Expansion and contraction of resource allocation in sensory bottlenecks

Laura R Edmondson[1,2], **Alejandro Jiménez Rodríguez**[2,3], **Hannes P Saal**[1,2*]

***For correspondence:** h.saal@sheffield.ac.uk (HPS)

[1]Active Touch Laboratory, Department of Psychology, University of Sheffield, Sheffield S1 2LT, UK; [2]Sheffield Robotics, University of Sheffield, Sheffield S1 3JD, UK; [3]Department of Computer Science, University of Sheffield, Sheffield S1 3JD, UK

Abstract Topographic sensory representations often do not scale proportionally to the size of their input regions, with some expanded and others contracted. In vision, the foveal representation is magnified cortically, as are the fingertips in touch. What principles drive this allocation, and how should receptor density, e.g. the high innervation of the fovea or the fingertips, and stimulus statistics, e.g. the higher contact frequencies on the fingertips, contribute? Building on work in efficient coding, we address this problem using linear models that maximize information transmission through decorrelation. We introduce a sensory bottleneck to impose constraints on resource allocation and derive the optimal neural allocation. We find that bottleneck width is a crucial factor in resource allocation, inducing either expansion or contraction. Both receptor density and stimulus statistics affect allocation and jointly determine convergence for wider bottlenecks. Furthermore, we show a close match between the predicted and empirical cortical allocations in a well-studied model system, the star-nosed mole. Overall, our results suggest that the strength of cortical magnification depends on resource limits.

Introduction

In many sensory systems, receptors are arranged spatially on a sensory sheet. The distribution of receptors is typically not uniform, but instead, densities can vary considerably. For example in vision, cones are an order of magnitude more dense in the fovea than in the periphery (***Goodchild et al., 1996***; ***Wells-Gray et al., 2016***). In the somatosensory system, mechanoreceptors are denser in the fingertips than the rest of the hand (***Johansson and Vallbo, 1979***). Alongside the density of receptors, the statistics of the input stimuli can also vary. For example, the fingertips are much more likely to make contact with objects than the palm (***Gonzalez et al., 2014***). Subsequent sensory areas are typically arranged topographically, such that neighbouring neurons map to nearby sensory input regions, for example, retinotopy in vision and somatotopy in touch. However, the size of individual cortical regions is often not proportional to the true physical size of the respective sensory input regions, and instead, representations might expand (often called magnification) or contract. For example, both the fovea and the fingertips exhibit expanded representations in early visual and somatosensory cortices respectively, compared to their physical size (***Azzopardi and Cowey, 1993***; ***Engel et al., 1997***; ***Sereno et al., 1995***; ***Martuzzi et al., 2014***). What determines such cortical magnification? For somatotopy, it has been proposed that cortical topography might directly reflect the density of sensory receptors (***Catani, 2017***). On the other hand, receptor density alone is a poor predictor of magnification (***Corniani and Saal, 2020***) and work on plasticity has

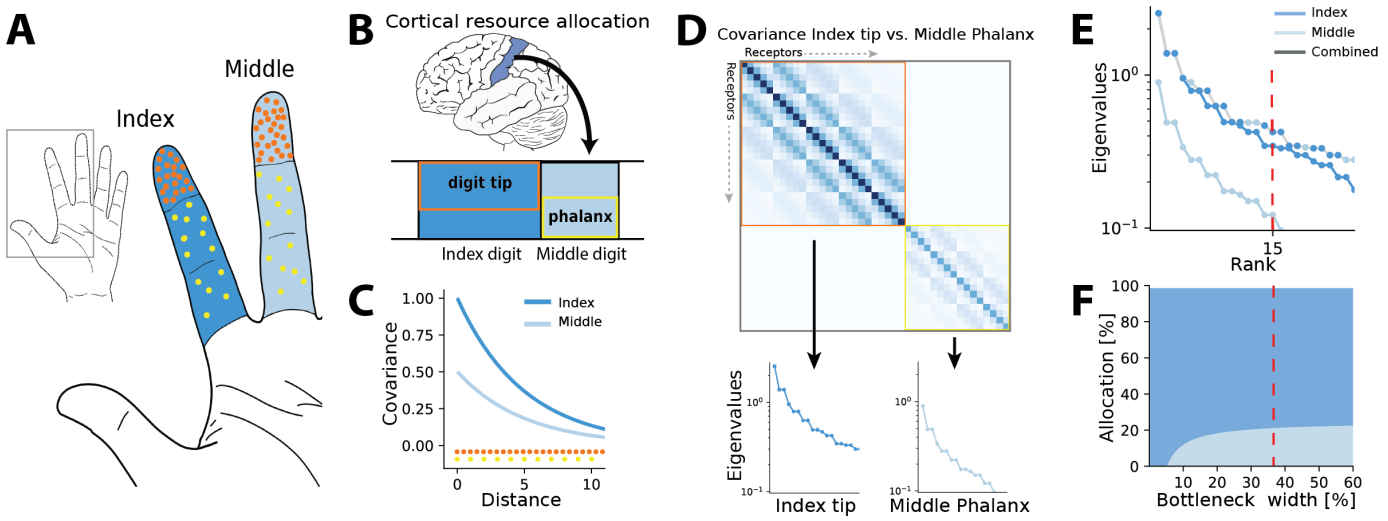

Figure 1. Illustration of the resource allocation problem and solution outline. A. Example of problem setup. The density of touch receptors differs across regions of the hand (e.g. fingertips, shown in orange, versus finger phalanx, yellow). Different fingers make contact with objects at different rates (dark blue versus light blue shading, darker colors indicating higher contact rates). **B.** How should cortical resources be divided between the different input regions? In the example shown here, the index digit might be allocated more resources than the middle digit due to its higher activation and the more densely innervated fingertips occupy an outsized area compared to the finger phalanxes. **C.** Sensory inputs are correlated according to a covariance function that decays with distance between receptors on the sensory sheet. This function is evaluated at different receptor distances depending on the density of sensory receptors (orange versus yellow dots at bottom). Regions with higher probability of activation exhibit greater variance (dark versus light blue curves). **D.** Decorrelation of sensory inputs in the presence of a bottleneck is achieved by calculating and retaining the eigenvectors and eigenvalues of the receptor covariance matrix. Here, this matrix is approximated as a block matrix, which allows calculating the eigenvalues for each region individually. For simplicity, the example compares two regions only, the index tip (orange box) and the proximal phalanx of the middle digit (yellow box). **E.** The combined set of eigenvalues from all regions is then sorted; the region where each successive eigenvalue in the combined sorted set originates from determines where that output neuron's receptive field will fall. **F.** Counting which input regions successive eigenvalues belong to results in the allocation breakdown for different bottlenecks. For certain forms of the covariance function, this allocation can be calculated analytically.

established that cortical regions can expand and contract dynamically depending on their usage, suggesting that expansion and contraction might be driven by the statistics of the sensory stimuli themselves (***Coq and Xerri, 1998***; ***Merzenich and Jenkins, 1993***; ***Xerri et al., 1996***).

Here, we tackle this problem from a normative viewpoint, employing efficient coding theory, which has been widely used to model and predict sensory processing. Efficient coding theory includes a number of different approaches such as sparse coding, redundancy reduction and predictive coding, depending on the constraints of the proposed problem (***Chalk et al., 2018***). Here we focus on efficient coding via redundancy reduction (***Barlow, 1961***), which suggests that neural populations are tuned to maximize the information present in the sensory input signals by removing redundant information (***Atick, 1992***; ***Atick and Redlich, 1990***; ***Attneave, 1954***; ***Graham and Field, 2009***; ***Chechik et al., 2006***). Efficient coding models have been most prominent in vision (***Kersten, 1987***; ***Karklin and Simoncelli, 2011***; ***Atick, 1992***; ***Atick and Redlich, 1990***; ***Doi et al., 2012***; ***Olshausen and Field, 1996***, ***1997***, ***2004***; ***Bell and Sejnowski, 1997***) and audition (***Smith and Lewicki, 2006***; ***Lewicki, 2002***). This prior work has mostly focused on predicting the response properties and receptive field structure of individual neurons. In contrast, here we ask how receptive fields—independent of their precise structure—should tile the sensory sheet when the receptors themselves differ in density and activation levels.

Some aspects of magnification in topographic representations have been qualitatively reproduced using self-organising maps (***Ritter et al., 1992***). However, these models generally lack a clear cost function and the magnification factor can be determined exactly only in rare cases, while a general expression is lacking (***Ritter and Schulten, 1986***). Assuming a uniform density of output

neurons, cortical maps may be optimizing for the spread of incoming information to be equally distributed over these neurons ***Plumbley*** (***1999***).

In contrast to receptor density, there has been some work on how populations of neurons should encode non-uniform stimulus statistics using Fisher information (***Ganguli and Simoncelli, 2010***, ***2014***, ***2016***; ***Yerxa et al., 2020***), an approximation of mutual information that can be used to calculate optimal encoding in a neural population (***Yarrow et al., 2012***). Rather than receptive fields uniformly tiling the input space, the optimal population should be heterogeneous, with receptive fields placed more densely over high probability inputs, at detriment to low probability regions (***Ganguli and Simoncelli, 2014***). Our approach differs from this prior work: rather than maximizing information between the neural population and the stimulus itself, we instead consider information between the neural population and an initial population of receptor neurons. This places a limit on the total amount of information that can be represented.

The need for information maximization is often motivated by resource constraints. These can take the form of an explicit bottleneck, where the number of receptor neurons is greater than the number of output neurons. This is the case in the early visual system, where photoreceptors in the retina are much more numerous than the retinal ganglion cells to which they project (***Wells-Gray et al., 2016***; ***Goodchild et al., 1996***). Other sensory systems might lack such explicit bottlenecks, but still place limits on the amount of information that is represented at a higher-order processing stage.

How then should resource allocation change for different sized bottlenecks, given varying densities of receptors and different stimulus statistics? Here, we derive optimal neural allocations for different bottlenecks, while systematically varying receptor density and stimulus statistics. A preliminary version of these results restricted to the effects of receptor density in a one-dimensional space were previously presented as a conference paper (***Edmondson et al., 2019***).

Results

We consider linear models which maximise information by decorrelating the sensory inputs, or, equivalently, minimize the mean-squared reconstruction error, under a sensory bottleneck where the number of output neurons is smaller than the number of receptors. Specifically, $\boldsymbol{r} = \boldsymbol{W}\boldsymbol{x}$, where $\boldsymbol{r}$ is the m-dimensional output vector, $\boldsymbol{x}$ is the n-dimensional input vector (with $n > m$) containing sensory receptor responses, and $\boldsymbol{W}$ contains the receptive field. Optimal decorrelation and minimal error can be achieved, if $\boldsymbol{W}$ takes the following form:

$$\boldsymbol{W} = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\Phi^T, \tag{1}$$

where $\boldsymbol{P}$ is an orthogonal matrix, and $\boldsymbol{\Lambda}$ and Φ contain the (top m) eigenvalues and eigenvectors, respectively, of the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{X}^T\boldsymbol{X}$ (***Doi et al., 2012***; ***Doi and Lewicki, 2014***). $\boldsymbol{P}$ is not uniquely determined and can be used to enforce additional constraints (see Methods for details). Importantly, these additional constraints will only affect $\boldsymbol{P}$ (and thereby change the precise receptive field structure), while the retained eigenvalues and their associated eigenvectors are unaffected.

In the specific setup we consider here, receptors are placed in non-uniform densities across the sensory sheet, with some regions having a higher density than others. For example, in ***Figure 1***A, both digit tips are tiled more densely than the rest of the finger. The receptor activation covariance decays with distance (see ***Figure 1***C). Thus, directly neighbouring units in denser regions covary more strongly than those from less dense regions. The activation level of receptors can also vary between regions, which is modelled through scaling the covariance function (see ***Appendix 1*** for further details). We initially focus on negative exponential covariance functions, for which the eigenvalues and the resulting allocation can be calculated analytically. The covariance between receptors x_i and x_j for two regions R_1 and R_2, differing in density and activation, can thus be expressed as:

$$\boldsymbol{\Sigma}_{ij}^{(R1)} = e^{-\gamma|x_i - x_j|} \quad \text{and} \quad \boldsymbol{\Sigma}_{ij}^{(R2)} = ae^{-d\gamma|x_i - x_j|}, \tag{2}$$

where a denotes the receptor activation ratio and d the receptor density ratio between both regions. It can be seen that changing the density affects the decay of the covariance function and thereby 'stretches' the space, while changing the activation scales the function.

We approximate the regions as separate, such that no receptors are activated from both regions simultaneously. The covariance matrix across all receptors then forms a block matrix, where the covariance between separate regions is zero (Fig. 1D):

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}^{(R1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}^{(R2)} \end{bmatrix} \tag{3}$$

Thus, eigenvalues can be calculated separately from the covariance matrices of each region (inset panels in ***Figure 1***D) and for negative exponential covariance functions, this can be done analytically (see Methods for mathematical derivations):

$$R_1 : \quad \lambda_{l,m}^{(R1)} = \frac{2\gamma}{\pi^2 L^{-2}(l^2 + m^2) + \gamma^2} \quad \text{and} \quad R_2 : \quad \lambda_{n,o}^{(R2)} = \frac{2\gamma a\sqrt{d}}{\pi^2 L^{-2}(n^2 + o^2) + \gamma^2} \tag{4}$$

Finally, the resulting eigenvalues are ordered by magnitude for both regions combined (grey line in ***Figure 1***E). A bottleneck is introduced by restricting the total number of output neurons. This is done by selecting from the combined set of ordered eigenvalues until the limit is reached. The proportion of eigenvalues originating from each input region determines its allocation for the chosen bottleneck width (see red dashed line ***Figure 1***E,F).

In the following, we will present results for the common case of two-dimensional sensory sheets, while results for one-dimensional receptor arrangements are summarized in the appendix. For ease of analysis, all cases discussed will assume two sensory input regions of equal size, which are differing in receptor density, input statistics, or both. In all our examples, receptors in the low-density (baseline) region are always spaced on a grid with a distance of 1.

Resource limits determine amount of magnification

First, we investigated resource allocation in bottlenecks for heterogeneous density of receptors and heterogeneous stimulus statistics separately, whilst keeping the other factor uniform across the input regions.

Heterogeneous density

For two regions with different receptor densities, we found that the higher density region could either expand or contract relative to its input density, depending on the width of the bottleneck. Specifically, for narrow bottlenecks (smaller than approximately 10% of input neurons), the higher density region is either exclusively represented or its representation is expanded compared to a proportional density allocation (see example in ***Figure 2***A). Mathematically, this can be explained by a multiplicative scaling of the eigenvalue function for the higher density region (see illustration in ***Figure 2–Figure Supplement 3***A). In contrast, for intermediate bottlenecks, low density regions expand their representation, beyond what is expected for a proportional mapping (see dashed red line in ***Figure 2***A denoting proportional allocation), leading to a contraction of the high density region. For negative exponential covariance functions, this allocation converges to a fixed ratio at wider bottlenecks of $1/(1+\sqrt{d})$, where d is the density ratio between both regions (dashed yellow line in ***Figure 2***A; see Methods for derivation), as neurons are allocated to either region at a fixed ratio (see inset in ***Figure 2–Figure Supplement 3***A). Finally, for wide bottlenecks, all information arising from the low density region has now been captured, and any additional output neurons will therefore be allocated to the high density region only. The over-representation of the low density region thus decays back to the original density ratio. The overall nonlinear effect of bottleneck width is present regardless of the ratio between the densities (see ***Figure 2***B). The spatial extent of the correlations over the sensory sheet (controlled by the decay constant γ in the covariance function, see Methods) determines allocation at narrow bottlenecks, and how fast the allocation converges,

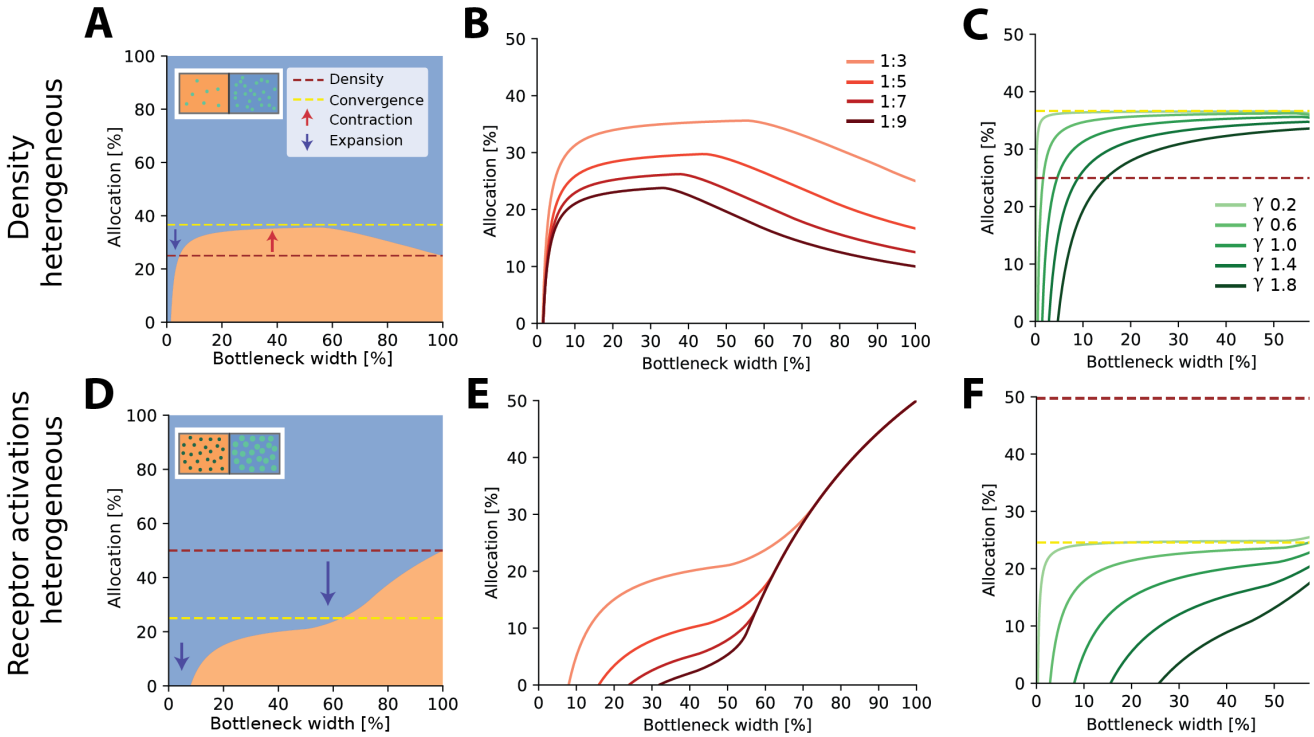

Figure 2. Optimal resource allocation for heterogeneous receptor densities or input statistics leads to complex trade-offs. A. Illustration of resource allocation for heterogeneous receptor density but homogeneous stimulus statistics over all bottleneck sizes. Orange denotes the lower density region and blue the higher density region, with a ratio of 1:3. Dotted lines show proportional representation according to receptor numbers (red) and convergence of the optimal allocation in the limit (yellow). Arrows indicate contraction (up) and expansion (down) of the higher density region representation. Inset demonstrates change in density between the regions. **B.** Bottleneck allocation boundaries for different density ratios (given as low:high). The area below each line corresponds to the low density representation, while the area above corresponds to the high density representation, as in panel A. **C.** Effect of changing the extent of the spatial correlations (parameterized by the decay value γ, see Methods for details and ***Figure 2–Figure Supplement 1*** for an illustration of the covariance function for different values of γ). Density ratio is set at 1:3 for all γ. Increasing γ leads to expansion of the higher density region for a larger initial portion of the bottleneck. **D-F.** Same as in row above but for homogeneous density and heterogeneous receptor activation ratios. **D.** Illustrative example with the blue region having higher receptor activation. Note that the representation of the higher activation region is expanded for all bottleneck widths. Inset demonstrates difference in activation between the regions. Larger, brighter colored points indicate higher activation for that region compared to the other. **E** Allocation boundaries for different activation ratios. The representation of the higher-activation regions is expanded for all bottlenecks. As activation ratio increases, the highly active region allocation is expanded for wider bottlenecks. **F.** Changing the extent of spatial correlations (γ) has larger effects when the activation ratio is heterogeneous (set at 1:3 for all γ) compared to heterogeneous density (panel **C**). See ***Figure 2–Figure Supplement 2*** for an equivalent figure considering one-dimensional receptor arrangements.

Figure 2–Figure supplement 1. Effect of different values of γ on the covariance function decay.

Figure 2–Figure supplement 2. Resource allocation for heterogeneous receptor densities and variations in input statistics in 1D.

Figure 2–Figure supplement 3. Illustration of eigenvalue sorting and resulting allocation.

Figure 2–Figure supplement 4. Limit on information rather than number of neurons.

but does not affect the convergence limit itself (***Figure 2***C). As γ increases, and therefore spatial correlations decrease, the convergence limit is approached only at increasingly wider bottlenecks. The extent of magnification thus heavily depends on the correlational structure of the stimuli for narrower bottlenecks, while receptor densities are more important for wider bottlenecks.

Heterogeneous statistics

Aside from receptor densities, the stimulus statistics can also vary over the input space, leading to differences in receptor activation levels between the regions and their associated response variance. Overall, allocations for heterogeneous receptor activation are similar to those found with

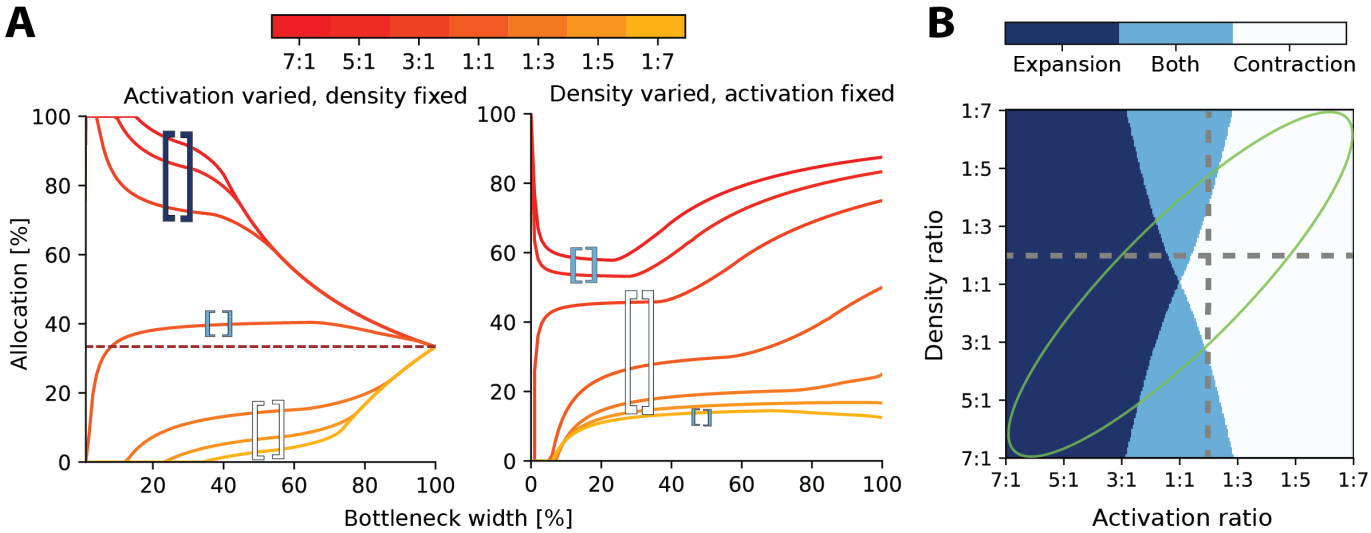

Figure 3. Interactions between heterogeneous statistics and density. A. Allocations with both heterogeneous density and activation ratios. Expansion and contraction for a baseline region where relative density and activation is varied over the other region. All ratios given are baseline:other region. Left: Fixed density ratio of 1:2, while activation ratio is varied between 7:1 and 1:7 (see colorbar). Purple dashed line indicates allocation proportional to density ratio. Right: Fixed activation ratio of 1:2, while density ratio is varied between 7:1 and 1:7. Whether the baseline region is expanded or contracted is denoted by the colored bracket symbols. Bracket range shows all combinations with expansion, contraction or both: *Dark blue brackets* denotes expansion of the baseline region. *Light blue brackets*, expansion and contraction of the baseline region. *White brackets* indicates contraction of the baseline region. **B.** Possible expansion/contraction regimes for the baseline region based on combinations of density and activation ratios. Colors correspond to bracket colors in A. Grey dashed lines show all possible allocation regimes for an example with wither fixed density ratio (horizontal, corresponding plot in panel A, left) or activation ratio (vertical, panel A, right). When activation is fixed and density is varied the allocations can be either expanded/both or contracted/both across the full bottleneck width. In contrast, when the density ratio is fixed and activation is varied, the allocation of a region could be any of the three regimes. The green ellipse highlights parameter combinations where activation and density ratios are correlated. See ***Figure 3–Figure Supplement 1*** for a comparison of how receptor density and activation interact between one-dimensional and two-dimensional receptor arrangements.

Figure 3–Figure supplement 1. Comparison between 1D and 2D results

heterogeneous density. However, while the allocations are again a nonlinear function of bottleneck width, the representations are solely expanded for the higher activation region across the entire bottleneck (see example in ***Figure 2***D). The extent of this expansion depends on the width of the bottleneck and is again more extreme for narrower bottlenecks (see ***Figure 2***E). The convergence limit is $1/(1+a)$, where a is the activation ratio, implying that the level of expansion and contraction in intermediate bottlenecks is more extreme than in the heterogeneous density case (see difference between red and yellow dashed lines in ***Figure 2***A,D). Finally, the effect of spatial correlations is also more pronounced (***Figure 2***F).

In the cases described above, the bottleneck was constrained by the number of output neurons. Alternatively, the limiting factor might be set as the amount of information (total variance) captured. Doing so results in allocation curves that retain the nonlinear behaviour described here (see ***Figure 2–Figure Supplement 4*** for illustrations).

In summary, we find that representations of different input regions can contract or expand, depending on the bottleneck width. This effect plays out similarly for differences in receptor density and receptor activation, however with some crucial differences. Finally, for narrow bottlenecks, the spatial extent of the correlations across the sensory sheet becomes an important driver.

Interplay between stimulus statistics and receptor density

In sensory systems, such as in touch, both the density and input statistics vary across regions and will therefore jointly determine the resulting allocation. As a consequence, the convergence for

intermediate bottlenecks will depend on both density and activation ratios, and can be calculated as $\frac{1}{(1+a\sqrt{d})}$, where a is the activation and d is the density ratio (see Methods for derivation).

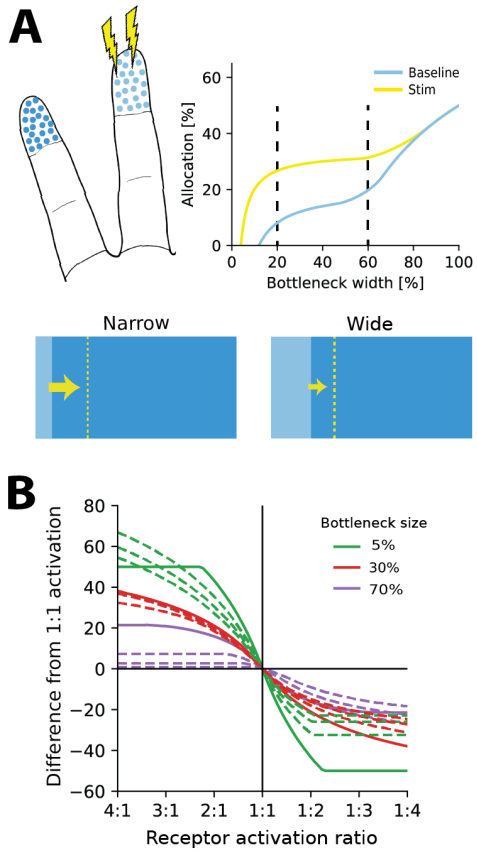

Figure 4. Re-allocation to account for changes in stimulus statistics. A. Top left: Illustration of problem setup. Increased stimulation is applied to the middle digit (yellow symbols), leading to changes in optimal allocations. Top right: Optimal allocations for baseline (blue) and stimulation (yellow) conditions across all bottleneck widths. Stimulation of the middle finger increases its representation, but the relative magnitude of the effect depends on the bottleneck width. Bottom: Changes in allocation of the middle digit for two bottleneck widths (indicated by dashed lines above). The increase is proportionally larger for the narrow compared to the wide bottleneck. **B**. Change in allocation when receptor activation for an input region increases (left half) or decreases (right half). Drastic changes in cortical allocation are observed for narrow bottlenecks (green lines), while wider bottlenecks (red and purple lines) induce only moderate change. Solid lines denote equal receptor density across both regions, while dashed lines show examples with varying density.

At narrow bottlenecks, the spread of possible allocations is much wider for varying the activation ratio (see left panel of ***Figure 3***A) compared to varying the density ratio. Thus, for two-dimensional sensory sheets, the activation ratio more strongly determines the possible allocation than does the density ratio. Specifically, this means that the allocation regime, i.e. whether the allocation expands, contracts or exhibits both behaviours across all bottleneck widths, is more dependent on relative receptor activation than densities (***Figure 3***B).

Finally, it is likely that regions with higher receptor densities will also show greater activation than lower density regions. For example, in touch, the regions on the hand with the highest receptor densities also are the most likely to make contact with objects (***Gonzalez et al., 2014***). In these cases, both effects reinforce each other and drive the resulting allocation further from proportional allocation (see orange lines in ***Figure 3***A, and green ellipse in ***Figure 3***B).

Resource limits determine the strength of plasticity under changes in stimulus statistics

Over most of the lifetime of an organism, the resources available for processing sensory information and the density of sensory receptors should be relatively constant. The stimulus statistics, on the other hand, can and will change, for example when encountering a new environment or learning new skills. These changes in stimulus statistics should then affect sensory representations, mediated by a variety of plasticity mechanisms. For example, increased stimulation of a digit will lead to an expansion of that digit's representation in somatosensory cortex (***Jenkins et al., 1990***).

We asked how representations should adapt under the efficient coding framework, and whether resource limits would affect the resulting changes. To answer this question, we calculated optimal allocations for different bottleneck widths, receptor densities, and stimulus statistics. We then introduced a change in stimulus statistics and re-calculated the resulting allocations (see illustration in ***Figure 4***A). As expected,

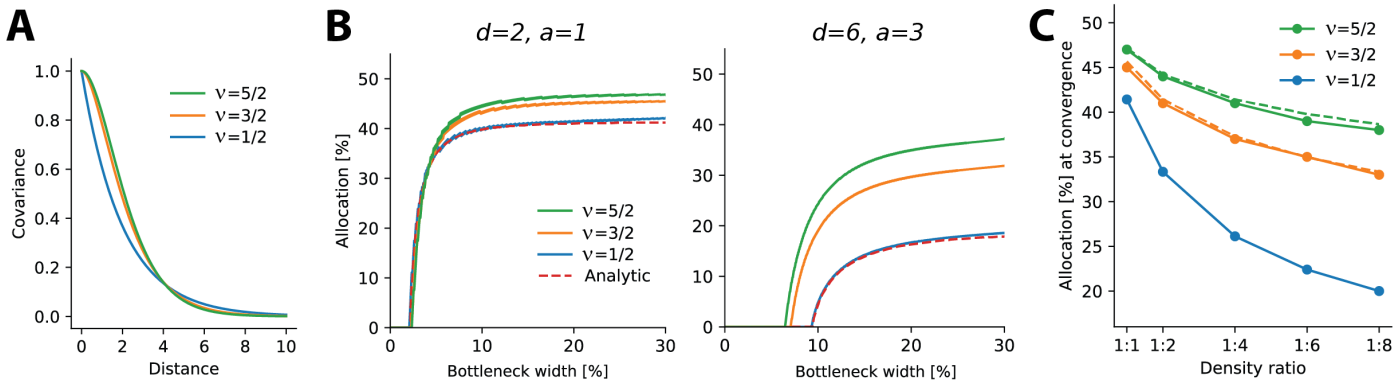

Figure 5. Allocations for other monotonically decreasing covariance functions. A. Three covariance functions of different smoothness taken from the Matérn class, differing in the parameter ν (see Methods). **B.** Examples of numerically determined allocation for the covariance functions shown in A for two scenarios. Left: density ratio 1:2, equal activation. Right: density ratio 1:6 activation ratio 1:3. The red dashed line shows the analytic solution for the negative exponential covariance ($\nu = 1/2$). **C**. Estimated allocations at convergence for different density ratios (horizontal axis) and a fixed activation ratio of 1:2. Solid lines denote numerical convergence, dashed lines refer to fitted functions (see main text).

we found that when increasing the receptor activation over a region (for example, by increasing stimulation of the region), more neurons would be allocated to this region than the other. Interestingly, however, this effect was dependent on the width of the bottleneck. The largest effects are seen for smaller bottlenecks and then diminish as the bottleneck size increases. ***Figure 4***B demonstrates such allocation changes for three different bottleneck widths for a range of receptor densities and activation ratios. This suggests that plasticity should be relatively stronger under extreme resource constraints than in cases where limits on the information are generous.

Generalisation to other covariance functions

Our results above relate to negative exponential covariance functions, where allocations can be solved analytically. How might these findings generalize to other covariance functions? We previously demonstrated non-linear allocations for the covariance structure derived from natural images (***Edmondson et al., 2019***). Here, to test how optimal allocations depend systematically on the covariance function, we numerically calculated allocations for three covariance functions of the Matérn class, whose smoothness depends on a single parameter ν. The negative exponential considered so far is a special case of the Matérn class when $\nu = 1/2$ and we also considered $\nu = 3/2$ and $\nu = 5/2$ (see Methods). As ν increases the decay function becomes smoother, yielding larger correlations at short distances and smaller ones at farther distances (see ***Figure 5***A).

We numerically calculated allocations for different density and activation ratios in 1D (***Figure 5***B). We focused on smaller bottleneck widths, where solutions are numerically stable, and observed a close match between the numerical and analytical solutions for the negative exponential covariance function (see dashed lines in ***Figure 5***B).

This analysis yielded two main findings. First, allocation curves are qualitatively similar across all tested covariance functions in that regions with higher density, activations, or both are systematically over-represented. Furthermore, this effect is more extreme at smaller bottleneck widths. Second, the resulting allocations depend systematically on the smoothness of the covariance function, such that smoother covariances yield less extreme differences in allocation. This effect is most obvious when considering the convergence points at larger bottlenecks: smoother covariance functions induce a more uniform allocation (closer to a 50:50 split) than does the negative exponential (***Figure 5***C). We also noticed that the convergence points appeared to follow a simple

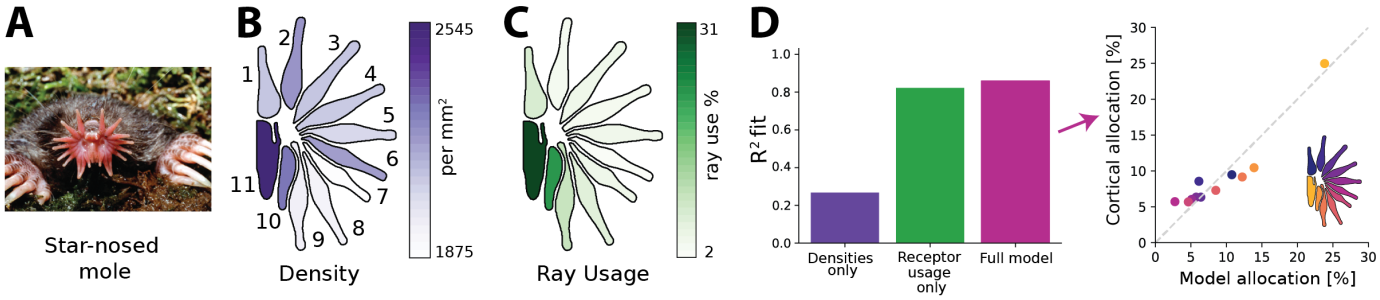

Figure 6. Resource allocation in the star-nosed mole. A. Star-nosed moles have two sets of 11 tactile rays used for detecting and capturing prey. Image by Kenneth Catania and reprinted with permission. **B.** Fiber innervation densities for each ray. **C.** Typical usage percentages for each ray during foraging. Higher usage corresponds to greater average activation of receptors located on the corresponding ray. Typically prey is funnelled from distal rays towards ray 11, which is located next to the mouth. Ray outlines reproduced from ***Catania and Kaas*** (***1997***). **D.** Left: Explained variance (R^2) between model predictions and cortical allocation for three different models: restricted to receptor density only (purple), restricted to receptor activation only (green), and a full model (pink) that incorporates accurate values for both factors. Results confirm previous findings that ray usage is a better predictor of cortical allocation than receptor densities alone. Additionally, we show that including both of these factors provides a marginal improvement to the fit, with the highest R^2 (86%). Right: predicted versus empirical cortical allocations for all rays. When including both density and activation parameters, the model provides a good fit to empirical measurements. Scatter plots for all models are available in ***Figure 6–Figure Supplement 1***B.

Figure 6–Figure supplement 1. Additional figures for the star-nosed mole.

function, namely $\frac{1}{1+\sqrt[6]{ad}}$ for $\nu = 5/2$, $\frac{1}{1+\sqrt[4]{ad}}$ for $\nu = 3/2$ (see dashed lines in ***Figure 5***C), and $\frac{1}{1+\sqrt{ad}}$ for $\nu = 1/2$ (which is the result obtained analytically, see ***Appendix 6***).

In summary, then, monotonically decreasing covariance functions result in qualitatively similar allocations to the negative exponential studied earlier. However, precise allocations and convergence are determined by the correlational structure: smoother covariance functions lead to less extreme differences in allocation.

Predicting cortical magnification in the star-nosed mole

Finally, we investigated to what extent the procedure outlined in the previous sections might predict actual resource allocation in the brain. While our model is relatively simple (see Discussion), decorrelation has been found to be a strong driver in early sensory processing, including in vision (***Atick and Redlich, 1992***; ***Graham et al., 2006***; ***Vinje and Gallant, 2000***), audition (***Clemens et al., 2011***; ***Smith and Lewicki, 2006***), and touch (***Ly et al., 2012***), and one might therefore expect the approach to at least yield qualitatively valid predictions. As currently available empirical data makes it difficult to test the impact of different bottleneck widths on the resulting allocation directly (see Discussion), we instead focused on another predicted outcome of the proposed model: the precise interaction between receptor density and receptor activation in driving resource allocation that we presented earlier. We picked the star-nosed mole as our model system, because stimulus statistics, receptor densities, and cortical allocations have been precisely quantified. Moreover, the star-nosed mole displays considerable variation in all these parameters, presenting a good opportunity to put the model to the test.

The star-nosed mole is a mostly underground dwelling creature relying on active tactile sensing while foraging for prey (***Catania and Kaas, 1997***; ***Catania, 2020***). This process is facilitated by two sets of eleven appendages arranged in a star-like pattern that make up the mole's nose (***Figure 6***A). Individual rays are tiled with tactile receptors known as Eimer's organs that detect prey, which is then funneled towards the mouth (***Catania and Kaas, 1997***; ***Sawyer and Catania, 2016***). The density of fibres innervating the Eimer's organs differs across the rays, with rays closer to the mouth exhibiting higher densities (***Figure 6***B). The rays also vary in size as well as in their location with respect to the mouth. This affects their usage, as rays closest to the mouth encounter tactile stimuli much more frequently than other rays (***Figure 6***C). In cortex, a clear topographic represen-

tation of the rays can be found (***Catania and Kaas, 1997***). However, the extent of the cortical ray representations is not proportional to the physical size of the rays. Ray 11, which sits closest to the mouth, is cortically magnified several-fold and is considered the tactile equivalent of the visual fovea (***Catania and Remple, 2004***).

Previous work by ***Catania and Kaas*** (***1997***) found that cortical sizes are correlated more strongly with the different patterns of activation across the rays, rather than their innervation densities. Using the empirical data quantifying receptor densities and stimulus statistics from this study, we investigated whether the efficient coding model could predict typical cortical representation sizes for each ray (see Methods for details), and whether innervation densities or usage would lead to more accurate allocation predictions. As the bottleneck size between periphery and cortex is unknown for the star-nosed mole, we calculated the optimal allocations over all possible bottleneck sizes. Using the model described above we calculated allocations considering three different scenarios. The first, 'density only', included accurate afferent densities, but with activations set to the mean over all rays. The second model, 'usage only', included accurate activation ratios but mean receptor densities. Finally, the 'full model' included accurate values for both factors. We found that the empirical cortical representation sizes are most accurately predicted by the models that include receptor activation- the 'receptor usage only' and 'full models' (***Figure 6***D), suggesting the star-nosed mole could be employing an efficient coding strategy based on decorrelation in the neural representation of somatosensory inputs.

Discussion

We examined efficient population coding under limited numbers of output neurons in cases of non-uniform input receptor densities and stimulus statistics. Instead of focusing on precise receptive field structure, we asked which coarse region on the sensory sheet the receptive field would fall. We showed that the resulting allocations are nonlinear and depend crucially on the width of the bottleneck, rather than being proportional to the receptor densities or statistics. Specifically, narrow bottlenecks tend to favour expansion of a single region, whereas for larger bottlenecks, allocations converge to a constant ratio between the regions that is closer to a proportional representation. Whether, across all possible bottlenecks, allocations are always expanded, contracted or show both expansion and contraction depends on the relative density and activation ratios, but receptor activation plays a bigger role. When allocation changes due to novel stimulus statistics, a larger fraction of output neurons will switch their receptive field to another region for narrow compared to wide bottlenecks. Finally, we demonstrated that in the star-nosed mole, a model including both accurate innervation densities and statistics of contact provides a better fit to somatosensory cortical region sizes than considering each of these factors alone.

Comparison with previous approaches

A key feature of efficient coding population models is the non-uniform allocation of output neurons, whereby stimuli occurring at higher probabilities are represented by a greater number of neurons. A common approach is to use Fisher Information as a proxy for mutual information, enabling the calculation of optimal output neuron density and tuning curve placement given a distribution of sensory stimuli (***Ganguli and Simoncelli, 2010***, ***2014***, ***2016***; ***Yerxa et al., 2020***). However, these approaches maximize information between the output population and the stimulus distribution itself, such that allocating additional neurons to a given region of the input space will always lead to an increase in the overall amount of information represented. In contrast, our approach assumes a finite number of input receptors in each region. We are thus asking a different question than previous research: once information about a sensory stimulus has been captured in a limited population of receptors, what is the most efficient way of representing this information? This framing implies that once all information from a given region has been fully captured in the output population, our method does not allocate further neurons to that region. Therefore, this

also places a limit on the total size of the output population, as this cannot exceed the total number of input receptors.

There has also been prior work on how bottlenecks affect sensory representations, though mostly focused on how different levels of compression affect receptive field structure. For example, ***Doi and Lewicki*** (***2014***) predicted receptive fields of retinal ganglion cells at different eccentricities of the retina, which are subject to different convergence ratios. More recently, such direct effects on representations have also been studied in deep neural networks (***Lindsey et al., 2019***). Finally, some approaches employ a different cost function than the mean-squared reconstruction error inherent to the PCA method used here. For example, the Information Bottleneck method (***Tishby et al., 2000***; ***Tishby and Zaslavsky, 2015***) aims to find a low-dimensional representation that preserves information about a specific output variable, while compressing information available in the input. It is quite likely that the choice of cost function would affect the resulting allocations, a question that future research should pursue.

Implications for sensory processing

The quantitative comparison of the model results with the cortical somatosensory representation in the star-nosed mole provided limited evidence that one of the model predictions—namely that receptor density and activation statistics should jointly determine cortical organisation—is borne out in a biological model system. Further direct tests of the theory are hampered by a lack of reliable quantitative empirical data. Nevertheless, the model makes a number of qualitative predictions that can be directly compared to available data or tested in future experiments.

First, there is additional evidence that both receptor density and stimulus statistics drive allocation in neural populations. Similar to the star-nosed model, the primate somatosensory system also exhibits non-uniform distributions of both stimulus statistics and receptor density. Both of these factors are also broadly correlated, for example, receptor densities are higher on the fingertips, which are also more likely to be in contact with objects (***Gonzalez et al., 2014***). Importantly, while receptor densities alone can explain some of the magnification observed in the cortical somatosensory homunculus, they cannot account for this effect fully (***Corniani and Saal, 2020***). Indeed, evidence from non-human primates shows that cortical magnification also critically depends on experience (***Xerri et al., 1996***, ***1999***). Similar results have recently been obtained from the brainstem in mice (***Lehnert et al., 2021***). Empirically reported differences in sizes of receptive fields between the digit-tips and palm have also been recreated from typical contact across the hand during grasping (***Bhand et al., 2011***), suggesting an influence of hand use statistics. Both mechanoreceptor densities (***Verendeev et al., 2015***, e.g.) and hand use statistics (***Fragaszy and Crast, 2016***) differ across primates, forming the basis for a potential cross-species study. The model presented here demonstrates how both effects can be treated within a single framework driven by information maximization.

Second, it appears that allocation in sensory systems with severely constrained resources is qualitatively in agreement with model predictions. As our results demonstrated, magnification should be more extreme the tighter the bottleneck. The best characterized and most clearly established bottleneck in sensory processing is likely the optical nerve in vision. Given that the optic nerve serves as a narrow bottleneck (approx. 12-27%)[1], and that the fovea contains a much higher density of cone receptors than the periphery (***Wells-Gray et al., 2016***; ***Curcio et al., 1990***), the model would predict a large over-representation of the fovea, agreeing with experimental observations (see ***Edmondson et al., 2019***, for further qualitative evidence). In order to further test the proposed model, comparisons could be made between individuals to study variations within a population. Specifically, it would be expected that optic nerves containing a high number of fibers would devote proportionally fewer of them to the fovea than optic nerves containing smaller numbers of fibers (assuming equal receptor densities and numbers in the retina). These comparisons could be

[1] Assuming 0.71 - 1.54 million retinal ganglion cells with 80% Midget cells (***Curcio and Allen, 1990***; ***Perry et al., 1984***) and 4.6 million cones (***Curcio et al., 1990***)

extended across species, making use of the fact that photoreceptor densities and optic nerve fiber counts differ across many primate species (***Finlay et al., 2008***). Along these lines, recent computational work has demonstrated that the amount of neural resources allocated to the optic nerve can be expected to affect the receptive field structure itself ***Lindsey et al.*** (***2019***). Finally, the extent of cortical areas can be controlled by experimental interventions in animals (***Huffman et al., 1999***), which would constitute a direct manipulation of the bottleneck.

Re-allocation during development, learning, and ageing

Changing receptor densities, stimulus statistics, or resource limits over the lifespan of an organism should lead to a dynamic re-allocation of the available resources. The most common case will be changes in stimulus statistics, as both receptor densities and resources should be relatively stable. In such cases, representations should adapt to the new statistics. For example in touch, changing the nature of tactile inputs affects cortical representations (***Coq and Xerri, 1998***; ***Merzenich and Jenkins, 1993***; ***Xerri et al., 1996***). Increasing statistics of contact over a region typically leads to expansion of that region in cortex. Our method suggests that the precise level of expansion would be dependent on the bottleneck width with larger effects observed for narrower bottleneck sizes.

For the other cases, changes in fiber numbers during development and ageing might be interpreted as a change in resources. For example, the optic nerve undergoes a period of rapid fiber loss early during development (***Sefton et al., 1985***; ***Provis et al., 1985***). Similarly, fiber counts in the optic nerve decrease during ageing (***Dolman et al., 1980***; ***Jonas et al., 1990***; ***Sandell and Peters, 2001***). In this case, the model would predict a decrease in the size of peripheral representation in the bottleneck compared to the fovea. It is also possible that the receptor densities themselves may change. In touch, ageing leads to reductions in the densities of receptors in older adults (***García-Piqueras et al., 2019***). In such cases, we have effectively increased the bottleneck width relative to our receptor population, which again should lead to the re-allocation of resources.

Expansion and contraction along the sensory hierarchy

Magnification of specific sensory input regions can be observed all along the sensory hierarchy, from the brainstem and thalamus up to cortical sensory areas. Often, the amount of expansion and contraction differs between areas. For example, the magnification of the fovea increases along the visual pathway from V1 to V4 (***Harvey and Dumoulin, 2011***). Which of these representations might be best addressed by the model presented here? The main component of the model is decorrelation, which has been shown to be a driving principle for efficient coding in settings where noise is low (***Chalk et al., 2018***). This is generally the case in low-level sensory processing, for example in touch (***Goodwin and Wheat, 2004***). Our results might therefore best match early sensory processing up to perhaps low-level cortical representations. Beyond this, it is likely that noise will be much higher (***Shadlen and Newsome, 1998***) and for efficient codes to shift away from decorrelation (***Hermundstad et al., 2014***). Furthermore, while distinct bottlenecks, whether on the number of neurons or the amount of information, are common in low-level processing, it is less clear whether such restrictions constrain cortical processing. Whether and how such different regimes should affect neural allocations remains an open question.

Perceptual consequences

Does the allocation of output neurons lead to testable perceptual consequences? While we do not model neurons' receptive fields directly, allocating more neurons to a given region would increase perceptual spatial acuity for that region. Indeed, cortical magnification and perceptual acuity are correlated in both vision (***Duncan and Boynton, 2003***) and touch (***Duncan and Boynton, 2007***). At the same time, the absolute limits on spatial acuity are determined by the density of receptors in each input region. A naive allocation scheme that assigns output neurons proportional to the density of receptors would therefore result in perceptual spatial acuity proportional to receptor distance. Instead, as our results have shown, the allocation should not be proportional in most cases.

Specifically, for narrow bottlenecks we would expect relatively higher spatial acuity for regions with high receptor density than might be expected from a proportional allocation. Conversely, for wider bottlenecks this relationship should be reversed and spatial acuity should be better than expected for lower density regions. In agreement with these results, it has been found that in vision spatial resolution declines faster than expected with increasing eccentricity, suggesting a narrow bottleneck in the optic nerve (***Anderson et al., 1991***).

A second consequence is that spatial acuity should be better on regions with higher activation probability even when receptor densities are equal. Indeed, spatial discrimination in touch improves with training or even just passive stimulation (***Van Boven et al., 2000***; ***Godde et al., 2000***), up to a limit that is presumably related to receptor density (***Wong et al., 2013***; ***Peters et al., 2009***). Assuming a fixed resource limit, training may offer improvements to some digits at the detriment of others. Whether this is indeed the case has to our knowledge not yet been empirically tested.

Finally, previous work has shown that non-uniform tuning curves across a population will lead to characteristic biases in perceptual tasks (***Wei and Stocker, 2015***). While the original formulation assumed that this heterogeneous allocation of output neurons was driven by stimulus statistics alone, we have shown here that it can also be a consequence of receptor densities. Thus, perceptual biases might also be expected to arise from neural populations that efficiently represent sensory inputs sampled by non-uniform receptor populations.

Limitations and future work

We considered simple linear models based on decorrelation and demonstrated that even such seemingly straightforward models exhibit surprising complexity in how they manage trade-offs in resource allocation under constraints. Specifically, we found that output neurons were not generally allocated proportionally to input neurons or according to some other fixed rule. It therefore stands to reason that similarly complex trade-offs would manifest in more complex models, even though the precise allocations might differ. Nevertheless, since principal component analysis is widely employed for decorrelation and dimensionality reduction, and therefore incorporated into many other algorithms, our results immediately generalize to several other methods. For example, it is straightforward to extend the model with additional constraints (for example response sparsity or power constraints) that would not affect the resulting allocations (see Methods for details), and therefore the model presented here already covers a number of related models. This includes independent component analysis, which considers higher-order rather than second-order statistics, but relies on a whitened signal, which in the undercomplete (bottleneck) case is obtained via PCA (***Hyvärinen and Oja, 2000***). Similarly, some models that do incorporate sensory noise and maximize reconstruction accuracy also use an undercomplete set of principal components to reduce the dimensionality of the sensory signal (***Doi and Lewicki, 2014***). In both of these examples the resulting receptive field structure will differ, but their allocation—where on the sensory sheet they will fall—will be governed by the same principles described earlier. However, we did not consider nonlinear models and previous work has demonstrated that response nonlinearities can make important contributions to decorrelation (***Pitkow and Meister, 2012***). Additionally, the precise structure and origin of noise in nonlinear models has been demonstrated to affect efficient coding strategies (***Brinkman et al., 2016***) and might therefore also influence the resulting allocations. Future work should explore resource allocation in such complex models.

We considered allocations for monotonically decreasing covariance functions. Many sensory systems will likely obey the intuitive notion that receptor correlations decrease with distance, however there are exceptions, for example in the auditory system (***Terashima and Okada, 2012***). While it is possible in principle to solve the allocation problem numerically for arbitrary covariance functions, in practice we noticed that the presence of small numerical errors can affect the sorting process and caution is therefore warranted, especially when considering non-monotonic covariance functions.

Methods and Materials

Our main goal was to derive a method for efficiently allocating output neurons to one of several input regions with different correlational response structure, in the presence of constraints on the number of output neurons or amount of information being transmitted. In the following, we focus on the main rationale and equations, while specific proofs can be found in the appendix. First, we outline the framework for combined whitening and dimensionality reduction that is employed. Next, we demonstrate how this framework can be applied to multiple input regions with different statistics and densities of receptors, and how calculation of the eigenvalues of region-specific covariance matrices solves the problem of resource allocation. Finally, we demonstrate how the problem can be solved analytically for a certain choice of covariance function. Code implementing the equations that can be used to recreate the figures in this manuscript is available on Github (***Edmondson, 2021***).

Combined whitening and dimensionality reduction

We assume that receptors are arranged on a two-dimensional sensory sheet. Correlations in the inputs occur as receptors that are nearby in space have more similar responses. Information is maximised in such a setup by decorrelating the sensory inputs. Here we decorrelate using a simple linear model. To model the bottleneck, we restrict the number of outputs to $m < n$, where n is the total number of receptors.

If the inputs are represented as a matrix $\boldsymbol{X}$ of dimensions $n \times z$ (where z is the number of sensory input patterns), then our goal is to find an $m \times n$ dimensional matrix $\boldsymbol{W}$ such that $\boldsymbol{W}\boldsymbol{X}$ is uncorrelated:

$$\boldsymbol{X}^T \boldsymbol{W}^T \boldsymbol{W} \boldsymbol{X} = \boldsymbol{I}. \tag{5}$$

This is achieved by setting $\boldsymbol{W} = \boldsymbol{\Sigma}^{-\frac{1}{2}}$, where $\boldsymbol{\Sigma} = \boldsymbol{X}^T \boldsymbol{X}$. Solutions can then be expressed in terms of the diagonal matrix of eigenvalues, $\boldsymbol{\Lambda}$, and eigenvectors, Φ, of the covariance matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{W} = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\Phi^T. \tag{6}$$

Whitening filters are not unique and obtained with any orthogonal matrix $\boldsymbol{P}$. If $\boldsymbol{P} = \boldsymbol{I}$, the receptive fields $\boldsymbol{W}$ are the principal components of the inputs (standard PCA). Localized receptive fields are obtained by setting $\boldsymbol{P} = \boldsymbol{\Phi}$, which is known as ZCA (Zero-Phase Component Analysis). In cases with a bottleneck the solution involves solving an Orthogonal Procrustes problem (***Doi and Lewicki, 2014***) to find $\boldsymbol{P}^*$, an m-dimensional orthogonal matrix (where m is the size of the bottleneck) which minimizes the reconstruction error of the inputs and a set of ideal local receptive fields $\boldsymbol{W}_{opt}$:

$$\boldsymbol{P}^* = \min_{\boldsymbol{P}} \left\| \boldsymbol{W}_{opt} - \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\Phi^T \right\|_F^2, \tag{7}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$ are as above but retaining only those components with the m largest eigenvalues. Because the minimum-error solution is independent of $\boldsymbol{P}$, additional constraints on the solution can be enforced. For example, power constraints are often added, which either limit the total variance or equalize the variance across output neurons, and additional sparsity constraints can be placed on the output neuron's activity or their receptive field weights (***Doi et al., 2012***; ***Doi and Lewicki, 2014***). Importantly for our problem, such additional constraints will affect the precise receptive field structure (through $\boldsymbol{P}$), but not the eigenvalues and eigenvectors included in $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$, respectively. It is also possible to extend the model to include certain simple forms of noise, for example independent Gaussian noise on the output neurons (***Doi and Lewicki, 2014***), but again this would not affect the eigendecomposition. As we will see in the following section, our solution relies on the eigenvalues only, and we can therefore solve the allocation problem irrespective of the precise receptive field structure or additional constraints on the solution.

Extension to multiple input regions

For our specific problem, we are interested in the case of multiple input regions with different correlational structure (i.e. due to differing receptor density or activation). To simplify the derivations we approximate different input regions as independent, such that the overall covariance matrix will be a block diagonal matrix. The covariance $\mathbf{\Sigma}$ for two input regions, R_1 and R_2, can then be expressed as follows:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}^{(R1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}^{(R2)} \end{bmatrix} \tag{8}$$

This assumption turns out to be a reasonable approximation when region sizes are relatively big and correlations typically do not extend far across the sensory sheet (see ***Edmondson et al., 2019***, for a comparison between block and non-block region covariance matrices in 1D). Furthermore, in many sensory systems, the borders between regions of differing density tend to be relatively narrow. For example, in touch, the digits of the hand are spatially separated, and regions of differing densities, for example between the digit tips and proximal phalanges, neighbour along the short rather than long axis of the digit. In the star-nosed mole, rays of different innervation densities are separated and neighbour only along their connection to the rest of the nose. However, the block matrix approximation might be problematic in cases with many very small adjacent regions with strong, far-ranging correlations.

The eigenvalues and eigenvectors of a block diagonal covariance matrix also follow the block diagonal form and can be calculated from the individual region covariances alone, by a simple application of the Cauchy interlacing theorem. Thus, the corresponding eigenvalues and eigenvectors are:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}^{(R1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}^{(R2)} \end{bmatrix} \quad \text{and} \quad \Phi = \begin{bmatrix} \Phi^{(R1)} & \mathbf{0} \\ \mathbf{0} & \Phi^{(R2)} \end{bmatrix}. \tag{9}$$

Due to the imposed bottleneck, only the m largest eigenvalues from the combined set $\mathbf{\Lambda}$ will be retained. If receptive fields are localised such that they are constrained to fall within a single input region, then an eigenvalue selected from $\mathbf{\Lambda}^{(R1)}$ indicates that the receptive field of the corresponding output neuron will fall onto region R_1, and analogously for R_2. This fact holds independent of the structure of $\boldsymbol{P}$ that is chosen in ***Equation 7***, because in order to preserve decorrelation a given region cannot contain more output neurons than eigenvalues retained from this region. In the following, we show how the eigenvalues can be calculated analytically for certain covariance functions.

Calculation of eigenvalues for negative exponential covariance functions

We model the covariance between receptors as a negative exponential function. The covariance matrix is then calculated as a function of distance between pairs of receptors (see ***Figure 1***C). We will first go through the calculation of eigenvalues for the baseline region R_1, and then continue with the derivation for R_2, which exhibits different receptor density, receptor activation, or both.

For region R_1 the covariance between receptors is calculated as:

$$\mathbf{\Sigma}_{ij}^{(R1)} = e^{-\gamma|x_i - x_j|}, \tag{10}$$

where x_i and x_j are the locations of the ith and jth receptors, and γ is the decay constant.

The corresponding eigenvalues for an exponential covariance function in the continuous domain can be calculated analytically. The eigenvalue-eigenvector problem is expressed as an integral homogeneous equation, such that for R_1 we get:

$$\lambda_k \phi_k(x) = \int_0^L e^{-\gamma|x-y|} \phi_k(y) dy, \tag{11}$$

where $\phi_k(x)$ is the kth eigenfunction and λ_k its corresponding eigenvalue. The domain length L is the input region size for one of the dimensions.

It can be shown that solutions to this problem can be related to the Laplacian operator (see ***Appendix 2*** and ***Appendix 3*** for proofs), such that:

$$\lambda_k = \frac{2\gamma}{\mu_k + \gamma^2}, \tag{12}$$

where μ_k are the eigenvalues of the Laplacian operator.

The general solution for the Laplacian eigenvalue problem for a two-dimensional rectangle with Dirichlet boundary conditions is (***Strauss, 2007***):

$$\mu_k = \mu_{l,m} = \frac{l^2\pi^2}{L_1^2} + \frac{m^2\pi^2}{L_2^2}, l, m = 1, 2, ..., \tag{13}$$

where L_1 an L_2 are the size of the domain for each dimension.

To calculate the covariance for R_2, we need to take into account the potentially different receptor density and response variance for this region. Denoting the ratio of the response variances between both regions by a, and the ratio of receptor densities by d, the covariance for R_2 can be expressed as:

$$\mathbf{\Sigma}_{ij}^{(R2)} = ae^{-d\gamma|x_i - x_j|}. \tag{14}$$

It can be seen that a scales the overall covariance matrix (see also ***Figure 1***C), while d changes the spatial extent of the correlations and thereby implicitly accounts for the different receptor density.

The calculation of the eigenvalues for R_2 proceeds analogously. The eigenvalue-eigenvector problem is given as:

$$\lambda_k \phi_k(x) = \int_0^{L\sqrt{d}} ae^{-d\gamma|x-y|}\phi_k(y)dy \tag{15}$$

Since the receptor density ratio d causes an implicit stretching of the space for the higher density region, the region length L needs to be adjusted in order to keep the effective size of the region constant. In 2D, each axis is therefore scaled by $\sqrt{d}$, resulting in an upper integration limit of $L\sqrt{d}$.

Allocation in the bottleneck

Given a sensory system with limited representational capacity, different regions may be allocated different amounts of resources. Here we calculate the allocations over different bottleneck widths for two regions, while the extension to multiple regions is given in ***Appendix 5***. In the following, we assume two-dimensional square regions of equal size for ease of analysis (see ***Appendix 6*** for the equivalent solution in 1D). A single variable L is therefore used to denote the lengths of the squares. Following ***Equation 12***, the eigenvalues for regions R_1 and R_2 can now be calculated as:

$$R_1 : \lambda_{l,m}^{(R1)} = \frac{2\gamma}{\pi^2 L^{-2}(l^2 + m^2) + \gamma^2} \tag{16}$$

$$R_2 : \lambda_{n,o}^{(R2)} = \frac{2\gamma a\sqrt{d}}{\pi^2 L^{-2}(n^2 + o^2) + \gamma^2} \tag{17}$$

where l, m and $n, o \in \mathbb{N}$ enumerate different eigenvalues for regions R_1 and R_2, respectively.

In order to calculate how many output neurons are allocated to R_1 and R_2 for different bottleneck widths, we will need to establish an ordering of the eigenvalues, such that for each pair (l, m) we can determine the sorted rank of the eigenvalues. In contrast to the 1D case (see ***Appendix 6***), there is no natural ordering of the eigenvalues in two dimensions, however, a close approximation can be obtained by calculating the number of lattice points enclosed by a quarter circle with radius $p = l^2 + m^2$ (see ***Appendix 4*** for full details). Denoting this function as $N(p)$ and setting $p^{(R1)} = l^2 + m^2$ and $p^{(R2)} = n^2 + o^2$, we can then calculate the number of eigenvalues allocated to R_1 as a function of the number of neurons allocated to R_2, by setting $\lambda^{(R1)} = \lambda^{(R2)}$ and solving for $p^{(R2)}$. This yields:

$$p^{(R2)} = a\sqrt{d}p^{(R1)} + \frac{L^2\gamma^2 a\sqrt{d} - L^2\gamma^2}{\pi^2}. \tag{18}$$

As we allocate more neurons to region R_1, the ratio $\frac{N(p^{(R1)})}{N(p^{(R2)})}$ simplifies to: $\lim_{R_1\to\infty} \frac{N(p^{(R1)})}{N(p^{(R2)})} = a\sqrt{d}$. The fraction of neurons allocated to each region therefore depends on the size of the bottleneck and converges to $\frac{1}{1+a\sqrt{d}}$ and $\frac{a\sqrt{d}}{1+a\sqrt{d}}$ for R_1 and R_2 respectively.

Table 1. Model parameters for the star-nosed mole.

Ray	1	2	3	4	5	6	7	8	9	10	11
Size	1.14	1.21	1.21	1.17	1.08	1.02	1.00	1.13	1.05	0.87	1.10
Dens.	45.78	47.14	45.82	45.69	45.2	46.91	43.3	44.01	44.26	47.92	50.46
Act.	0.03	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.04	0.08	0.11
γ	0.99	1.00	1.02	1.01	1.00	1.01	0.99	1.04	1.10	1.16	1.27

Alternative covariance functions

To test generalisation to other covariance functions that decrease monotonically with receptor distance, we tested a number of functions from the Matérn class, in which the parameter ν controls the smoothness of the function. The negative exponential covariance function we employed in previous sections is equivalent to a Matérn function with $\nu = 1/2$. Larger values of ν lead to progressively higher correlations for smaller distances and lower correlations for larger distances. Specifically, we tested a Matérn function with $\nu = 3/2$:

$$\mathbf{\Sigma}_{ij}^{\nu=3/2} = \left(1 + \gamma\sqrt{3}|x_i - x_j|\right) e^{-\gamma\sqrt{3}|x_i - x_j|}, \tag{19}$$

and with $\nu = 5/2$:

$$\mathbf{\Sigma}_{ij}^{\nu=5/2} = \left(1 + \gamma\sqrt{5}|x_i - x_j| + \frac{\gamma^2 5|x_i - x_j|^2}{3}\right) e^{-\gamma\sqrt{5}|x_i - x_j|} \tag{20}$$

For simplicity and to obtain numerically stable solutions, all calculations were performed for one-dimensional regions only. For all cases tests, the baseline region contained 500 receptors, while the number of receptors for the other region was determined by the density ratio. Eigenvalues were calculated from the covariance matrix using the *eigh* method of the *numpy* Python package (***Harris et al., 2020***). Comparing the numerically obtained allocation for the negative exponential covariance function with its analytical solutions showed a close match for most bottleneck widths. However, numerical errors increased for wider bottlenecks, where eigenvalues became very small and their decay flattened out, affecting the sorting process. For the analyses here, we therefore restricted ourselves to bottleneck widths where a close match between the numerical and analytical solutions could be obtained. This range was sufficient to demonstrate allocation at narrow bottlenecks and to estimate the convergence point for larger ones.

Calculations for star-nosed mole

The eleven rays were approximated as 2D square regions with areas set to their reported sizes (***Sawyer and Catania, 2016***). Receptor densities for each ray were calculated as the sensory fiber innervation per mm^2 (***Catania and Kaas, 1997***). Approximations of receptor activation on each ray were calculated from empirical data of prey foraging interactions recorded by ***Catania and Kaas*** (***1997***). Contact statistics were converted to receptor activation probabilities with receptors following a Bernoulli distribution. Finally, activation variance was calculated as the variance of the Bernoulli distribution (see ***Appendix 1***). The decay rate γ of the negative exponential covariance function was determined for each ray using a model of the typical extent of receptor activation during interaction with prey stimuli of varying sizes. Each ray interacts with varying prey sizes at different frequencies. For example, ray 11 is typically contacted by smaller stimuli more often than other rays. A 2D model of the rays was used to simulate average responses to each stimulus size. Each model ray was tiled with receptors, and circular stimuli of different sizes were then randomly placed over the ray. The radii and frequencies of each stimulus size were based on the prey model (***Catania and Kaas, 1997***)). A ray receptor was marked as active if its coordinate position was within the bounds of the stimuli. Response covariance between receptors was then calculated and an exponential function was fit to find the γ decay parameter. See ***Table 1*** for the full set of parameters. Code implementing the receptor model is available on Github (***Edmondson, 2021***). To

determine allocations, eigenvalues associated with each ray were calculated analytically, resulting in allocations for each ray at all bottleneck widths. Three models were compared: first, a 'density only' model, which includes accurate receptor density ratios, but receptor activation ratio remains uniform across all rays; second, an 'activation only' model, which includes heterogeneous receptor activation ratios, but uniform receptor density ratios across all rays; finally, the 'full model' combines both accurate densities and receptor activation ratios. Model allocations for each ray were compared to the cortical allocation empirical data from ***Catania and Kaas*** (***1997***). As the bottleneck size for the star-nosed mole is unknown, the root-mean-square error (RMSE) was calculated for each model at all bottleneck widths. The bottleneck resulting in the lowest error was then selected for each. Allocations for rays at each bottleneck can be found in ***Figure 6–Figure Supplement 1***A. For the 'activation only' and 'full' models, the lowest RMSE values were for bottleneck widths of between 37-45%; for the 'density only' model, the RMSE was similar over all bottlenecks widths (see ***Figure 6–Figure Supplement 1***C). We also tested a 'baseline model 'where densities and activation were randomly selected for each ray within the possible range of parameters. The aim was to determine how much explained variance within the cortical allocations was due to the selection of the best-fitting bottleneck and how much was due to the specific density and activation parameters. A total of 20 random models were run, and the average R^2 was -0.09 (see ***Figure 6–Figure Supplement 1***D).

Acknowledgments

This work was supported by the Wellcome Trust [209998/Z/17/Z] and by the European Union Horizon 2020 program as part of the Human Brain Project [HBP-SGA2, 785907].

References

Anderson SJ, Mullen KT, Hess RF. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. J Physiol. 1991 Oct; 442:47–64.

Atick JJ. Could information theory provide an ecological theory of sensory processing? Network: Computation in neural systems. 1992; 3(2):213–251.

Atick JJ, Redlich AN. Towards a theory of early visual processing. Neural Computation. 1990; 2(3):308–320.

Atick JJ, Redlich AN. What does the retina know about natural scenes? Neural Comput. 1992; 210:196–210.

Attneave F. Some informational aspects of visual perception. Psychological review. 1954; 61(3):183.

Azzopardi P, Cowey A. Preferential representation of the fovea in the primary visual cortex. Nature. 1993; 361(6414):719–721.

Barlow HB. Possible principles underlying the transformation of sensory messages. Sensory communication. 1961; 1:217–234.

Bell AJ, Sejnowski TJ. The "independent components" of natural scenes are edge filters. Vision Research. 1997; 37(23):3327–3338.

Bhand M, Mudur R, Suresh B, Saxe A, Ng AY. Unsupervised learning models of primary cortical receptive fields and receptive field plasticity. In: Shawe-Taylor J, Zemel RS, Bartlett PL, Pereira F, Weinberger KQ, editors. *Advances in Neural Information Processing Systems 24* Curran Associates, Inc.; 2011.p. 1971–1979.

Brinkman BAW, Weber AI, Rieke F, Shea-Brown E. How Do Efficient Coding Strategies Depend on Origins of Noise in Neural Circuits? PLoS Comput Biol. 2016; 12(10):e1005150–e1005150.

Catani M. A little man of some importance. Brain. 2017; 140(11):3055–3061.

Catania KC. All in the family–touch versus olfaction in moles. The Anatomical Record. 2020; 303(1):65–76.

Catania KC, Kaas JH. Somatosensory fovea in the star-nosed mole. Journal of Comparative Neurology. 1997; 387(2):215–233.

Catania KC, Remple FE. Tactile foveation in the star-nosed mole. Brain, behavior and evolution. 2004; 63(1):1–12.

Chalk M, Marre O, Tkačik G. Toward a unified theory of efficient, predictive, and sparse coding. Proceedings of the National Academy of Sciences. 2018; 115(1):186–191.

Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I. Reduction of information redundancy in the ascending auditory pathway. Neuron. 2006; 51(3):359–368.

Clemens J, Kutzki O, Ronacher B, Schreiber S, Wohlgemuth S. Efficient transformation of an auditory population code in a small sensory system. Proc Natl Acad Sci U S A. 2011 Aug; 108(33):13812–13817.

Coq JO, Xerri C. Environmental enrichment alters organizational features of the forepaw representation in the primary somatosensory cortex of adult rats. Experimental brain research. 1998; 121(2):191–204.

Corniani G, Saal HP. Tactile innervation densities across the whole body. J Neurophysiol. 2020 Oct; 124(4):1229–1240.

Curcio CA, Allen KA. Topography of ganglion cells in human retina. Journal of comparative Neurology. 1990; 300(1):5–25.

Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. Human photoreceptor topography. Journal of comparative neurology. 1990; 292(4):497–523.

Doi E, Gauthier JL, Field GD, Shlens J, Sher A, Greschner M, Machado TA, Jepson LH, Mathieson K, Gunning DE, et al. Efficient coding of spatial information in the primate retina. Journal of Neuroscience. 2012; 32(46):16256–16264.

Doi E, Lewicki MS. A simple model of optimal population coding for sensory systems. PLoS computational biology. 2014; 10(8):e1003761.

Dolman CL, McCormick AQ, Drance SM. Aging of the optic nerve. Archives of ophthalmology. 1980; 98(11):2053–2058.

Duncan RO, Boynton GM. Cortical magnification within human primary visual cortex correlates with acuity thresholds. Neuron. 2003; 38(4):659–671.

Duncan RO, Boynton GM. Tactile hyperacuity thresholds correlate with finger maps in primary somatosensory cortex (S1). Cerebral Cortex. 2007; 17(12):2878–2891.

Edmondson LR, expansion_contraction_sensory_bottlenecks; 2021. https://github.com/lauraredmondson/expansion_contraction_sensory_bottlenecks. Github. https://github.com/lauraredmondson/expansion_contraction_sensory_bottlenecks 5483347.

Edmondson LR, Rodriguez AJ, Saal HP. Nonlinear scaling of resource allocation in sensory bottlenecks. In: *Advances in Neural Information Processing Systems*; 2019. p. 7545–7554.

Engel SA, Glover GH, Wandell BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cerebral cortex (New York, NY: 1991). 1997; 7(2):181–192.

Finlay BL, Franco ECS, Yamada ES, Crowley JC, Parsons M, Muniz JAP, Silveira LCL. Number and topography of cones, rods and optic nerve axons in New and Old World primates. Visual Neuroscience. 2008; 25(3):289.

Fragaszy DM, Crast J. Functions of the hand in primates. In: *The evolution of the primate hand* Springer; 2016.p. 313–344.

Ganguli D, Simoncelli E. Implicit encoding of prior probabilities in optimal neural populations. Advances in neural information processing systems. 2010; 23:658–666.

Ganguli D, Simoncelli EP. Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. Neural computation. 2014; 26(10):2103–2134.

Ganguli D, Simoncelli EP. Neural and perceptual signatures of efficient sensory coding. arXiv preprint arXiv:160300058. 2016; .

García-Piqueras J, García-Mesa Y, Cárcaba L, Feito J, Torres-Parejo I, Martín-Biedma B, Cobo J, García-Suárez O, Vega JA. Ageing of the somatosensory system at the periphery: age-related changes in cutaneous mechanoreceptors. Journal of anatomy. 2019; 234(6):839–852.

Godde B, Stauffenberg B, Spengler F, Dinse HR. Tactile coactivation-induced changes in spatial discrimination performance. Journal of Neuroscience. 2000; 20(4):1597–1604.

Gonzalez F, Gosselin F, Bachta W. Analysis of hand contact areas and interaction capabilities during manipulation and exploration. IEEE transactions on haptics. 2014; 7(4):415–429.

Goodchild AK, Ghosh KK, Martin PR. Comparison of photoreceptor spatial density and ganglion cell morphology in the retina of human, macaque monkey, cat, and the marmoset Callithrix jacchus. Journal of Comparative Neurology. 1996; 366(1):55–75.

Goodwin AW, Wheat HE. Sensory signals in neural populations underlying tactile perception and manipulation. Annu Rev Neurosci. 2004 Dec; 27:53–77.

Graham DJ, Chandler DM, Field DJ. Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields? Vision Research. 2006; 46(18):2901–2913.

Graham D, Field D. Natural images: Coding efficiency. Encyclopedia of Neuroscience. 2009; 6:19–27.

Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, Fernández del Río J, Wiebe M, Peterson P, Gérard-Marchant P, et al. Array programming with NumPy. Nature. 2020; 585:357–362. doi: 10.1038/s41586-020-2649-2.

Harvey BM, Dumoulin SO. The Relationship between Cortical Magnification Factor and Population Receptive Field Size in Human Visual Cortex: Constancies in Cortical Architecture. Journal of Neuroscience. 2011; 31(38):13604–13612.

Hermundstad AM, Briguglio JJ, Conte MM, Victor JD, Balasubramanian V, Tkacik G. Variance predicts salience in central sensory processing. Elife. 2014; .

Huffman KJ, Molnár Z, Van Dellen A, Kahn DM, Blakemore C, Krubitzer L. Formation of cortical fields on a reduced cortical sheet. J Neurosci. 1999 Nov; 19(22):9939–9952.

Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. Neural networks. 2000; 13(4-5):411–430.

Jenkins WM, Merzenich MM, Ochs MT, Allard T, Guic-Robles E. Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. Journal of neurophysiology. 1990; .

Johansson RS, Vallbo A. Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. The Journal of physiology. 1979; 286(1):283–300.

Jonas JB, Müller-Bergh J, Schlötzer-Schrehardt U, Naumann G. Histomorphometry of the human optic nerve. Investigative ophthalmology & visual science. 1990; 31(4):736–744.

Karklin Y, Simoncelli EP. Efficient coding of natural images with a population of noisy linear-nonlinear neurons. Advances in neural information processing systems. 2011; 24:999.

Kersten D. Predictability and redundancy of natural images. J Opt Soc Am. 1987 dec; 4(12):2395.

Lehnert BP, Santiago C, Huey EL, Emanuel AJ, Renauld S, Africawala N, Alkislar I, Zheng Y, Bai L, Koutsioumpa C, Hong JT, Magee AR, Harvey CD, Ginty DD. Mechanoreceptor synapses in the brainstem shape the central representation of touch; 2021, unpublished.

Lewicki MS. Efficient coding of natural sounds. Nature neuroscience. 2002; 5(4):356–363.

Lindsey J, Ocko SA, Ganguli S, Deny S. A unified theory of early visual representations from retina to cortex through anatomically constrained deep CNNs. International Conference on Learning Representations (ICLR). 2019; .

Ly C, Middleton JW, Doiron B. Cellular and circuit mechanisms maintain low spike co-variability and enhance population coding in somatosensory cortex. Front Comput Neurosci. 2012 Mar; 6:7.

Martuzzi R, van der Zwaag W, Farthouat J, Gruetter R, Blanke O. Human finger somatotopy in areas 3b, 1, and 2: a 7T fMRI study using a natural stimulus. Human brain mapping. 2014; 35(1):213–226.

Merzenich MM, Jenkins WM. Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience. Journal of hand therapy. 1993; 6(2):89–104.

Olshausen BA, Field DJ. Wavelet-like receptive fields emerge from a network that learns sparse codes for natural images. Nature. 1996; 381:607–609.

Olshausen BA, Field DJ. Sparse coding with an overcomplete basis set: A strategy employed by V1? Vision research. 1997; 37(23):3311–3325.

Olshausen BA, Field DJ. Sparse coding of sensory inputs. Current opinion in neurobiology. 2004; 14(4):481–487.

Perry V, Oehler R, Cowey A. Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. Neuroscience. 1984; 12(4):1101–1123.

Peters RM, Hackeman E, Goldreich D. Diminutive digits discern delicate details: fingertip size and the sex difference in tactile spatial acuity. Journal of Neuroscience. 2009; 29(50):15756–15761.

Pitkow X, Meister M. Decorrelation and efficient coding by retinal ganglion cells. Nat Neurosci. 2012 Apr; 15(4):628–635.

Plumbley MD. Do cortical maps adapt to optimize information density? Network. 1999 Feb; 10(1):41–58.

Provis JM, Van Driel D, Billson FA, Russell P. Human fetal optic nerve: overproduction and elimination of retinal axons during development. Journal of Comparative Neurology. 1985; 238(1):92–100.

Ritter H, Martinetz T, Schulten K. Neural Computation and Self-organizing Maps: An Introduction. Addison-Wesley; 1992.

Ritter H, Schulten K. On the stationary state of Kohonen's self-organizing sensory mapping. Biological cybernetics. 1986; 54(2):99–106.

Sandell JH, Peters A. Effects of age on nerve fibers in the rhesus monkey optic nerve. Journal of Comparative Neurology. 2001; 429(4):541–553.

Sawyer EK, Catania KC. Somatosensory organ topography across the star of the star-nosed mole (Condylura cristata). Journal of Comparative Neurology. 2016; 524(5):917–929.

Sefton AJ, Horsburgh GM, Lam K. The development of the optic nerve in rodents. Australian and New Zealand journal of ophthalmology. 1985; 13(2):135–145.

Sereno MI, Dale A, Reppas J, Kwong K, Belliveau J, Brady T, Rosen B, Tootell R. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science. 1995; 268(5212):889–893.

Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci. 1998 May; 18(10):3870–3896.

Simmons GF. Differential equations with applications and historical notes. CRC Press; 2016.

Smith EC, Lewicki MS. Efficient auditory coding. Nature. 2006; 439(7079):978–982.

Strauss WA. Partial differential equations: An introduction. John Wiley & Sons; 2007.

Terashima H, Okada M. The topographic unsupervised learning of natural sounds in the auditory cortex. In: F Pereira and C J C Burges and L Bottou and K Q Weinberger, editor. *Advances in Neural Information Processing Systems 25* Curran Associates, Inc.; 2012. .

Tishby N, Pereira FC, Bialek W. The information bottleneck method. arXiv preprint physics/0004057. 2000; .

Tishby N, Zaslavsky N. Deep learning and the information bottleneck principle. In: *2015 IEEE Information Theory Workshop (ITW)* IEEE; 2015. p. 1–5.

Van Boven RW, Hamilton RH, Kauffman T, Keenan JP, Pascual-Leone A. Tactile spatial resolution in blind Braille readers. Neurology. 2000; 54(12):2230–2236.

Verendeev A, Thomas C, McFarlin SC, Hopkins WD, Phillips KA, Sherwood CC. Comparative analysis of Meissner's corpuscles in the fingertips of primates. Journal of anatomy. 2015; 227(1):72–80.

Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during natural vision. Science. 2000 Feb; 287(5456):1273–1276.

Wei XX, Stocker AA. A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. Nature Neuroscience. 2015; 18(10):1509.

Wells-Gray E, Choi S, Bries A, Doble N. Variation in rod and cone density from the fovea to the mid-periphery in healthy human retinas using adaptive optics scanning laser ophthalmoscopy. Eye. 2016; 30(8):1135.

Wong M, Peters RM, Goldreich D. A Physical Constraint on Perceptual Learning: Tactile Spatial Acuity Improves with Training to a Limit Set by Finger Size. Journal of Neuroscience. 2013 May; 33(22):9345–9352.

Xerri C, Coq J, Merzenich M, Jenkins W. Experience-induced plasticity of cutaneous maps in the primary somatosensory cortex of adult monkeys and rats. Journal of Physiology-Paris. 1996; 90(3-4):277–287.

Xerri C, Merzenich MM, Jenkins W, Santucci S. Representational plasticity in cortical area 3b paralleling tactual-motor skill acquisition in adult monkeys. Cerebral cortex. 1999; 9(3):264–276.

Yarrow S, Challis E, Seriès P. Fisher and Shannon information in finite neural populations. Neural computation. 2012; 24(7):1740–1780.

Yerxa TE, Kee E, DeWeese MR, Cooper EA. Efficient sensory coding of multidimensional stimuli. PLoS computational biology. 2020; 16(9):e1008146.

Appendix 1

Stimulus statistics and response variance

Decorrelation works on second-order statistics and therefore stimulus statistics would only be taken into account by the model if they affect the covariance matrix. One way this can happen is through the extent of the spatial correlations (parameter γ in the covariance function). For example, in touch the size distribution of stimuli that would typically make contact with a given skin region might differ, leading to a different correlational structure. While we calculate allocations for different values of γ, we keep this value fixed across different input regions, for simplicity.

A more common case is that of receptors in one region being more active than receptors in another region. For example, in touch, the fingertips make contact with objects much more frequently than does the palm. An increased probability of making contact would translate into higher receptor response rates. In turn, higher response rates imply higher response variance, which would be directly reflected in the covariance matrix as a multiplicative scaling of the covariance function. For example, assuming that each receptor follows a simple Bernoulli distribution (either on or off, with a probability of p being on), then the response variance can be calculated as $p(1-p) = p - p^2$. Assuming that the likelihood of any receptor being active is generally low, the variance scales almost linearly with receptor activation. Differences in activation between two regions are represented by the activation ratio a throughout the paper.

Appendix 2

Relationship between PCA and Laplacian eigenvalue problem

Rationale

Let Ω be a region with a density of receptors ρ. In a one dimensional region $\Omega = [0, L]$, the density can be expressed as $\rho = \frac{N}{L}$ or number of receptors per unit length. Assuming an exponential decay of correlations, the covariance between receptors i and j is:

$$C(i, j) = e^{-\gamma|i\Delta x - j\Delta x|}, \tag{21}$$

where $\Delta x = 1/\rho$ is the distance between receptors. Subsampling the space by taking a fraction N/d of the original receptors, $d > 1$, the covariance for positions i, j, becomes:

$$\tilde{C}(i, j) = e^{-\gamma|id\Delta x - jd\Delta x|} = e^{-d\gamma|i\Delta x - j\Delta x|}. \tag{22}$$

Therefore, we encode the changes in receptor density in a scaling of the exponential decay rate. For a given distribution of receptors, there is an induced partition of the interval $[0, L]$, therefore, for a fixed $x = i\Delta x$, the covariance in the jth bin is approximately equal to the area of the exponential covered in that bin:

$$\int_{(j-1)\Delta x}^{j\Delta x} C(x, y)dy \approx C(i\Delta x, j\Delta x)\Delta x,$$

Summing over all the bins, we arrive at the PCA problem:

$$\sum_{j=0}^{n} C(i\Delta x, j\Delta x)\phi(j\Delta x)\Delta x = \lambda\phi(i\Delta x). \tag{23}$$

The continuum limit is found formally as $\Delta x \to 0$.

Derivation

In order to find the optimal assignment for a given receptor density, we are interested in solutions to the following equation which can be seen as a continuous version of the traditional PCA problem with an exponentially decaying covariance matrix:

$$\lambda\phi(x) = \int_0^L e^{-\gamma|x-y|}\phi(y)dy, \tag{24}$$

where γ is the decay rate. We are interested in solutions $\phi \in C^2(\mathbb{R})$, that is, twice differentiable solutions that satisfy appropriate boundary conditions.

Theorem 1 *If ϕ is a solution of* **Equation 24**, *then it is an eigenfunction of the Laplacian operator with eigenvalues:*

$$\mu = \frac{2\gamma}{\lambda} - \gamma^2. \tag{25}$$

That is, in one dimension, solutions ϕ satisfy:

$$-\frac{d^2}{dx^2}\phi(x) = \mu\phi(x). \tag{26}$$

Proof: Differentiating ***Equation 24*** twice using the Leibniz rule we obtain:

$$\frac{d}{dx}\phi(x) = \frac{\gamma}{\lambda}\left\{-\int_0^x e^{-\gamma(x-y)}\phi(y)dy + \int_x^L e^{\gamma(x-y)}\phi(y)dy\right\}, \quad \text{and} \tag{27}$$

$$\frac{d^2}{dx^2}\phi(x) = \frac{\gamma}{\lambda}\left\{-2\phi(x) + \gamma\int_0^L e^{-\gamma|x-y|}\phi(y)dy\right\}. \tag{28}$$

The second term on the right hand side can be replaced using ***Equation 24*** obtaining the desired result:

$$\frac{d^2}{dx^2}\phi(x) = -\frac{2\gamma}{\lambda}\phi(x) + \gamma^2\phi(x), \quad \text{or} \tag{29}$$

$$-\frac{d^2}{dx^2}\phi(x) = \left(\frac{2\gamma}{\lambda} - \gamma^2\right)\phi(x). \tag{30}$$

■

The previous is a sufficient condition on the solutions to ***Equation 24***. A necessary condition is given in the following theorem:

Theorem 2 *A solution to* **Equation 26** *is a also solution to* **Equation 24** *if it satisfies the following boundary conditions:*

$$\phi'(0) = \gamma\phi(0) \tag{31}$$

$$\phi'(L) = -\gamma\phi(L). \tag{32}$$

Proof: Assume ϕ is a solution to ***Equation 26***. We proceed by convolving ***Equation 26*** on both sides with the kernel $e^{-\gamma x}$:

$$\int_0^x e^{-\gamma(x-y)}\phi''(y)dy = \mu\int_0^x e^{-\gamma(x-y)}\phi(y)dy. \tag{33}$$

Integrating by parts twice we get:

$$\phi'(x) - e^{-\gamma x}\phi'(0) - \gamma\phi(x) + \gamma e^{-\gamma x}\phi(0) + \gamma^2\int_0^x e^{-\gamma(x-y)}\phi(y)dy = \mu\int_0^x e^{-\gamma(x-y)}\phi(y)dy. \tag{34}$$

Using ***Equation 25*** and ***Equation 32***, we obtain:

$$-\phi'(x) + \gamma\phi(x) = \frac{2\gamma}{\lambda}\int_0^x e^{-\gamma(x-y)}\phi(y)dy. \tag{35}$$

Repeating the procedure with the kernel $e^{\gamma x}$ in the interval $[x, L]$, yields:

$$\phi'(x) + \gamma\phi(x) = \frac{2\gamma}{\lambda}\int_x^L e^{-\gamma(x-y)}\phi(y)dy. \tag{36}$$

Adding ***Equation 35*** and ***Equation 36*** we recover ***Equation 24*** which finalizes the proof.

■

Scaling

In some instances of our problem, the exponential covariance will be scaled by the activation ratio, a. In general, the same reasoning applies to any linear combination of solutions, therefore, our results extend to that case. In particular, we have the following result:

Theorem 3 *The eigenvalues of the scaled covariance matrix, $C'(x, y) = aC(x, y)$ are:*

$$\lambda_s = a\lambda, \tag{37}$$

where λ is an eigenvalue of the original problem.

Proof: Let $\phi(x)$ be a solution of ***Equation 24***. By linearity of the integral we have:

$$\int_0^L ae^{-\gamma|x-y|}\phi(y)dy = a\int_0^L e^{-\gamma|x-y}\phi(y)dy = a\lambda\phi(x) \tag{38}$$

■

Appendix 3

Solutions

In the previous section, we saw that solutions of the PCA problem (***Equation 24***) and the Laplacian eigenvalue problem (***Equation 26***) coincide if boundary conditions specified in ***Equation 32*** and ***Equation 31*** are met. Here, we show how these solutions relate to solutions of the boundary value problem of ***Equation 26*** with $\phi(0) = \phi(L) = 0$, which correspond to the eigenmodes of an idealized vibrating string fixed at the extremes. Such modes are considerably simpler than the exact ones and, as we show, are sufficient for our analysis.

Eigenmodes of a vibrating string

Solving equation ***Equation 26***, for $\kappa = \sqrt{\mu}, \mu > 0$, using standard methods (see ***Simmons, 2016***, p. 355 onwards), we find the general expression for the eigenfunctions:

$$\phi(x) = A\sin(\kappa x) + B\cos(\kappa x), \tag{39}$$

with coefficients to be determined using the boundary conditions. The first boundary condition ($\phi(0) = 0$) implies that $B = 0$. The second boundary condition ($\phi(L) = 0$) gives the equation $\sin(\kappa L) = 0$, which is satisfied for $\kappa L = n\pi$, or:

$$\mu = \frac{n^2\pi^2}{L^2}, \tag{40}$$

where $n = 1, 2...$ is the index of the eigenvalue. The value of A is arbitrary and is usually selected so the solution is normalized.

Exact eigenvalues

In order to find the exact analytical eigenvalues of ***Equation 24***, we again assume $\mu > 0$ and a solution of the form given in ***Equation 39***. Using the first boundary condition (***Equation 31***, ***Equation 32***) we get the following relationship:

$$A = \frac{\gamma}{\kappa}B, \tag{41}$$

and with the second boundary condition, we obtain

$$\tan\kappa L = \frac{\kappa A + \gamma B}{\kappa B - \gamma A}; \tag{42}$$

replacing ***Equation 41***, we find the transcendental equation

$$\tan\kappa L = \frac{2\gamma\kappa}{\kappa^2 - \gamma^2}, \tag{43}$$

whose solutions lead to the exact eigenvalues of the continuous PCA problem.

Relationship between exact and approximate eigenvalues

The main results presented in this study rely on the ordering of the eigenvalues only, rather than their precise magnitude. Any divergence between the exact and approximate eigenvalues is therefore relevant only if it affects this ordering. Furthermore, since allocation of eigenvalues to regions proceeds cumulatively as more eigenvalues are added, localized disturbances in a few eigenvalues will lead to only minor errors that are quickly corrected as more eigenvalues are added (since the correct allocation of eigenvalues to regions will be restored). In the following, we demonstrate that, comparing the ordering of the exact eigenvalues where the boundary conditions are derived from the integral equation (***Equation 43***)

with that of the approximate ones (***Equation 40***), this order is altered only in a localized fashion that does not affect the analytical results of this paper. To understand the changes in the ordering, consider the functions:

$$f(\kappa) = \tan \kappa L, \tag{44}$$

and

$$g(\kappa) = \frac{2\gamma\kappa}{\kappa^2 - \gamma^2}, \tag{45}$$

whose intersection gives the solutions to ***Equation 43***. The function g has two hyperbolic-like branches, one to each side of the singularity $\kappa = \gamma$. We can distinguish two cases for said intersection:

1. For $|\kappa| \gg \gamma$, g approaches 0 and there is clearly only one intersection with f, which happens to be increasingly close to the approximate eigenvalues. Moreover, the function is also monotonic in this regime. This implies that the order is preserved.
2. For $|\kappa| \approx \gamma$, that is, close to the singularity, an additional root is inserted as f cuts both branches of g and the order is disturbed.

Altogether, this implies that, for the cases studied in this paper that consider $\gamma < \pi$, only the first two eigenvalues might differ from the order given by the approximate solution. In general, in the intervals $(n\pi/2, (n+1)\pi/2)$ there will be one root for both equations except in the one closest to the singularity which will contain two. Moreover, from ***Equation 25*** we see that, for two regions with densities γ and $d\gamma$ and Laplacian eigenvalues κ_1 and κ_2, $\lambda_1 > \lambda_2$ implies $\kappa_2 - d\kappa 1 > (1-d)d\gamma^2$. This inequality is satisfied by the exact eigenvalues, again, in all intervals except the one in which the singularity lies (as can be confirmed by expanding $f - g = 0$ around the approximate eigenvalues and solving for κ). We have compared the exact and approximate ordering for two regions for a number of (manageable) cases, and have found the relationship described above to hold.

Appendix 4

Ordering in the 2D square case

For rectangle regions, the ordering can be solved by calculating the number of lattice points enclosed by a quarter ellipse (***Strauss, 2007***). Here we use square regions and therefore the solution is the number of points enclosed in a quarter circle. The Gauss circle problem determines the number of integer lattice points which lie within a circle with radius $p \geq 0$, with its center at the origin:

$$N(p) = \#\{(l, m) \in \mathbb{R} | l^2 + m^2 \leq p^2\}. \tag{46}$$

The number of lattice points within the circle is approximately equal to its area. The number of points within a square region can be approximated by calculating the area of the upper quarter of the circle (positive values only):

$$N(p) = \frac{\pi p^2}{4} \tag{47}$$

The number of eigenvalues in each region is therefore the area of the intersection of the circle and region.

For each region we calculate the number of lattice points enclosed by a quarter circle; for R_1 we set the radius equal to $l^2 + m^2$ and for R_2 to $n^2 + o^2$ (the solution of ***Equation 18***), where $l, m, n, o = 1, 2...$ This number is approximated as the area of the quarter circle. For values of $l^2 + m^2$ or $n^2 + o^2$ greater than the total number of eigenvalues in each dimension ($L\sqrt{d}$), the approximation diverges from the true ordering as the area of the quarter circle becomes larger than the area of the lattice (region). In this case, a correction term is added:

$$N(p) = \begin{cases} \frac{\pi p}{4} - p \arccos\left(\frac{k}{\sqrt{p}}\right) - k\sqrt{p - k^2}, & \text{if } \frac{k}{\sqrt{p}} < 1. \\ \frac{\pi p}{4}, & \text{otherwise.} \end{cases} \tag{48}$$

where p is either $l^2 + m^2$ or $n^2 + o^2$ for R_1 and R_2 respectively, and k is the total number of eigenvalues in each region. Assuming a region size of L, with each receptor spaced one unit apart, $k = L^2$ for R_1, and $k = L^2 d$ for R_2.

Appendix 5

Allocation for multiple regions in 2D

For more than two regions, density and activation ratios for each additional region are calculated relative to a chosen baseline region. This leads to the following general form for calculation of the eigenvalues of any Region x:

$$R_x\colon\ \lambda_{l,m} = \frac{2\gamma a\sqrt{d_x}}{l^2\pi^2 L^{-2}\sqrt{d_b} + m^2\pi^2 L^{-2}\sqrt{d_b} + \gamma^2\sqrt{d_b}} \tag{49}$$

where a is the region activation scaling ratio, d_b is the density of the baseline region, and d_x the density of region x. $l, m \in \mathbb{N}$ enumerate different eigenvalues for region x.

Appendix 6

Allocation for the 1D case

The 1D case for changes in density has previously been addressed in ***Edmondson et al.*** (***2019***). Here we extend this to include changes in activation. For two regions R_1 and R_2, we can calculate their eigenvalues as:

$$R_1 : \quad \lambda_l^{(R1)} = \frac{2\gamma}{l^2\pi^2 L^{-2} + \gamma^2} \tag{50}$$

$$R_2 : \quad \lambda_m^{(R2)} = \frac{2\gamma ad}{m^2\pi^2 L^{-2} + \gamma^2} \tag{51}$$

where d is the ratio of higher and lower densities, a is the ratio of receptor activation, L is the length of the region, and $l, m \in \mathbb{N}$ denote successive eigenvalues for regions R_1 and R_2, respectively.

To calculate how many output neurons are allocated to region R_2 as a function of the number of neurons allocated to region R_1, we set $\lambda_l^{(R1)} = \lambda_m^{(R2)}$ and solve for m. This yields

$$m = \frac{\sqrt{ad(l^2\pi^2 + L^2\gamma^2) - L^2\gamma^2}}{\pi}. \tag{52}$$

It becomes apparent that for $l = 1$, i.e. the first neuron allocated to region R_1, we have already assigned $m = \frac{\sqrt{ad(\pi^2 + L^2\gamma^2) - L^2\gamma^2}}{\pi}$ neurons to region R_2. As we allocate more neurons to region R_1, the ratio $\frac{m}{l}$ simplifies to: $\lim_{l\to\infty} \frac{m}{l} = \sqrt{ad}$. The fraction of neurons allocated to each region therefore depends on the size of the bottleneck and converges to $\frac{1}{1+\sqrt{ad}}$ and $\frac{\sqrt{ad}}{1+\sqrt{ad}}$ for R_1 and R_2 respectively.

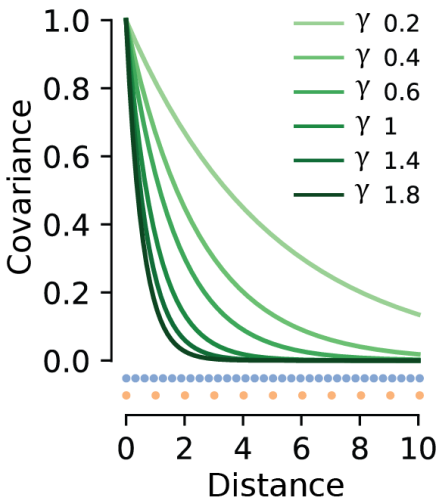

Figure 2–Figure supplement 1. Effect of different values of γ on the covariance function decay. Smaller γ leads to a slower decay and therefore a larger extent of spatial correlations. When the sigma is large, receptors have less similar responses to their neighbours, and therefore co-vary less. Orange points denote the baseline spacing of receptors along one dimension. When receptor density is varied, the other region will have higher density (e.g. 1:3, indicated by blue dots).

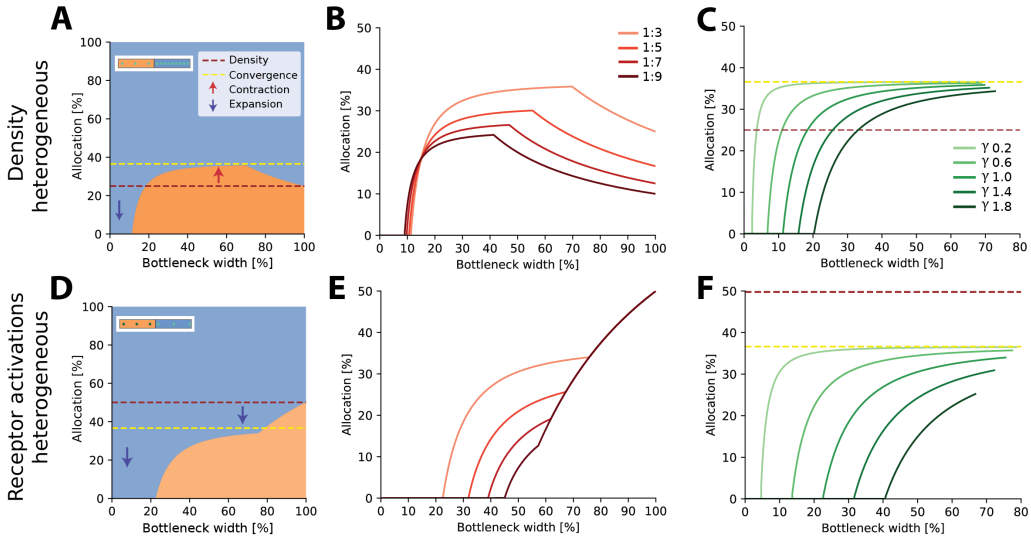

Figure 2–Figure supplement 2. Resource allocation for heterogeneous receptor densities and variations in input statistics in 1D. A. Examples of allocation for two input regions with differing receptor densities and activation (see insets) for different bottleneck widths, demonstrating complex trade-offs in resource allocation. Here ratio for low:high density = 1:3, γ = 1. Dotted lines show representation proportional to receptor density (red) and convergence in the limit (yellow). Arrows indicate expansion (down) and contraction (up) of the higher density (blue) region. Both expansion and contraction of the higher density region is seen, similar to the 2D case. **B.** Allocation boundaries for different density ratios. Ratios are given as low:high density. **C** Effect of changing the decay γ parameter. Density ratio is set at 1:3 for all sigmas. Panels **D-F**, same as above but for density constant across the two regions and receptor activation manipulated. Activation ratio is set at 1:3 in panels D and F. Note that in panel D, the representation of the higher activation (blue) region is expanded for all bottlenecks.

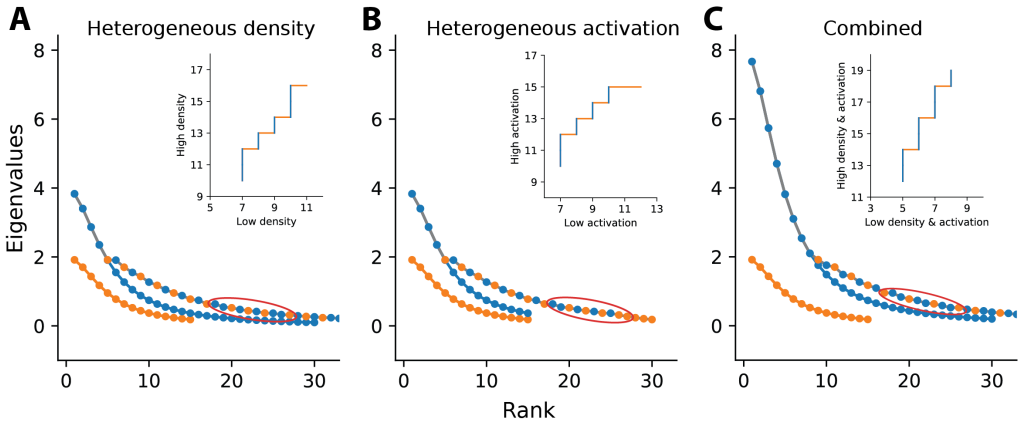

Figure 2–Figure supplement 3. Illustration of eigenvalue sorting and resulting allocation. Main panels: Eigenvalues for heterogeneous density (A), heterogeneous activation (B), or combined (C). For each, the manipulated ratio is set as 1:2. For simplicity, the example considers 1D regions. Orange markers show eigenvalues from the baseline region, which is identical for all panels. Blue markers show eigenvalues from the second region with higher density, activation, or both, respectively. Markers connected by the grey line show the combined set of sorted eigenvalues from both regions. Insets: Cumulative allocation for the set of eigenvalues marked by the red ellipse in the main panel. Horizontal orange lines indicate eigenvalues from and therefore allocation to the orange region, whilst vertical blue lines indicate allocation to the blue region. **A.** Heterogeneous density. Increased receptor density in the blue region causes higher spatial correlations between neighbouring receptors and therefore larger eigenvalues, leading to a scaling of the eigenvalue curve. Increased density also leads to a larger number of receptors in the blue region, increasing the number of eigenvalues from this region and the total number of eigenvalues considered. **B.** Heterogeneous activation. Increased variance in the blue region leads to a scaling of the eigenvalue curve. However, the number of receptors in the blue region is not affected, explaining differences observed compared to A. **C.** Heterogeneous density and activation. Effects from A and B are combined, leading to a double-scaling of the eigenvalue curve, plus an increase in the total number of eigenvalues.

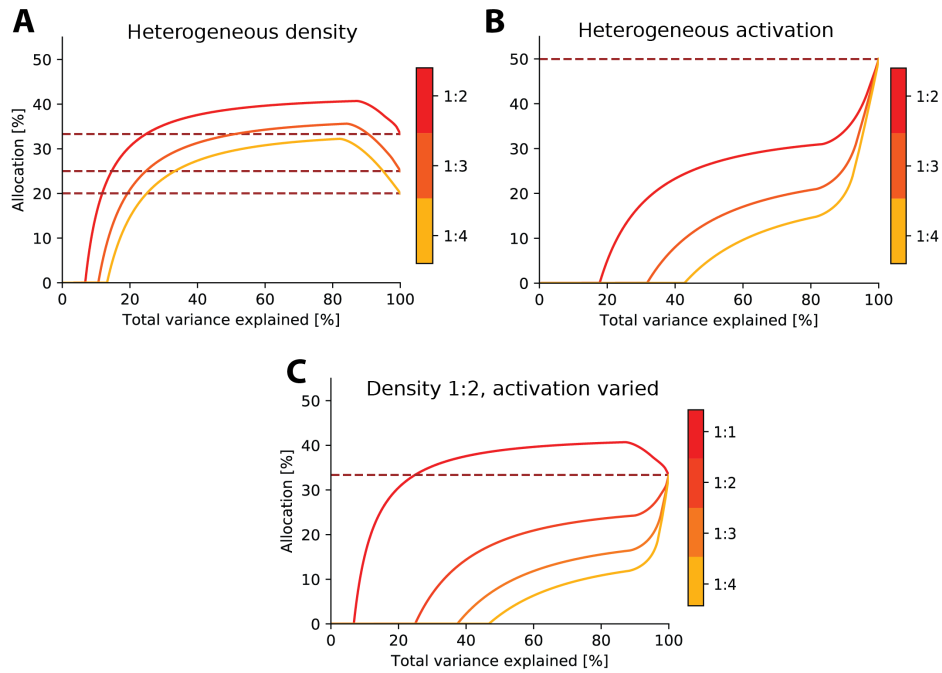

Figure 2–Figure supplement 4. Limit on information rather than number of neurons. Resource allocations where the bottleneck is expressed as variance explained. Because eigenvalues decrease dramatically in size, this re-expression results in a 'squashing' of the allocation curve towards higher bottleneck sizes, as the initial eigenvalues in the sorted set explain a much larger amount of the variance. **A.** Examples with heterogeneous density. **B.** Examples with heterogeneous activation. **C.** Examples with heterogeneous density and activation. Dashed red line in each denotes allocation proportional to density.

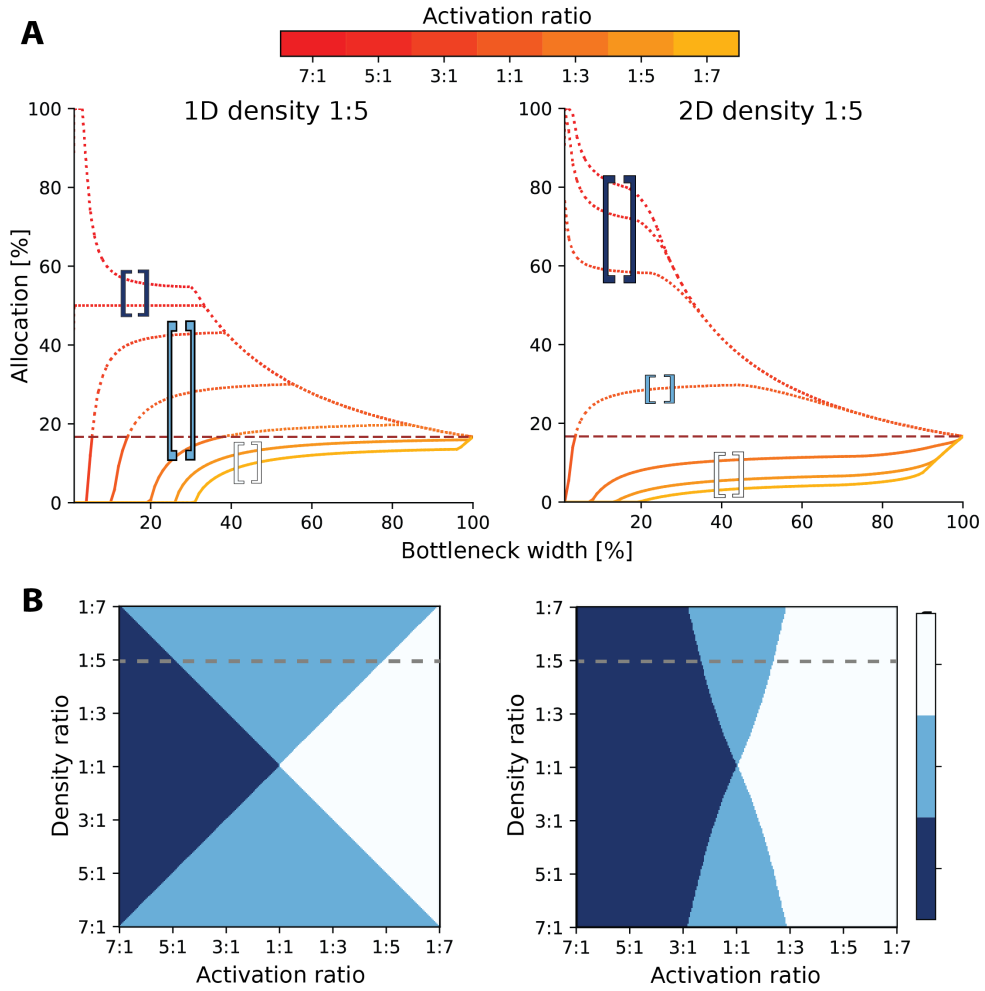

Figure 3–Figure supplement 1. Comparison between 1D and 2D results for heterogeneous activation and density. A. Effect of changing both the density and activation ratios, and possible resource allocations for two regions. Plots show the same density ratio, 1:5, considering 1D (left) and 2D (right) regions. Allocation % is for baseline region (region 1). Each line denotes the variance ratios between Region 1:Region 2 (see colorbar). Solid lines denote contraction of baseline region representation, whilst dotted lines are expansion. The horizontal red dashed line shows the proportional density representation. Across both 1D and 2D, different patterns of expansion and contraction occur. The baseline region is more likely to either be fully expanded or contracted for the whole bottleneck in 2D, whereas in 1D both expansion and contraction occur (denoted by light blue brackets, which covers three activation ratios in 1D, but only 1:1 in 2D as indicated by the brackets). In 1D, the same activation and density ratio leads to convergence at the proportional density, whereas for 2D, the convergence is always lower. **B.** Comparison of the possible allocation schemes between 1D (left) and 2D (right). Dotted grey line denotes density ratio from panel A. In 2D, the likelihood of both expansion and contraction (light blue region) of the representations decreases.

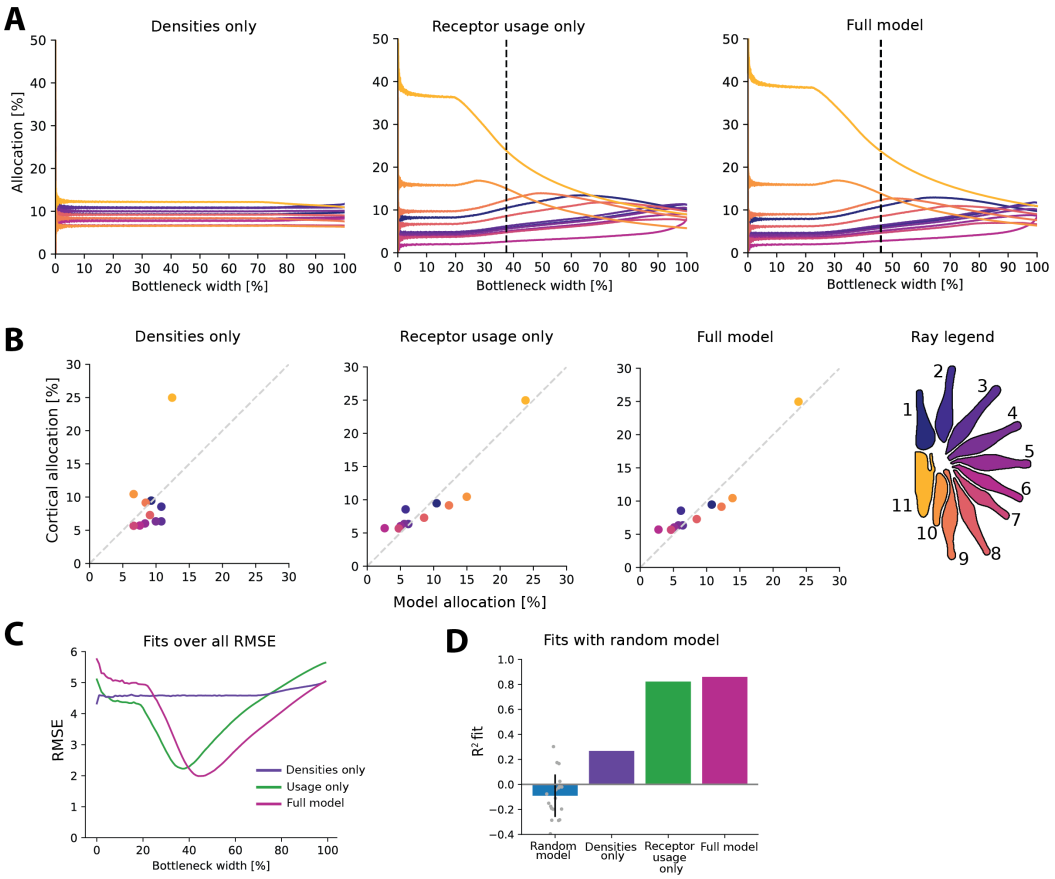

Figure 6–Figure supplement 1. Additional figures for the star-nosed mole. A. Fits across all bottleneck sizes for each model. Lowest RMSE is indicated for each model. Dashed line indicates lowest RMSE (excluded for Densities only model as RMSE is fairly constant across bottlenecks, see C). **B.** Corresponding fits at each selected bottleneck size, compared to cortical data. **C.** RMSE values over all bottleneck sizes for each model. **D.** Can any values be fit given the range of bottleneck sizes? Here we test whether a good fit can be obtained for a series of 'random models', where the random values are selected for each ray's density and variance within the range of parameters. The same fitting process is then applied, where the lowest RMSE fit to the cortical data is selected based on the range of possible bottlenecks. Note that R^2 values are calculated with fixed intercept, resulting in some negative values. Most random models perform poorly, suggesting that the model cannot be fit successfully without using the precise parameter values from the empirical star-nosed mole data.