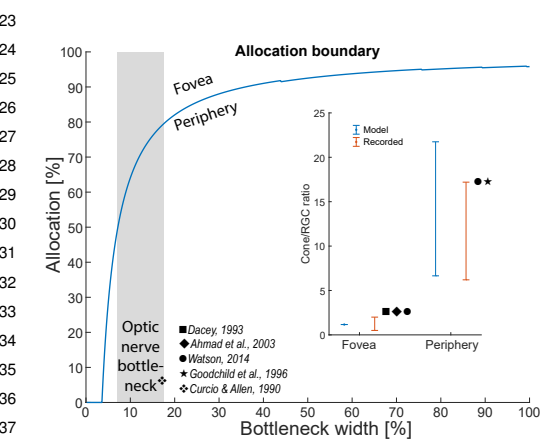


1 We would like to thank the reviewers for their positive and constructive comments. We address the major points below.

2 **Cost function** Reviewers asked to clarify the cost function we used, and whether/how changing the cost function  
3 would alter the results. We agree that the cost function should be clearly specified in the text and we will do so in  
4 the final version. Here, we did not include an explicit noise model; our result can be derived by either minimizing  
5 the reconstruction error (under mean-squared error) or by maximizing information, which restricted to second-order  
6 statistics equates to maximizing explained variance. We agree with the reviewers that it would be very interesting to  
7 explore alternative cost functions. We chose our approach for the following reasons:

- 8 1. We wanted to demonstrate that even the simplest and most commonly employed model yields an interesting  
9 and (to us) surprising relationship between the resulting allocation and the bottleneck width. The presence of  
10 this relationship is therefore not dependent on noise or nonlinearities.
- 11 2. Often, more complex models include a projection to PC space as the first step, and thereby yield the same  
12 allocation described in the paper. For example, this applies to a common variant of ICA (fastica) in the  
13 undercomplete case, which we have confirmed numerically. Similarly, the approach taken by Doi & Lewicki  
14 (2014) that includes an explicit noise model contains the same projection. Of course, alternative formulations  
15 of the noise model or cost function that would affect allocation directly are conceivable.
- 16 3. Finally, resource allocation can be solved analytically under our cost function. This paves the way for future  
17 analyses of more complex models, for example introducing nonlinearities by means of kernel PCA.

18 **Generalization of results** First, regions of different sizes can be incorporated by adjusting the model’s length scale,  
19 *a*. Second, the solution can be extended to  $n > 2$  regions: eigenvalues are calculated for each region independently,  
20 and can be sorted (and interlace) irrespective of the number of regions. Third, the 2D case can also be solved by a  
21 straightforward adaptation of the Laplacian eigenvalue problem. We will clarify these points in the final version. Finally,  
22 we have confirmed numerically that the block-matrix approach provides a valid approximation (see paper appendix).



38 Figure 1: Resource allocation in the optic nerve  
39 calculated from natural image statistics.

40  
41 should receive input from only 1 to 2 cones; the same ratio has been well established experimentally. For the periphery,  
42 we predict on average inputs from 10-20 cones onto a single RGC, again in agreement with the literature.

43  
44 As the density of cones varies smoothly across the retina, future work should build on these results to explore the  
45 principles of RGC allocation on a finer scale; however, such more complex analysis is beyond the scope of the current  
46 paper.

47 **Implications for machine learning** Bottlenecks are common in ML models to help with generalization and have  
48 recently attracted renewed interest in the field of deep neural networks. Non-uniform inputs have not been studied in  
49 detail, however appear useful for robotics applications, where power and size constraints are important. In both cases,  
50 our work suggests that the size of the bottleneck is critically important in shaping the resulting representations.

**Implications for neuroscience** Several comments asked about the  
match between our analysis of natural image statistics and neuro-  
physiological data. We initially included this example simply to  
demonstrate the importance of the covariance function on resource  
allocation. The original analysis was based on equal region sizes and  
therefore precluded a direct comparison with data from the visual  
literature, as the fovea is much smaller than the retinal periphery.  
We have now re-expressed our results to take this size difference  
into account. Importantly, the data set and methods are the same as  
before, the only difference being the number of inputs for each of the  
regions to account for the difference in size. We took the central 5  
degrees of the retina as the fovea, yielding a size ratio of 1:166. We  
assumed 260,000 cones in the fovea, and the rest in the periphery,  
containing around 5,940,000 cones (values extrapolated from Curcio,  
1990; Wells-Gray et al., 2016).

The resulting allocation boundary is shown in Fig. 1. Importantly, if  
we use this result to calculate the implied ratio between cones and  
(midnet) retinal ganglion cells (RGCs) in the optic nerve, we find a  
predicted ratio for the fovea of just over 1, indicating that each RGC

should receive input from only 1 to 2 cones; the same ratio has been well established experimentally. For the periphery,  
we predict on average inputs from 10-20 cones onto a single RGC, again in agreement with the literature.

As the density of cones varies smoothly across the retina, future work should build on these results to explore the  
principles of RGC allocation on a finer scale; however, such more complex analysis is beyond the scope of the current  
paper.

**Implications for machine learning** Bottlenecks are common in ML models to help with generalization and have  
recently attracted renewed interest in the field of deep neural networks. Non-uniform inputs have not been studied in  
detail, however appear useful for robotics applications, where power and size constraints are important. In both cases,  
our work suggests that the size of the bottleneck is critically important in shaping the resulting representations.