gamma} \mathscr{K}_{B_{S_\gamma}}(x,y) \right), \tag{12}$$

where we have set $U_{\beta}(t) \equiv \left[e_{\beta} \overset{t}{*} \mathscr{K}_{B_T} \right](t)$. The response of cell α is thus expressed as a convolution of the stimulus with a spatio-temporal kernel $\mathscr{K}_{\alpha}(x,y,t)$, an expected result from the linear response. Nevertheless, it's important to point out that the expression (12) holds for all cell types, not only RGCs and that it contains the network effects induced by the BCs-ACs network. Thus, for $\alpha = 1 \ldots N$, equation (12) characterizes the indirect (network induced) response of BCs to the stimulus drive, in addition to the direct response (1). For $\alpha = N+1 \ldots 2N$, equation (12) represents the RF of ACs. Finally, for $\alpha = 2N+1 \ldots 3N$ we obtain the RF of RGCs. We focus on this last case from now on, essentially because this predicted RF can be confronted to experiments, whereas we have no experimental access to BCs or ACs RF.

The Receptive Field of RGCs. Henceforth, we will refer to $\mathscr{K}_{\alpha}(t)$ as $\mathscr{K}_{G_\alpha}(t)$, to make explicit that we are dealing with RGCs. The RF of RGCs can be often written as a product of a space dependent term and a time dependent term (separability). In our case, this would correspond to write $\mathscr{K}_{G_\alpha}(x,y,t)$ in the form of a product $\mathscr{K}_{G_\alpha}(x,y,t) = \mathscr{K}_{G_{T_\alpha}}(t) \times \mathscr{K}_{G_{S_\alpha}}(x,y)$ where:

$$\mathscr{K}_{G_{T_\alpha}}(t) = \sum_{\beta=1}^{3N} \mathscr{P}_{\alpha\beta} U_{\beta}(t), \tag{13}$$

is a temporal kernel and:

$$\mathscr{K}_{G_{S_\alpha(x,y)}} = \sum_{\gamma=1}^{N} \mathscr{P}^{-1}_{\beta\gamma}\, \varpi_{\beta\gamma} \mathscr{K}_{B_{S_\gamma}}(x,y), \tag{14}$$

is a spatial kernel.

This separation is not strictly possible in eq. (12), because there is a dependency on β on the term $\sum_{\gamma=1}^{N} \mathscr{P}^{-1}_{\beta\gamma}\, \varpi_{\beta\gamma} \mathscr{K}_{B_{S_\gamma}}(x,y)$. Nevertheless, depending on the structure of the matrix $\mathscr{P}$ constrained by connectivity, we can neglect this dependence, ergo separability holds with a very good accuracy (Cessac, 2022).

The spatial part of the RGCs RF. Equation (14) appears as an overlap of spatial RFs of BCs. In such naive overlaps approximations, spatial RFs of BCs are just summed up with a uniform weight. However, here the contribution of each RFs of BCs is weighted by the term $\mathscr{P}^{-1}_{\beta\gamma}\, \varpi_{\beta\gamma}$ which is constrained by ACs lateral connectivity. In particular, equation (14) is not necessarily circular even if BCs RFs are, and the center of the RGC cell RF is not necessarily at the barycentre of connected BCs RFs. This holds, for example, if AC connectivity is not invariant by rotation.

The temporal part of the RGCs RF, (13). As we consider monophasic temporal kernels $\mathscr{K}_{B_T}$ of BCs with the form (2) we have:

$$U_\beta(t) = A_0 \left(\frac{\left[2\tau_{RF}^2 e^{\lambda_\beta t} - \left(t^2 \left(\lambda_\beta \tau_{RF} + 1 \right)^2 + 2t\tau_{RF} \left(\lambda_\beta \tau_{RF} + 1 \right) + 2\tau_{RF}^2 \right) e^{-\frac{t}{\tau_{RF}}} \right]}{2\left(\lambda_\beta \tau_{RF} + 1 \right)^3 \tau_{RF}^2} \right) H(t) \tag{15}$$

U_β, and, thereby, $\mathscr{K}_{G_{T_\alpha}}(t)$ is the temporal part of the RGC receptive field, that changes their shape due to variations in the eigenvalues of $\mathscr{L}$, who are themselves

controlled by model parameters. A striking effect arises when some eigenvalues become complex, leading to temporal oscillations of U_β. This remark is at the core of the analysis exposed in section 3.2.2.

3.1.2 Fitting the RFs of ganglion cells

In order to assess the validity of the model, we have first fitted the recorded RGCs in CTL and CNO conditions. We have a data base of 117 cells sensitive to CNO, i.e. exhibiting increase or decrease in firing rate beyond a certain threshold. Experimentally, RGCs RFs were reconstructed from Spike Triggered Average (STA) in response to Shifted White Noise (SWN) (Pamplona et al., 2021). This resulted in temporal traces with duration 600 ms sampled with a rate $33/4 = 8.25$ ms.

These reconstructed RFs provide the linear response of a RGC to a spatially uniform flashed stimulus, mathematically corresponding to a Dirac distribution. As we wanted to compare our model's output, the RGC voltage, V_G, to this experimental RF, computed from firing rates, we neglected the effect of non linearities and assumed that the experimental response is proportional to the RGC voltage. We considered a one dimensional model (chain) with $N = 60$ cells of each type, with nearest neighbours connectivity where the cells at the boundaries have a fixed, zero, voltage (zero boundary conditions), corresponding to the reference rest state. To reduce the boundaries effect, we made the fit for the RGC in the center of the network.

We perform simulations of the model (3) using a spatially uniform Dirac pulse as the stimulus and compute the cell responses, using two modalities: simulation of the differential equations (3) (green traces labelled "Sim" in Figure 3) and analytic computation (12) (black traces labelled "Th" in Figure 3). We observe that these two traces are always identical confirming the goodness of the simulation scheme.

We recall that the parameters shaping this response are: w^+, controlling the synaptic intensity from BCs to ACs; w^-, controlling the synaptic intensity from ACs to BCs; w_G^B, controlling the synaptic intensity from BCs to RGCs; w_G^A, controlling the synaptic intensity from ACs to RGCs; τ_B, the characteristic membrane time scale of BCs; τ_A, the characteristic membrane time scale of ACs; τ_G, the characteristic membrane time scale of RGCs; τ_{RF}, the characteristic membrane time scale of the OPL drive;

We have two additional parameters: A_0, the intensity of the OPL input and b_0, controlling a slight, residual, depolarization/hyper polarization observed in experimental responses. This last parameter encompasses the effect of CNO on the polarization of the RGCs. This effect results from the conjugated action of CNO on ACs (term ζ_A in (3)), action of CNO on RGCs (term ζ_G in (3)) and entangled network effects appearing in equations (8) and (9). The introduction of a single parameter b_0 to mimic this intricate effect is therefore a sharp simplification, making the estimation easier. We note $\vec{\eta}$ the set of all the parameters shaping response. $\vec{\eta}$ is therefore a point in a 10-dimensional space.

The fit was then done by a gradient descent to minimize the L^2-distance $D_2(\vec{\eta})$ between the experimental trace of the time STA, $STA(s)$ and the theoretical temporal RF (13) which depends on $\vec{\eta}$. The minimization is done by iterating the differential equation:

$$\frac{d\vec{\eta}}{du} = -\vec{\nabla}_{\vec{\eta}} D_2.$$

The gradient of $\vec{\nabla}_{\vec{\eta}} D_2$ involves $\vec{\nabla}_{\vec{\eta}} \mathscr{K}_{G_{T\alpha}}(s)$ which can be explicitly computed when we have the analytic form of RF, or numerically. Note that having the analytic form gives better results especially because it allows second order corrections (Hessian). Although the theoretical result provides both the temporal and the spatial RF (including

the surround), it is difficult, experimentally, to fit the spatial part. That is why our minimisation is done only on the temporal trace. Nevertheless, the simulation allows us to draw the corresponding spatial RF.

Although the experimental temporal RFs were quite diverse among cells, we were able to fit all of them with a very good accuracy (final error smaller than 1%). We rejected fits where some parameters became unrealistic (e.g. τ_A larger than 1 s or $|w^-| > 1$ kHz). We rejected about 4% of the fits. An example of fit is shown in Figure 3. The complete results displayed in the same form as Figure 3 can be found on the web page https://gitlab.inria.fr/biovision/dreadds where the C-code used for simulations can also be found, for all experimental cells.

In this synthetic representation the simulated responses of the OPL, BCs and ACs connected to the RGC located at the centre of the network appears in the top left figure. We also show the simulated response of the RGC vs the experimental temporal STA of this cell (bottom left). On the top right we see the numerical spatio-temporal RF using a color map. Finally, at the bottom right, we show the power spectrum (modulus of the Fourier transform) of the temporal response. As developed below, this spectrum provides important information on the cell response.

RGCs RFs in the presence of CNO could be fit equally well with our model. All fitted cells, in CTL and CNO conditions, can be found on the web page `https://gitlab.inria.fr/biovision/dreadds`.

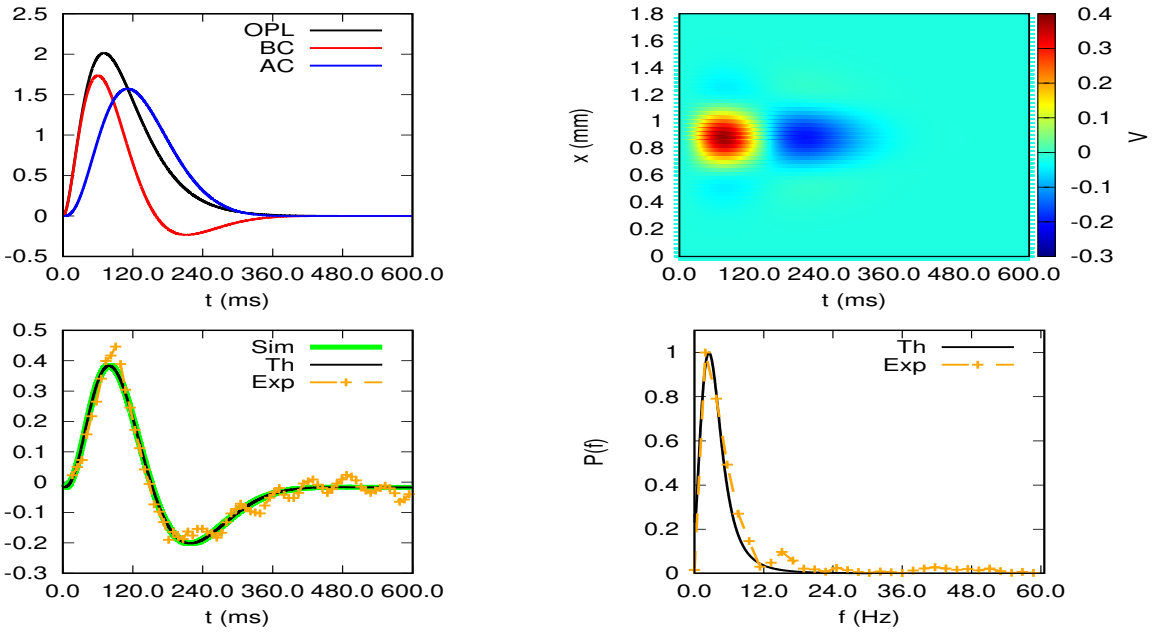

Figure 3: **Summary panel.** The top left panel illustrates the simulated responses of the OPL (term $\mathcal{V}_{\alpha}^{(drive)}(t)$ in (1), black trace), BCs (red trace), ACs (blue trace) connected to the RGC located at the centre of the network. The bottom left panel shows the simulated response of the RGC (green, "Sim", and black, "Th", trace) vs the experimental temporal STA of this cell (orange dots). The green trace ("Sim") is the result of a numerical simulation of the dynamical system (3) under a spatially uniform flashed stimulus, whereas the black trace ("Th") is the result given by the analytic expression (12). The top right panel shows the spatio-temporal RF of the RGC, time in abscissa, space in ordinate. The bottom right panel displays the power spectrum of the time response, experimental (orange dots) and theoretical (black lines).

3.2 Network connectivity shapes the receptive fields of ganglion cells

Throughout our analysis of the experimental data, we observed great variability in the effect of DREADD activation with CNO on the RGCs responses. Some of them are expected due to the underlying mechanism of the excitatory DREADD activation that causes an increase of intracellular concentration of free calcium which leads to membrane depolarisation. Consequently, the ionic diffusion across the cell membrane through ion channels increases and this, in turn, affects the permeability and electrical conductance of these channels to the respective ions. A larger conductance would yield a shorter time constant, which would make the cell faster. Thus, we should notice an increase of the baseline activity and decrease of the time constant. However, we didn't observe any systematic trend (see Figure 5). This can be justified if we consider that DREADDs are not only expressed in RGCs, but also ACS, thus CNO activation has a direct impact on the ACs response and a potential indirect effect on RGCs response. In a case where a RGC receives input from DREADDs-expressing inhibitory ACs and does not express DREADDs itself, we would expect to notice a decrease in the baseline activity of its response and perhaps an increase in the time constant, meaning that the cell will become more sluggish. If, though, this RGC is also susceptible to CNO, the reinforced inhibitory effect of the ACs competes with the direct effect of CNO activation on the RGC's response. In this scenario, the net effect of CNO might vary widely.

In this section, we develop the consequences of our mathematical analysis in an attempt to explain the observed diversity of CNO effects on RF features. We propose here an explanation purely based on network effects. There are certainly other possible interpretations based on single cell characteristics such as non linear effects due to

changes in conductance etc, discussed in the discussion section. The main advantages of our analysis is that it determines network effects on the RGCs RF, controlled by two main parameters, and that it predicts the response to more complex stimuli than full field flashes.

3.2.1 Two main parameters constrain the RF shape of ganglion cells

The entangled, feedback effects of ACs-BCs can be characterized by two a-dimensional parameters. The first one, $r = \frac{\tau_A}{\tau_B}$, characterizes the ratio between the ACs and BCs membrane integration times. The second, $s = \frac{w^-}{w^+}$, characterizes the ratio between the ACs $\rightarrow$ BCs interaction (w^-) and the BCs $\rightarrow$ ACs interaction (w^+). Of course, the other parameters play an important role when fitting a specific RF. But, what we argue here is that the shape of RF and its space-time scaling essentially depend on the value of r, s.

The theoretical explanation is that the RF of a RGC is given by the formula (11), which is a cascade of convolutions involving the BC response to the stimulus (OPL input) and the network effects expressed in terms of eigenvalues λ_β and eigenvectors components $\mathscr{P}_{\alpha\beta}$ appearing in equation (12). As explained in the supplementary section 4.1.4, these eigenvalues and eigenvectors are essentially tuned by the two parameters r, s. There is also a dependence on other parameters discussed in section 3.2.3. Depending on the location in the space r, s, some eigenvalues are real, some others are complex. All eigenvalues have a negative real part, ensuring the stability of the linear system. Imaginary parts in eigenvalues introduce oscillations in the response, whereas the real part fixes a characteristic decay time. The RF formula (12), involving a sum of exponentials $e^{\lambda_\beta t}$ mixes these effects. As we considered monophasic OPL response here, the time RF of RGCs is monophasic when all eigenvalues are real. In contrast,

oscillations in this RF can appear when some eigenvalues are complex. However, the shape of this RF depends in more detail on the period of oscillations, brought by the imaginary part of complex eigenvalues, and on the characteristic decay times, brought by the real parts.

When moving in the (r,s) plane, the eigenvalue n switches from real to complex conjugate pair when crossing a critical line, depending on n, whose equation (36) is given in the supplementary section 4.1.4. There are $2N$ eigenvalues associated with the BCs-ACs network each one determining a critical line in the plane (r,s). The set of all these lines is what we call the "skeleton". An example of this skeleton is shown in Figure 4, where we only show some of the critical lines. These lines delimit color regions corresponding to the number of complex eigenvalues (see colorbar legend on the right of the figure).

3.2.2 The RFs map

The existence of this skeleton determines regions in the (r,s) plane with specific shapes for the temporal RF, given by eq. (13), a linear combination of functions $U_\beta(t)$ given by (15). The Fourier transform $\hat{U}_\beta(\omega)$ of $U_\beta(t)$ is:

$$\hat{U}_\beta(\omega) = \frac{1}{(1+i\omega\tau_{RF})^3}\,\frac{1}{i\omega-\lambda_\beta}. \tag{16}$$

Thus, the Fourier transform of (13) is:

$$\mathscr{K}_{G_{T\alpha}}(\omega) = \sum_{\beta=1}^{3N} \mathscr{P}_{\alpha\beta}\,\hat{U}_\beta(\omega), \tag{17}$$

a linear combination of rational fractions. Extending to complex ωs, $\hat{U}_\beta(\omega)$ has two poles: $\omega = \frac{i}{\tau_{RF}}$ and $\omega = -i\lambda_\beta$, corresponding to complex resonances. The contributions of all these poles (for $\beta = 1 \ldots 3N$) are combined in eq. (17) with weights $\mathscr{P}_{\alpha\beta}$.

As we move in the (r,s) plane, we notice the following. When (r,s) are small, eigenvalues are real and the terms $\mathscr{P}_{\alpha\beta}$ are close to diagonal. In this case, the dominant pole contribution in (17) is the pole $\omega = i\frac{1}{\tau_{RF}}$ corresponding to the OPL contribution. Equation (17) has a single peak centered at $\omega = 0$, corresponding to a monophasic response. For larger values of r,s some eigenvalues become complex, giving potential additional peaks in the power spectrum. Actually, we observed two cases mutually compatible. First, the central peak at $\omega = 0$ switches to a non zero value. This corresponds to the appearance of an exponentially damped oscillation in the RF, giving a biphasic response. However, secondary peaks may appear leading to residual oscillations, in addition to the main trend (monophasic or biphasic). This gives what we call a polyphasic response. Such residual oscillations were observed in our experiments and were relatively numerous (about 40%). There are, of course, other hypotheses explaining these residual oscillations, but here, we will support the hypothesis that they are generated by a network effect. An example is given in Figure 3 where we observe, at the bottom left, residual oscillations after the main biphasic response, and, at the bottom right, the power spectrum with a main peak not centered at zero and a secondary peak corresponding to the residual oscillations. Note that this secondary peak is observed on experimental data but we failed to reproduce it in the fits. This is further explained in the discussion section.

This analysis lead us to broadly decompose the (r,s) plane into 3 regions corresponding to cells response phases: monophasic, biphasic, polyphasic. One switches

from one phase to the other when some peaks in the power spectrum appear or disappear, driven by the spectrum of $\mathscr{L}$.

The corresponding "phase diagram", obtained numerically, constitutes what we call the "RFs map" shown in Figure 4 (top right). There are four points, labelled A,B,C,D, on this map, each representing a different cell response phase. For each point, we have plotted the RGC temporal RF, as computed with the model (bottom panels). A more general representation of what is going on when moving along a specific pathway in this map can be found at the web page `https://team.inria.fr/biovision/cno_paper_supplementary/`, where one can see movies showing how the network effects shape the RGCs RFs when (r,s) vary.

Figure 4 illustrates how the BCs-ACs network shapes the RF of a RGC by a subtle balance between BCs-ACs, BCs-ACs interactions (parameter s) and the time scale of their response (parameter r). For simplicity, we consider here ON BCs, but the explanation holds also for OFF BCs.

The OPL drive (black line) induces a depolarisation of BCs (red line) within a time scale of order τ_B. This excites the connected ACs (blue line), with an intensity w^+. The excitation of ACs hyperpolarises BCs with an intensity w^- within a time scale of order τ_A. RGCs receive a combination of excitatory and inhibitory inputs from their afferent circuit with respective weights w_G^B and w_G^A.

In the monophasic region, ACs respond in the same time scale as BCs. One observes, for RGCs a monophasic response (Panel "A", middle left) whose intensity depends on the ratio between excitation, provided by BCs, with a weight w_G^B and inhibition, with a weight w_G^A. When moving to the biphasic region, ACs respond with a longer time scale as BCs, leading to the biphasic response illustrated with panel "B". When moving upward in the RFs map (increasing s) this biphasic response becomes polypha-

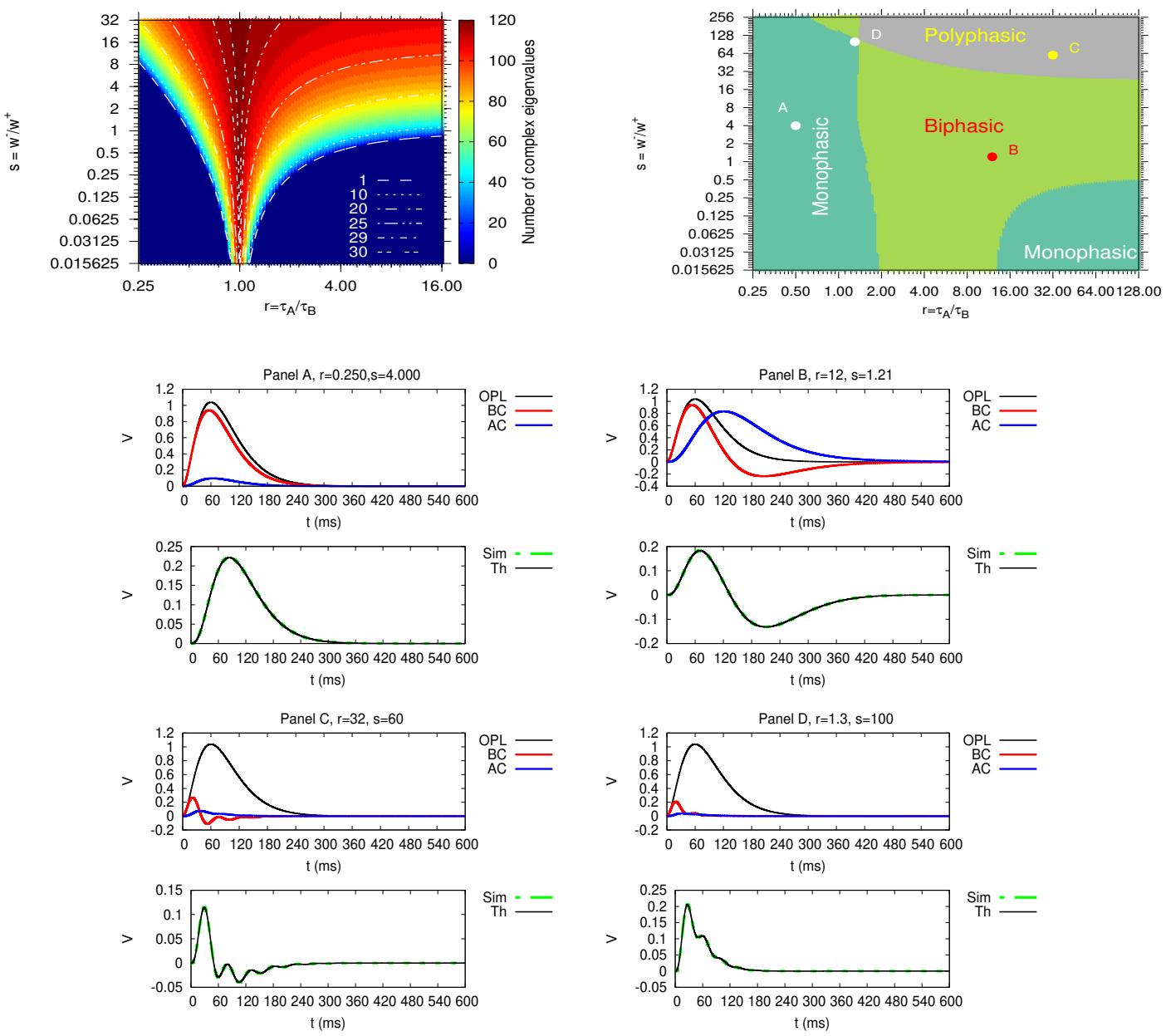

Figure 4: **Top Left. The skeleton.** Map in the (r,s) plane showing the "skeleton" of eigenvalues structure (white lines). When crossing line n the eigenvalue n switches from real to complex. The color corresponds to the number of complex eigenvalues (see the color bar on the right side). **Top right. The RFs Map.** This figures summarizes how the entangled effects of the BCs-ACs network act on the RF of RGCs. In this map in the (r,s) plane we distinguish three main regions (see text for their determination): monophasic, biphasic, polyphasic. The points labelled "A", "B, "C", "D" correspond to the temporal RF plotted in the bottom panels respectively called "Panel A", "Panel B", "Panel C", "Panel D". We use the same representation as in Figure 3.

sic with oscillations. This is illustrated in panel "C". BCs start to raise due to the OPL drive, leading to a rising of ACs, slower than BCs, leading to a hyper-polarisation of BCs. This leads to a decrease of ACs voltage, thereby, to a rising of BCs which still respond to the OPL drive. This cycle can be repeated several times, depending again on the parameters r, s. Note that the amplitude is always decreasing exponentially fast. The period of the observed oscillations and the damping characteristic time depend on the location in the Map. Finally, panel "D" is the point at the intersection of the 3 phases regions. We show it for completeness. Note that increasing s leads to a decrease of the RGCs response. When s is too high, the response becomes too weak to be observed experimentally.

3.2.3 Experimental cells spread in the RFs map

To confront our theoretical insight with experiments, we have placed the recorded cells in the RFs map as shown in Figure 5. That is, for each experimentally recorded RGC, we fit the model parameters $\vec{\eta}$ as explained in section 3.1.2) thereby providing an estimation of r, s. This defines a virtual network, made of identical cells in each layer, where all RGCs are responding like the experimental RGC. Thus, the map is *only a projection* of $\vec{\eta}$, which exists in a 10 dimensional space, in the two dimensional plane r, s. Some parameters are linked together though. The mathematical analysis in the supplementary section 4.1.4 shows us that the skeletons obtained for a fixed value of τ_B, w^+, can be extrapolated to other values τ'_B, w'^+ by the simple rescaling $r' = r, s' = \left(\frac{w'^+ \tau'_B}{w^+ \tau_B}\right)^2 s$. Using this property, the map of Figure 5 has been drawn for a specific value of $w^+ = 8.5$ Hz and $\tau_B = 30$ ms. This corresponds to mean values of these parameters, averaged over the set of experimental cells (in CTL conditions). In this two dimensional representation, extra information coming from the other parame-

ters $\tau_G, w_G^B, w_G^A, A_0, b_0$ is lost.

Figure 5, left, shows us the repartition of cells in the map, in CTL conditions. A few of them are monophasic, but many of them are biphasic, with a significant proportion close to the polyphasic region and showing residual oscillations. The figure 5, right, shows the same cells in CNO conditions.

The main observation is that CNO (right panel), does not dramatically change the repartition of cells in the RFs map. This is made more explicit in the bottom panels of figure 5. We show the mean and standard deviation of the main model parameters: $\tau_A, \tau_B, w^-, w^+, b_0$ in CTL and CNO conditions, separating the two subclasses of investigated genes: Grik4 and Scnn1a, and separating ON or OFF cells. These parameters are essentially constant showing that there is no statistical trend induced by CNO.

The situation is radically different when investigating the effect on *individual* cells. Indeed, the application of CNO makes some cells to move their representative point from one region in the RF map to the other, thereby drastically changing the cell's response. Two examples are shown in Figure 6 and 7. In Figure 6, the application of CNO induces a tiny motion of the representative point in the RFs map. However, as the cell is close to the area separating the monophasic from the biphasic phase, this motion impacts dramatically the shape of the time response. Actually, the RFs map can be refined by plotting the value of the main period T_1 (corresponding to the main peak in the power spectrum) as shown in the top right figure. In CTL conditions, the cell is located in yellow-green region with a high T_1 of order 600 ms, in the limit of experimental resolution. With CNO, the cell switches to a region where T_1 is of order 300 ms. At the bottom of the figures, the synthesis panel (same representation as in Figure 3) for CTL (left) and CNO conditions (right) is presented.

In Figure 7 we show a motion inducing a switch from biphasic to polyphasic. Here,

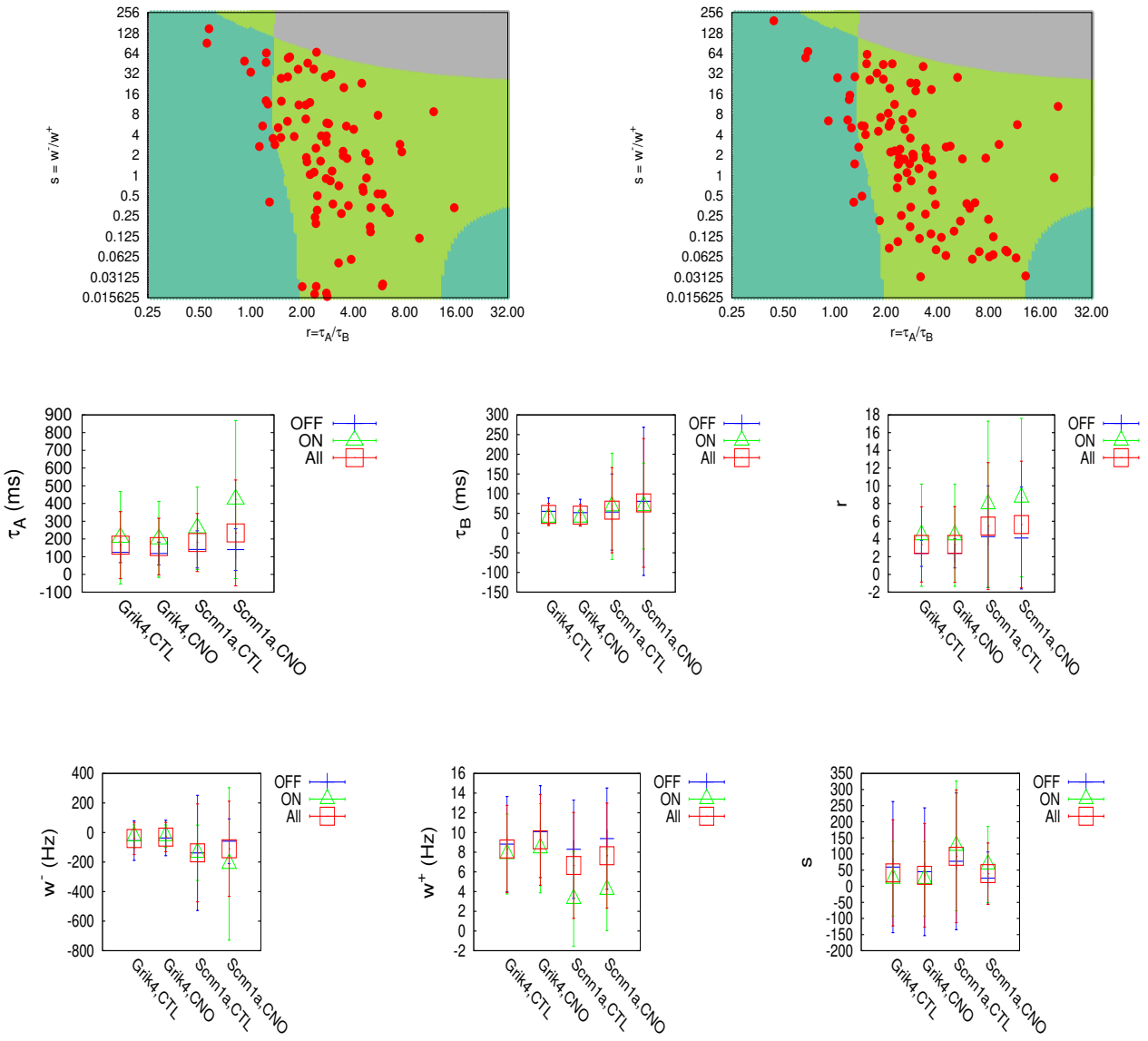

Figure 5: **Top. Repartition of fitted experimental cells in the RFs map.** Each point corresponds to an experimental cell (upper case corresponds to cell number while the index is the experience index). **Top left.** CTL conditions. **Top right.** CNO conditions **Middle.** Mean and standard deviation of τ_A (left), τ_B (center), r (right), fitted from experiments, for genes Grik4 and Scnn1a, in CTL and CNO conditions. We have separated the estimation for OFF cells (blue), ON Cells (green) and all cells (red). **Bottom.** Mean and standard deviation of w^- (left), w^+ (center), s (right), fitted from experiments. The representation is the same as the previous row.

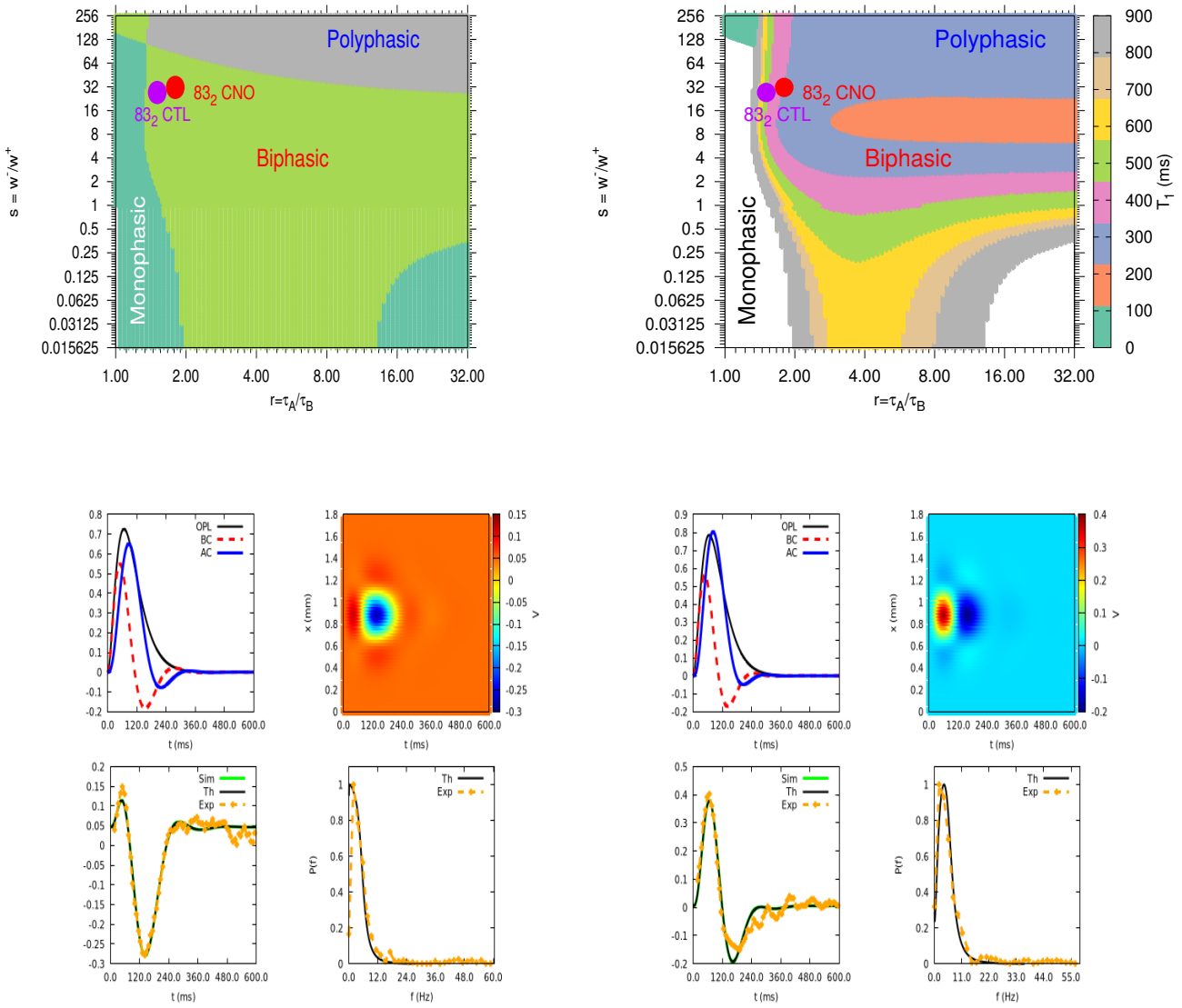

Figure 6: **CNO may change the cell response. Left top.** Application of CNO moves the representative point of the cell (here label 83_2) in the RFs map. Although this motion looks small, it corresponds to a switch from monophasic to biphasic region (**Bottom panel**). More precisely, as shown **Right top**, there is a displacement from a region where the oscillation period is quite high (larger than the observation time) to a region where it is of order 300 ms.

CNO drives the cell from the boundary of the polyphasic region to the biphasic region. The representation is the same as for Figure 6 except that the top right figure displays the period of the second period T_2 (secondary peak in the power spectrum).

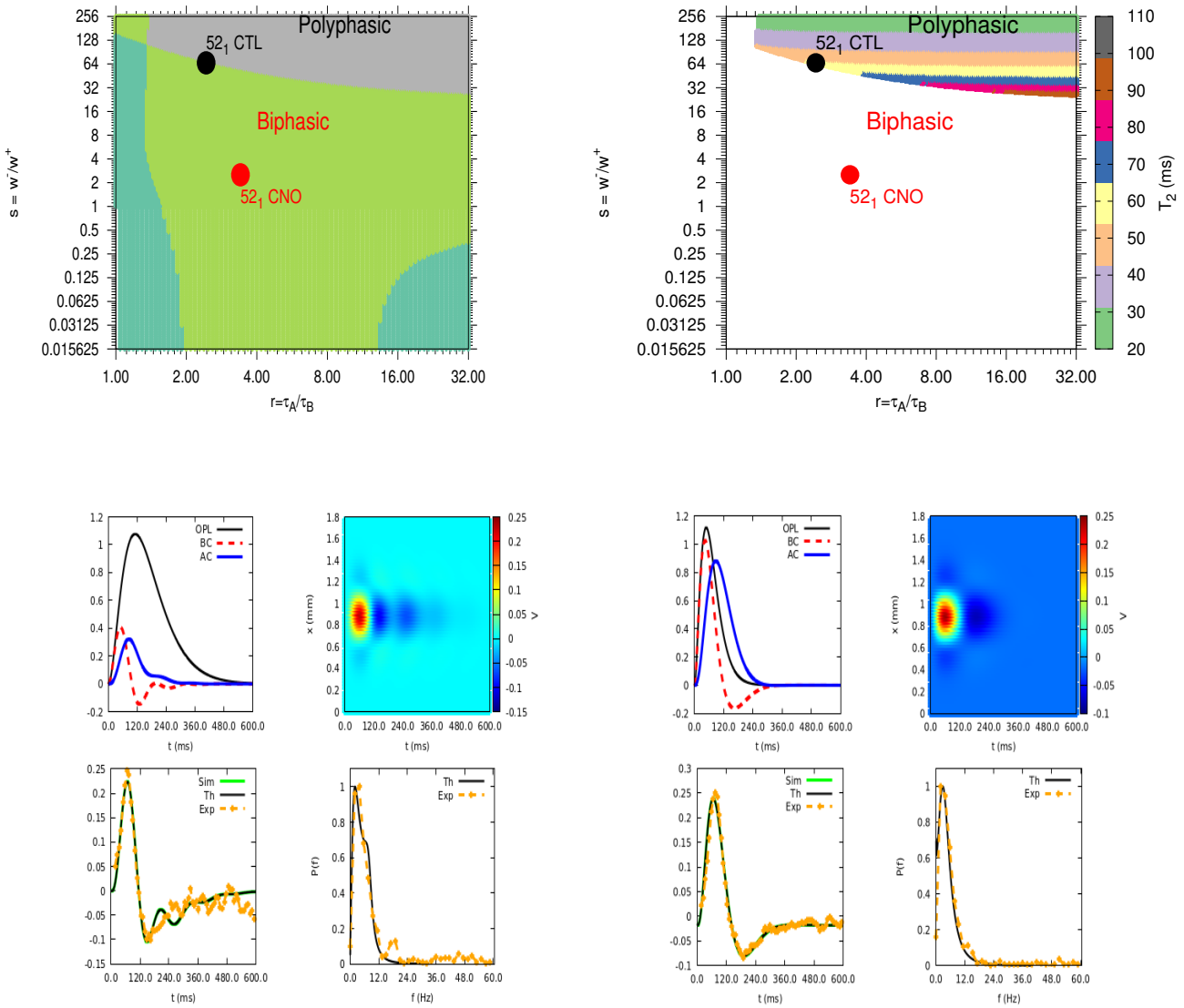

Figure 7: **CNO may change the cell response. Left top.** Application of CNO moves the representative point of the cell (here label 53_1) in from polyphasic to biphasic region resulting in a change in the cell response (**Bottom panel**). The figure **Right top** shows how the period of polyophasic oscillations depends on the place in the RFs map.

4 Discussion

In this paper we investigated the role of AC-mediated lateral connectivity in the response of RGCs to visual stimuli. Our conjecture was that these responses are strongly constrained by such lateral connectivity. Based on the mathematical analysis of a network featuring the interaction of BCs-ACs-RGCs we were able to produce an analytic form for the spatio-temporal response (receptive fields) of all cell types in the model. This finding has significant implications for the usefulness, identifiability (i.e. its parameters could be obtained from experimental data) and interpretability of the model. First, it provides an algorithmic way to fit the model parameters to the light responses recorded from mouse RGCs, using the analytical formula of the RF. This means that we are able not only to find the parameters that best fit the variables concerning the RGCs responses, but also to infer the possible behaviour of ACs and BCs leading to the RGCs responses, even if we don't measure them experimentally. Second, it provides an intuitive understanding of the role of various model variables and highlights the impact of two phenomenological parameters (with a physical meaning) on the spatio-temporal response (i), the intensity of the interactions BCs-ACs, and, (ii), the characteristic time scale of these cells response. This can be summarized in the two dimensional RF maps, where one observes phases corresponding to different modalities in the response. We were able to validate experimentally these modelling results, based on the ability to pharmacologically modify the level of ACs and RGCs neural activity using pharmacogenetics (DREADD-CNO). We would like now to comment some caveats and potential extensions of this work.

Polyphasic phase. Although we observe about 40% of polyphasic cells in the experimental plots (characterized by secondary peaks in the power spectrum) the model has

difficulty to properly fit them. This is visible in figure 5 where no cell is within the polyphasic region whereas the secondary peaks are clearly visible in the experimental power spectra (see Figure 7). This can be explained by several factors. First, the fitting method, trying to estimate a model with 10 parameters from a trace with a few hundreds of points, has clearly limits. In particular, the secondary peaks, having a few points in the power spectrum, are hard to capture and require a patient fine tuning. The main limitations may also come from the model itself, as developed in the following.

Connectivity. The nearest neighbors connectivity has been essentially chosen for mathematical convenience, but the realistic connectivity is certainly quite more complex. In particular, it depends on the cell subclasses and it remains a challenge to determine the effective connectivity experimentally. Note that, in addition to chemical synapses, roughly featured in the model, gap junctions are also present. Including them would essentially not change the structure of the model (see (Souihel and Cessac, 2019)).

Inhomogeneities. The model assumes that there is only one subclass for each cell type, represented by a unique set of parameters, while the real retinal network is quite more complex. The intrinsic heterogeneity of cell types and physiology can in particular explain the widespreading variety of parameters observed in the model and the fact that there is no systematic trend when applying CNO.

Rectification. The model includes weak non linearities (rectification) that were neglected in the mathematical computations. The effects of such rectification can be mathematically investigated (Cessac, 2022). Mainly, rectification projects dynamics on the subspace of non rectified cells. This means that the dimensionality of the dynam-

ical system changes in time, depending on the stimulus and network parameters, with strong consequences on the spectrum of $\mathscr{L}$, and thereby, on the power spectrum briefly discussed in the supplementary section 4.1.4.

Local non linearities vs network effects. Additional non linearities take place in retinal dynamics. Ion channels have a non linear behaviour inducing phenomena such as bifurcation and bursting, essential, for example, in the development of the retina were bursting Starburst Amacrine Cells generate retinal waves (Hennig et al., 2009; Karvouniari et al., 2019; Cessac and Matzakou-Karvouniari, 2022). In addition, gain control plays also a central role in the response to spatio-temporal stimuli inducing, for example, retinal anticipation (Berry et al., 1999; Chen et al., 2013; Souihel and Cessac, 2021).

All these key assumptions and simplifications, regarding the model structure and its parameters, are certainly crucial in shaping the RGC RFs. Yet, they made the model tractable mathematically and numerically, and aided its identifiability, i.e. its parameters could be obtained from experimental data, thus providing interesting insights in the retinal *network* dynamics.

Although our study was limited to responses to full-field flashes we would like to extend the consequence of our analysis to more complex stimuli. First, the presence of peaks in the power spectrum implies the existence of *resonances*, that is preferred frequencies for the RGCs. Exciting a cell with a resonance frequency will produce a maximal response. When applying a stimulus like the Chirp stimulus (Hilgen et al., 2022; Baden et al., 2016) there is a phase where periodic flashes, with constant contrast but increasing frequency, are applied. One observes frequently a bump in the experi-

mental RGCs response that might correspond to such a resonance.

In addition to preferred time frequencies, our analysis also shows that the response of RGCs, induced by the network, may also involve specific *space* scales. Mathematically, these scales appear in the eigenvectors of the transport operator (see eq. (37) in the Supplementary section 4.1.5). The practical implication would be that, presenting a local time periodic stimulus at a resonant frequency and with small radius, and increasing slowly this radius, one may observe scales where the response is maximal. One of these scales may correspond to the size of the RF but we conjecture that there should be other, larger, scales where this phenomenon appears. This would actually be a way to disentangle local, intrinsic responses of cells, to network induced responses, by blocking the ACs synapses (e.g. strychnine for glycinergic cells). More generally, the existence of time resonances and preferred space scales would also induce resonant response to moving objects with the appropriate speed. Such resonances effects could be involved in the mechanism generating anticipatory waves (Menz et al., 2020).

To conclude, this research supports the view that RGCs do not act independently, just to convey local spatio-temporal information, but rather encode visual information at the population level as well. We argue that this capability is due to the lateral inhibitory circuitry provided by the population of ACs, which creates a dense network connecting BCs, ACs and RGCs locally and globally. This has two implications for information processing in the retina. Firstly, RGCs diverse responses arise from a unique combination of network excitation and inhibition, which relies on the activities and interactions of upstream neurons projecting onto RGCs. This means that natural variation in RGC types cannot be only attributed to single cell characteristics, such as morphology, genetics, physiology, etc. Future studies should therefore include individual and

population properties, when exploring the functional diversity in RGCs. Second, altering the activity in any one neuron might affect the activity of any other neuron belonging to this network. Consequently, disrupting the balance of excitatory and inhibitory inputs within a given circuit, using e.g. genetic-based tools, allows to dismantle inner retinal circuits and understand how ACs shape retinal output. Our results are part of a rapidly growing body of literature, arguing that ACs hold a more universal role in retinal encoding, like parallel processing or motion anticipation (Franke and Baden, 2017; Souihel and Cessac, 2021).

Our work could easily be used in future studies to explore the role of other RGCs subclasses or other retinal neurons and their interactions. In addition, it could be used to disassemble the components of other retinal circuits, by manipulating the activity of specific neurons. It could also potentially benefit research in other parts of the nervous system, as fundamental properties of the inner retina are shared with other parts of the brain.

Supplementary

4.1 Mathematical analysis of network dynamics

4.1.1 Joint dynamics

The joint dynamics of all cells type is given by the dynamical system (3), that we rewrite here for convenience:

$$\begin{cases} \frac{dV_{B_i}}{dt} &= -\frac{1}{\tau_{B_i}} V_{B_i} + \sum_{j=1}^{N_A} W_{B_i}^{A_j} \mathscr{N}^{(A)}\left(V_{A_j}\right) + F_{B_i}(t), \quad i = 1 \dots N \\ \\ \frac{dV_{A_j}}{dt} &= -\frac{1}{\tau_{A_j}} V_{A_j} + \sum_{i=1}^{N_B} W_{A_j}^{B_i} \mathscr{N}^{(B)}\left(V_{B_i}\right) + \zeta_A, \quad j = 1 \dots N, \\ \\ \frac{dV_{G_k}}{dt} &= -\frac{1}{\tau_{G_k}} V_{G_k} + \sum_{i=1}^{N_B} W_{G_k}^{B_i} \mathscr{N}^{(B)}(V_{B_i}) + \sum_{j=1}^{N_A} W_{G_k}^{A_j} \mathscr{N}^{(A)}(V_{A_j}) + \zeta_G, \quad k = 1 \dots N. \end{cases}$$

We use Greek indices $\alpha, \beta, \gamma = 1 \dots 3N$ and define the state vector $\vec{\mathscr{X}}$ with entries:

$$\mathscr{X}_\alpha = \begin{cases} V_{B_i}, & \alpha = i, & i = 1 \dots N; \\ V_{A_j}, & \alpha = N + j, & j = 1 \dots N; \\ V_{G_k}, & \alpha = 2N + k, & k = 1 \dots N. \end{cases}$$

We introduce $\vec{\mathscr{F}}$ with entries:

$$\mathscr{F}_\alpha = \begin{cases} F_{B_i}, & \alpha = i, & i = 1 \dots N; \\ \zeta_A, & \alpha = N + j, & j = 1 \dots N; \\ \zeta_G, & \alpha = 2N + k, & k = 1 \dots N; \end{cases}$$

and the rectification vector $\vec{\mathscr{R}}(\vec{\mathscr{X}})$ with entries:

$$\mathscr{R}_\alpha(\vec{\mathscr{X}}) = \begin{cases} \mathscr{N}^{(B)}(V_{B_i}), & \alpha = i, \quad i = 1\ldots N; \\ \mathscr{N}^{(A)}\left(V_{A_j}\right), & \alpha = N+j, \quad j = 1\ldots N; \\ 0, & \alpha = 2N+k, \quad k = 1\ldots N; \end{cases}$$

We introduce the $N \times N$ matrices:

$$\mathscr{T} = \begin{pmatrix} -\operatorname{diag}\left[\tau_{B_i}\right]_{i=1\ldots N} & 0_{NN} & 0_{NN} \\ 0_{NN} & -\operatorname{diag}\left[\tau_{A_j}\right]_{j=1\ldots N} & 0_{NN} \\ 0_{NN} & 0_{NN} & -\operatorname{diag}\left[\tau_{G_k}\right]_{k=1\ldots N} \end{pmatrix}, \quad (18)$$

characterizing the characteristic integration times of cells,

$$\mathscr{W} = \begin{pmatrix} 0_{NN} & W_B^A & 0_{NN} \\ W_A^B & 0_{NN} & 0_{NN} \\ W_G^B & W_G^A & 0_{NN} \end{pmatrix}, \quad (19)$$

summarizing chemical synapses interactions. Then, the dynamical system (3) reads, in vector form:

$$\frac{d\vec{\mathscr{X}}}{dt} = \mathscr{T}^{-1}.\vec{\mathscr{X}} + \mathscr{W}.\vec{\mathscr{R}}(\vec{\mathscr{X}}) + \vec{\mathscr{F}}(t). \quad (20)$$

We remark that eq. (20) has a specific product structure: the dynamics of RGCs is driven by BCs and ACs with no feedback. This means that one can study first the coupled dynamics of BCs and ACs and then the effect on RGCs.

4.1.2 Linear evolution

We consider the evolution of eq. (20) from an initial time t_0. Typically, t_0 is a reference time where the network is at rest, before the stimulus is applied. The dynamical system has almost the form of a non-autonomous linear system driven by the term $\vec{\mathscr{F}}(t)$. There is however a weak non linearity, due to the piecewise linear rectification appearing in the term $\vec{\mathscr{R}}(\vec{\mathscr{X}})$. Therefore, when all cells voltage are large enough the system is linear. Mathematically, there is a domain of $\mathbb{R}^{3N}$:

$$\mathscr{D} = \left\{ V_{B_i} \geq \theta_B, V_{A_j} \geq \theta_A, i, j = 1 \dots N \right\}, \tag{21}$$

where $R\left(\vec{\mathscr{X}}\right)$ is linear so that eq. (20) is linear too (check (Cessac, 2022) for more details).

From on we consider this linear case. We write $\mathscr{L} = \mathscr{T}^{-1}.\vec{\mathscr{X}} + \mathscr{W}$ so that:

$$\mathscr{L} = \begin{pmatrix} -\operatorname{diag}\left[\frac{1}{\tau_{B_i}}\right]_{i=1\dots N} & W_B^A & 0_{NN} \\ W_A^B & -\operatorname{diag}\left[\frac{1}{\tau_{A_j}}\right]_{j=1\dots N} & 0_{NN} \\ W_G^B & W_G^A & -\operatorname{diag}\left[\frac{1}{\tau_{G_k}}\right]_{k=1\dots N} \end{pmatrix}, \tag{22}$$

We introduce the N dimensional vector $\vec{1}_N = \vec{(}1)_{i=1}^N$, and the $3N$ dimensional vector $\vec{\mathscr{C}} = \begin{pmatrix} -\theta_A W_B^A.\vec{1}_N \\ -\theta_B W_A^B.\vec{1}_N \\ -\left(\theta_B W_G^B.\vec{1}_N + \theta_A W_G^A\right).\vec{1}_N \end{pmatrix}$ and (20) reads $\frac{d\vec{\mathscr{X}}}{dt} = \mathscr{L}.\vec{\mathscr{X}} + \vec{\mathscr{F}}(t) + \vec{\mathscr{C}}$.

We assume that $\mathscr{L}$ is invertible. This assumption, and more generally, the spectrum

of $\mathscr{L}$ is further discussed in section 4.1.4. The general solution of eq. (20) is:

$$\vec{\mathscr{X}}(t) = e^{\mathscr{L}(t-t_0)}.\vec{\mathscr{X}}(t_0) + \int_{t_0}^{t} e^{\mathscr{L}(t-s)}\,\vec{\mathscr{F}}(s)\,ds - \mathscr{L}^{-1}.\left[\mathscr{I}_{3N,3N} - e^{\mathscr{L}(t-t_0)}\right].\vec{\mathscr{C}}. \tag{23}$$

where $\mathscr{I}_{3N,3N}$ is the $3N$ dimensional identity matrix.

Although this equation is general, it actually stands when one can define a notion of asymptotic regime. That is, when $\mathscr{L}$ has stable eigenvalues (eigenvalues with a strictly negative real part). The spectrum of $\mathscr{L}$ is studied below and conditions ensuring the stability of eigenvalues are given. Here, we are going to assume that eigenvalues are all stable and that $t - t_0$ is large so that we can remove the transient term $e^{\mathscr{L}(t-t_0)}.\vec{\mathscr{X}}(t_0)$ depending on the initial condition $\vec{\mathscr{X}}(t_0)$. In addition, the last term converges to:

$$\vec{\mathscr{X}}^{*} = -\mathscr{L}^{-1}.\vec{\mathscr{C}}, \tag{24}$$

the rest state of the linear system, which vanishes whenever the thresholds θ_A, θ_B are set to 0.

4.1.3 Derivation of eq. (7)

We note the eigenvalues of $\mathscr{L}$, $\lambda_\beta, \beta = 1 \ldots 3N$ and its eigenvectors, $\mathscr{P}_\beta$ (the columns of the matrix $\mathscr{P}$ transforming $\mathscr{L}$ in diagonal form). We consider first the case $\vec{\mathscr{C}} = \vec{0}$. We have then, from (23):

$$\mathscr{X}_\alpha(t) = \sum_{\beta=1}^{3N}\sum_{\gamma=1}^{3N} \mathscr{P}_{\alpha\beta}\,\mathscr{P}^{-1}_{\beta\gamma}\int_{-\infty}^{t} e^{\lambda_\beta(t-s)}\,\mathscr{F}_\gamma(s)\,ds,$$

where $\mathscr{F}_\gamma = F_{B_i}$, $\gamma = i = 1 \ldots N$ (BCs).

We recall that, from (5), $F_{B_\gamma}(t) = \frac{\mathscr{V}_\gamma^{(drive)}}{\tau_{B_\gamma}} + \frac{d\mathscr{V}_\gamma^{(drive)}}{dt}$, so that:

$$\int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \mathscr{F}_\gamma(s)\, ds = \mathscr{V}_\gamma^{(drive)}(t) + \varpi_{\beta\gamma} \int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \mathscr{V}_\gamma^{(drive)}(s)\, ds, \quad \gamma = 1 \ldots N,$$

for B cells, with $\varpi_{\beta\gamma} = \frac{1}{\tau_{B_\gamma}} + \lambda_\beta$ and using $\mathscr{V}_\gamma^{(drive)}(-\infty) = 0$.

For $\gamma = N+1 \ldots 2N$, $\mathscr{F}_\gamma = \zeta_A$ (ACs) we have:

$$\int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \zeta_A\, ds = -\frac{1}{\lambda_\beta} \left[1 - e^{\lambda_\beta (t-t_0)} \right] \zeta_A.$$

Finally, for $\gamma = 2N+1 \ldots 3N$, $\mathscr{F}_\gamma = \zeta_G$ (RGCs):

$$\int_{-\infty}^{t} e^{\lambda_\beta (t-s)}\, \zeta_G\, ds = -\frac{1}{\lambda_\beta} \left[1 - e^{\lambda_\beta (t-t_0)} \right] \zeta_G.$$

We remark that:

$$\begin{aligned}
\sum_{\beta=1}^{3N} \sum_{\gamma=1}^{3N} \mathscr{P}_{\alpha\beta}\, \mathscr{P}_{\beta\gamma}^{-1}\, \mathscr{V}_\gamma^{(drive)}(t) &= \sum_{\gamma=1}^{3N} \mathscr{V}_\gamma^{(drive)}(t) \left(\sum_{\beta=1}^{3N} \mathscr{P}_{\alpha\beta}\, \mathscr{P}_{\beta\gamma}^{-1} \right) \\
&= \sum_{\gamma=1}^{3N} \mathscr{V}_\gamma^{(drive)}(t)\, \delta_{\alpha\gamma} \\
&= \mathscr{V}_\alpha^{(drive)}(t).
\end{aligned}$$

It follows that:

$$\mathscr{X}_\alpha(t) = \mathscr{V}_\alpha^{(drive)}(t) + \mathscr{E}_\alpha^{(drive)}(t) + \mathscr{E}_\alpha^{(CNO_A)} + \mathscr{E}_\alpha^{(CNO_G)}, \quad \alpha = 1 \ldots 3N.$$

with:

$$\mathscr{E}_\alpha^{(drive)}(t) = \sum_{\beta=1}^{3N}\sum_{\gamma=1}^{N} \mathscr{P}_{\alpha\beta}\,\mathscr{P}^{-1}_{\beta\gamma}\varpi_{\beta\gamma}\int_{-\infty}^{t} e^{\lambda_\beta(t-s)}\,\mathscr{V}_\gamma^{(drive)}(s)\,ds, \quad \alpha = 1\ldots 3N,$$

$$\mathscr{E}_\alpha^{(CNO_A)} = -\,\zeta_A \sum_{\beta=1}^{3N}\sum_{\gamma=N+1}^{2N} \frac{\mathscr{P}_{\alpha\beta}\,\mathscr{P}^{-1}_{\beta\gamma}}{\lambda_\beta}, \alpha = N+1\ldots 2N;$$

$$\mathscr{E}_\alpha^{(CNO_G)} = \zeta_G\,\tau_{G_\alpha}, \quad \alpha = 2N+1\ldots 3N.$$

which is (7).

When $\mathscr{C} \neq \vec{0}$, there is an additional term corresponding to the rest state (24).

4.1.4 Eigenvalues and eigenvectors of $\mathscr{L}$

In the following, the term $\mathrm{diag}\,(x_n)_{n=1}^N$ denotes a diagonal $N \times N$ matrix with diagonal entries x_n.

Linear case. We start from the eq. (22) of the linear operator ruling the dynamics in the set $\mathscr{D}$ defined by (21). We consider, as in the main text, the case where all characteristic times τ_{B_i} are equal to τ_B, all characteristic times τ_{A_j} are equal to τ_A and all characteristic times τ_{G_k} are equal to τ_G. Using the same notations as the main text we have:

$$\mathscr{L} = \begin{pmatrix} -\frac{\mathscr{I}_{NN}}{\tau_B} & -w^-\,\Gamma_B^A & 0_{NN} \\ w^+\,\Gamma_A^B & -\frac{\mathscr{I}_N}{\tau_A} & 0_{NN} \\ w_G^B\,\Gamma_G^B & w_G^A\,\Gamma_G^A & -\frac{\mathscr{I}_{NN}}{\tau_G} \end{pmatrix},$$

where 0_{NN} is the $N \times N$ 0 matrix and I_{NN} the $N \times N$ 0 identity matrix.

Eigenvalues and eigenvectors. We consider the case where $\Gamma_B^A = \Gamma_A^B$. We note $\kappa_n, n = 1 \ldots N$, the eigenvalues of Γ_B^A ordered as $|\kappa_1| \leq |\kappa_2| \leq \cdots \leq |\kappa_n|$ and note the normalized eigenvectors $\vec{\psi}_n$, $n = 1 \ldots N$.

We seek the eigenvalues, λ_β, and eigenvectors, $\vec{\mathscr{P}}_\beta$, $\beta = 1 \ldots 3N$, of $\mathscr{L}$. It is evident, from the form of $\mathscr{L}$, that there are N eigenvalues $\lambda_\beta = -\frac{1}{\tau_G}, \vec{\mathscr{P}}_\beta = \vec{e}_\beta$ where $\vec{e}_\beta$ is the canonical basis vector in direction β. We attribute them the indices $\beta = 2N+1 \ldots 3N$ as this indexing corresponds to the form of $\mathscr{L}$ when $w^- = w^+ = 0$. We seek the $2N$ remaining eigenvalues-eigenvectors assuming that $\vec{\mathscr{P}}_\beta$s has the form:

$$\vec{\mathscr{P}}_\beta = \begin{pmatrix} \vec{\psi}_n \\ \rho_n \vec{\psi}_n \\ \vec{\phi}_n \end{pmatrix}, \quad n = 1 \ldots N, \tag{25}$$

where ρ_n is an unknown parameter and $\vec{\phi}_n$ a N dimensional vector, to be determined, from the characteristic equation:

$$\mathscr{L} . \vec{\mathscr{P}}_\beta = \lambda_\beta \vec{\mathscr{P}}_\beta .$$

This leads to the system of equations:

$$\begin{cases} \lambda_\beta & = & -\frac{1}{\tau_B} - w^- \rho_n \kappa_n ; \\ w^+ \kappa_n - \frac{\rho_n}{\tau_A} & = & \rho_n \lambda_\beta ; \\ \lambda_\beta \vec{\phi}_n & = & \left(w_G^B \Gamma_G^B + w_G^A \rho_n \Gamma_G^A \right) \vec{\psi}_n - \frac{1}{\tau_G} \vec{\phi}_n , \end{cases} , \tag{26}$$

We first assume that $w^-, w^+ > 0$ and later discuss the limit when these quantities tend to zero. Combining the two first equations leads to a second-order polynomial in the

variable ρ_n,

$$w^- \kappa_n \rho_n^2 - \frac{1}{\tau}\rho_n + w^+ \kappa_n = 0, \tag{27}$$

giving 2 solutions for each n:

$$\rho_n^{\pm} = \begin{cases} \frac{1}{2\tau w^- \kappa_n}\left(1 \pm \sqrt{1-4\mu\,\kappa_n^2}\right), & \kappa_n \neq 0, \frac{1}{\tau} \neq 0; \\ w^+ \tau\, \kappa_n, & \kappa_n = 0, \frac{1}{\tau} \neq 0; \\ \pm\sqrt{-\frac{w^+}{w^-}}, & \frac{1}{\tau} = 0. \end{cases} \tag{28}$$

where:

$$\frac{1}{\tau} = \frac{1}{\tau_A} - \frac{1}{\tau_B}. \tag{29}$$

and:

$$\mu = w^- w^+ \tau^2 \geq 0. \tag{30}$$

The $2N$ first eigenvalues of $\mathscr{L}$ are therefore given by:

$$\lambda_n^{\pm} = \begin{cases} -\frac{1}{2\tau_{AB}} \mp \frac{1}{2\tau}\sqrt{1-4\mu\,\kappa_n^2}, & \frac{1}{\tau} \neq 0; \\ \\ -\frac{1}{\tau_A} \mp \sqrt{-w^- w^+ \kappa_n^2}, & \frac{1}{\tau} = 0. \end{cases} \tag{31}$$

where:

$$\frac{1}{\tau_{AB}} = \frac{1}{\tau_A} + \frac{1}{\tau_B}. \tag{32}$$

We finally obtain $2N$ vectors $\vec{\phi}_n$:

$$\vec{\phi}_n^{\pm} = \frac{1}{\lambda_n^{\pm} + \frac{1}{\tau_G}}\left(w_G^B \Gamma_G^B + w_G^A \rho_n^{\pm} \Gamma_G^A\right)\vec{\psi}_n. \tag{33}$$

Let us now discuss the limit when w^- or w^+ or both tend to 0. If $w^- = 0$, $\rho_n =$

$w^+\kappa_n\tau$ from (27). If $w^+=0$ there are two solutions of (27), $\rho_n=0$ or $\rho_n=\frac{1}{\tau w^-\kappa_n}$. Finally, when $w^-=w^+=0$, $\rho_n=0$ and the ansatz (25) does not apply. Actually, in this case, $\mathscr{L}$ is diagonal, the N first eigenvalues are $-\frac{1}{\tau_B}$, the N next eigenvalues are $-\frac{1}{\tau_A}$. We have, in this case: $\lambda_n^+=-\frac{1}{\tau_B}$ and $\lambda_n^-=-\frac{1}{\tau_A}$. Therefore, we order eigenvalues and eigenvectors of $\mathscr{M}$ such that the N first eigenvalues are $\lambda_n^+, n=1\dots N$, and the N next are $\lambda_n^-, n=N+1\dots 2N$.

We finally end up with the following form for the eigenvalues and eigenvectors of $\mathscr{L}$:

$$
\begin{aligned}
\lambda_\beta=\lambda_n^+, \vec{\mathscr{P}}_\beta &= \begin{pmatrix} \vec{\psi}_n \\ \rho_n^+\vec{\psi}_n \\ \frac{1}{\lambda_n^+ + \frac{1}{\tau_G}}\left(w_G^B\Gamma_G^B + w_G^A\rho_n^+\Gamma_G^A\right)\vec{\psi}_n \end{pmatrix}, \quad \beta=n=1\dots N, \\
\lambda_\beta=\lambda_n^-, \vec{\mathscr{P}}_\beta &= \begin{pmatrix} \vec{\psi}_n \\ \rho_n^-\vec{\psi}_n \\ \frac{1}{\lambda_n^- + \frac{1}{\tau_G}}\left(w_G^B\Gamma_G^B + w_G^A\rho_n^-\Gamma_G^A\right)\vec{\psi}_n \end{pmatrix}, \quad \beta=N+1\dots 2N, \quad n=1\dots N, \\
\lambda_\beta=-\tfrac{1}{\tau_G}, \vec{\mathscr{P}}_\beta &= \vec{e}_\beta, \qquad\qquad \beta=2N+1\dots 3N.
\end{aligned}
\tag{34}
$$

Skeleton. The eigenvalues $\lambda_n^\pm$ in 31 can be real or complex conjugated. By increasing μ, they become complex when:

$$
\mu > \frac{1}{4\,\kappa_n} \equiv \mu_{n,c}. \tag{35}
$$

In this case the real part is $-\frac{1}{2\,\tau_{AB}}$, the imaginary part is $\pm\frac{1}{2\,\tau}\sqrt{1-4\,\mu\,\kappa_n^2}$. If $\mu\leq\mu_{n,c}$, eigenvalues λ_β are real with a negative real part. This implies that the linear dynamical

system (20) is stable.

The N equations (35) define what we call the "skeleton of the RFs map". In the main text, we introduced the quantities $r = \frac{\tau_A}{\tau_B}, s = \frac{w^-}{w^+}$. In these variables, the critical condition (35) reads:

$$s_{n_c} = \frac{1}{w^{+2}\,\tau_B^2}\,\frac{1}{4\,\kappa_n^2}\,\frac{(1-r)^2}{r^2}. \tag{36}$$

This defines two critical lines symmetric with respect to $r = 1$. These lines are invariant by the variable change $r' = r, s' = \left(\frac{w'^+\,\tau'_B}{w^+\,\tau_B}\right)^2 s$. This allows to map the skeleton obtained from a set of values τ'_B, w'^+ to the skeleton obtained with references value τ_B, w^+.

Rectification. In this paper we have essentially considered a situation where cells are not rectified, whereas the full model (eq. (3)) considers rectification, in agreement with realistic biological systems. The mathematical effect of rectification of a cell A_j, is to set to zero the corresponding row in the matrix W_B^A. This has several consequences. First, we cannot apply anymore the useful Ansatz used in the section, that is $\Gamma_B^A = \Gamma_A^B$. In addition, the vanishing of only one row in $\mathscr{L}$ completely modifies its spectrum. However, thresholding in rectification corresponds to partition the phase space of the model, a compact subset of $\mathbb{R}^{3N}$, into convex subdomains delimited by hyperplanes. In each of these domains the matrix $\mathscr{W}.\vec{\mathscr{R}}(\vec{\mathscr{X}})$ appearing in eq. (20) is linear with a number of zero eigenvalues corresponding to the number of rectified cells. This matrix acts as a projector on the complementary subspace of its kernel. In each of these subdomains eq. (7) applies. One can actually compute, for a given stimulus, the time of entrance and exit in a new subdomain with the effect of modifying the eigenvalues and eigenvectors appearing in eq. (7). The resulting equation is quite complex though and will require further investigations. See (Cessac, 2020) for more details.

4.1.5 Nearest neighbours connectivity

We consider the case where the connectivity matrices $\Gamma_A^B = \Gamma_B^A$ have nearest neighbours symmetric connections. We also assume that the dynamics hold on a square lattice with null boundary conditions. We define $\alpha = i_x \in \{1 \dots L = N\}$ in one dimension and $\alpha = i_x + (i_y - 1)L \in \{1 \dots L^2 = N\}$ in two dimensions. We also set $n = n_x \in \{1 \dots L = N\}$ in one dimension and $n = n_x + (n_y - 1)L \in \{1 \dots L^2 = N\}$ in two dimensions. Then, the eigenvectors and eigenvalues of these matrices have the form:

$$\begin{aligned} \psi_{\alpha,n} &= \left(\tfrac{2}{L+1}\right)^{\frac{d}{2}} \prod_l \sin\left(\tfrac{n_l \pi}{L+1} i_l\right), \\ \kappa_n &= 2 \sum_l \left[\cos\left(\tfrac{n_l \pi}{L+1}\right)\right]; \end{aligned} \tag{37}$$

with $l = x$ for $d = 1$ and $l = x, y$ for $d = 2$. Especially, in one dimension:

$$\psi_{\alpha,n} = \sqrt{\frac{2}{L+1}} \sin\left(\frac{\alpha\pi}{L+1} n\right),$$

The quantum numbers (n_x, n_y) define a wave vector $\vec{k}_n = \left(\tfrac{n_x \pi}{L+1}, \tfrac{n_y \pi}{L+1}\right)$ corresponding to wave lengths $\left(\tfrac{L+1}{n_x}, \tfrac{L+1}{n_y}\right)$. Hence, the first eigenmode $(n_x = 1, n_y = 1)$ corresponds to the largest space scale (scale of the whole retina) with the smallest eigenvalue (in absolute value) $s_{(1,1)} = 2\left(\cos\left(\tfrac{\pi}{L+1}\right) + \cos\left(\tfrac{\pi}{L+1}\right) - 2\right)$. Each of these eigenmodes is related to a characteristic time $\tau_n = \frac{1}{\lambda_n}$.

Eigenvalues κ_n can be positive or negative. However, from eq. (31) this has no impact on the eigenvalues as what matters is κ_n^2. This induces however a symmetry $\kappa_n \to -\kappa_n$ that can be seen in the skeleton Figure 4, not forgetting that this figure is in log scale.

References

Asari, H. and Meister, M. (2012). Divergence of visual channels in the inner retina. Nature Neuroscience.

Baden, T., Berens, P., Franke, K., Román Rosón, M., Bethge, M., and Euler, T. (2016). The functional diversity of retinal ganglion cells in the mouse. Nature, 529(7586):345–350.

Berry, M., Brivanlou, I., Jordan, T., and Meister, M. (1999). Anticipation of moving stimuli by the retina. Nature, 398(6725):334—338.

Besharse, J. and Bok, D. (2011). The Retina and its Disorders. Elsevier Science.

Cessac, B. (2020). Recent Trends in Chaotic, Nonlinear and Complex Dynamics, in honour of Prof. Miguel A.F. Sanjuán on his 60th Birthday., chapter The retina as a dynamical system. World Scientific.

Cessac, B. (2022). Retinal processing: Insights from mathematical modelling. Journal of Imaging, 8(1).

Cessac, B. and Matzakou-Karvouniari, D. (2022). The non linear dynamics of retinal waves. Physica D: Nonlinear Phenomena, 439:133436.

Chalupa, L. M. and Werner, J., editors (2004). The visual neurosciences. MIT Press. Two volumes.

Chen, E. Y., Marre, O., Fisher, C., Schwartz, G., Levy, J., da Silviera, R. A., and Berry, M. (2013). Alert response to motion onset in the retina. Journal of Neuroscience, 33(1):120–132.

Daw, N. (2012). How Vision Works: The Physiological Mechanisms Behind What We See. Oxford University Press, USA.

Demb, J. B. and Singer, J. H. (2015). Functional Circuitry of the Retina, volume 1.

Diamond, J. S. (2017). Inhibitory Interneurons in the Retina: Types, Circuitry, and Function. Annual Review of Vision Science, 3:1–24.

Euler, T., Haverkamp, S., Schubert, T., and Baden, T. (2014). Retinal bipolar cells: Elementary building blocks of vision.

Franke, K. and Baden, T. (2017). General features of inhibition in the inner retina. Journal of Physiology, 595(16):5507–5515.

Gollisch, T. and Meister, M. (2010). Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. Neuron, 65(2):150–164.

Hennig, M., Adams, C., Willshaw, D., and Sernagor, E. (2009). Early-stage waves in the retinal network emerge close to a critical state transition between local and global functional connectivity. J Neurosci.

Hilgen, G., Kartsaki, E., Kartysh, V., Cessac, B., and Sernagor, E. (2022). A novel approach to the functional classification of retinal ganglion cells. Open Biology, 12(3).

Karvouniari, D., Gil, L., Marre, O., Picaud, S., and Cessac, B. (2019). A biophysical model explains the oscillatory behaviour of immature starburst amacrine cells. Scientific Reports, 9:1859.

Marr, D. (1982). Vision. W.H. Freeman and Co.

Masland, R. H. (2012). The tasks of amacrine cells. Visual neuroscience.

Menz, M. D., Lee, D., and Baccus, S. A. (2020). Representations of the amacrine cell population underlying retinal motion anticipation. bioRxiv.

Pamplona, D., Hilgen, G., Hennig, M., Cessac, B., Sernagor, E., and Kornprobst, P. (2021). Large visual neuron assemblies receptive fields estimation using a super-resolution approach.

Roth, B. L. (2016). DREADDs for Neuroscientists. Neuron, 89(4):683–694.

Sanes, J. R. and Masland, R. H. (2015). The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification. Annual Review of Neuroscience, 38(1):221–246.

Schröder, C., Klindt, D., Strauss, S., Franke, K., Bethge, M., Euler, T., and Berens, P. (2020). System identification with biophysical constraints: A circuit model of the inner retina. bioRxiv, 2020.06.16.154203.

Souihel, S. and Cessac, B. (2019). Anticipation in the retina and the primary visual cortex : towards an integrated retino-cortical model for motion processing. In ICMNS 2019 - The 5th International Conference on Mathematical NeuroScience, Copenhague, Denmark.

Souihel, S. and Cessac, B. (2021). On the potential role of lateral connectivity in retinal anticipation. J. Math. Neurosc., 11(3).

Urban, D. J. and Roth, B. L. (2015). DREADDs (Designer Receptors Exclusively Activated by Designer Drugs): Chemogenetic Tools with Therapeutic Utility. Annual Review of Pharmacology and Toxicology, 55(1):399–417.

Wässle, H. (2004). Parallel processing in the mammalian retina.

Werblin, F. S. and Dowling, J. E. (1969). Organization of the retina of the mudpuppy, Necturus maculosus. II. Intracellular recording. Journal of neurophysiology.