POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Stimulus Optimization for Retinal Ganglion Cell Characterization

Tesi di Laurea Magistrale in
Biomedical Engineering - Ingegneria Biomedica

Author: **Melide Crippa**

Student ID: 992772
Advisor: Prof. Riccardo Barbieri
Co-advisors: Dr. Matthew Chalk
Academic Year: 2022-23

Abstract

Retinal Ganglion Cells (RGCs) are the output neurons of the retina and send signals down the optic nerve. There are many different types of RGC, which respond in a diverse way to visual stimuli. Their classification based on the responses to a visual stimulus takes the name of functional RGCs typing. Previous methods demonstrated that a good typing can be performed by exploiting the responses of the RGCs to a stimulus composed of disks of varying sizes centered in the receptive field of the cell. This method requires presenting a different stimulus to each neuron, which is difficult to do in high-throughput experiments, where many cells are recorded simultaneously using a multi-electrode array. An alternative solution proposed here is to show the same stimulus to many cells simultaneously and use these data to train a model of the RGCs able to predict responses to the disks. The specific challenge we address is to understand which stimulus to use to best learn the model. We use gaussian processes to fit cells responses to different stimuli and we quantify how good they are at predicting the responses to disks. The performance of the model reflects the goodness of the stimulus, and this measure can be used to optimize the stimulus. The procedure has been applied on the commonly used checkerboard stimulus and we were able to identify the optimal stimulus for the characterization of the responses of RGCs to the disks as a temporal sequence of checkerboards containing different check sizes in equal proportion. In the future this method can be used to improve the functional typing of RGCs.

Keywords: Gaussian Process, Stimulus Optimization, Cell Typing, Retinal Ganglion Cells

Abstract in lingua italiana

Le cellule gangliari retiniche (RGC) sono i neuroni di uscita della retina e inviano segnali lungo il nervo ottico. Esistono svariati tipi di RGC, i quali rispondono in modo diverso agli stimoli visivi. La loro classificazione sulla base delle risposte a uno stimolo visivo prende il nome di tipizzazione funzionale delle RGC. Studi precedenti hanno dimostrato che è possibile eseguire una buona tipizzazione analizzando le risposte delle RGC a uno stimolo composto da dischi di varie dimensioni centrati nel campo recettivo della cellula. Questo metodo richiede la presentazione di un diverso stimolo a ogni neurone, ma ciò è difficilmente realizzabile in esperimenti ad alto rendimento, in cui si effettua la registrazione simultanea di svariate cellule utilizzando un array a multi-elettrodi. Una soluzione alternativa qui proposta è quella di mostrare lo stesso stimolo a molte cellule contemporaneamente e utilizzare questi dati per addestrare un modello delle RGC in grado di predire le risposte ai dischi. Lo specifico problema qui affrontato è capire quale stimolo utilizzare per apprendere al meglio il modello. Abbiamo utilizzato dei processi gaussiani per modellizzare le risposte delle cellule a diversi stimoli e in seguito abbiamo quantificato la loro capacità di predire le risposte ai dischi. Le prestazioni del modello riflettono la bontà dello stimolo, e questa misura può essere utilizzata per ottimizzare lo stimolo. Questa procedura è stata applicata a uno stimolo a scacchiera e siamo stati in grado di identificare lo stimolo ottimale per la caratterizzazione delle risposte delle RGC ai dischi in una sequenza temporale di scacchiere contenente quadrati di dimensioni diverse nella stessa proporzione. In futuro questo metodo potrà essere utilizzato per migliorare la tipizzazione funzionale delle RGC.

Parole chiave: Processi Gaussiani, Ottimizzazione dello Stimolo, Tipizzazione Cellulare, Cellule Gangliari della Retina

Contents

Introduction

Retinal Ganglion Cells (RGCs) are a type of neuron located in the the retina, near the optic nerve head. They play a crucial role in the process of vision: in fact, RGCs serve as the output neurons of the retina, transmitting visual signals to the brain, integrating and conveying information about the visual scene. There are different subtypes of retinal ganglion cells, each sensitive to specific aspects of the visual input, such as contrast, motion, or spatial details. This diversity allows the visual system to process a wide range of visual information. Understanding the properties and functions of retinal ganglion cells is essential for unraveling the mechanisms of vision and the processing of visual information in the brain. A particularly significant task in this regard is the classification of these cells based on their functional types, representing a major undertaking in computational neuroscience. Considerable research efforts have been dedicated to this pursuit in recent years, aiming to uncover the ways in which retinal ganglion cells contribute to visual processing and perception.

An exemplary study on cell typing is the work by Baden et al. in 2016 [1]. The approach taken involved typing multiple cells simultaneously by stimulating the retina with various full-field stimuli, i.e. stimuli covering the entire field of view of the retina. However, a limitation of this method is that it tends to obscure differences in the spatial selectivity of individual cells. In contrast, Goetz et al. in 2022 [2] addressed this limitation by employing disks of varying sizes, centered within the receptive field of each single cell as stimuli. This approach allowed them to explore distinctions that were not apparent when eliciting cells with a full-field stimulus. For more details about these studies, refer to Section 1.3.

The first crucial takeaway from these examples is the impact of using different stimuli on the outcomes of cell typing. This raises the important question of determining the optimal stimulus for effective cell typing. Additionally, in both studies, stimuli were manually selected based on prior knowledge of retinal ganglion cell characteristics and the specific task. What if we want to perform the choice of this stimulus automatically? Automating stimulus selection can potentially enhance the objectivity and reproducibility of cell typing, allowing for broader applicability across diverse datasets and experimental

conditions.

A second important fact to highlight is that Goetz et al. demonstrated in their study that the precise localization of stimuli in relation to the receptive field of a cell is crucial for effective cell typing. They emphasized the importance of studying cell responses to disk stimuli of varying sizes centered within their respective receptive fields. However, a challenge arises with this approach, as it necessitates a multitude of distinct stimuli to elicit responses from different neurons. This becomes impractical in the context of the experimental setup employed in our project, i.e. large-scale recordings with a Multi-Electrode array [3], that allows to simultaneously record the activity of many different neurons. A potential solution is to show the same stimulus, such as a checkerboard, to all the cells and record all the responses simultaneously. Subsequently, based on the eliciting stimuli and the responses observed during the experiment, it becomes feasible to train a model that mimics the behavior of the RCGs. This trained model can then be utilized to predict responses to disk stimuli, eliminating the need to collect them in highly time-consuming experiments.

The idea of this thesis work originates from the challenges outlined above, and aims at developing a method for the identification of the optimal temporal checkerboard stimulus to be shown to the retina for the characterization of RGCs responses to the disks stimulus. I have combined the advantages emerged in the two studies, namely the higher throughput in the first one [1] and the maintenance of the spatial selectivity in the second [2] developing a model of the retina, trained on experimental data, that is able to accurately predict responses to checkerboards and disks stimuli. Then, I have addressed the problem of stimulus optimization, using the first model to generate synthetic responses to a discrete set of temporal checkerboard stimuli, exploring different variations, such as the dimension of the checks in the checkerboards and the percentage of presence of black checks. I have then exploited these synthetic pairs stimuli-responses in the training of several new models of the cells. Evaluating the performance of these models, that are dependent on the stimuli used for the training, it is possible to gain insights on the impact of the stimuli on the performance in the characterization of RCG responses to the disks.

1 | Background and Context

In this chapter, a general overview of the context in which we are operating is presented. First of all, a description of some common RCG neural modelling methods is provided, followed by a focus on some stimulus optimization techniques developed in other studies. Finally, the functional cell typing problem is addressed through a broader analysis of the studies performed by Baden et al. in 2016 [1] and by Goetz et al. in 2022 [2].

1.1. Neural Modelling of Retinal Ganglion Cells

Neural modelling is a multidisciplinary field that employs computational and mathematical frameworks to simulate and understand the complex behavior of biological neurons. It involves the development of mathematical models that mimic the functioning of the nervous system, capturing the interactions between individual neurons. Neural models aim to explain how biological systems process information and ultimately give rise to various cognitive functions. Regarding more specifically neural modelling in the retina, this kind of models has the goal to replicate the behavior of retinal neurons, such as photoreceptors, bipolar cells and ganglion cells, to understand how visual information is encoded and processed.

One first family of models that is extensively employed in neuroscience is the one of Linear - Non Linear Poisson (LNP) models, a type of Generalized Linear Model (GLM) specifically utilized in neuroscience for describing the encoding of sensory information in the firing patterns of neurons. The model comprises three main components: a linear filter, a pointwise nonlinearity, and a Poisson spike generation process. The linear filter accounts for the neuron's receptive field, capturing how the cell responds to different features in the input. The pointwise nonlinearity introduces a non-linear transformation, accounting for the neuron's thresholding behavior. Finally, the Poisson spike generation process models the probabilistic nature of spike generation, assuming that the neuron's firing rate follows a Poisson distribution. The LNP model has proven effective in capturing the spiking activity of neurons in response to various stimuli and, despite its simplicity, is commonly employed for tasks such as neural encoding [4].

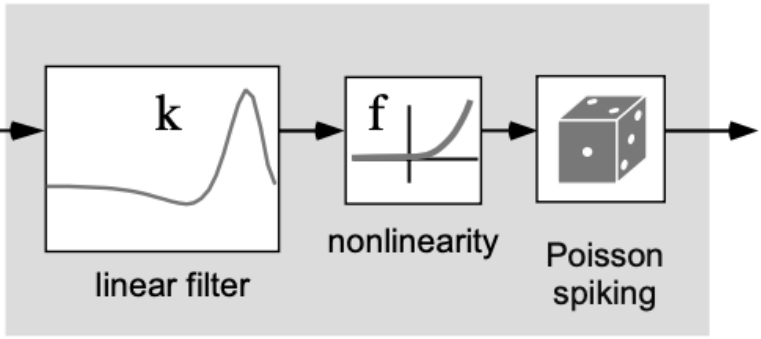

Figure 1.1: Illustration of Linear - Non Linear Poisson model by [5], showing the general structure for a LNP model, encompassing a linear filter, a nonlinearity and finally a Poisson spiking generation process.

However, LNP models have as a founding hypothesis the fact that times of single spikes are statistically independent of each other, and this assumption limits their power: they do not account for any spiking history when computing the current probability of firing. However, it is well known that neurons exhibit a biophysical history dependence through, for example, their refractory periods or the spike-rate adaptation [6]. Instead, by employing the GLM, which is a generalization of the LNP model, we can circumvent this issue, including dependencies on spike history. A GLM is a statistical model that generalizes the linear regression model and can be used for modelling relationships between a response variable and a set of predictors. An example of their utilization for retinal modeling can be found in [7], a study in which Pillow et al. used it as a model of multi-neuron spike responses for the analysis of the functional significance of correlated firing in a population of macaque parasol retinal ganglion cells. The schematic of their model can be seen in Figure 1.2.

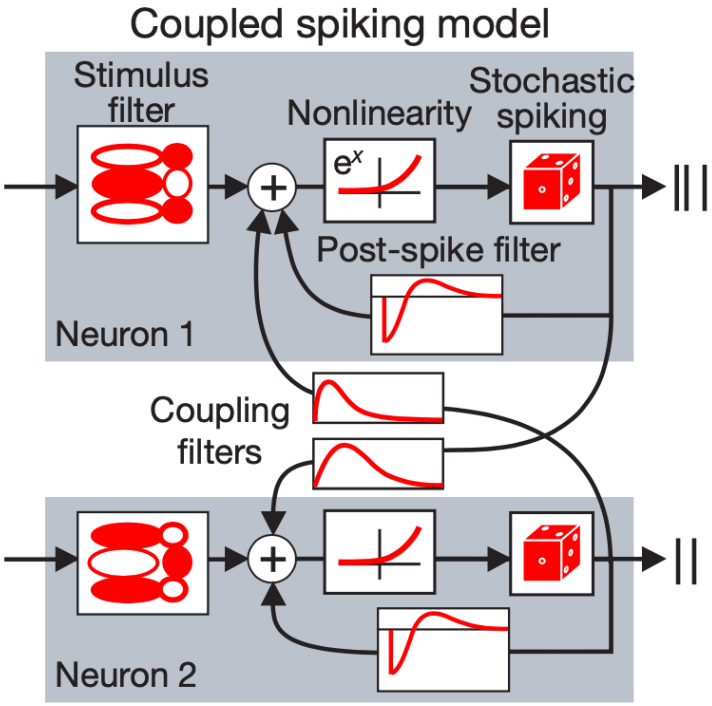

Figure 1.2: Illustration of the Generalized Linear Model employed in [7], shown for two coupled neurons. It encompasses three different types of filters: a stimulus filter, that mimics the spatio-temporal receptive field of the neuron; a post-spike filter, that takes into account the dependencies on spikes history; coupling filters, to account for the interactions between different neurons. The sum of the outputs of the filters is then passed through an exponential nonlinearity. The model gives as output the predicted spike trains.

The advantage of the models just described is the fact that they are highly interpretable, in the sense that it is easy to comprehend their mathematical components, but at the same time one drawback is that they are unable to capture the complicated visual processing of a natural image. One potential solution involves employing advanced deep neural network models, capable to grasp very complicated phenomena. For instance, this kind of model can clarify the process of object recognition carried out by the ventral visual stream [8–10] or unravel the mechanisms underlying sound discrimination in the auditory cortex [11]. Nevertheless, the challenge lies in the interpretation of these models, both from a computational and a biological point of view.

In a work performed in 2023, Maheswaranathan et al. [12] propose an approach that succeeds in modelling the neural code for natural scenes but is also able to fulfill the goals of being computationally and mechanistically interpretable. They achieved this objective exploiting a three-layer convolutional neural network (CNN) trained only on

natural stimuli (Figure 1.3).

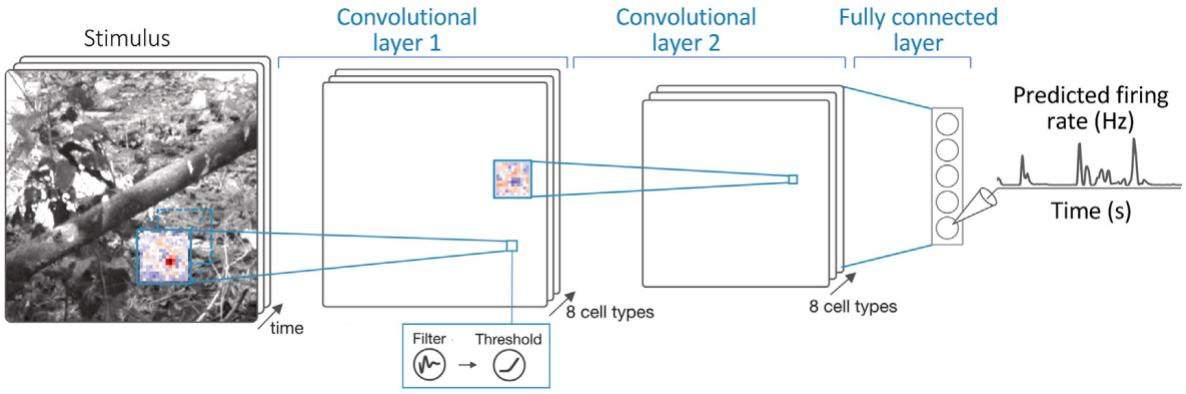

Figure 1.3: Schematic of the three-layer convolutional neural network (CNN), illustration by [12]. The CNN takes as input natural images and gives as output the predicted firing rate of the RGCs. The initial layer involves spatio-temporal convolution, is followed by a spatial convolution in the second layer, and finally by a fully connected layer. Rectifying nonlinearities are incorporated between the different layers.

The CNN, even when trained on ganglion cell responses, demonstrates a degree of internal correlation with interneuron recordings from different experiments, providing a level of mechanistic interpretability in its functioning. The objective of computational interpretability was reached using a new and versatile strategy that allows to interpret the computations performed by the retinal model by attributing them to the actions of model interneurons.

However, a problem of CNNs in the context of stimulus optimization is that they do not give in output an estimate of the uncertainty over the performed prediction. Uncertainty estimates guide the design of experiments by helping the identification of areas where the current knowledge is less certain: this information aids in selecting stimuli or conditions that are more likely to provide valuable information and reduce uncertainty. A family of models that give estimates of their own uncertainty are the Gaussian processes (GPs), statistical models that define a distribution over functions. A recent example of their use can be found in [13], a study in which Goldin et al. use GPs for the forecasting of the firing rate of the neurons in response to stimuli with high dimensionality, such as natural images. A schematic of their model is shown in Figure 1.4. This approach not only provides an estimate of prediction uncertainty but also can facilitate the creation of a closed-loop experiment based on it. Furthermore, in contrast to CNNs, which typically require extensive training with large datasets, the Gaussian process approach demonstrates the ability to adapt even with limited amounts of data. Nevertheless, in this study, this methodol-

ogy has solely been applied to static images. In my research, I extend its application to dynamic settings, specifically incorporating its use in the analysis of movies.

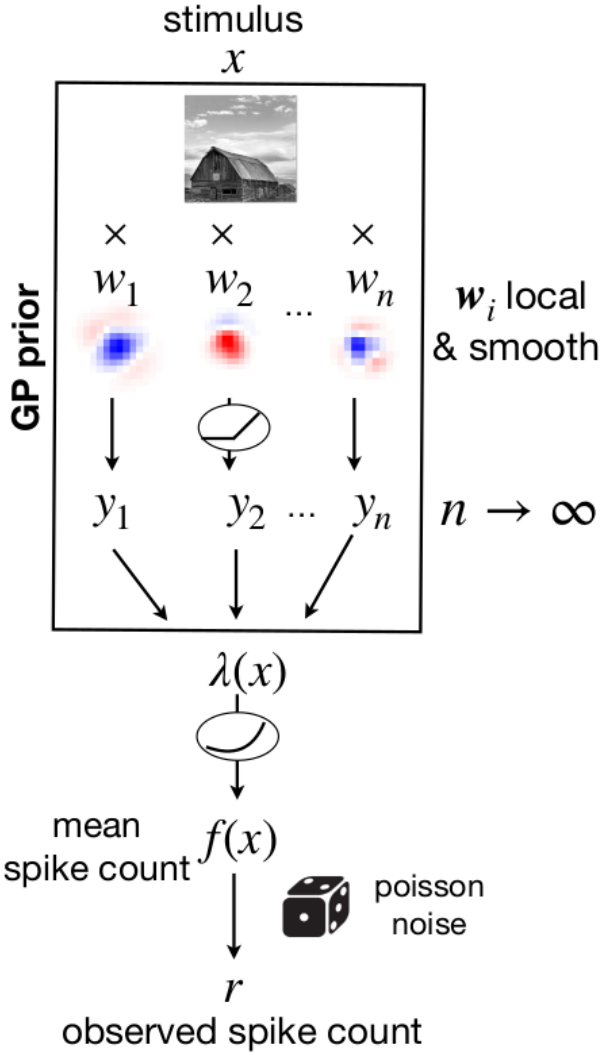

Figure 1.4: Schematic of the GP model from [13]. The input stimulus, a static natural image, is initially processed through a Gaussian process. The output from the GP undergoes a non-linear transformation before the spike count is derived, assuming a Poisson distribution hypothesis.

1.2. Stimulus Optimization

The term stimulus optimization refers to the strategic selection of stimuli in experiments. While this field has been extensively explored in statistics, its investigation in neuroscience is still relatively limited. Optimizing the stimulus plays a main role in neural modeling, aiming to enhance the understanding of how neurons respond to various inputs. The

process of stimulus optimization involves systematically selecting stimuli to maximize the information gained about neural responses or to achieve specific objectives, such as eliciting maximal firing rates. Researchers employ sophisticated algorithms, for example genetic algorithms and Gaussian process models, to efficiently explore the stimulus space. By designing experiments that incorporate optimal stimuli, scientists can investigate the complex relationships between stimuli and neural response, yielding to more accurate neural models.

One first example of stimulus optimization strategy is the one designed by Park et al. in 2011 [14]. In this research, the authors developed an algorithm to estimate the nonlinearity and the hyperparameters for a Linear - Non Linear Poisson model within the framework of a Gaussian process prior (Figure 1.5).

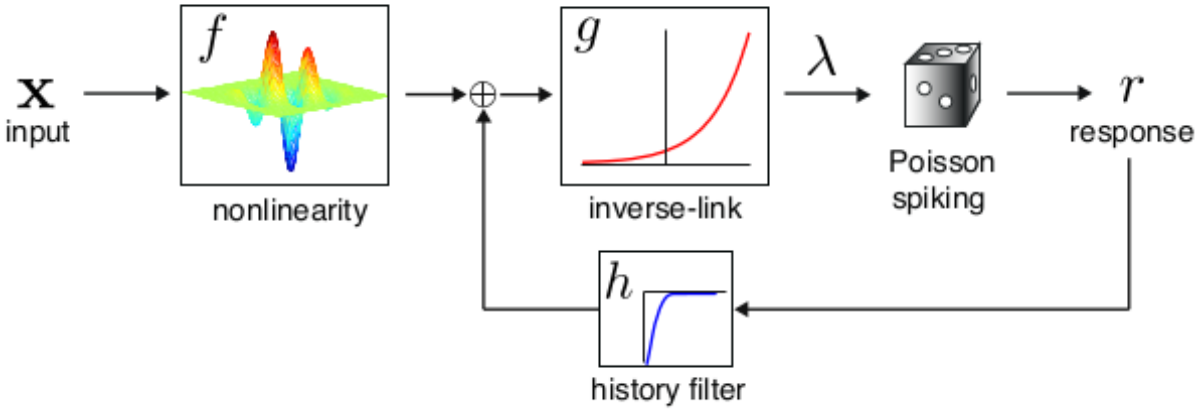

Figure 1.5: Illustration of the model (Figure by [14]). The input vector first passes through a nonlinearity f, that transforms it in a scalar. Then, the inverse-link g assures that the output is greater than 0. This value will be the mean spike rate in the Poisson spiking process, that gives as output the predicted spike count. The h is a history filter, added to include the effects of recent spike history on the response.

Thanks to this algorithm, the researchers were able to perform optimal experimental design, commonly referred to as "active learning" because stimuli are chosen dynamically during the ongoing experiment. This approach is based on uncertainty sampling, that means selecting the stimulus x for which the inverse link $g(f(x))$ exhibits the highest level of uncertainty given the data collected in the experiment up to the current point in time. Basically, they aim at the maximization of the expected information acquired on the model parameters with the new selected stimulus. In the end, this strategy allows

the nonlinearity f to be characterized even from a reduced quantity of data. The method has been used in the context of the non-linear color-tuning properties of macaque V1 neurons, with the final results of showing that the GP-Poisson model can properly model these responses, and that optimal design can significantly diminish the number of stimuli needed to comprehensively characterize these responses. The main weakness of this study is that it is only dealing with low dimensional stimuli.

In another exemplary study by Ponce et al. in 2019 [15] the focus was on exploring neuronal selectivity in the monkey inferotemporal cortex through the hypothesis space of a generative deep neural network. In contrast to the previous study by Park, in this case the "optimal stimuli" were not related to the concept of uncertainty reduction, but rather they were the stimuli that achieved maximal neuronal firing. A genetic algorithm was employed to identify stimuli that maximized the firing activity of neurons, letting the cell under study to guide its own stimulus selection. The approach of the genetic algorithm was already proposed in the past to study selectivity in macaque V4 and IT cortex (Yamane et al., 2008 [16]; Carlson et al., 2011 [17]). The method proposed by Ponce et al. basically builds upon the previous approach, in order to delve into the tuning properties of neurons in the inferior temporal cortex of macaque monkeys. The overall pipeline, repeated for around 250 times, can be broken down into three steps. Neuronal responses are recorded from macaques engaged in a passive fixation task. Subsequently, image codes derived from these responses are ranked and subjected to selection, recombination and mutation within a genetic algorithm, producing new image codes. Finally, these image codes are given as input to a generative neural network, which synthesizes images based on the provided codes. The overall approach has been named XDREAM (EXtending Deep-Dream with Real-time Evolution for Activity Maximization in real neurons) and its schematic is shown in Figure 1.6.

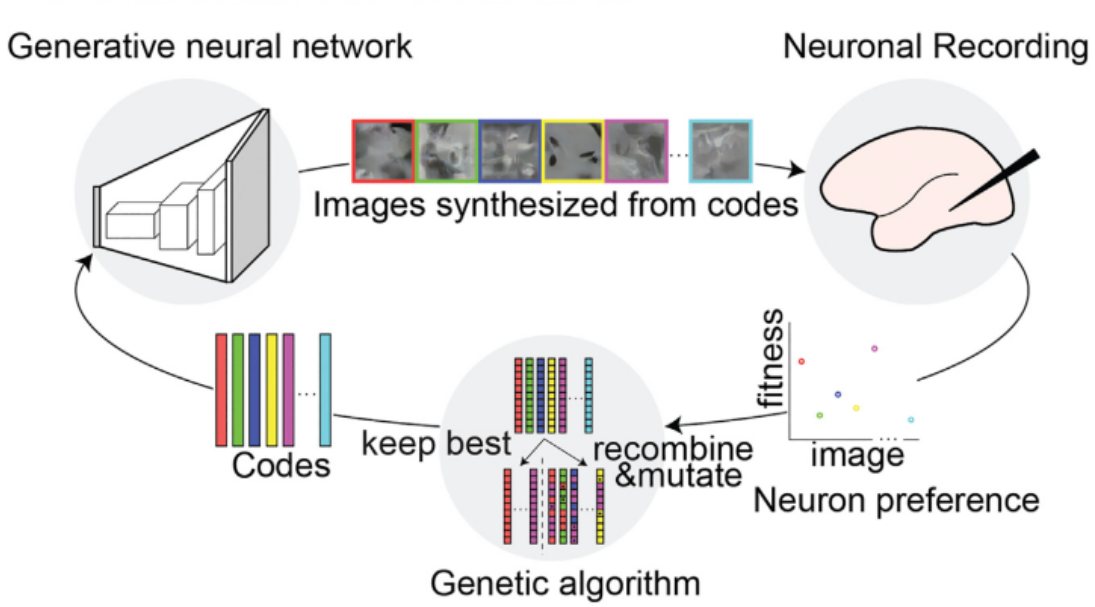

Figure 1.6: Illustration of XDREAM approach (Figure by [15]). First, image codes are passed through the deep neural network to create images. These images are shown to the monkey under study, and based on the firing rate in the responses the ranking of image codes is computed. According to the ranking the codes undergo a process of selection and are then recombined and mutated to produce new image codes. The loop restarts passing again the codes to the generative network.

1.3. Retinal Ganglion Cell typing

Functional retinal ganglion cell (RGC) typing is defined as the classification of different types of retinal ganglion cells based on their functional properties, such as their response to visual stimuli. This categorization helps to understand the diverse roles and functions of different types of retinal ganglion cells in visual processing. Techniques like electrophysiology and imaging are often used to study the responses of retinal ganglion cells to visual stimuli, allowing for the identification and classification of different functional types. The goal is to gain insights into the neural circuits and pathways involved in visual processing and to better understand how the visual information is encoded and transmitted through the retina to the brain.

In the work by Baden et al. in 2016 [1], they merged two-photon calcium imaging for comprehensive retinal recordings and performed on them unsupervised clustering, finding that the mouse retina encompasses significantly more than 30 distinct functional output channels. This array comprises both already recognized and novel ganglion cell types, and

has been confirmed through genetic and anatomical assessments. Thus, they discovered that the level of diversity of channels conveying information from the mouse's eye to its brain surpasses previous anatomical studies. Going more specific, to perform the typing they registered responses to the cells elicited with 4 different type of stimuli, chosen to investigate different properties: for the characterization of polarity, kinetics, and temporal frequency and contrast preferences a full-field "chirp" stimulus was used; then, to assess direction and orientation selectivity, a moving bar was leveraged; to estimate receptive fields they exploited binary dense noise, and finally, to investigate chromatic preference, an alternating long-/short-wavelength full-field stimulus was used. The pipeline followed by the researchers for the classification is now described. In the initial stage, an automated unsupervised clustering approach was employed: sparse Principal Component Analysis (sPCA) was performed on light-driven Ca^{2+} signals of ganglion cell layer cells to extract features, followed by the application to the derived feature set of a Mixture of Gaussian model for clustering purposes. In the subsequent step, post-processing of the clustered data was conducted and the clusters were organised in accordance with a hierarchical tree, considering their functional similarities. Additionally, they propose a grouping scheme based on available domain knowledge. In the end, they found that a minimum of 32 functional types can be distinguished. The results of the clustering are shown in Figure 1.7.

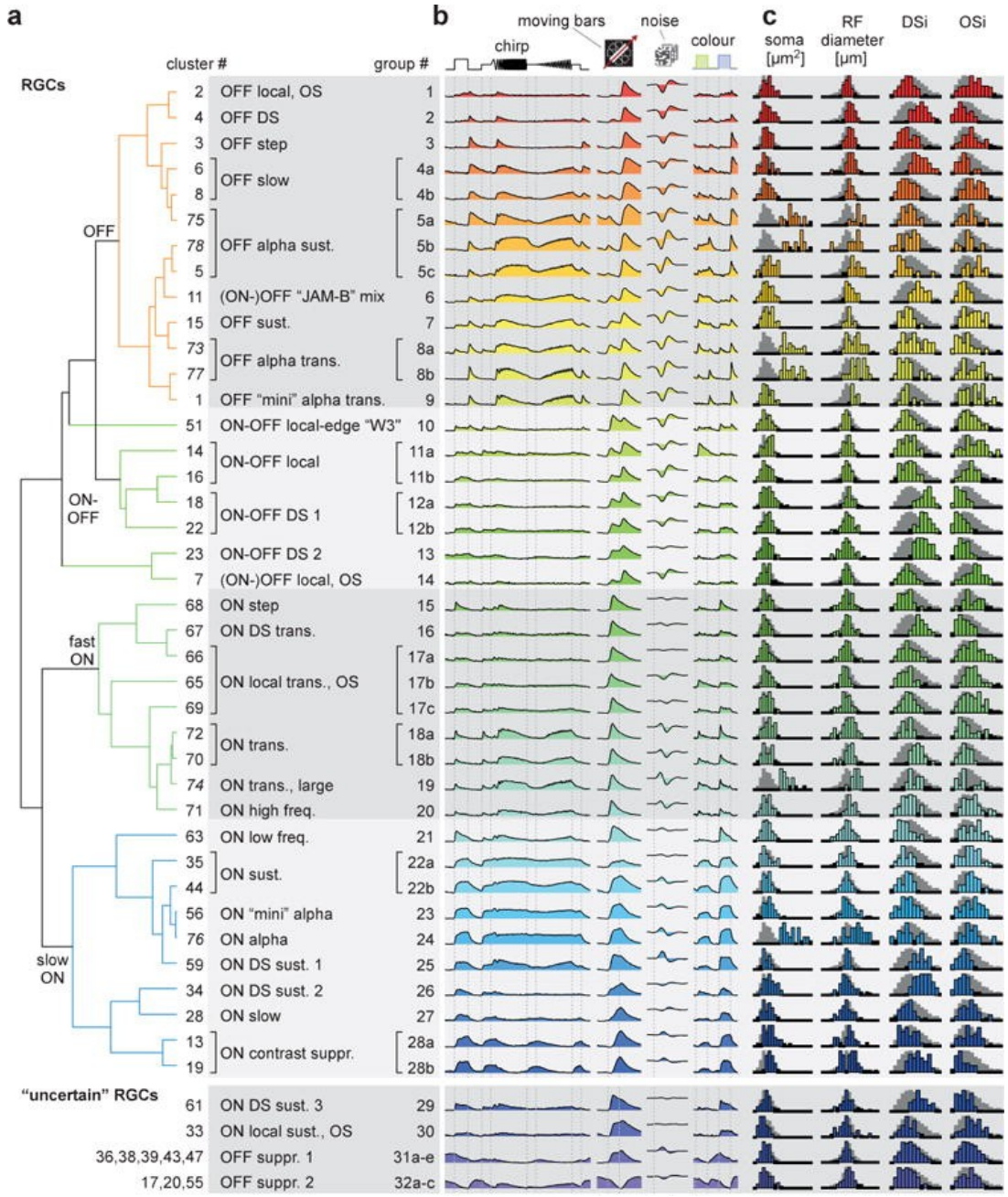

Figure 1.7: The figure shows the typing results obtained with the clustering, together with some properties of the specific cell type. **a**. Cluster-dendrogram. **b**. Cluster-mean responses to the 4 stimuli (chirp, moving bars, noise, colour). **c**. Histograms for selected metrics, from left to right: soma area, receptive field (RF) diameter, DS-index and OS-index (Figure by [1]).

Another investigation on the same subject was conducted by Goetz et al. in 2022 [2]. In this study, the researchers developed a machine learning classifier designed to enable the inference of a RGC functional type based on a standardized set of spike measure-

ments. The main difference from the previous study lies in the stimuli chosen to elicit cell responses: Goetz et al. (2022) employed spots with diameters ranging from 30 to 1200 μm, covering a spectrum from below the smallest receptive field up to the largest surround. The disks were each time centered in the receptive field of the single cells. Their experimental setup, utilizing cell-attached electrodes, allowed for a single-cell at a time measurement, making it feasible to center the stimuli on the receptive field of each cell, even if limiting the throughput. To create a gold standard for classification, researchers manually assigned RGCs to 42 different functional types based on their responses to the stimuli (see Figure 1.8). Subsequently, this dataset was employed to train a machine learning classifier, able of assigning cells to one of the 42 identified types. In their work they highlight the importance of centering the stimuli on individual RGCs: in fact, several RGC types exhibit poor responses to full-field stimuli or spatiotemporal white noise. Moreover, in certain types fundamental properties are dependent on spot size and so responses may vary according to different dimension of the disks. Additionally, in this way we can assess surround suppression behavior of the cells, an information that is essential for distinguishing types that show a similar functional behavior.

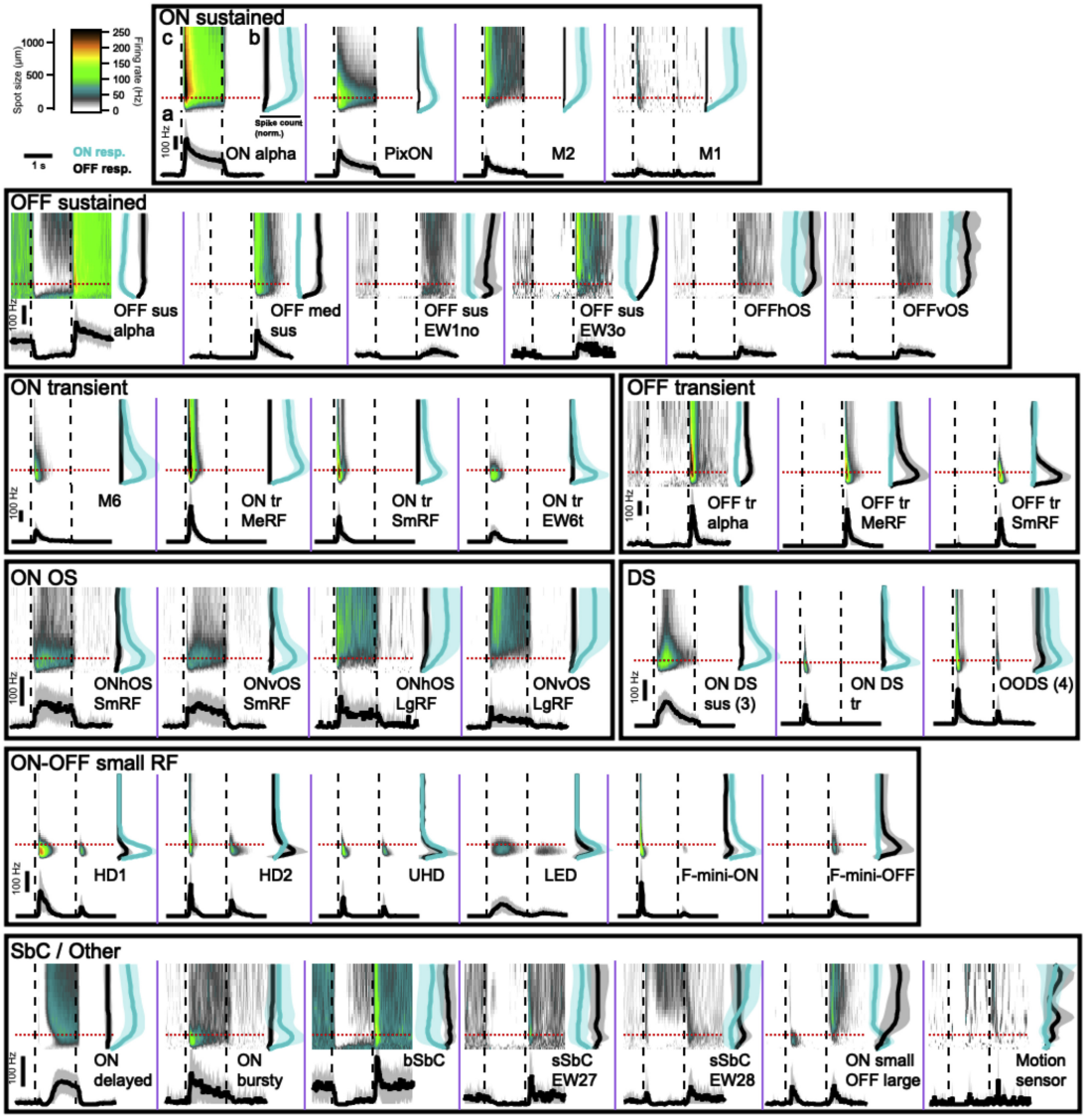

Figure 1.8: In each panel three graphs are shown, describing the response of a retinal ganglion cell type to disks of varying dimension. The top left graph is a 2D plot of average firing rate with time on the x-axis and spots dimensions on y-axis. The stimulus onset and offset are represented by the dashed lines. The top right graph shows the total spike count for all the spot sizes. The bottom graph shows PSTH plots, averaging the response of each cell type to 200 μm spots (Figure by [2]).

2 | Data Collection and Description

The data used in this project have been experimentally gathered from a mouse retina. Specifically, the responses of the retinal ganglion cells to various stimuli were captured using a Multielectrode Array (MEA) system. An interesting aspect of this recording method lies in its ability to simultaneously capture signals from numerous neurons: this capability is crucial for comprehending how the neurons collectively encode information. The experimental setup is extensively described in [3]. The raw traces registered from the array encode the activity from more than one cell, thus a spike-sorting algorithm has to be used to separate the contributes coming from different neurons. The source separation was performed using as sorting algorithm SpyKING CIRCUS [18]. The stimuli that have been projected onto the mouse retina during the experiment are of two kinds: multiscale checkerboards and multisize spots. In the following paragraphs a detailed description is provided. The responses to the two stimuli have been registered for the same 22 cells, thus allowing a parallel analysis of the neurons in the two cases.

2.1. Multiscale Checkerboards

The multiscale checkerboard stimulus, also called spatiotemporal white noise [19], comprises a series of images displayed over time, each presenting a checkerboard pattern. As the name suggests, these checkerboards include squares of various sizes across different temporal frames. This stimulus design is suggested because, while maintaining the same stimulus for each cell, it effectively captures the non-linear behavior in RGC responses when spatial scales vary, as observed in phenomena like surround suppression. Instead, using only checkerboards with the same check size would capture only small ranges of the possible RGC responses, failing to emphasize their non-linear behavior. Some illustrative examples of single checkerboard frames are presented in Figure 2.1.

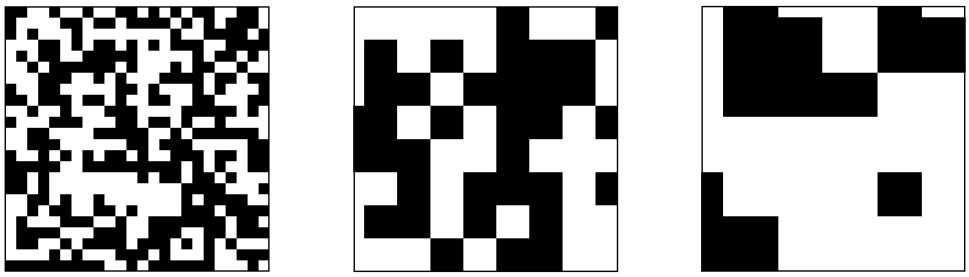

Figure 2.1: Example of checkerboard stimulus frames with three different check sizes.

Five sizes of checks were present, spanning a range from 50 to 250 μm. 90 sequences including checkerboards frames with different check sizes were generated, with every check size present in the same proportion in each of them. Prior to the utilization in the computational model, the raw data from the experiment underwent preprocessing to make both spatial and temporal structures more suitable to be given as input to the model.

Regarding the spatial structure transformations, individual images were initially rescaled from 864x864 to 108x108 pixels, and then further downsampled with a factor of 2 and cropped at the center. After the preprocessing, each individual image has a dimension of 24x24 pixels. This approach was adopted to provide the model with lighter input images without sacrificing information: in fact, the cropped region was strategically chosen to encompass the receptive fields of all cells. The excluded part of the image, although presented during the experiment, did not impact on RGC responses: including this portion might even be detrimental to the model, by introducing nonexistent correlations.

Regarding instead the temporal considerations, when analysing the responses we exploited time bins, in particular we considered temporal windows that were 50 ms long. Considering that the frequency with which stimulus frames were projected during the experiment was 4 Hz and the total duration of the stimulus was 22.5 s, it implies that each frame was displayed for 250 ms. In our framework, since the temporal window chosen to analyze the responses was 50 ms, we have the same stimulus frame for $250/50 = 5$ temporal bins. After the rebinning process, each sequence comprises 450 frames. This stimulus is the same for all 22 cells, resulting in a final dimensionality of 24x24x450x90, that respectively are the height and width of the image, the number of temporal frames and the number of different sequences. For every sequence a response that has the same temporal length as the stimulus is registered from each cell. The responses are different for each cell, thus the dimensionality of the responses is 22x450x90, respectively the number of cells, the number of temporal frames and the number of different sequences (Figure 2.2).

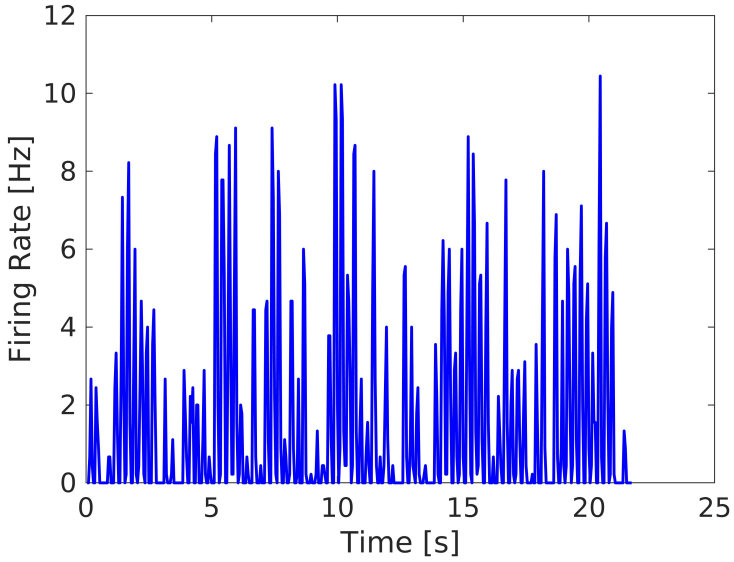

Figure 2.2: Example of the response of one cell to the Multiscale Checkerboard stimulus. The x-axis corresponds to the time in seconds, while the y-axis shows the firing rate in hertz. The response, just like the stimulus, has an overall length of 22.5 s.

2.1.1. Spike Triggered Average

It is important to assess how retinal ganglion cells respond to checkerboards also because this evaluation enables the computation of the spike-triggered average (STA), one of the first techniques used in neuroscience to characterize the receptive field of the cells. In fact, this is the simplest way to characterize the linear filter of a neuron response, and it can be easily computed evaluating the responses to a checkerboard stimulus. As the name suggests, it involves calculating the average stimulus that precedes a spike in the neural activity of a cell. In mathematical terms, the STA is computed by summing the stimuli that occurred before each spike and dividing this sum by the total number of spikes recorded. This process is expressed by the following formula:

$$\text{STA} = \frac{\sum_{t=1}^{T} s_t f_t}{\sum_{t=1}^{T} f_t} \tag{2.1}$$

where T is the length of the recording period, s_t is the stimulus at time t and f_t are the spikes at time t [4]. By examining the STA, researchers can gain insights into the receptive field of the cell, helping to understand the types of stimuli that tend to evoke neural responses. In our specific case, having an estimate of the receptive field is crucial, since the information about its position is essential for the Gaussian process models used in this project, and should be provided as input.

The STA can be considered an approximation of the initial linear component within a polynomial series expansion representing the system response function [20]. While the STA delivers a comprehensive characterization for linear systems, it is important to acknowledge that neural responses, being highly non-linear, deviate from this conventional behavior, but still this technique can be used as a first step in a more complex cascade model, like the Linear - Non Linear Poisson model described in Section 1.1. Apart from some very specific cases, however, it is important to consider that the STA incorporates also some structures that may arise from the correlations between different stimulus inputs: in such cases, a pre-multiplication by the inverse of the stimulus autocorrelation matrix becomes necessary for unbiased analysis.

Up to this point the STA as been referred to as an estimate of the spatial receptive field, but it can be generalized to be an estimate of the spatio-temporal receptive field, taking into account not only the sensitivity of the neuron to inputs at different spatial points but also at inputs at different points in the stimulus history [6]. To do that, we consider in the computation a window of stimulus values ranging from the current instant to some maximal delay in the past. The result will have a single filter for each time-delay, each one with the same dimensionality as the input stimulus, as illustrated in Figure 2.3 and 2.4.

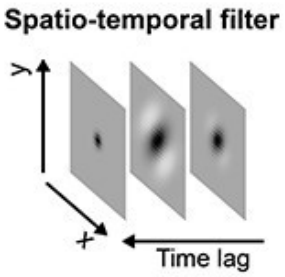

Figure 2.3: Illustration of the spatio-temporal filter provided by the spike-triggered average [6]. For each different time lag, there is a filter with the same spatial dimension as the input stimulus. In this example a maximum delay of 3 timestamps in the past is assumed.

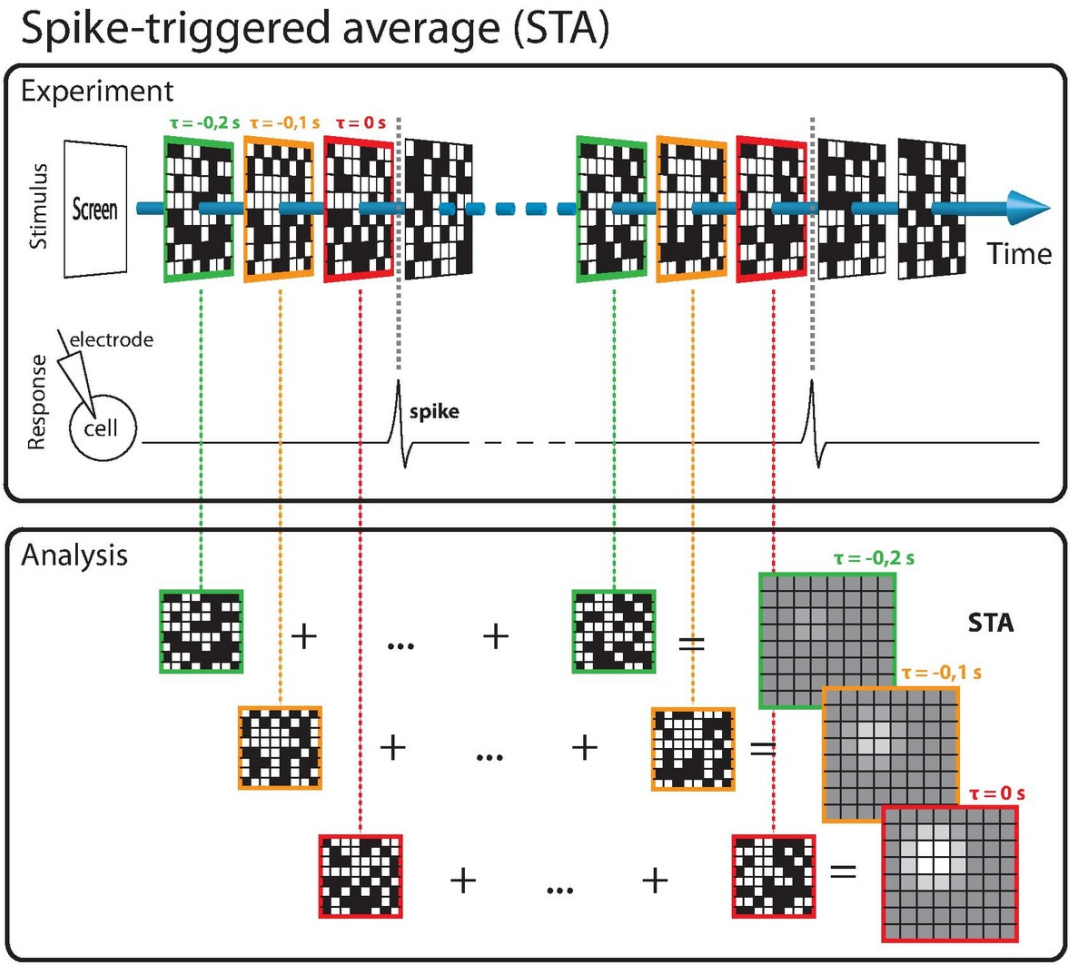

Figure 2.4: Spatio-temporal spike-triggered average computation schema by [21]. Each colour corresponds to a different time delay in the past; in the final STA, the filters for the n_{th} delay are computed taking the average of all the images occurred n timestamps before the spike.

2.2. Multisize Spots

The multisize spots stimulus consists in a series of frames with light spots on a dark background shown over time. As the name suggests, the spots have varying size and for each cell they present the following six diameters [μm]: 100, 200, 300, 400, 600 and 800.

Figure 2.5: Example of spots frames with different disk sizes and centered in different receptive field positions.

In this case the stimulus differs across cells, since the disks are consistently centered in the receptive field of each individual neuron. This characteristic, as elucidated by Goetz et al. in [2], proves to be critical for cell typing. Also this stimulus, like the checkerboard one, underwent preprocessing before being given as input to the model. The disk stimulus, in its post-processed version, presents the same spatial dimension of the checkerboard (24x24) and the same temporal frequency of 20 Hz. The stimulus is constructed in time such that, for each size, the disk is displayed for 1 second, and is followed by dark frames. Consequently, the overall dimension of the stimulus is 24x24x40x6, with the last two dimensions denoting respectively the number of timestamps and the number of different sizes chosen for the spots. The same disk stimulus, with the dimensionality just described, is presented in 30 different trials, resulting in 30 responses per cell. For each cell, these responses can be averaged to generate the Peristimulus Time Histogram (PSTH) in Figure 2.6. A PSTH is a graphical representation of the average firing rate of a neuron over time in response to a repeated stimulus.

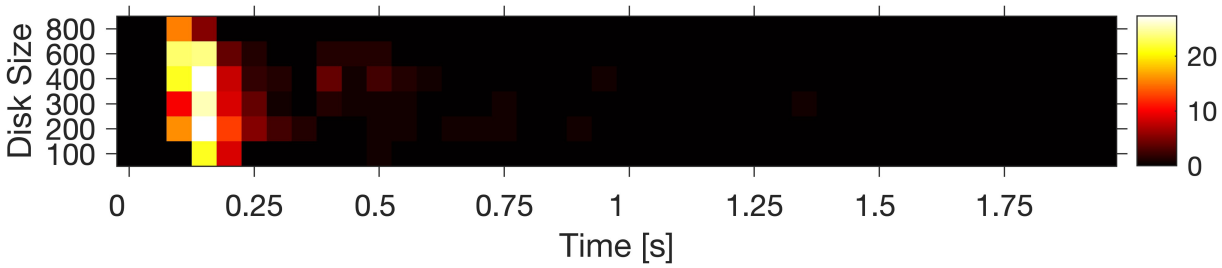

Figure 2.6: Example of the response of one cell to the multisize spots stimulus. On the x-axis there is the time in seconds and on the y-axis the spot sizes in μm. The firing rate in hertz is colour coded with the values shown in the righthand colorbar.

3 | Modeling: Gaussian Processes

As computational model in this project Gaussian processes have been selected. This family of models will be described in details in the following sections, and their main advantages in the context of stimulus optimization will be elucidated.

3.1. Introduction to Gaussian Processes

Gaussian processes (GPs) are powerful mathematical models that, unlike traditional parametric models with a fixed number of parameters, are non-parametric and define distributions over functions. In particular, a GP is a collection of random variables, any finite set of which is jointly Gaussian-distributed, making it a flexible tool for capturing complex relationships in data. GPs are completely defined by their mean and covariance function, which determine the overall shape and smoothness of the functions drawn from the process. Drawing functions from a Gaussian process means generating random samples that conform to the distribution specified by the mean and the covariance function.

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \tag{3.1}$$

It can be seen that the covariance between the outputs is written as a function of the inputs: the correlation structure between outputs depends on the specific values of the input point. The choice of the covariance function is crucial because it defines the type of relationships that the GP can capture. Different covariance functions model different types of dependencies, such as smooth trends, periodic patterns, or abrupt changes. Furthermore, a GP must satisfy consistency requirement, that basically means that examination of a larger set of variables does not change the distribution of the smaller set. Notice that the consistency requirement is automatically fulfilled if the covariance function specifies entries of the covariance matrix. [22]

3.1.1. Prediction with Gaussian Processes

Gaussian processes perform predictions by leveraging their probabilistic nature to estimate the distribution of possible outcomes. Of course, before being used to perform prediction the model must be trained on a set of input-output pairs (training data), where each input is associated with a corresponding output. So, the mean and the covariance functions are optimized based on the training data to best fit the observed input-output relationships. Given a set of new input points (test data) for which predictions are needed, the GP estimates the distribution of possible function values at those points. The joint distribution of the observed training data and the test data can be written as:

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right)$$

with $K(X_*, X)$ being the covariance between the test input and the training input, $K(X, X)$ the covariance of the training set itself and $K(X_*, X_*)$ the covariance of the test set. The training output is indicated with $\mathbf{f}$ and the (unknown) test output with $\mathbf{f}_*$. The conditional distribution of the test data given the observed training data can then be calculated. This is a Gaussian distribution that provides predictions along with uncertainty estimates.

$$\mathbf{f}_*|X_*, X, \mathbf{f} \sim \mathcal{N}(\bar{\mathbf{f}}_*, \text{cov}(\mathbf{f}_*))$$

The mean and the covariance of the conditional distribution can be computed as:

$$\bar{\mathbf{f}}_* = K(X_*, X)K(X, X)^{-1}\mathbf{f},$$
$$\text{cov}(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*)$$

The mean of the conditional distribution $\bar{\mathbf{f}}_*$ represents the predicted values, while the covariance matrix $\text{cov}(\mathbf{f}_*)$ provides information about the uncertainty associated with these predictions. [22]

3.2. Features of Gaussian Processes

As described earlier in 1.1, gaussian processes possess several key attributes that make them effective models in the context of neural modelling. First of all, GPs can be effectively trained even when the available dataset is relatively small, making them adaptable

to situations with limited training samples. In fact, they are able to intrinsically implement Occam's razor, adjusting the inferred model complexity according to the available data to prevent overfitting, and this feature is useful especially in situations where the dataset size is small [23], [13].

Furthermore, these processes have the ability to estimate their own uncertainty, and this information is valuable for evaluating the robustness of their predictions, in particular in closed-loop applications where understanding the model's confidence is crucial.

Another advantage, as stated also previously, is that Gaussian processes are non-parametric, meaning that they do not impose strong assumptions about the specific form of the response function. This flexibility allows them to capture a wide range of complex relationships without being constrained by predefined functional forms.

One drawback of Gaussian processes, however, is the fact that they may perform suboptimally in high-dimensional stimuli scenarios. This limitation can be mitigated by incorporating prior knowledge about visual neuron responses, specifically adding some constraints on the receptive field of the cells, i.e. that they are spatially local and smooth, as demonstrated by Park and Pillow [24].

4 | Stimulus Optimization

The utilization of the Gaussian process models within the specific framework of my project will be elucidated in the subsequent paragraphs. The discussion will commence with an exploration of the distinctive features of the employed Gaussian models. Then, a comprehensive depiction of the overall pipeline for stimulus optimization will be provided.

4.1. Model Specifics

The models employed in this project have some peculiar characteristics that it is important to emphasize. A framework similar to the one described by Goldin et al. in [13] has been used. This previous research has validated the effectiveness of the GPs in handling static images. A novel aspect in my work lies in its extension to the domain of movies, broadening the scope of application. In this case the gaussian model is used to forecast the firing rate of a cell in reaction to a spatio-temporal visual stimulus, such as checkerboards or spots. Notably, the output of the Gaussian process is continuous and can even encompass negative values. Considering that spike counts are intrinsically positive, a positive monotonic nonlinearity must be applied to transform the model's output. This transformed value serves as the mean spike count in a Poisson probability distribution, subsequently utilized to generate the observed spike count.

$$\begin{gathered}\lambda(x) \sim GP(m(x), k(x, x')) \\ \downarrow \\ f(x) = e^{A\lambda(x)+\lambda_0} \\ \downarrow \\ r(x) \sim p_{poisson}(r(x)|f(x))\end{gathered}$$

The distribution of $\lambda(x)$ is the output of the Gaussian process, $f(x)$ is the firing rate obtained as an exponential transformation of $\lambda(x)$ and finally $r(x)$ is the spike count obtained from a Poisson distribution with parameter $f(x)$. A and λ_0 are hyperparameters deriving from the training of the Gaussian process.

As regards the selection of the kernel employed in computing the covariance function, it must be kept in mind that this decision holds significant importance, as opting for a different kernel means introducing distinct prior assumptions about the output function. Typically, the conventional choice is the exponential kernel, which implies the smoothness of the function, i.e. that similar inputs yield similar outputs. However, this kernel suffer from the curse of dimensionality and fails with high dimensional inputs. In this case, the chosen one is the arc-cosine kernel proposed by Cho et al. in [25]. Nevertheless, relying solely on the kernel choice is insufficient to address the curse of dimensionality issue. Additional assumptions regarding the spatial and temporal localization and smoothness of the receptive field become imperative to further constrain the inference process.

Furthermore, to facilitate inference in scenarios where the number of stimulus-response pairs is high, the inducing points framework proposed by Hensman et al. in [26] has been employed. This approach entails conditioning the variational posterior over a restricted set of inducing points, enabling more efficient computations when facing massive data volumes.

4.2. Methodology for Stimulus Optimization

The method implemented for the stimulus optimization is explained in detail in the following paragraphs.

In summary (Figure 4.1), a dedicated gaussian model is trained for each cell, utilizing the entirety of available data encompassing checkerboards and disks. This model, denoted as the "generative" model, serves the purpose of generating responses to various sets of synthetic checkerboard stimuli created ad hoc. It is designed with the specific aim of being the best possible in predicting responses from both checkerboard stimuli and multisize spots stimuli. Subsequently, these synthetic stimulus-response pairs are employed to train new models, referred to herein as the "fitted" models.

To assess the efficacy of the fitted model in predicting responses to the multisize spots stimulus, a comparative evaluation against the generative model is conducted. The Kullback-Leibler divergence between the two models is computed, with lower values indicating superior performance in replicating the generative model's behavior. Given that the fitted model's performance is dependent on the synthetic training stimulus, this analysis provides insights into the checkerboard stimulus that yields to the fitted model that is the best in the characterization of RGC responses to the multisize spots stimulus. We aim to achieve a good prediction of the responses to the disk stimulus because these are the ones on which the typing will be performed: having accurate fit of the responses allows

to maintain the distinctive features of the cells and thus increases the probabilities of achieving a good classification in functional types.

The Kullback-Leibler divergence, besides being evaluated only on a multisize spots test set, is also calculated for the prediction of responses to a test checkerboard stimulus, with the sole aim of offering a broader perspective on the generalization capabilities of the method on a different task.

Let's now analyze each passage of the procedure in greater detail.

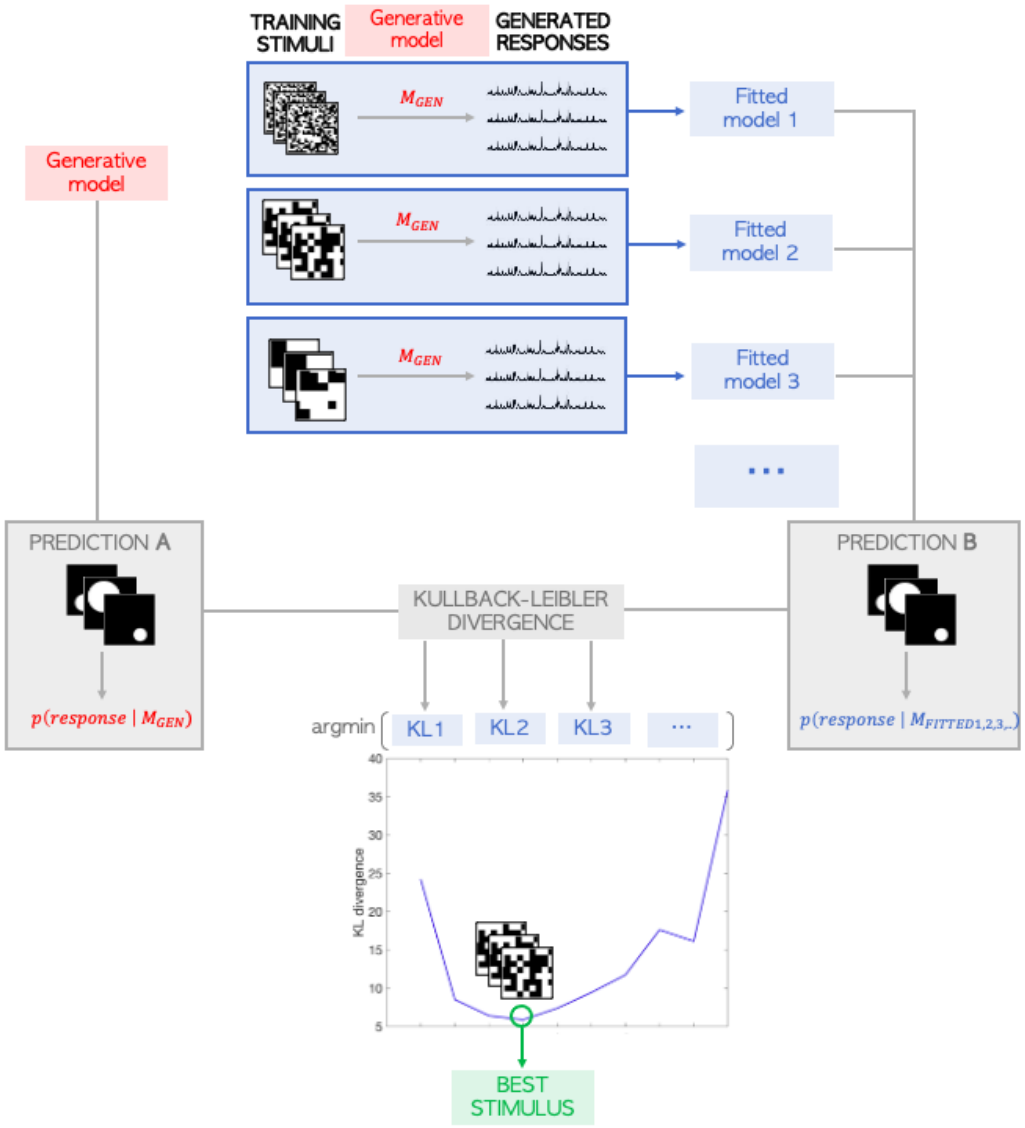

Figure 4.1: Schematic of the procedure to be followed to perform stimulus optimization. The response to different synthetic stimuli is generated exploiting the generative model. These new pairs stimulus-response are used to train different fitted models. Each of them is evaluated in comparison to the generative model, computing the KL divergence between the probability distributions predicted for the response to the multisize spots. The fitted model with the lowest KL corresponds to the best stimulus for the characterization of RCG responses to the disks.

4.2.1. Generative Model Training

A model for each neuron is trained exploiting all the available data, including both multisize spots and multisize checkerboards. The temporal correspondence stimulus - response firing rate is not 1 by 1, instead for the prediction of each temporal point of the response

stimulus history is taken into account, considering a window of values ranging from the current instant t to a maximum delay of 16 timestamps in the past. To achieve this, the stimulus has to be brought in a suitable format for being taken as input from the model. In particular, the input to the model at time t is [6]:

$$\mathbf{s}_{\text{input}}(t) = (\mathbf{s}(t-1),\ \mathbf{s}(t-2),\ \mathbf{s}(t-3), ...,\ \mathbf{s}(t-16))$$

As mentioned earlier, the model requires as input the position of the receptive field: it is estimated based on the Spike-Triggered Average (Section 2.1.1) evaluated over the checkerboard data and their corresponding responses.

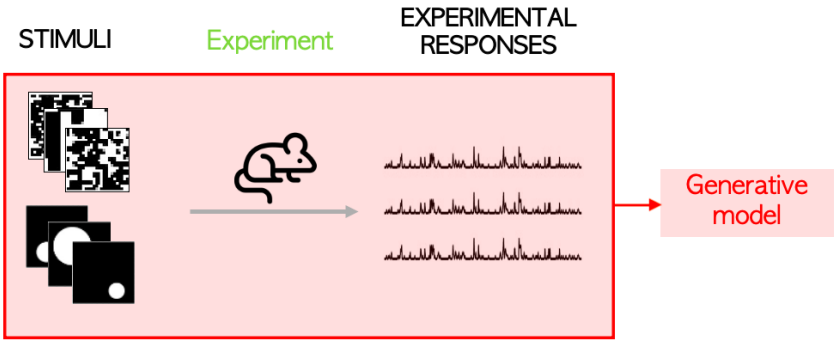

Figure 4.2: The generative model is trained exploiting the data used in the experiment, i.e. the multisize checkerboard stimulus and the multisize spots stimulus.

4.2.2. Synthetic Stimulus Generation

As said, our aim is to find a checkerboard stimulus that is optimal in the sense that it allows to train a model that behaves well in the prediction to a multisize spot test stimulus. What differs from one stimulus to another? In our framework, they change in terms of dimension of the checks, percentage of presence of a certain check size in the temporal stimulus and percentage of presence of black tiles. Since it would be computationally impossible to test all the possible combinations of check sizes, our approach was to test a discrete set of possible synthetic stimuli, starting from simple stimuli such as sequences containing only one check size, also as a tool to gain insight about approximately which check sizes are "preferred" by the cells, and then moving to more complex ones. The synthetic sets of checkerboards created underwent identical preprocessing as the authentic ones, maintaining an equivalent time length. This ensures that the results remain consistent and comparable with those obtained during the actual experiment.

So, various sets of stimuli were created, each with distinct characteristics. Each stimulus comprised 20 different sequences, each of which had a duration of 90 timestamps (450 after temporal rebinning). The sizes in μm considered for the tiles are the following:

Size 1	Size 2	Size 3	Size 4	Size 5	Size 6	Size 7	Size 8	Size 9	Size 10
28	56	84	168	252	378	504	756	1008	1512

Table 4.1: Checks sizes.

The discrete set of considered stimuli can be categorized as follows:

- **Single sizes**: each stimulus consisted of checkerboards with tiles of only one size, resulting in a total number of stimuli equal to the one of the sizes, so 10 different stimuli.
- **Uniform stimulus**: all sizes are present with equal proportions.
- **Small sizes**: only the first five sizes are present with equal proportions.
- **Big sizes**: only the last five sizes are present with equal proportions.
- **Single size with different densities**: a medium size (Size 5) was chosen and maintained fixed. What varies in this case is the probability of the checks of being black, resulting in a different percentage of presence of black checks in the final image. Nine different percentages of presence were explored, from 0.1 to 0.9, thus in this case nine different stimuli were generated. Examples of the generated images are shown in Figure 4.3.

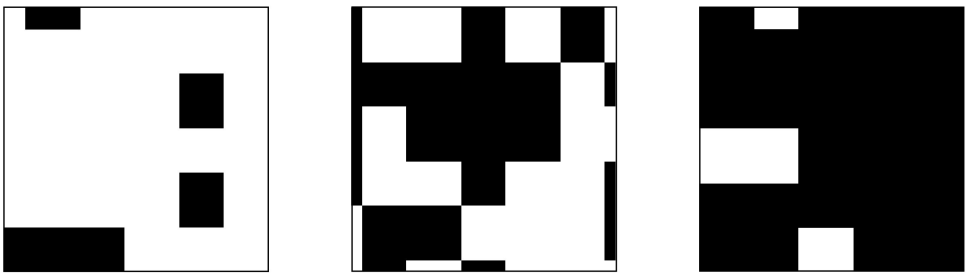

Figure 4.3: Example of checkerboard stimulus frames with different probability of having a black check. The first one shows a percentage of black checks of the 10%, the second one of the 50% and the third one of the 90%.

For these stimuli sets experimental results for the responses are not available. They are mathematically inferred exploiting for the prediction the generative model mentioned in the previous paragraph, trained on all the experimental data.

Consequently, at the conclusion of this stage, numerous pairs of synthetic stimuli and predicted responses are generated.

4.2.3. Fitted Model Training

In this phase, the new stimulus-response pairs from the preceding step are utilized to train other gaussian models. The same formatting described in Section 4.2.1 is applied also to these inputs, so as to take into account the stimulus history in the responses. To expedite the training process, the majority of the model's hyperparameters are maintained at the values previously computed for the generative model. Given the extensive number of models to be trained (22 for every generated stimulus, corresponding to each cell), this approach helps in the reduction of the training time. When referring to hyperparameters, we denote the values influencing the size, smoothness, and position of the receptive field. Even if these parameters can be learned from data, utilizing the ones derived from the training on true data is a reasonable approximation.

4.2.4. Performance Evaluation: Kullback-Leibler Divergence

The time has come to assess the optimal stimulus set, seeking the one that yields the best model based on a specific performance measure. What exactly is this measure? Our criterion for optimality lies in selecting the stimulus that results in a fitted model producing outcomes most closely aligned with those provided by the generative model, making the assumption that this model is reasonably good in predicting the responses to the multisize spots stimulus. We used as a metric to assess the goodness of each stimulus the Kullback-Leibler divergence, measuring for each cell the probabilistic distance between the distribution of the responses provided by the generative model and the one provided by the models fitted on the synthetic data. This measure has been chosen for its advantage of taking into consideration the uncertainty given in output by the gaussian model. To compute it, first of all it is necessary to generate with the generative model responses to a test stimulus. Then, the log-likelihood of the generated data under the generative model (M_{gen}) and under all the different fitted models (M_{fitted}), trained on the synthetic data, is computed. The Kullback-Leibler divergence corresponds to the difference between these two quantities averaged over the generative model parameters. A value close to 0 means that the second model is as good as the first one in modelling the responses. Basically, we aim at minimising the KL-divergence between the "true" distribution generating the data and the distribution modelled with the fitted model. A better model will result in a lower KL value.

Let's now see how this is mathematically computed.

$$\mathrm{KL} = \langle \log p(R|X, M_{gen}) \rangle_{p(R|X,M_{gen})} - \langle \log p(R|X, M_{fitted}) \rangle_{p(R|X,M_{gen})} \tag{4.1}$$

with M_{gen} being the generative model, M_{fitted} one of the models fitted on the synthetic data, X a test dataset and finally R the spike counts sampled from the generative model in response to the test stimulus.

The first term in 4.1 is the same for all the synthetic stimuli tested, while the second is different for each of them. It represents the average log-likelihood of the generated responses R under the fitted model:

$$\langle \log p(R|X, M_{fitted}) \rangle_{p(R|X,M_{gen})} = \sum_R \log p(R|X, M_{fitted}) \cdot p(R|X, M_{gen}) \tag{4.2}$$

In the resolution of this sum, rather than computing the exact value, which might be analytically challenging, we exploit Monte Carlo method and approximate the result through random sampling, generating N_1 samples of responses R from the distribution described by the Gaussian generative model.

$$\sum_R \log p(R|X, M_{fitted}) \cdot p(R|X, M_{gen}) \sim \frac{1}{N_1} \sum_{R \sim p(R|X,M_{gen})} \log p(R|X, M_{fitted}) \tag{4.3}$$

The right-hand term can be further developed as:

$$\frac{1}{N_1} \sum_{R \sim p(R|X,M_{gen})} \log \int p(R|\lambda) p(\lambda|X, M_{fitted}) d\lambda \tag{4.4}$$

It is possible to approximate the integral in 4.4 using again the Monte Carlo integration, by utilizing random samples of λ drawn from the distribution $p(\lambda|X, M_{fitted})$. This distribution is the one inferred from the fitted Gaussian process.

$$\frac{1}{N_1} \sum_{R \sim p(R|X,M_{gen})} \log \left(\frac{1}{N_2} \sum_{\lambda \sim p(\lambda|X,M_{fitted})} p(R|\lambda) \right) \tag{4.5}$$

Let's further simplify equation 4.5. Knowing that R has a behavior guided by a Poisson distribution and assuming independence across different time instants, we can express its

probability as follows:

$$p(R|\lambda) = \prod_{time} \frac{f^R e^{-f}}{R!} \tag{4.6}$$

with f equal to $e^{A\lambda+\lambda_0}$.

The logarithm of this quantity is:

$$\log p(R|\lambda) = \sum_{time} (R \log f - f - \log R!) \tag{4.7}$$

So, the final result is:

$$\frac{1}{N_1} \sum_{R \sim p(R|X, M_{gen})} \log \left(\frac{1}{N_2} \sum_{\lambda \sim p(\lambda|X, M_{fitted})} \exp \sum_{\text{time}} (R \log f - f - \log R!) \right) \tag{4.8}$$

Similarly, we can compute the first term of the divergence, that also in this case represents a average log-likelihood, but this time according to the distribution provided by the generative model.

$$\frac{1}{N_1} \sum_{R \sim p(R|X, M_{gen})} \log \left(\frac{1}{N_2} \sum_{\lambda \sim p(\lambda|X, M_{gen})} \exp \sum_{\text{time}} (R \log f - f - \log R!) \right) \tag{4.9}$$

The difference from the previous term is that in this case we draw samples of λ from the distribution described by the generative gaussian process, trained on experimental data.

Essentially, the computation involves evaluating the probability of observing the given data under the distribution postulated by the model. The likelihood estimation provides insight into how effectively the fitted model captures and describes the data inferred from the generative one. This measure has been computed for all the 22 cell, using the models fitted on the synthetic stimuli mentioned in Section 4.2.2. In the end, for each couple cell-model, two different KL divergences were evaluated, varying the test stimulus X in the equation. Specifically, in the first case we are considering the log-likelihoods of the generated spike counts in response to a multisize spots stimulus:

$$KL_{\text{disk}} = \langle \log p(R|X_{\text{disk}}, M_{gen}) \rangle_{p(R|X_{\text{disk}}, M_{gen})} - \langle \log p(R|X_{\text{disk}}, M_{fitted}) \rangle_{p(R|X_{\text{disk}}, M_{gen})}$$

In the second case we evaluate the likelihoods in response to a multisize checkerboard stimulus:

$$KL_{\text{check}} = \langle \log p(R|X_{\text{check}}, M_{gen}) \rangle_{p(R|X_{\text{check}}, M_{gen})} - \langle \log p(R|X_{\text{check}}, M_{fitted}) \rangle_{p(R|X_{\text{check}}, M_{gen})}$$

In the end, optimization strategy lies in selecting as optimal the stimulus that resulted in the fitted model yielding the lowest KL. This stimulus corresponds to the best stimulus for the characterization of RCG responses. It is possible to identify the best stimulus according to different subsets: for a single cell, for a specific test set (disks, i.e. our main objective, but also for the checkerboards), for all the cells together or for different subsets of stimuli.

5 | Results

5.1. Gaussian Process Performances

In this section the results of the trained gaussian processes are presented. They are evaluated in terms of performance in the prediction from test sets for which the true experimental response is known. As a measure of goodness-of-fit of the models the noise-corrected correlation, also called explained variance, was computed [13]. It was evaluated for each cell both on a set of held-out checkerboard frames and on a set of disks. The formula for the computation is the following:

$$R^2_{nc} = \left(\frac{\frac{1}{2}\left(c_{\hat{r}_o,\bar{r}_o} + c_{\hat{r},\bar{r}_e}\right)}{\sqrt{c_{\bar{r}_o,\bar{r}_e}}} \right)^2 \tag{5.1}$$

Basically, the responses are divided into even and odd trials, their mean values $\bar{r}_o$ and $\bar{r}_e$ are estimated and the reliability is defined as their correlation $c_{\bar{r}_o,\bar{r}_e}$. The computation is performed with respect to a model prediction $\hat{r}$. The metric has been computed for all the 22 cells, predicting the responses for both a checkerboard test set and on a multisize spots set.

5.1.1. Generative GP

We first wanted to see whether the gaussian processes trained on checkerboards and disks experimental data performed well. The results of explained variance achieved by the 22 RGCs models are summarized in Figure 5.1. The model shows high values for the noise-corrected correlation in the predictions on both the checkerboard and the disks test set, respectively with a value of 0.75 ± 0.11 and 0.84 ± 0.08 on average for all the cells. The inference of the firing rates in response to the multisize spots stimulus is not an actual prediction, since the model has already been exposed to these data and this will bias the results towards an higher value. However, it is important to assess that the explained variance is high, as the evaluation of the KL divergence assumes that the generative model accurately reproduces the actual experimental behavior of RGCs in response to

this stimulus. Example predictions are shown in Figures 5.2 and 5.3.

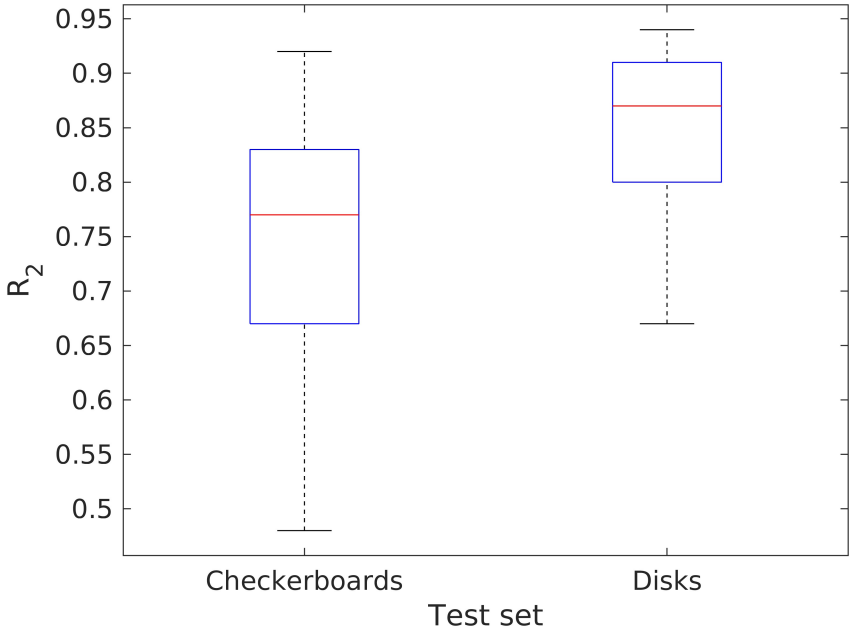

Figure 5.1: The two boxplots summarize the performance of the gaussian processes on the test sets. The left one shows the results obtained on the checkerboards test set, for which the model achieves a mean R_2 of 0.75 and a variance of 0.11 across all the cells. An analogous boxplot is shown on the right, this time considering multizise spots stimulus as test set. In this case we obtain a mean R_2 of 0.84 and a variance of 0.08.

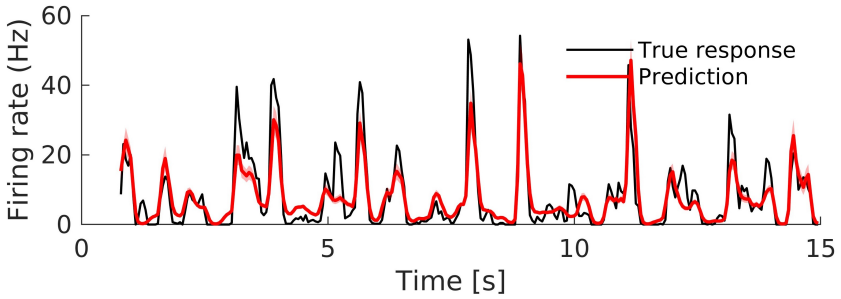

Figure 5.2: Prediction of the response to a checkerboard stimulus for cell 14. The prediction has been performed using a Gaussian model trained on experimental data, specifically on multisize spots and multisize checkerboard stimuli with the relative responses. The model is tested on an held-out checkerboard set, for which a response gathered experimentally is available. The true response is shown in black, while the prediction is in red. The continuous red line corresponds to the mean prediction of the model, while the shaded area represents the variance around the mean for each time point. The explained variance R_2 in this case is 0.88.

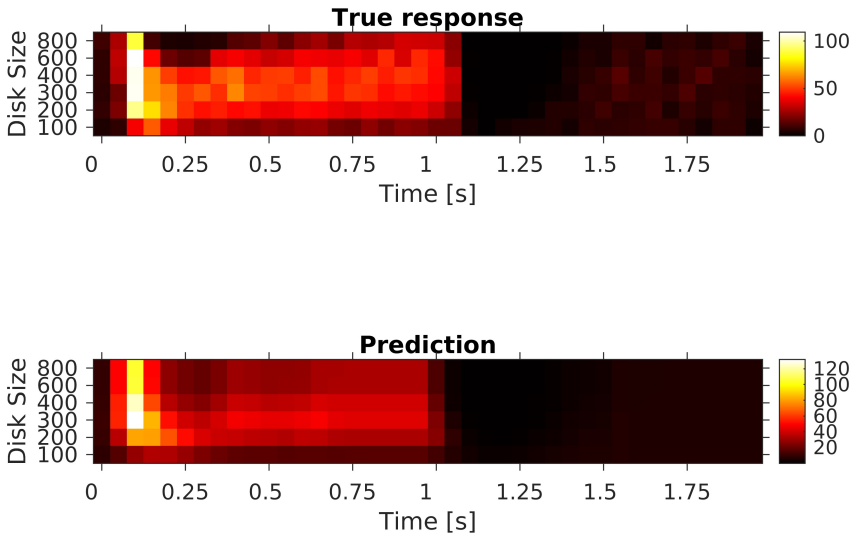

Figure 5.3: Prediction of the response to a disk stimulus for cell 14, performed using the generative gaussian model. The model is tested on a multisize spots set, for which the true response is known. The true response is shown in the above heatmap, while the prediction is shown in the down one. Along the x-axis there is the time evolution of the response, the y-axis refers to the different spot sizes in μm shown to the cell. The firing rate is colour coded. The model achieved for this cell an explained variance R_2 of 0.93.

5.1.2. Fitted GPs

Then, we passed to the assessment of the performance of the gaussian models fitted on the synthetic data. The results of explained variance for the models trained on different single sizes are summarized in Figure 5.4. It is possible to see that almost never the results achieved with the model trained on the smallest size are good.

Regarding instead the performances of the models trained on checkerboards with varying percentage of presence of black checks, summarized in figure 5.5, the main observation that can be done is that the models that achieve the best performances are the ones with a medium presence of black, thus the ones in which the two colours are more balanced.

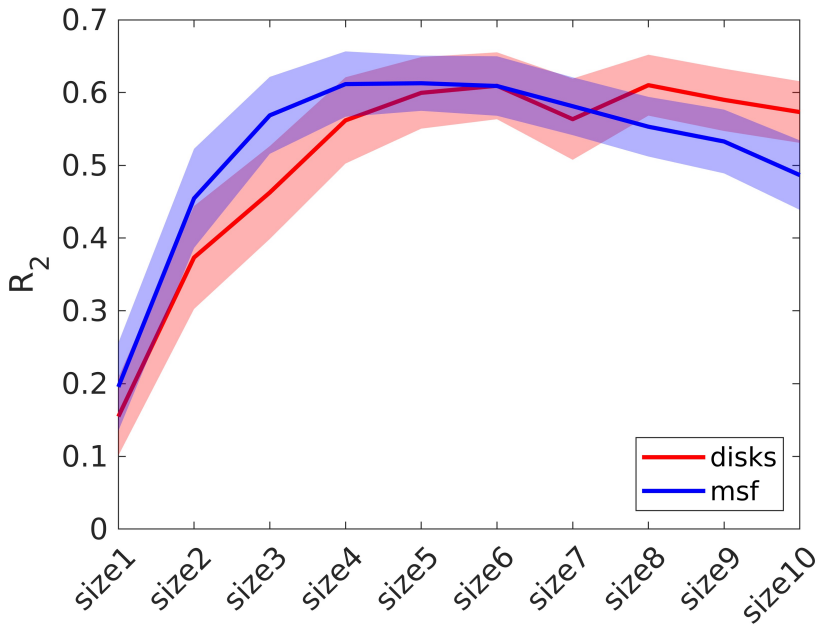

Figure 5.4: Results of explained variance obtained for models trained on different synthetic data, specifically on ten single-sized checkerboards. The models have been tested on a multisize spots test set (in red) and on a checkerboards test set (in blue). The continuous line represents the mean value across all the cells, while the shaded area refers to the standard error around the mean.

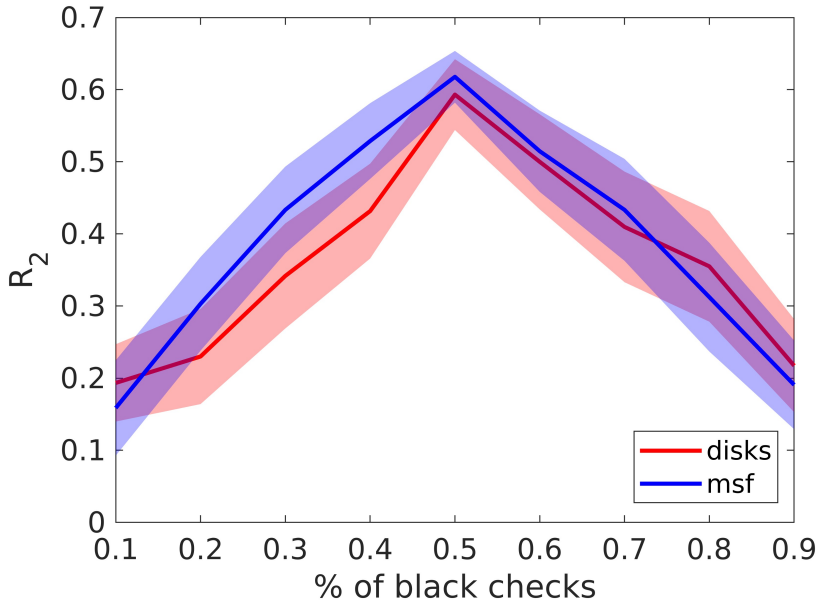

Figure 5.5: Results of explained variance obtained for models trained on different synthetic data, specifically with different percentage of presence of black tiles, which is labelled along the x-axis. The models have been tested on a multisize spots test set (in red) and on a checkerboards test set (in blue).

Finally, also the performances of the models trained on combinations of sizes have been evaluated, so the models fitted on uniform stimulus, only on the small sizes and only on the big sizes. The obtained explained variances are summarised in Figure 5.6.

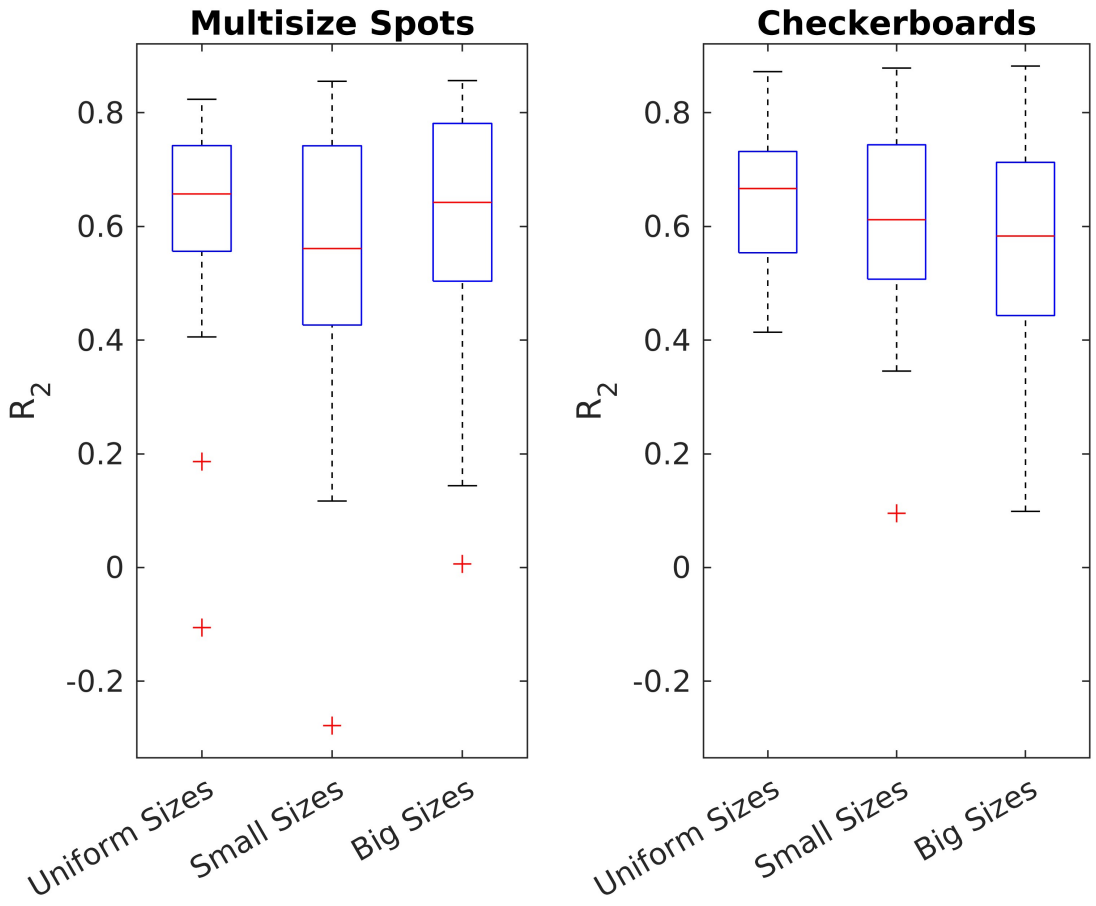

Figure 5.6: Boxplots summarizing results of explained variance obtained for models trained on the synthetic stimuli labelled along the x-axis. In the left panel results for the multisize spots test set are shown, while in the right panel the ones for checkerboard test set are illustrated.

5.2. Kullback-Leibler Divergence

After having verified that the Gaussian models trained on the experimental data (the generative models) performed well, we proceeded with the evaluation of the Kullback-Leibler divergence.

5.2.1. Single Sizes

Upon reaching the stage when models are trained using synthetic stimuli of individual sizes, described in Section 4.2.2, an analysis of predictions from various models allows to make an initial observation. Specifically, it becomes evident that the models' performances vary depending on the training stimulus. To illustrate this, Figure 5.7 presents prediction outcomes for a specific toy cell using two distinct models.

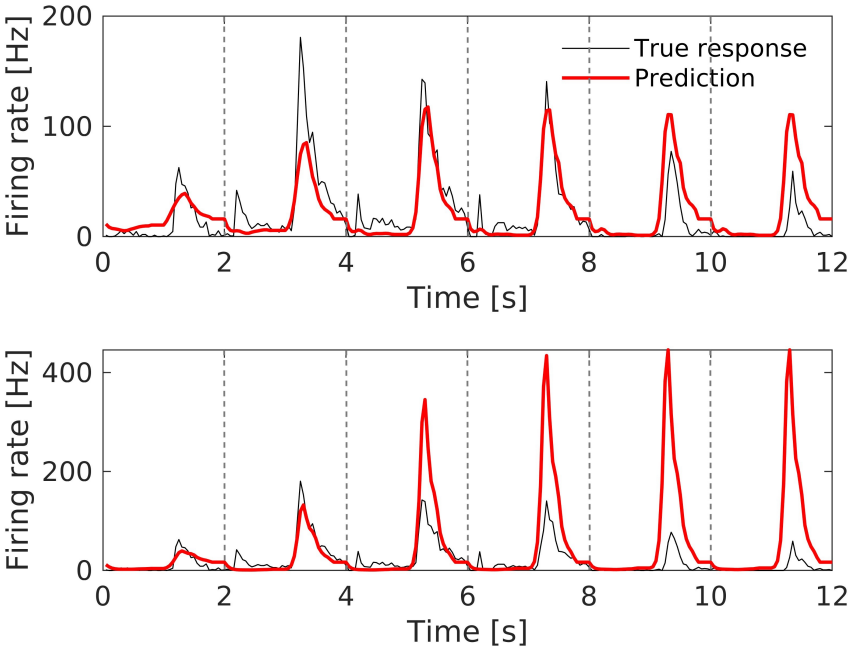

Figure 5.7: In the upper panel, the true response of cell 20 to a multisize spots stimulus is depicted in black, while the red line represents the prediction of the response to the same stimulus generated by a model exclusively trained on checkerboards of a single size. In the lower panel, a similar graph is presented, with the distinction that the model used for the prediction has been trained on a different size of checkerboards. The vertical dashed lines separate the responses to various spot sizes, ranging from the smallest to the largest.

Notably, the prediction in the first case appears to align more closely with the actual data, whereas the prediction in the second case is less accurate. The key distinction between these two models lies in their training stimuli – the first model was exclusively trained on checkerboards with size 6, while the second model was trained on checkerboards with size 1. So the training stimulus, and specifically the sizes of the checks, influence the development of a high-quality model. Is there a way to quantitively assess the goodness of the training stimulus? To answer this question, the Kullback-Leibler divergence, extensively described in 4.2.4, has been introduced.

The metric has been computed for all the cells, each time considering models trained on ten single-sized checkerboards stimuli. Results for cell 20 are shown in Figure 5.8. We can see that the metric has its minimum value in correspondence of size 6 and its maximum for size 1. This two sizes are exactly the ones employed in the training of the models in Figure 5.7: the upper panel shows the prediction of the model trained on checkerboard with size 6, the lower one instead the prediction of the model trained on checkerboard with size 1.

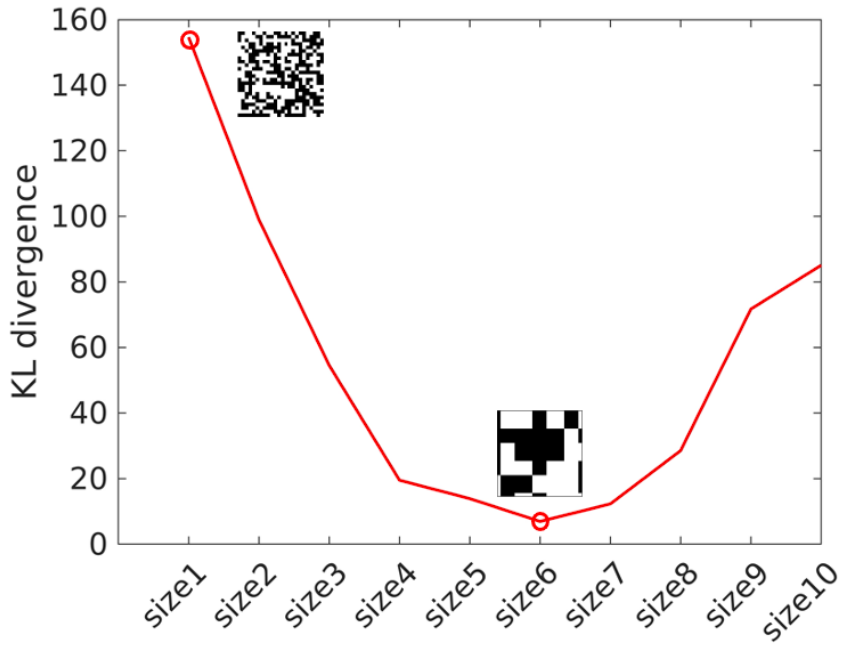

Figure 5.8: Kullback-Leibler divergence for the prediction to the multisize spots stimulus for cell 20, computed for models trained on 10 single-sized stimuli. It has a minimum in correspondence of size 6, meaning that this size of checkerboards allows to train a model that achieves a good fit on the disks test set; in contrast, it has a maximum in size 1, meaning that this size is worse in training a model able to achieve a good fit.

To delve deeper into the investigation, the KL divergence was also computed using a sequence of checkerboards as test set. This test set is composed by checkerboards presenting all the same check size, in particular each check has a dimension of 56 μm. The results are shown in Figure 5.9. In this case the stimuli identified as the best and worst differ from the previous analysis, highlighting that the optimal stimulus is also influenced by the choice of the test set. This suggests that the most suitable stimulus can vary based on the specific task. To illustrate this point, let's compare the metrics in the two situations (see Figure 5.10). The two trends exhibit notable differences. We can conclude

that, depending on the intended task, the optimal stimulus may change. Therefore, in stimulus optimization procedures, it is crucial to clearly define the ultimate objective as it will influence the final result.

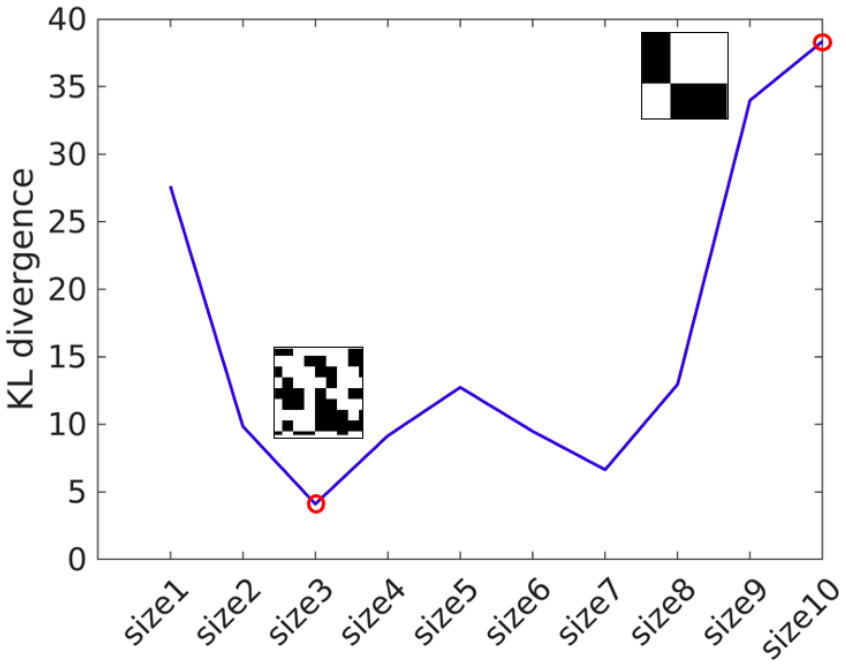

Figure 5.9: Kullback-Leibler divergence for the prediction to the checkerboard test set for cell 20, computed for models trained on 10 single-sized checkerboard stimuli. It has a minimum in correspondence of size 3, meaning that this size of checkerboards allows to train a model that achieves a good fit on this test set; in contrast, it has a maximum in size 10, meaning that this size is worse in training a model able to achieve a good fit.

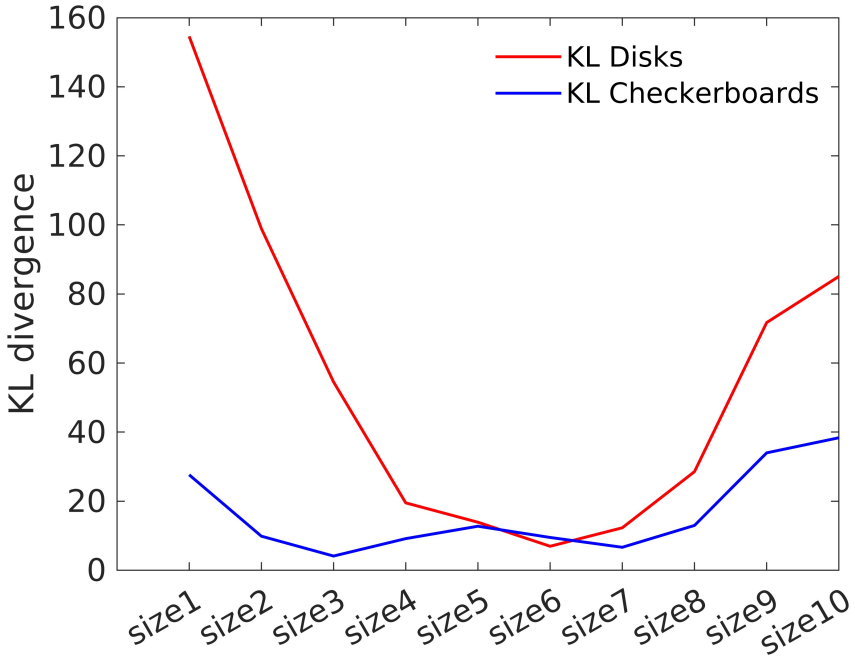

Figure 5.10: Kullback-Leibler divergence for the prediction to both the checkerboard (in blue) and the disk stimulus (in red) for cell 20, computed for models trained on 10 single-sized stimuli. The two curves have different minimum and maximum and exhibit a different trend.

Taking a more comprehensive view, let's shift the focus to the results obtained for all the cells. In Figure 5.11, the averaged Kullback-Leibler divergence for all cells is presented. Once again, the graph reveals distinct results when altering either the training set (i.e., the single-sized checkerboard stimulus) or the test set (i.e., the stimulus for which the model's fitting quality is assessed). The method is able to identify between the single-sized stimuli the best one for the characterization of RGC responses to the disks: in particular, checkerboards from size 4 on, and in particular size 6 with a mean KL of 12.1, allow to well characterize these responses. We can also see that instead the response to the checkerboards is better characterized by the small sizes, in particular the best stimulus turns out to be the single-sized checkerboard set with size 3. Numerical results of Kullback-Leibler divergence obtained for these stimuli are shown in Table 5.1.

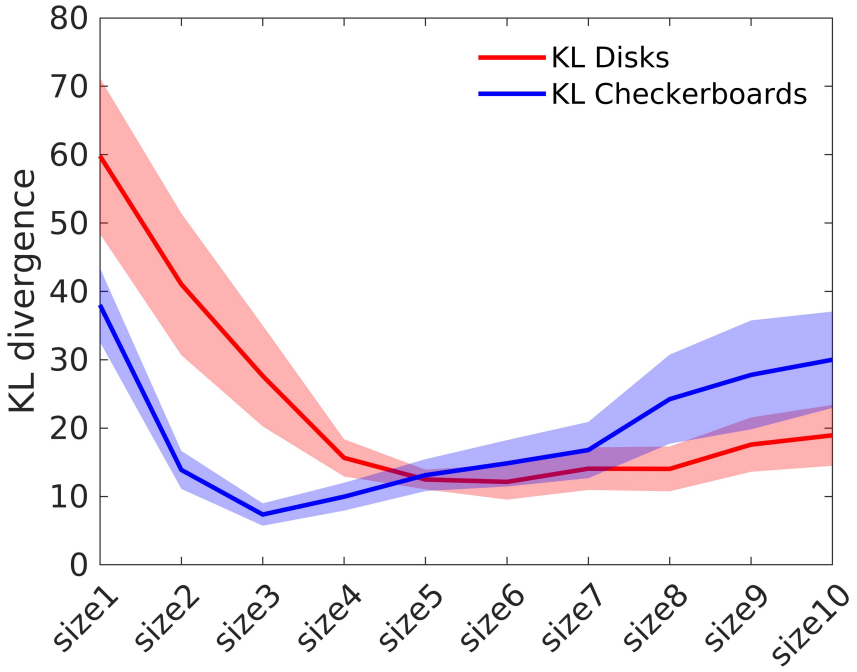

Figure 5.11: The graph displays the average KL divergence for all cells in response to the checkerboard stimulus in blue, and in response to the multisize spots in red. The shaded areas correspond to the standard error among different cells.

5.2.2. Uniform Stimulus, Small Sizes & Big Sizes

The KL divergence was also assessed on models trained on multiple sizes. Specifically, it was computed for stimuli containing all sizes in equal proportions, stimuli with only the five smallest sizes, and stimuli with only the largest ones. The results averaged on all the cells are depicted in Figure 5.12.

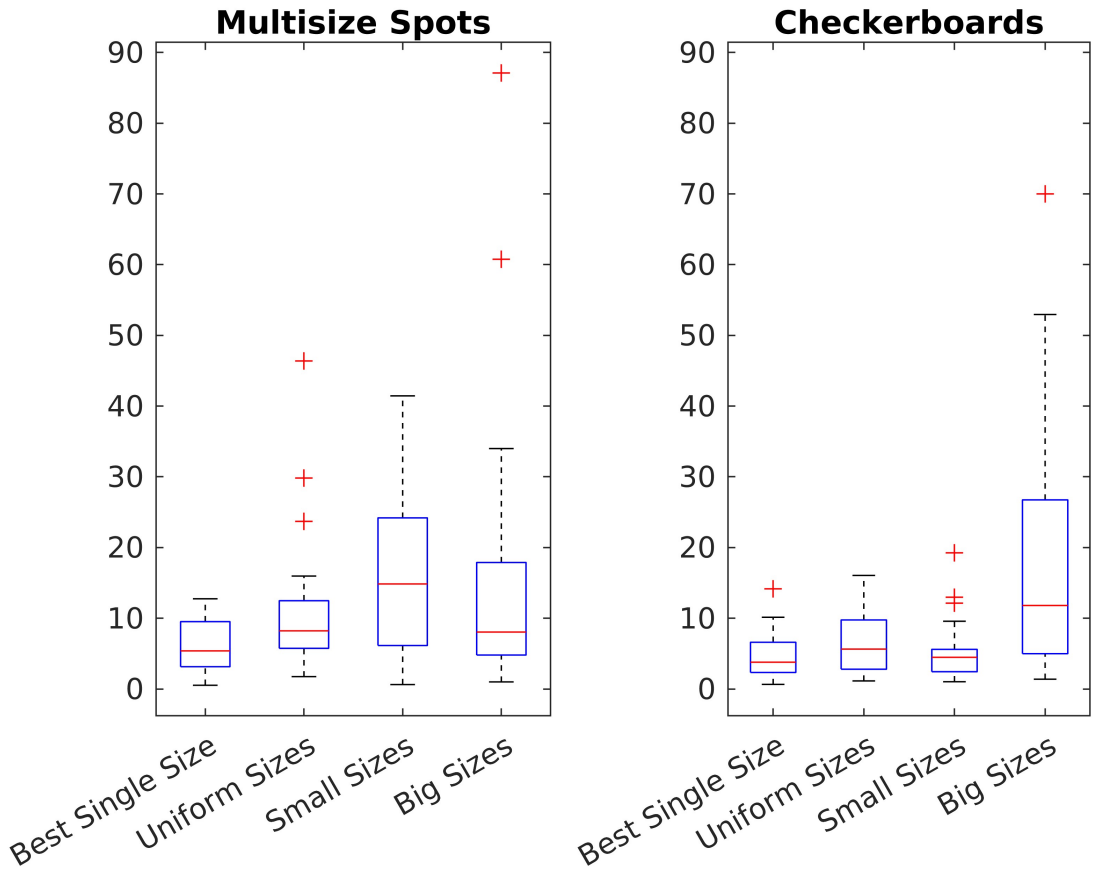

Figure 5.12: Boxplots summarizing results of KL divergence obtained for models trained on the synthetic stimuli labelled along the x-axis. In the left panel results for the multisize spots test set are shown, while in the right panel the ones for checkerboard test set are illustrated.

Observing the boxplots obtained from the multisize spot stimulus test set, it is possible to notice that, on average for all the cells, models trained on big sizes achieve slightly better fits with respect to the ones trained on the smalls, even if a lower KL is obtained when all the sizes are equally present in the dataset. Analyzing the boxplot for the checkerboards test set reveals a distinct behavior: in this case, checkerboards with smaller checks yield a better performance. This reinforces the conclusion that the results are indeed dependent on the specific characteristics of the chosen test set. To maintain a link with what was

explained in Section 5.2.1, a boxplot referencing the single sizes has been included to facilitate a direct comparison with the other stimuli. Specifically, for each cell, only the size that resulted in the lowest KL value was considered. It is evident that models trained on a carefully chosen single-sized stimulus can produce similar and even better results compared to combinations of sizes. Numerical results of Kullback-Leibler divergence obtained for these stimuli, together with the ones for the single-sized checkerboards, are shown in Table 5.1. We can see that the overall best stimulus for the characterization of RGC responses to the disks is the one presenting all the sizes in a uniform way. Analyzing instead the results for the checkerboards, the optimal stimulus is the combination of small sizes.

STIMULUS	$\mathbf{KL}_{disks}$	$\mathbf{KL}_{checks}$
Size 1	59.8 ± 11.4	38 ± 5.4
Size 2	41 ± 10.4	13.9 ± 2.8
Size 3	27.6 ± 7.4	7.4 ± 1.6
Size 4	15.6 ± 2.7	10 ± 2
Size 5	12.5 ± 1.4	13.1 ± 2.3
Size 6	12.1 ± 2.6	14.8 ± 3.4
Size 7	14.1 ± 3.1	16.8 ± 4.1
Size 8	14 ± 3.3	24.2 ± 6.5
Size 9	17.6 ± 4	27.8 ± 8
Size 10	18.9 ± 4.5	30 ± 7.1
Uniform Sizes	**11.6 ± 2.2**	6.6 ± 1
Small Sizes	16.6 ± 2.5	**5.4 ± 0.9**
Big Sizes	16.1 ± 4.5	18.9 ± 4

Table 5.1: Numerical results obtained for the Kullback-Leibler divergence for stimuli of varying sizes. The values are presented in the format Mean ± Standard Error, evaluated on all the cells. For both the test sets, the lowest value of KL is highlighted in bold.

5.2.3. Different Percentages of Black Checks

The Kullback-Leibler divergence was also assessed by varying the probability of presence of black tiles in the training sets. The averaged results for all cells are presented in Figure 5.13 and in Table 5.2. It can be noticed that the cells tend to manifest a preference for checkerboards with 50% presence of white and black, while worse results are achieved going towards the extremes, i.e. a predominance of white or black. This behavior is more pronounced in the evaluation of the multisize spots stimulus test set, even if it is present also for the checkerboard set. This can be seen as a confirmation of the choice of exploring stimuli with varying sizes or combinations of sizes while consistently maintaining a 50%

balance between white and black components.

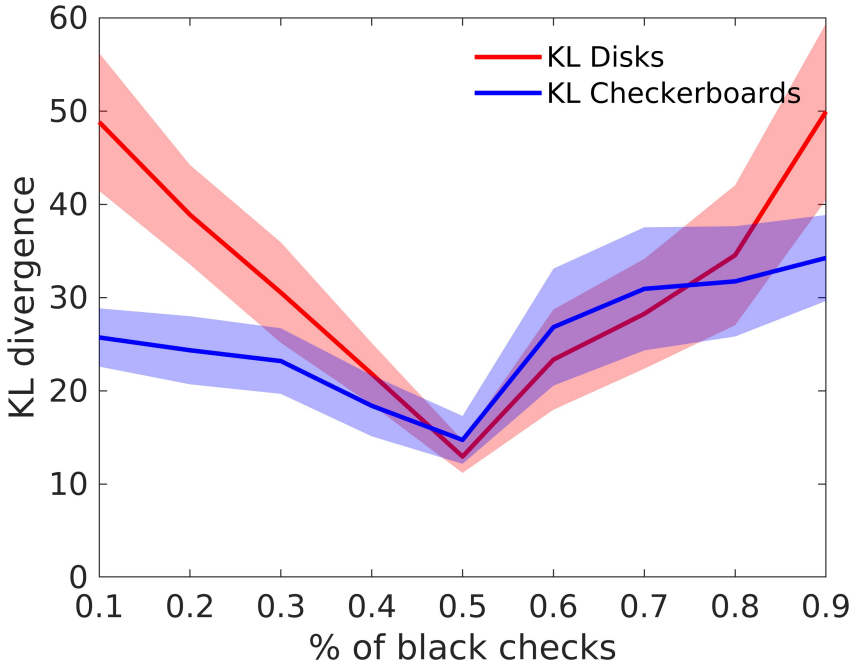

Figure 5.13: The results of KL obtained for models trained on synthetic stimuli with varying presence of black tiles are here shown, averaged on all the cells. The red line refers to the measure computed on a multisize spots test set, the blue one on a checkerboards test set. The shaded area refers to the standard error around the mean among all the cells.

% of Black	KL_{disks}	KL_{checks}
10%	48.8 ± 7.4	25.7 ± 3.1
20%	38.9 ± 5.3	24.3 ± 3.7
30%	30.5 ± 5.4	23.2 ± 3.5
40%	21.8 ± 3.3	18.4 ± 3.3
50%	**12.9 ± 1.8**	**14.7 ± 2.6**
60%	23.3 ± 5.4	26.8 ± 6.3
70%	28.3 ± 5.9	30.9 ± 6.6
80%	34.5 ± 7.5	31.7 ± 5.9
90%	49.9 ± 9.5	34.2 ± 4.6

Table 5.2: Numerical results obtained for the Kullback-Leibler divergence for stimuli with varying percentages of black tiles. The values are presented in the format Mean ± Standard Error, evaluated on all the cells. For both the test sets, the lowest value of KL is highlighted in bold.

5.3. Summary of Results

The family of models chosen to replicate the behavior of the RCGs is the one of the gaussian processes. First of all, the performances of the models are evaluated in the prediction from test sets for which the true experimental response is known, computing the noise-corrected correlation. Regarding the GPs trained on experimental data, it was possible to assess that they are able to accurately forecast the response to the presented stimuli, achieving high values for the noise-corrected correlation in the predictions on both the checkerboard and the disks test set, respectively with values of 0.75 ± 0.11 and 0.84 ± 0.08 on average for all the cells. Regarding instead the GPs trained on the synthetic data, the key takeaways are the fact that almost never the results achieved with the model trained on the smallest size are good and that, varying the percentage of black in the training stimulus, the models that achieve the best performances are the ones with 50% of presence of black checks.

Then, Kullback-Leibler divergence was computed to assess the performance of the synthetic stimuli in RCG characterization. Analyzing this measure, we understood that the best stimulus for characterizing RGC responses to the disks, among checkerboards with varying percentages of black checks, is the one that exhibits a balanced presence of black and white checks. Fixing the percentage of black checks at 50% and changing the size of the tiles, it was evaluated that the best single-sized checkerboard stimulus is the one entailing checks of size 6, i.e. 378 μm, with a KL of 12.1 ± 2.6. Incorporating also combination of sizes in the analysis, the overall lowest KL was achieved by the stimulus containing all check sizes, with a value of 11.6 ± 2.2, thus making it the optimal one for the characterization of RCGs responses to the disks.

A parallel analysis was performed also for the characterization of RGC responses to a checkerboards test set, to see whether the method is able to generalize also for different tasks. In this case the overall lowest KL was achieved by the stimulus containing the combination of the five smallest sizes, with a value of 5.4 ± 0.9, making it the best one for the characterization of RCGs responses to the checkerboards test set.

6 | Discussion & Conclusions

The implemented method successfully enables the assessment of the performance of different temporal checkerboard stimuli in achieving a well-fitting model response to multisize spots stimuli. The objective is to obtain accurate predictions for the disks because, according to the findings in [2], investigating the firing rates evoked by multisize spots stimuli can provide valuable insights for an accurate and meaningful functional typing of retinal ganglion cells.

First of all, we have verified that the generative GPs were able to accurately predict the response to the stimuli, both the checkerboard and the multisize spots. Firstly, this aspect is crucial as we depend on accurate predictions for checkerboards when employing the model to forecast the firing rates of RGCs in response to synthetic temporal checkerboard stimuli. Secondly, we also rely on its capability to predict responses to the multisize spot stimulus. This reliance stems from our assumption when computing the Kullback-Leibler divergence, namely that the generative GP can accurately reproduce the behavior of a real RCG. Consequently, we regard its modeled distribution as the gold standard to be attained by the fitted GPs, thus it is essential for the generative models to achieve good performances in this task.

Moving to the KL divergence computation, one first observation that can be done about the results is that the accuracy in the fit varies in dependence on the training data, and this difference is quite evident in some of the cells (Figure 5.7). To quantify this diversity in goodness of fit, the KL divergence of the fitted model with respect to the generative model was exploited. The metric quantified for these cells large changes in quality of the fit. However, for some other neurons the difference in quality of the fit with dependence to the stimulus is quite small, and this somehow limits how far we can go with optimization.

It was further deduced that the fitted models performances also depend on the test set, as can be seen observing the different trends for the disks and the checkerboards test sets in Figures 5.10 and 5.11. This suggests that the optimal stimulus can vary based on the specific task to be performed. Therefore, in stimulus optimization procedures, it is crucial to clearly define the ultimate objective as it will influence the final result.

Also, the best stimulus identified by the method is not the same for each cell. It may be overly optimistic to expect to find a stimulus that is simultaneously optimal for each individual cell. Instead, a more realistic approach is to search for the best trade-off that accommodates the characteristics of all cells collectively. Analyzing the results achieved on average for all the cells for the single-sized stimuli, we can see that the lowest KL for the disks is reached by the model trained on checkerboards with checks of size 378 μm (Figure 5.11), and more in general low KL values are obtained with sizes growing from size 4 on. This was not the case with the checkerboard test set, for which lower KL values were achieved with smaller sizes of checks. Intuitively, we have better results if the training stimulus matches the test one: the stimuli yielding low KL values for disks are the ones characterized by bigger checks, more similar in size to the disks and thus leading to better models for their predictions. Instead for the checkerboards, characterized by checks of small sizes, better results for them are achieved with training stimuli of small size. However, including in the analysis also models trained on combinations of checkerboards with different sizes, we can see that the overall lowest KL is achieved by the stimulus containing all check sizes (Table 5.1), suggesting that training the model on a wide range of sizes empowers its generalization capability and allows to achieve a better accommodation of the characteristics of the different RGCs. Nevertheless, the difference in KL results between the best single-sized stimulus and the uniform stimulus is small, thus making both of them reasonable choices for RGCs characterization responses to disks.

Finally, it was assessed that if the checkerboard frames used in the training exhibit an equal presence of black and white elements, the model is more likely to achieve a good fit (Figure 5.13). This can be seen as a confirmation of the choice of exploring stimuli with varying sizes or combinations of sizes while consistently maintaining a 50% balance between white and black components.

6.1. Innovations of the Work

Previous studies on typing have utilized stimuli specifically chosen ad hoc by hand, typically based on prior knowledge about retinal ganglion cells. In this study, we introduce a novel model-based stimulus optimization approach that enables the automatic determination of the optimal stimulus for characterizing retinal ganglion cells. The method has been developed with the specific goal of optimizing the stimulus for the characterization of RGCs responses to multisize spots. This is grounded on the assumption that achieving accurate predictions for responses to the disks will enhance functional typing. However, the method is versatile and can be applied to other objectives, facilitating stimulus opti-

mization for various tasks beyond the scope of this study. For instance, here we provide an example of optimization for characterizing RGC responses to a checkerboard test set.

This method differs from many other stimulus optimization techniques in that we do not seek the maximally exciting stimulus, a common approach often adopted, for example by Ponce et al. in [15]. Rather we look for the best stimulus in an open loop manner, creating the stimuli to be tested and evaluating them with a model-based approach, without the need to perform further experiments beyond the ones needed for the training of the generative model.

Regarding the RGC modelling part, gaussian processes have been exploited. This modelling approach was also employed by Goldin et al. [13]. However, while their work focused on modelling neuronal responses to static images, our study extends this approach to model responses to temporal stimuli.

6.2. Limitations and Future Developments

As stated in the previous section, a peculiar feature of this method lies in its dependence on the performance of the model generating responses to synthetic stimuli. If the model struggles to accurately predict responses, the subsequent models trained on these data may also falter in making reliable predictions, and this will not depend on the training synthetic stimulus but on the generative model: the success of the overall approach heavily relies on the effectiveness of the model. Ensuring the generative model is optimized is crucial. The process of enhancing the generative model could involve exploring different hyperparameters in its training and for its approximations, or also operating on the training data, perhaps investigating different preprocessing techniques.

Also, the Kullback-Leibler divergence was chosen as a measure to assess the quality of the fitted models. However, alternative metrics can serve the same purpose, such as the noise-corrected correlation between the predictions provided by generative and fitted models. Although the chosen metric offers the advantage of considering the uncertainty given in output by the gaussian model, it comes with the drawback of being unbounded. This characteristic can sometimes pose challenges when evaluating various cells and sets of stimuli, making the comparison process less straightforward.

Furthermore, to enhance this approach, a more extensive exploration of the stimulus space is recommended, including more combinations of sizes than those covered in this analysis. Additionally, exploring other stimulus features, such as the rate at which frames are presented to the cells, can contribute helpful insights.

Finally, introducing a final typing step in the procedure, beyond solely aiming for the best multisize spots response prediction, could add valuable depth to the analysis. This additional layer in the method could contribute to a more comprehensive analysis of the stimulus effectiveness in achieving accurate typing outcomes.

6.3. Conclusions

In this thesis, I have worked on the development of a method capable of identifying the optimal stimulus for effectively learning a model of retinal ganglion cells (RCGs), enabling accurate prediction of responses to multisize spots stimuli. The primary motivation behind this project stems from the desire to conduct effective typing of retinal ganglion cells (RCGs). Literature has demonstrated that proficient typing can be achieved based on responses to multisize spots.

The stimuli considered for analysis comprised temporal sequences of checkerboards, featuring varying check sizes and percentages of black checks. The implemented method successfully enables the assessment of the performance of different temporal checkerboard stimuli in achieving a well-fitting model response to multisize spots stimuli.

Our results have demonstrated that the goodness-of-fit of the model depends on the stimulus used on its training. In particular, the best checkerboard stimulus, among checkerboards with varying percentages of black checks, turned out to be the one that exhibits a balanced presence of black and white checks (Figure 5.13). Fixing the percentage of black checks at 50% and changing the size of the tiles, it was evaluated that the best single-sized checkerboard stimulus is the one entailing checks of size 6, i.e. 378 μm (Figure 5.11). Incorporating also combination of sizes in the analysis, the overall lowest KL was achieved by the stimulus presenting all the checks sizes in equal proportion (Table 5.1), thus making it the optimal one for the characterization of RCGs.

Bibliography

[1] T. Baden, P. Berens, K. Franke, M. Román Rosón, M. Bethge, and T. Euler, "The functional diversity of retinal ganglion cells in the mouse," *Nature*, vol. 529, no. 7586, pp. 345–350, 2016.

[2] J. Goetz, Z. F. Jessen, A. Jacobi, A. Mani, S. Cooler, D. Greer, S. Kadri, J. Segal, K. Shekhar, J. R. Sanes, *et al.*, "Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression," *Cell reports*, vol. 40, no. 2, 2022.

[3] O. Marre, D. Amodei, N. Deshmukh, K. Sadeghi, F. Soo, T. E. Holy, and M. J. Berry, "Mapping a complete neural population in the retina," *Journal of Neuroscience*, vol. 32, no. 43, pp. 14859–14873, 2012.

[4] E. Chichilnisky, "A simple white noise analysis of neuronal light responses," *Network: computation in neural systems*, vol. 12, no. 2, p. 199, 2001.

[5] J. Pillow, "Likelihood-based approaches to modeling the neural code,"

[6] A. F. Meyer, R. S. Williamson, J. F. Linden, and M. Sahani, "Models of neuronal stimulus-response functions: elaboration, estimation, and evaluation," *Frontiers in systems neuroscience*, vol. 10, p. 109, 2017.

[7] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. Chichilnisky, and E. P. Simoncelli, "Spatio-temporal correlations and visual signalling in a complete neuronal population," *Nature*, vol. 454, no. 7207, pp. 995–999, 2008.

[8] D. L. Yamins, H. Hong, C. F. Cadieu, E. A. Solomon, D. Seibert, and J. J. DiCarlo, "Performance-optimized hierarchical models predict neural responses in higher visual cortex," *Proceedings of the national academy of sciences*, vol. 111, no. 23, pp. 8619–8624, 2014.

[9] S.-M. Khaligh-Razavi and N. Kriegeskorte, "Deep supervised, but not unsupervised, models may explain it cortical representation," *PLoS computational biology*, vol. 10, no. 11, p. e1003915, 2014.

[10] N. Kriegeskorte, "Deep neural networks: a new framework for modeling biological

vision and brain information processing," *Annual review of vision science*, vol. 1, pp. 417–446, 2015.

[11] A. J. Kell, D. L. Yamins, E. N. Shook, S. V. Norman-Haignere, and J. H. McDermott, "A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy," *Neuron*, vol. 98, no. 3, pp. 630–644, 2018.

[12] N. Maheswaranathan, L. T. McIntosh, H. Tanaka, S. Grant, D. B. Kastner, J. B. Melander, A. Nayebi, L. E. Brezovec, J. H. Wang, S. Ganguli, *et al.*, "Interpreting the retinal neural code for natural scenes: From computations to neurons," *Neuron*, vol. 111, no. 17, pp. 2742–2755, 2023.

[13] M. A. Goldin, S. Virgili, and M. Chalk, "Scalable gaussian process inference of neural responses to natural images," *bioRxiv*, pp. 2023–01, 2023.

[14] M. Park, G. Horwitz, and J. Pillow, "Active learning of neural response functions with gaussian processes," *Advances in neural information processing systems*, vol. 24, 2011.

[15] C. R. Ponce, W. Xiao, P. F. Schade, T. S. Hartmann, G. Kreiman, and M. S. Livingstone, "Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences," *Cell*, vol. 177, no. 4, pp. 999–1009, 2019.

[16] Y. Yamane, E. T. Carlson, K. C. Bowman, Z. Wang, and C. E. Connor, "A neural code for three-dimensional object shape in macaque inferotemporal cortex," *Nature neuroscience*, vol. 11, no. 11, pp. 1352–1360, 2008.

[17] E. T. Carlson, R. J. Rasquinha, K. Zhang, and C. E. Connor, "A sparse object coding scheme in area v4," *Current Biology*, vol. 21, no. 4, pp. 288–293, 2011.

[18] P. Yger, G. L. Spampinato, E. Esposito, B. Lefebvre, S. Deny, C. Gardella, M. Stimberg, F. Jetter, G. Zeck, S. Picaud, *et al.*, "Fast and accurate spike sorting in vitro and in vivo for up to thousands of electrodes," *BioRxiv*, p. 067843, 2016.

[19] S. Wienbar and G. W. Schwartz, "The dynamic receptive fields of retinal ganglion cells," *Progress in retinal and eye research*, vol. 67, pp. 102–117, 2018.

[20] E. P. Simoncelli, L. Paninski, J. Pillow, O. Schwartz, *et al.*, "Characterization of neural responses with stochastic stimuli," *The cognitive neurosciences*, vol. 3, no. 327-338, p. 1, 2004.

[21] StphTphsn, "Illustration diagram for the spike-triggered average," 2015.

[22] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, vol. 2. MIT press Cambridge, MA, 2006.

[23] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[24] M. Park and J. W. Pillow, "Receptive field inference with localized priors," *PLoS computational biology*, vol. 7, no. 10, p. e1002219, 2011.

[25] Y. Cho and L. Saul, "Kernel methods for deep learning," *Advances in neural information processing systems*, vol. 22, 2009.

[26] J. Hensman, N. Fusi, and N. D. Lawrence, "Gaussian processes for big data," *arXiv preprint arXiv:1309.6835*, 2013.

List of Figures

List of Tables