One of the primary goals in neuroscience is to figure out simple principles that explain how nervous systems are organized. One of the most successful theories in this vein states that sensory systems should remove redundancies in their inputs to optimize the information they process1. 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 vision2-7, audition8 and touch9. This hypothesis has been especially successful in the retina, where it can explain many features of retinal encoding such as center-surround receptive fields and ON-OFF pathways2, 3, 7, 10, 11. However, we still lack efficient coding predictions for how the retina processes many complex features of the visual world such as color. My work will tackle this problem by providing a theoretical account of how the retina integrates redundant inputs across different color channels. These results extend our understanding of how retinal physiology can be explained by efficient coding principles.

Retinal processing of visual information follows a well-known structure12, 13: First, photoreceptors transform light from the outside world into electrical activity. They then send this information to bipolar cells, and bipolar cells send this information to retinal ganglion cells (RGCs). These RGCs are the output layer of the retina. Their axons form the optic nerve and sends information to the thalamus, which then transmits it to the primary visual cortex. RGCs are separated into two different pathways (ON and OFF), and each neuron within a pathway processes a small region of visual space — its receptive field. These receptive fields form ‘mosaics’ (one per RGC type) that tile visual space. The receptive fields of RGCs have a center-surround organization: ON RGCs encode light in the center and dark in the surround, and vice-versa for OFF RGCs.

Early theoretical work on efficient coding explained how this center-surround organization arises from decorrelation, both for achromatic and for color inputs10, 11, but these made many simplifying assumptions, including an infinite number of neurons and linear output responses. While these assumptions help make the problem mathematically tractable, such models neglect key biological realities like strictly positive neural firing rates. More recent work has leveraged machine learning to make efficient coding models that incorporate non-linear output responses and a limited number of neurons4, 7, 14-16, accounting for the separation of RGCs into different subtypes, with neurons within a subtype forming a mosaic. However, the inputs to RGCs are much more complex than static achromatic images – RGCs receive input images from multiple color channels. While efficient coding can predict how the retina should process achromatic stimuli, its predictions for color involve simplifying assumptions (e.g. linearity17 and space-time separability15) which are difficult to relate to retinal physiology18, 19. To close this gap, I will use more general efficient coding models to model how RGCs process color information (Aim 1). In doing so, I will take advantage of the established collaboration between my supervisor and Dr. Greg Field (UCLA) on efficient coding models of the retina14, 15. This collaboration also opens the possibility of testing the predictions of our models against new experimental data.

Color is a crucial aspect of how we perceive the visual world. Encoding colors starts at the level of cone photoreceptors in the fovea of the retina, which come in three types –Long (L), Medium (M), and Short (S), roughly encoding red, green and blue stimuli, respectively20. This color information is ultimately encoded by RGCs, with different types integrating cone inputs differently. Most RGCs (~90% in the fovea and ~45% in the periphery) are midget cells13, which, in the fovea, are ‘red-green opponent’ and encode a contrast between L and M inputs13, 19, 20. In the periphery, these midget cells are not red-green opponent but instead sum L and M cones inputs13, 19, 21. 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 primate fovea, with approximately 3 RGCs for every cone22. Here we will test that hypothesis by building an efficient coding model for chromatic natural images (Figure 1). The inputs to this model will be the LMS cone responses to a series of images from the Kyoto Natural Images dataset23. We will then filter these responses through three different spatial filters (one for each cone), which are summed and then followed with an output non-linearity. The weights of the filter will be fit to optimize mutual information between the cone inputs and the RGC outputs:



Where represents the covariance matrix of the input natural images, and are diagonal matrices that represent the covariance of the input and output noise, respectively, and W is the weight matrix. G is a diagonal matrix that represents the local derivatives of the output responses for a specific set of input images. Since the output nonlinearity is a ReLU, the diagonal of G has binary values (1 if the neuron is firing and 0 if not). This function will be maximized using Adam optimization17. Consistent with efficient coding principles, each neuron will be restricted to have a fixed average firing rate using a Lagrange multiplier. 7, 10, 14, 15

Consistent with efficient coding principles, there will be a constraint on the total firing rate across all neurons7, 10, 14, 15. To model differences between the fovea and the periphery, we will change the number of outputs RGCs relative to the number of input cones. Preliminary results suggest that if we fix the RGC-cone ratio at 1:1, efficient coding predicts that L and M inputs should be summed. We will next test if we can replicate chromatic receptive fields in the fovea by building efficient coding models with a 3:1 RGC-cone ratio. We will also test whether we can replicate receptive fields of RGC types other than midget cells, such as parasol and bistratified cells. Completion of this project will allow us to assess whether efficient coding can accurately explain how the retina encodes chromatic information.

1. Barlow, H.B., *Possible principles underlying the transformation of sensory messages.* Sensory communication, 1961. **1**(01): p. 217-233.

2. Balasubramanian, V. and M.J. Berry II, *A test of metabolically efficient coding in the retina.* Network: Computation in Neural Systems, 2002. **13**(4): p. 531.

3. Doi, E., et al., *Efficient coding of spatial information in the primate retina.* Journal of Neuroscience, 2012. **32**(46): p. 16256-16264.

4. Ocko, S., et al., *The emergence of multiple retinal cell types through efficient coding of natural movies.* Advances in Neural Information Processing Systems, 2018. **31**.

5. Pitkow, X. and M. Meister, *Decorrelation and efficient coding by retinal ganglion cells.* Nature neuroscience, 2012. **15**(4): p. 628-635.

6. Soto, F., et al., *Efficient coding by midget and parasol ganglion cells in the human retina.* Neuron, 2020. **107**(4): p. 656-666. e5.

7. Karklin, Y. and E. Simoncelli, *Efficient coding of natural images with a population of noisy linear-nonlinear neurons.* Advances in neural information processing systems, 2011. **24**.

8. Lewicki, M.S., *Efficient coding of natural sounds.* Nature neuroscience, 2002. **5**(4): p. 356-363.

9. Miller, L.E., et al., *Somatosensory cortex efficiently processes touch located beyond the body.* Current Biology, 2019. **29**(24): p. 4276-4283. e5.

10. Atick, J.J. and A.N. Redlich, *Towards a theory of early visual processing.* Neural computation, 1990. **2**(3): p. 308-320.

11. Atick, J.J. and A.N. Redlich, *What does the retina know about natural scenes?* Neural computation, 1992. **4**(2): p. 196-210.

12. Kaneko, A., *Physiology of the retina.* Annual review of neuroscience, 1979. **2**(1): p. 169-191.

13. Dacey, D.M. *Physiology, morphology and spatial densities of identified ganglion cell types in primate retina*. in *Ciba Foundation Symposium 184‐Higher‐Order Processing in the Visual System: Higher‐Order Processing in the Visual System: Ciba Foundation Symposium 184*. 2007. Wiley Online Library.

14. Jun, N.Y., G.D. Field, and J. Pearson, *Scene statistics and noise determine the relative arrangement of receptive field mosaics.* Proceedings of the National Academy of Sciences, 2021. **118**(39): p. e2105115118.

15. Jun, N.Y., G. Field, and J. Pearson, *Efficient coding, channel capacity, and the emergence of retinal mosaics.* Advances in neural information processing systems, 2022. **35**: p. 32311-32324.

16. Maheswaranathan, N., et al., *Inferring hidden structure in multilayered neural circuits.* PLoS computational biology, 2018. **14**(8): p. e1006291.

17. Atick, J.J., Z. Li, and A.N. Redlich, *Understanding retinal color coding from first principles.* Neural computation, 1992. **4**(4): p. 559-572.

18. Kim, J.S., et al., *Space–time wiring specificity supports direction selectivity in the retina.* Nature, 2014. **509**(7500): p. 331-336.

19. Conway, B.R., et al., *Advances in color science: from retina to behavior.* Journal of Neuroscience, 2010. **30**(45): p. 14955-14963.

20. Crook, J.D., et al., *Horizontal cell feedback without cone type-selective inhibition mediates “red–green” color opponency in midget ganglion cells of the primate retina.* Journal of Neuroscience, 2011. **31**(5): p. 1762-1772.

21. Martin, P.R., et al., *Chromatic sensitivity of ganglion cells in the peripheral primate retina.* Nature, 2001. **410**(6831): p. 933-936.

22. Wässle, H., et al., *Retinal ganglion cell density and cortical magnification factor in the primate.* Vision research, 1990. **30**(11): p. 1897-1911.

23. Doi, E., et al., *Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes.* Neural computation, 2003. **15**(2): p. 397-417.