One of the primary goals in neuroscience is to figure out simple principles that explain how nervous systems are organized. Barlow (1961) proposed one of the most successful theories in this vein, which states that sensory systems should remove redundancies in their inputs to optimize the information they process. This *efficient coding hypothesis* provides us with a mathematical framework to understand how neurons *should* encode information, which can then be experimentally tested against how neurons *actually* do so. Over the past 60 years, efficient coding has successfully explained many experimental findings in different sensory modalities such as vision (Olshausen & Field, 1996; Rao & Ballard, 1999), audition (Lewicki, 2002) and touch (Miller et al., 2019). In particular, this hypothesis has been successful early on in the retina (Atick & Redlich, 1990, $$ more citations). Recent work has carried this even further by replicating the receptive fields of retinal ganglion cells (RGCs) across a population of neurons (Karklin & Simoncelli, 2011). My work will extend this even further by replicating how the retina integrates redundant inputs, either across different color channels (Aim 1) or due to internal noise (Aim 2). These results will allow us to understand how much of retinal physiology can be explained by efficient coding principles.

How the retina processes visual information follows a clear structure. Photoreceptors first transform light from the outside world into electrical activity. Photoreceptors then send this information to bipolar cells, and bipolar cells send this information to retinal ganglion cells (RGCs). These RGCs are the visual inputs to the brain, and send projections to the thalamus through the optic nerve. RGCs already pre-process the information they receive from photoreceptors, by having center-surround receptive fields. These receptive fields are circular and separated into two different regions, an excitatory center and an inhibitory surround. Early theoretical work (citation) has explained how this center-surround organization arises from efficient coding principles, both for achromatic and for color inputs

However, this early work made multiple simplifying assumptions to make the problem mathematically tractable, such as an infinite number of neurons and a linear output response. Because these models are too simple, they are not appropriate to try to explain how the retina works in more details. and are therefore of these assumptions, the model fails to explain why RGCs are separated into two different pathways: The ON pathway that processes increments of light in the center relative to the surround, and the OFF pathways that does the opposite. Early models also fail to explain why RGCs within a pathway encode distinct regions of visual space, repelling each other to form a ‘mosaic’ that tiles the entire retina.

More recent work has leveraged machine learning to make efficient coding models that are more general and more relatable to retinal physiology. These new models explain why RGCs are separated into ON and OFF pathways, and why each pathway forms its own mosaic. My lab recently expanded these models to have receptive fields in both space and time. We showed that new functional types emerge as we increase the number of neurons, with each functional type having its own mosaic.

Thus we showed that one input channel can result in multiple mosaics. However, we know that

By doing so, we were able to explain the spatiotemporal receptive fields of RGC subtypes that were discovered. One of these functional types encodes high spatial and low temporal frequencies, similar to midget RGCs. Another one of these subtypes encodes low spatial and high temporal frequencies, similar to parasol RGCs. However, the spatiotemporal receptive fields of midget and parasol RGCs have been discovered more than 30 years ago. In general, we know a lot more about retinal physiology than what efficient coding can currently explain. My research aims to close this gap by showing how efficient coding can explain retinal physiology in more details. To do so, I will expand the current efficient coding algorithms to explain how RGCs process color information (Aim 1) and replicate motion encoding (Aim 2). I will also collaborate with Greg Field from UCLA to use his expertise in retinal physiology for our efficient coding models, and in the prospect of testing the predictions of our model against new experimental data.

**Aim 1:** Test whether efficient coding can explain how the retina encodes chromatic information

Retinal ganglion cells integrate inputs from cone photoreceptors, which come in three types – Long (L), Medium (M) and Short (S) –which roughly encode red, green and blue stimuli, respectively. Most (~70%) of RGCs are midget cells, which encode a combination of L and M cones. In the fovea, these cells are ‘red-green opponent’ and encode a contrast between L and M inputs. In the periphery, these midget cells are not red-green opponent but instead sum L and M cones inputs. Why midget cells use different coding strategies for the fovea and for the periphery is still unclear. We hypothesize those different strategies have to do with the ratio between RGCs and cones in the fovea versus the periphery. While the periphery has more cones than RGCs, the opposite is true for the fovea, with approximately 3 RGCs for every cone. Here we will test that hypothesis by building an efficient coding model for chromatic natural images. The inputs to this model will be the LMS cone responses to a series of natural images. We will then filter these responses through three different spatial filters (one for each cone), followed with an output non-linearity. The weights of the filter will be fit to optimize the mutual information between the cone inputs and the RGC outputs. Consistent with efficient coding principles (citations), there will be a constraint on the total firing rate across all neurons. To model the center versus the periphery, we will change the number of outputs RGCs relative to the number of input cones. Preliminary results suggest that if we keep the RGCs-cone ratio to 1:1, efficient coding predicts that L and M inputs should be summed (Figure 2). We will next see if we can replicate chromatic receptive fields in the fovea by building efficient coding models with a 3:1 RGCs-cones ratio. We will also test whether we can replicate receptive fields from RGC types other than midget cells, such as parasol and bistratified cells ($$). test whether efficient coding can accurately explain how the retina encodes chromatic information. dfsx

**Aim 2:** Test whether efficient coding models can explain why some RGCs are motion-selective

There are multiple types of RGCs that integrate space and time differently. Midget cells have high acuity in space, but poor temporal resolution. Parasol cells are the opposite: They have great temporal acuity, but poor spatial resolution. Recent work from my lab used efficient coding to explain why it is optimal for RGCs to have either high spatial or high temporal acuity. Building such a spatiotemporal efficient coding model is challenging. To overcome this challenge, we made some assumptions about…

However, an important property of RGCs that this previous model fails to explain is that many RGCs are motion-selective: for example, some neurons will respond stronger if an object moves from left to right than from right to left.

Hypothesis: We can replicate how the retina encodes motion from efficient coding principles

Neuronal activity is not only correlated in both space and time, but spatial and temporal correlations also interact with each other. The most prominent example of this phenomenon is motion, where we can predict the future location of a moving object based on its current location and velocity. While it is clear that the efficient coding strategy for RGCs should include encoding motion, what exactly this strategy is – how many neurons should process motion and what should their spatiotemporal receptive fields be– is still unclear*. My working hypothesis is that the efficient coding strategy for encoding motion in natural images will replicate experimental findings about motion encoding in RGCs.* To answer this question, I will extend the previous spatiotemporal efficient coding model from my lab to be spatiotemporally inseparable; that is, the model will be able to learn a receptive field that changes across time, a crucial property to encode motion. Completion of this aim will enlighten us as to whether how the retina encodes motion can be fully explained by the efficient coding hypothesis.

Whole paragraph about what we will learn and why this matters

References

# References

1

Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*(3), 308–320. https://doi.org/10.1162/neco.1990.2.3.308

References

Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*(3), 308–320. https://doi.org/10.1162/neco.1990.2.3.308