AN EFFICIENT CODING ACCOUNT OF REPRESENTATIONAL DRIFT

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural responses to stimuli change continually over days even as perception and behavior remain stable, a phenomenon called representational drift. Drift has been modeled with unsupervised learning objectives that admit a degenerate solution space of representations that preserve input similarity structure. But such objectives are ill-suited for early sensory areas, where the code must allocate sensitivity according to input statistics rather than reproduce similarity structure. Here, we test the hypothesis that drift in these areas instead arises from mechanisms supporting efficient coding, formalized as maximizing Fisher information about the stimulus in the neural population. We show that a biologically plausible uncertainty-gated differential Hebbian plasticity rule, in which synaptic change depends on the time derivatives of pre- and post-synaptic activity, implements gradient ascent on Fisher information, and that adding synaptic noise produces substantial drift in individual neurons while population Fisher information remains stable. We confirm this prediction in longitudinal calcium imaging from the primary visual and auditory cortex in mice, where single-neuron tuning drifts over days while population Fisher information does not. These findings offer a novel, efficient coding account of drift in sensory cortex.

1 INTRODUCTION

Tuning of single neurons to sensory stimuli as well as abstract latent variables gradually changes over time. This phenomenon, called representational drift, has been reported in multiple brain regions, from sensory and associative cortices to the hippocampus (Bauer et al., 2024; Marks & Goard, 2021; Deitch et al., 2021; Schoonover et al., 2021; Zheng et al., 2026; Berners-Lee et al., 2023; Aschauer et al., 2022; Driscoll et al., 2017; Climer et al., 2025; Descamps et al., 2026; Ziv et al., 2013). Surprisingly, representational drift is seldom accompanied by behavioral changes. This has led to the view of representational drift as a diffusion process that explores a manifold of equally good solutions in the population activity space (Rule et al., 2019; Ratzon et al., 2024). However, the precise structure of this manifold and the biological mechanisms that cause representations to drift along it are yet to be fully worked out. Resolving this has important implications both for how downstream brain regions decode task variables from drifting representations and for the development of brain-machine interfaces that are robust to drift.

Changes in neural response naturally emerge in models that learn multiple tasks in sequence but are typically accompanied by catastrophic forgetting not seen in animal studies of representational drift. Therefore, recent studies have begun exploring representational drift in neural network models from the perspective of continual learning algorithms aimed at minimizing interference between tasks (van der Veldt et al., 2025). While some of those methods such as elastic weight consolidation indeed produce representational drift without affecting task performance, this usually comes at the cost of limiting later learning (Si & Qin, 2026; Kirkpatrick et al., 2017). Moreover, continual learning models are yet to provide concrete predictions that can be used to test whether drift in the brain arises from such mechanisms. A more promising line of work proposed instead that representational drift arises due to noisy synaptic updates in unsupervised learning models optimized for a similarity-matching objective (Qin et al., 2023). Cortical synapses are indeed highly dynamic, with spines turning over and synaptic strengths fluctuating over days even in stable environments (Ziv & Brenner, 2018; Holtmaat et al., 2005; Loewenstein et al., 2011). This account has some appealing aspects that make it a compelling explanation of representational drift. First,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

similarity-matching can be implemented using biologically plausible, Hebbian/anti-Hebbian plasticity (Pehlevan et al., 2018). Second, the theory makes a concrete prediction – that population-level representational similarity matrix is preserved despite drift in single neurons – that was tested and validated in experimental data from posterior parietal cortex and hippocampus (Qin et al., 2023).

But the mechanism underlying drift could vary depending on the function of the brain region, and the function of early sensory cortex, in particular, is very different from that of associative and subcortical regions. The sensory cortex transforms input representations into an expanded, high-dimensional neural space better suited for pattern separation or discrimination than for preserving the similarity structure of representations in the input (e.g. thalamus) (Babadi & Sompolinsky, 2014). Decades of works have shown that sensory responses are shaped by the statistics of the environment (Simoncelli & Olshausen, 2001; Benucci et al., 2013; Rabinowitz et al., 2011). Efficient coding theories formalize this as a resource allocation problem wherein a population with limited capacity should distribute its sensitivity across stimuli in proportion to how often they occur, devoting more coding resources to frequently encountered stimuli (Ganguli & Simoncelli, 2014; Wei & Stocker, 2016; Barlow, 1961; Laughlin, 1981; Wei & Stocker, 2015). Therefore, we asked whether an efficient coding objective can provide an alternative account of representational drift in the early sensory representation.

To address this, we first developed a two-layer network model in which efficient coding emerges from a biologically plausible, uncertainty-gated differential Hebbian plasticity rule which allocates Fisher information across stimuli in accordance with stimulus statistics. We demonstrated that in the presence of noisy synaptic updates, this model drives continual reorganization of single-neuron responses while leaving the population's Fisher information profile unchanged. Finally, we analyzed longitudinal neural recordings from the primary visual and auditory cortex of mice and found that they were well explained by the efficient coding account. Together, our findings offered new mechanistic insights into representational drift in the sensory cortex.

2 METHODS

Let stimulus $s \in \mathcal{S}$ be drawn from a prior $p(s)$ and let a fixed input population of N neurons represent it with tuning curves $\mathbf{x}(s)$. Suppose an output population of M neurons responds linearly through feedforward weights W, with additive noise:

$$\mathbf{y} = \boldsymbol{\mu}(s) + \boldsymbol{\xi}, \qquad \boldsymbol{\xi} \sim \mathcal{N}(0, \sigma^2 I),$$

where $\boldsymbol{\mu}(s) = W\mathbf{x}(s)$ denotes the tuning curves of the output population and $\mathbf{y}$ its activity on a single trial. The Fisher information about s in the population response $\mathbf{y}$ is

$$J(s; W) \;=\; \frac{1}{\sigma^2}\big\|\boldsymbol{\mu}'(s)\big\|^2 \;=\; \frac{1}{\sigma^2}\,\mathbf{x}'(s)^\top W^\top W\,\mathbf{x}'(s), \tag{1}$$

where primes denote differentiation with respect to s. For simplicity, we consider the linear case here. Corresponding results for nonlinear output neurons are included in Appendix A.3 and validated using simulations shown in Appendix A.4.

2.1 OBJECTIVE AND RESOURCE CONSTRAINT

The efficient coding principle states that a sensory population with limited resources should transmit as much information as possible about the stimuli it encounters, which requires devoting greater discriminability to stimuli that occur more frequently (Barlow, 1961; Laughlin, 1981). Fisher information equation 1 bounds the variance of any unbiased estimate of the stimulus and therefore measures how finely a population distinguishes nearby stimulus values (Seung & Sompolinsky, 1993; Brunel & Nadal, 1998; Wei & Stocker, 2016). When the population is large enough that the posterior $p(s \mid \mathbf{y})$ is approximately Gaussian with variance $J(s)^{-1}$, the transmitted information is determined by the expectation of log Fisher information (Brunel & Nadal, 1998),

$$I(s; \mathbf{y}) \;=\; H(s) \;+\; \tfrac{1}{2}\,\mathbb{E}_{s\sim p}\big[\log J(s)\big] \;-\; \tfrac{1}{2}\log 2\pi e \;+\; o(1),$$

such that maximizing transmitted information amounts to maximizing

$$\mathcal{L}(W) \;=\; \mathbb{E}_{s\sim p}\big[\log J(s;W)\big] \tag{2}$$

subject to a constraint on coding resources. Here, we take that constraint to be a fixed total excitatory input weight for each output neuron,

$$W \;\geq\; 0, \qquad \sum_j W_{ij} = 1 \quad \text{for all } i, \tag{3}$$

which restricts each row of W to a unit simplex. Constraint 3 is motivated by experimental evidence for homeostatic regulation of total synaptic input whereby cortical neurons multiplicatively rescale their excitatory synapses to maintain a target level of drive while preserving relative strengths (Turrigiano & Nelson, 2004), and potentiation of one input is accompanied by depression of others on the same cell (Chistiakova et al., 2014). Maximizing equation 2 subject to equation 3 gives (Appendix A.1-A.2):

$$J^\star(s) \;\propto\; p(s)^2, \tag{4}$$

matching the allocation obtained under a constraint on the number of neurons (Ganguli & Simoncelli, 2014).

2.2 Plasticity Rule

Differentiating equation 2 with respect to W and using $\boldsymbol{\mu}' = W\mathbf{x}'$ gives $\nabla_W \log J(s) = 2\sigma^{-2} J(s)^{-1}\boldsymbol{\mu}'(s)\mathbf{x}'(s)^\top$. Because the output noise is zero-mean and independent of the stimulus, $\mathbb{E}[\mathbf{y}'\mathbf{x}'^\top \mid s] = \boldsymbol{\mu}'(s)\mathbf{x}'(s)^\top$, so stochastic gradient ascent with $s \sim p$ can be written directly in terms of neural activity:

$$\Delta W \;=\; \eta\, U(s)\, \mathbf{y}'\, \mathbf{x}'^\top, \qquad U(s) \;=\; J(s)^{-1}, \tag{5}$$

subject to constraint equation 3, with constants absorbed into the learning rate η. The update features an outer product of presynaptic and postsynaptic derivatives, a form of differential Hebbian learning rule. Here, it is additionally gated by a stimulus-dependent global scalar signal $U(s)$ to achieve efficient coding.

2.3 Biological Plausibility

The plasticity rule in equation 5 requires only three quantities: the presynaptic and postsynaptic activities and a scalar $U(s)$, and we consider whether each is available to a synapse. First, by the Cramér–Rao bound, $U(s)$ is the minimum variance of an unbiased estimate of s from $\mathbf{y}$, and could be available in a downstream region that decodes s. The third factor is therefore a scalar uncertainty signal of the kind such an area could broadcast globally to the output neural population. Neuromodulators like acetylcholine have been proposed to signal uncertainty (Yu & Dayan, 2005), and pairing basal forebrain stimulation with a stimulus reorganizes sensory cortical maps (Kilgard & Merzenich, 1998), indicating that global neuromodulatory signals can gate plasticity in a way that reallocates representational resources across stimuli.

The remaining quantities are derivatives with respect to the stimulus, which may not be available at the synapse. However, since natural stimuli vary smoothly in time, the same update rule can also be implemented using derivatives of pre- and post-synaptic activities with respect to time, quantities that can be computed locally by high-pass filtering. This resembles the differential Hebbian learning rule introduced by Kosko (1986) on the grounds that concurrent change rather than concurrent activation captures functional dependence between units. Let the stimulus follow a trajectory s_t that is differentiable in time, so that $\dot{\mathbf{x}}_t = \mathbf{x}'(s_t)\dot{s}_t$ and $\dot{\mathbf{y}}_t = \mathbf{y}'(s_t)\dot{s}_t$. In simulations we use Langevin dynamics, $ds_t = v_t\,dt$ and $dv_t = [c\,\partial_s \log p(s_t) - \gamma v_t]\,dt + \sqrt{2\gamma c}\,dB_t$ where B_t is a standard Wiener process. This yields a stimulus that varies smoothly in time and has stationary density $p(s)$, so that $\dot{s}_t$ is well defined. Then:

$$\mathbb{E}\big[\dot{\mathbf{y}}_t\,\dot{\mathbf{x}}_t^\top \,\big|\, s_t = s\big] \;=\; \mathbb{E}\big[\dot{s}^2 \,\big|\, s\big]\;\boldsymbol{\mu}'(s)\,\mathbf{x}'(s)^\top.$$

Assuming $\mathbb{E}[\dot{s}^2 \mid s] = v^2$ is independent of s, the update

$$\Delta W \;=\; \eta\, U_t\,\dot{\mathbf{y}}_t\,\dot{\mathbf{x}}_t^\top \tag{6}$$

where $U_t \equiv U(s_t)$, therefore has the same effect, on average, as equation 5 with v^2 absorbed into the learning rate.

2.4 Synaptic noise and drift

We model stochastic synapses by adding Gaussian noise to equation equation 6,

$$\Delta W \;=\; \eta\, U_t\,\dot{\mathbf{y}}_t\,\dot{\mathbf{x}}_t^\top \;+\; \sigma_w\sqrt{\eta}\,E_t, \tag{7}$$

subject to constraint equation 3, where the entries of E_t are independent standard Gaussian variables and σ_w sets the amplitude of the fluctuations. The scaling by $\sqrt{\eta}$ makes the accumulated noise independent of the step size. In this case, the weights do not converge to a single fixed point but move between configurations that maximize $\mathcal{L}$. In the unconstrained case, there are numerous such configurations because from equation equation 1, $\mathcal{L}$ depends on W only through $W^\top W$, so $\mathcal{L}(OW) = \mathcal{L}(W)$ for any orthogonal O even if the tuning of individual neurons $\mu_i(s) = \mathbf{w}_i^\top \mathbf{x}(s)$ changes under such transformations. This yields a family of solutions in which the first term in equation 7 vanishes, so single-neuron tuning drifts due to weight diffusion while $J(s)$ is unchanged. Since rotation violates non-negativity, constraint equation 3 breaks this symmetry. Nonetheless, similar to Qin et al. (2023), many weight configurations give rise to the same allocation of Fisher information since J is determined by the arrangement of tuning curves rather than by the identity of the neurons. Therefore, we observe representational drift in simulations of learning dynamics in equation 7 even when W is constrained according to equation equation 3.

2.5 Comparison with similarity matching

A leading account of drift takes the representational objective to be preservation of input similarity structure, minimizing $\|X^\top X - Y^\top Y\|_F^2$ over the output representation (Qin et al., 2023; Pehlevan et al., 2018). This objective is also satisfied by a continuum of solutions related by orthogonal transformations, and thus produces drift under noisy learning. However, under similarity matching the output Gram matrix is fixed to that of the input, so the allocation of Fisher information is inherited from the input representation, whereas under equation 2 it reallocates Fisher information to track the stimulus distribution as shown in equation 4.

2.6 Neural Datasets

We analyzed two longitudinal two-photon calcium-imaging datasets from mouse auditory (Aschauer et al., 2022) and primary visual cortex (Bauer et al., 2024). In both sets of experiments, head-fixed animals passively viewed or listened to the same stimulus set on every recording day, and the same neurons were tracked across days.

Auditory cortex. 12 mice were presented with 19 pure tones spanning 2–45 kHz (50 ms duration, $\geq$20 repetitions per tone) during imaging of layer 2/3. Each animal was imaged in 8 ± 5 fields of view containing 222 ± 41 tracked neurons each, giving $1{,}792 \pm 1{,}217$ neurons per animal after pooling across fields. All animals were recorded on days 1, 3, 5 and 7.

Visual cortex. 6 mice were presented with drifting sinusoidal gratings in 12 directions spaced 30° apart (0.04 cycles per degree, 3 Hz, 5 s duration, $\sim$32 repetitions per direction), with 186 ± 178 tracked neurons per animal. Each animal was recorded on 13 ± 1 sessions spanning 26 ± 3 elapsed days, with consecutive recordings 2 ± 0.2 days apart.

2.7 Simulations and Data Analysis

We simulated networks for 10^6 learning steps with synaptic noise, discarding the first 10^5 steps allowing for the learning to converge. In the input layer, stimuli were represented by 50 units whose response was given by $x_j(s) \propto \exp[\kappa_{\text{in}} \cos 2(s - \phi_j)]$, with evenly spaced preferred stimuli ϕ_j and $\kappa_{\text{in}} = 8$ setting the input tuning width. These inputs projected to a population of 100 noisy output units through feedforward weights W, yielding response $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}(s), \sigma^2 I)$, where $\boldsymbol{\mu}(s) = W\mathbf{x}(s)$ and $\sigma = 0.25$. The uncertainty gate U was computed from the normalized likelihood $p(\mathbf{y} \mid s)$ evaluated under the network's own tuning curves $\boldsymbol{\mu}(s)$, so that U approximates $1/J(s)$ by the Cramér–Rao bound. The network therefore had no explicit access to the prior $p(s)$ or to a target information profile $J^\star(s)$, and learned only from the stimuli it happened to encounter. All statistics are reported as mean $\pm$ SD across 10 separate runs. For the mouse data, our analyses used only neurons tracked across all sessions of a given animal, with responses taken as trial-averaged change in fluorescence ($\Delta F/F$) over the stimulus window. Mean responses $\mu(s_k)$ and the residual covariance $\Sigma(s_k)$ were estimated separately within each session. We performed a set of three analyses both on the simulated model responses and neural responses from mice. In each case we compared a pair of time points a and b, where time points correspond to number of learning updates (for simulations) or the number of days (for mouse data), and quantified stability as the Pearson correlation between the quantity of interest measured at a and at b.

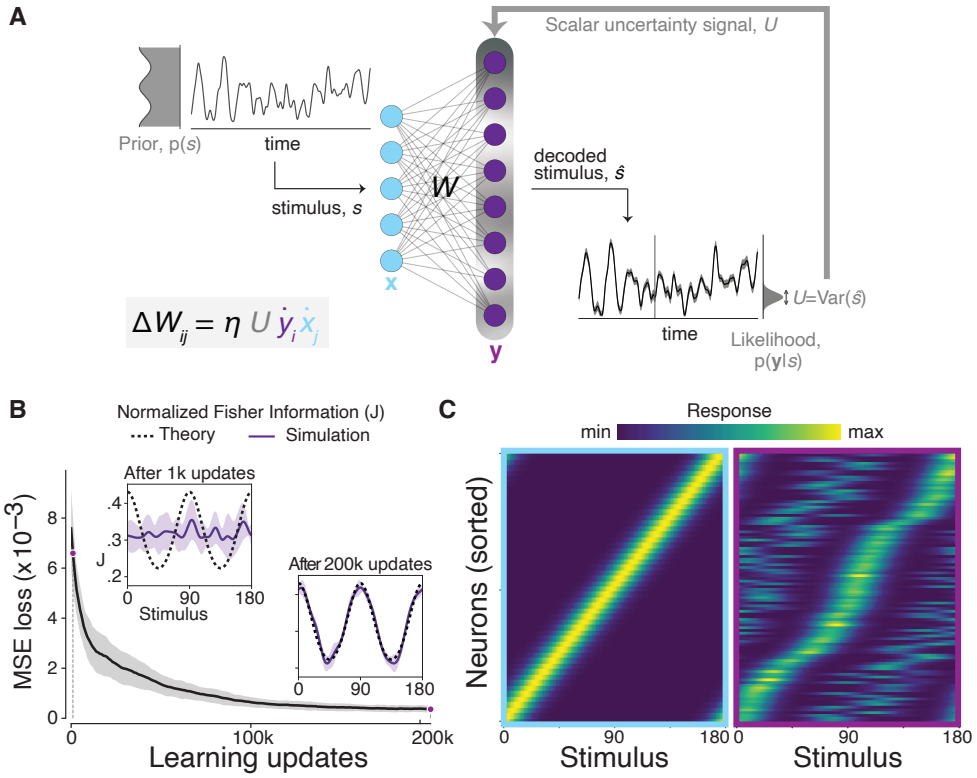

Figure 1: Theory vs simulation. (A) Model schematic. (B) Learning curve (black) and normalized Fisher information profile of the simulated output units before and after learning (inset). Dashed lines denote theory (equation 4). (C) Heatmap of tuning curves in the input (left) and output (right) populations after learning. Units have been sorted by preferred stimuli. See Appendix Figure A1.

Representational drift Drift was defined as the loss of similarity between a neuron's tuning curve at two time points, $r_i(a, b) = \text{corr}_k\big[\mu_i^{(a)}(s_k), \mu_i^{(b)}(s_k)\big]$. For visualizing the tuning heatmaps, simulated neurons were ordered by preferred stimulus i.e., the one evoking the largest mean response (Fig. 2A). Recorded neurons were included if their peak response in the reference session exceeded

their mean response across stimuli by at least 2 SD, and were normalized and ordered by preferred stimulus in that session (Fig. 3A,C).

Fisher information For simulations, Fisher information was computed from Eq. equation 1 using central differences of the learned $\boldsymbol{\mu}(s)$ across the stimulus grid and σ was a known constant. In the recordings, $J(s)$ was estimated as $\boldsymbol{\mu}'(s)^\top \Sigma^{-1} \boldsymbol{\mu}'(s)$ using the session's own residual covariance Σ. Because the number of tracked neurons exceeded the number of trials, Σ was inverted within the leading principal subspace accounting for 90% of the residual variance. Stability of the Fisher information profile was given by the correlation between $J^{(a)}(s)$ and $J^{(b)}(s)$ across stimuli.

Representational similarity Representational (dis-)similarity at each time point was estimated as the matrix of pairwise Euclidean distances between population responses, $D_{kl} = \|\boldsymbol{\mu}(s_k) - \boldsymbol{\mu}(s_l)\|_2$. Stability was the correlation between the off-diagonal entries of $D^{(a)}$ and $D^{(b)}$.

3 Results

We tested the theory by simulating the model shown in Figure 1A. Stimuli $s \in [0, \pi)$ were drawn from a prior $p(s)$ comprising a mixture of a uniform density and two von Mises components $\mathcal{V}(s)$ peaking at cardinal angles so that $p(s) = (1-\epsilon)/\pi + \epsilon/2\big[\mathcal{V}(s;0,\kappa) + \mathcal{V}(s;\pi/2,\kappa)\big]$, where $\mathcal{V}(s;\mu,\kappa) \propto \exp(\kappa \cos 2(s-\mu))$ is the von Mises density, $\epsilon = 0.4$ is the fraction of probability mass carried by the von Mises components and $\kappa = 1.5$ their concentration.

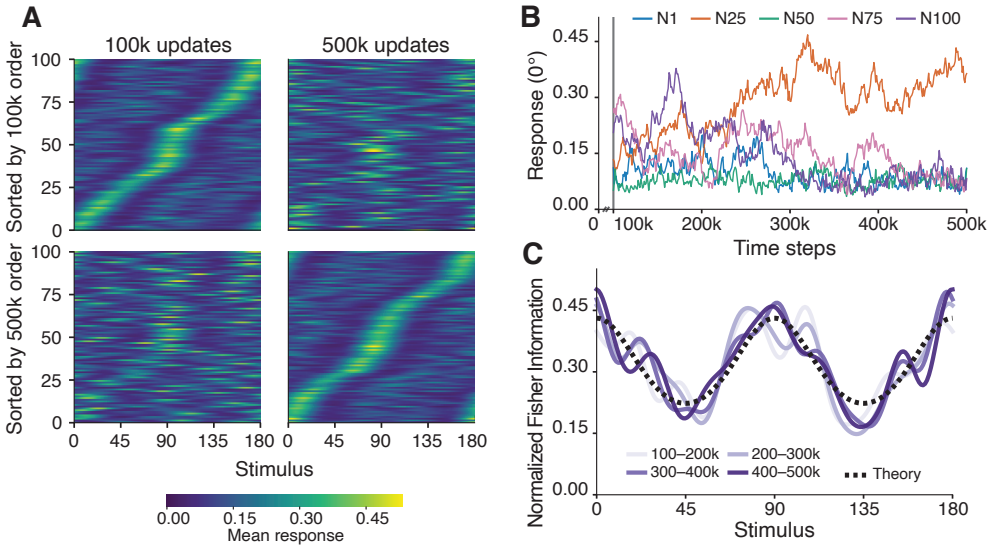

Figure 2: Model drift dynamics. (A) Tuning curves of the output units at two points in learning (columns), with neurons sorted by preferred stimuli at each of those points (rows). (B) Responses of five example output units to one of the stimuli, across time. (C) Normalized Fisher information of the output population at four different time points.

We first tested whether the uncertainty-gated differential Hebbian plasticity rule (Equation 5) can achieve efficient coding by initializing the model with random weights W and estimating the Fisher information (FI) profile (Equation 1) from the learned output response $\mathbf{y}$. We found that learning increased population level FI and reshaped the FI profile from roughly uniform early in learning to match the theoretically optimal profile upon convergence (Figure 1B). Regressing $\log J$ on $\log p$ across the stimulus range gave an exponent of 2.03 ± 0.17 (SD) over 10 independent runs, consistent with the predicted $J^\star \propto p(s)^2$. In contrast to input units which tiled stimuli uniformly, output neurons developed preferred stimuli concentrated near the cardinal angles (Figure 1C). We found very similar results under two variants that relaxed simplifying model assumptions. Replacing the linear output with saturating nonlinearity, $\boldsymbol{\mu}(s) = \tanh(W\mathbf{x}(s))$, and applying the corresponding

gain-modulated rule equation 13 gave a fitted exponent of 1.81 ± 0.21 (SD) (Figure A1A). Likewise, when the stimulus followed a Langevin trajectory with friction $\gamma = 1$ and velocity variance $c = 1$ (Methods 2.3), replacing stimulus derivatives with time derivatives (Equation 6) gave an exponent of 2.01 ± 0.24 (SD) (Figure A1B).

To test whether the model recapitulates representational drift, we modified the plasticity rule by adding i.i.d. Gaussian noise to every weight update according to equation 7. We found that tuning of the output units continued to change substantially even after the learning converged. Sorting units by preferred stimulus at any given time point in learning revealed a structure that was absent when the same ordering was applied to earlier or later time points (Figure 2A). Tracking single-unit responses to a fixed stimulus revealed continuous fluctuations over the course of learning (Figure 2B). Tuning similarity fell to 0.59 ± 0.01 (SD) after 10^5 noisy update steps, dropping further to 0.07 ± 0.04 (SD) by 9×10^5 steps.

Despite drift in neural tunings, population Fisher information profile was stable and approximately matched the theoretical optimum throughout learning (Figure 2C), with a correlation of 0.71 ± 0.12 (SD) even after 9×10^5 update steps. The model therefore predicts that the Fisher information profile should be conserved in early sensory cortex, even as single-neuron tuning changes.

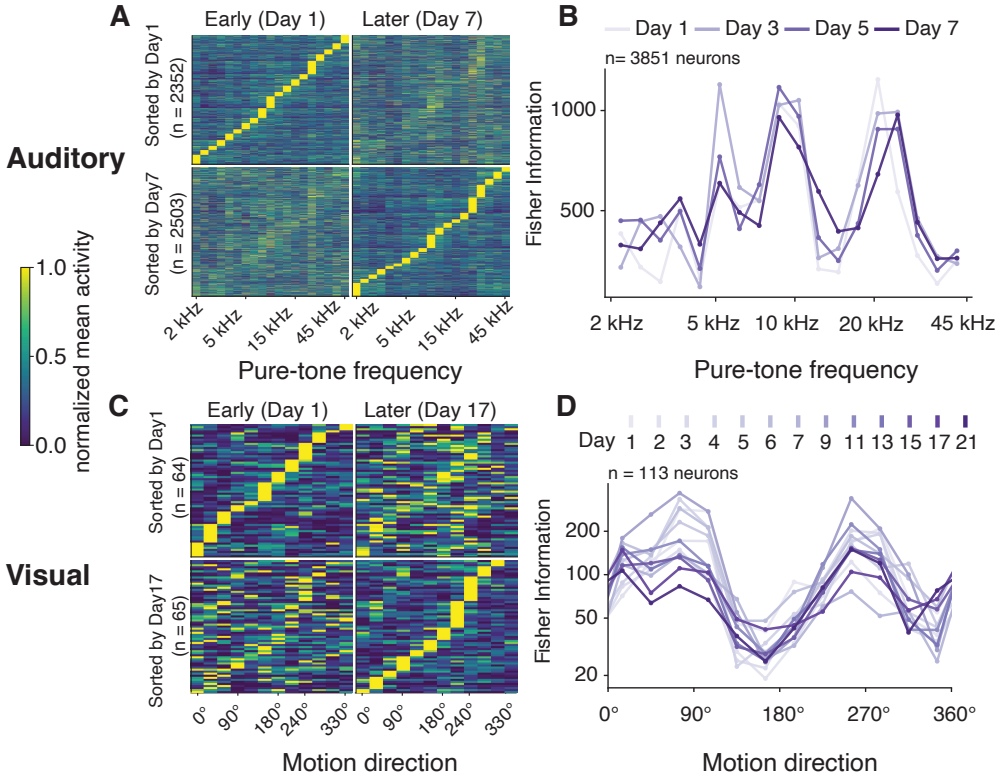

Figure 3: Representational drift in mouse data. (A) Tuning curves of auditory cortex neurons from one example mouse on two different days (columns), with neurons sorted by preferred frequency at each of those days (rows). (B) Population Fisher information profiles of the same mouse on different days. (C-D) Similar to A-B, but for neurons in the primary visual cortex of an example mouse sorted by preferred orientation. See Appendix Figures A2-A5 for all animals.

We tested this prediction in two longitudinal calcium-imaging datasets that identified and tracked the same set of neurons across days as described in the Methods 2.6 (Aschauer et al., 2022; Bauer et al., 2024). We found that single neurons drifted in both areas (Figure 3A,C; see Figure A2-A3 for all animals). Sorting neurons by preferred stimulus on one day revealed a structure that was largely absent when the same ordering was applied to recordings from a later day. Drift was faster in auditory cortex, where tuning similarity dropped to 0.30 ± 0.03 (SEM; 12 mice) following a two-day gap and to 0.26 ± 0.03 after six days. Drift in the visual cortex was more modest, with similarity

falling to 0.87 ± 0.01 (SEM; 6 mice) after a single session (median 2 days) and to 0.65 ± 0.04 (SEM) after 11 sessions (median 23 days).

The Fisher information profile was strongly non-uniform in both areas, with substantial heterogeneity across animals (Figure 3B,D; see Figure A4-A5 for all animals). Notably however, this structure was preserved across days. Fisher information profile similarity across sessions was 0.58 ± 0.09 (SEM; 12 mice, 72 session pairs) in auditory cortex and 0.70 ± 0.06 (SEM; 6 mice, 497 pairs) in visual cortex. Shuffling stimulus labels within session destroyed this correspondence (median $\rho \approx 0$), and stability exceeded the shuffled distribution in $51/72$ auditory and $409/497$ visual session pairs (permutation test, Benjamini–Hochberg corrected, FDR ≤ 0.05).

Preserved population structure under drift is also predicted by the similarity-matching account, which predicts that drift instead conserves the representational similarity between mean stimulus responses (Qin et al., 2023). To ask whether the two accounts can be distinguished, we trained a similarity-matching network with Hebbian feedforward and anti-Hebbian lateral plasticity on the same inputs and prior, with Gaussian noise added to its feedforward and lateral weights on every update, and measured both the representational similarity matrix and the Fisher information profile in each model. Both models drifted, and both conserved representational similarity almost perfectly (Figure 4A,B). In contrast, the Fisher information profile was stable in the efficient coding model across intervals up to 9×10^5 steps ($\rho = 0.80 \pm 0.02$, SD), but not in the similarity-matching model ($\rho = 0.014 \pm 0.04$, SD).

The recordings show the pattern produced by the efficient coding model. In both areas, representational similarity and Fisher information were preserved to a comparable degree across days (Mean $\pm$ SEM): 0.69 ± 0.07 and 0.58 ± 0.09 respectively in auditory cortex, 0.87 ± 0.03 and 0.70 ± 0.06 respectively in visual cortex, with the Fisher information profile in both areas far from the near-zero stability produced by similarity matching (Figure 4C,D; see Figure A6-A7 for representational similarity for all animals).

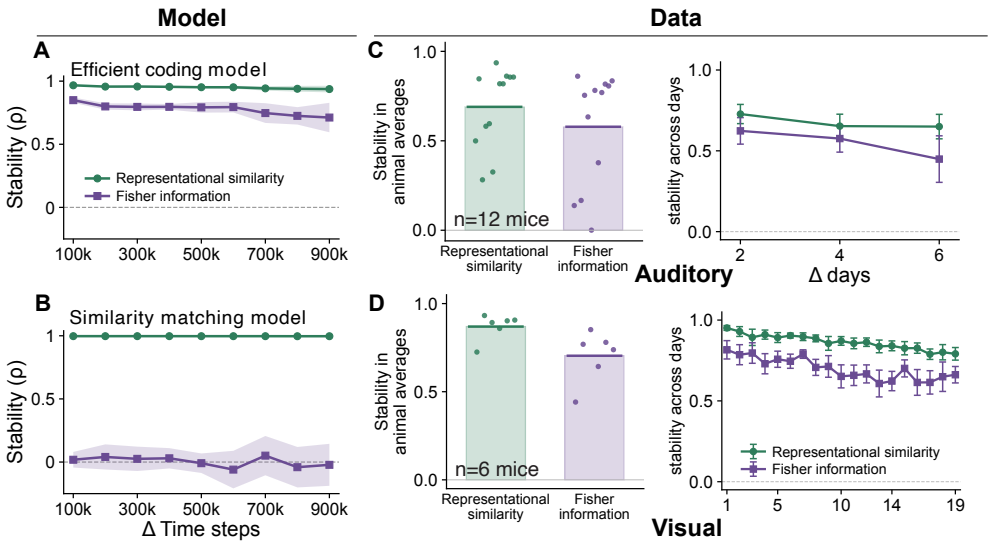

Figure 4: Stability of representational similarity and Fisher information. (A) Efficient coding model, across learning. (B) Similarity matching model, across learning. (C) Auditory cortex of individual mice (left: averaged over all session pairs; right: as a function of the gap between recordings). (D) Similar to C but for Visual cortex.

4 Discussion

The ubiquity of representational drift across neural systems has generated widespread interest in the neural mechanisms that produce it (Rule et al., 2019; Driscoll et al., 2022). Recent characterization of this phenomenon in animals has shown that drift is coordinated across neurons rather than

independent, prompting the development of theories in which it emerges from noisy synaptic updates that allow neural representation to explore the degenerate solution space of a computational objective. A leading account proposes that those solutions correspond to the ones that perform similarity matching i.e., preserve the similarity structure of the input population (Qin et al., 2023). Here, we developed and tested an alternative in which drift arises under a plasticity rule that implements efficient coding, providing a normative theory of drift in sensory cortex.

The rule takes the form of differential Hebbian plasticity, but differs from previous rules of this type in two respects. It combines the temporal derivatives of both pre- and postsynaptic activity, whereas rate-based reductions of spike-timing-dependent plasticity yield rules that are differential only in the postsynaptic factor, $\Delta W_{ij} \propto \dot{\nu}_i \nu_j$ (Roberts, 1999; Xie & Seung, 2000). The form used here was originally proposed by Kosko (1986) as a heuristic to capture concurrent change between units but here it emerges as a solution for performing gradient ascent on Fisher information. Second, the differential Hebbian term is gated by a scalar uncertainty signal which is essential to reallocate coding resources according to stimulus statistics. This gives a concrete role to diffuse neuromodulatory inputs in shaping sensory representations. Other related mechanistic models maximizing Fisher information have largely focused on recurrent spiking networks to derive optimal connectivity (Toyoizumi et al., 2006) and more recently, for synapses with short-term dynamics (Shimizu & Toyoizumi, 2026), whereas our work is aimed at finding the optimum through local plasticity.

In longitudinal recordings from mouse visual and auditory cortex, single-neuron tuning drifted while the Fisher information profile was preserved, consistent with the model prediction. Moreover, stable representational similarity, taken as evidence for a similarity-matching objective, also emerged in our model without being imposed as an objective. The present study has some limitations, however. First, although we estimated Fisher information in the neural recordings using the measured noise covariance, the plasticity rule was derived assuming independent noise. Extending the model to accommodate correlated noise would reduce the gap between theory and data. Second, while a stable Fisher information profile is consistent with efficient coding, we did not show that the mouse brain maximizes it. A direct test would require experimentally manipulating the stimulus distribution and analyzing whether the population Fisher information profile adapts to it. This limitation is not unique to our study. A direct test of the similarity-matching objective would require simultaneous neural recordings from input and output populations rather than the stability of output representational similarity matrix demonstrated by Qin et al. (2023). Experiments of this kind are needed to fully disentangle the two accounts. A further simplification concerns the uncertainty signal itself. In the model, U was computed from the normalized likelihood under the network's tuning curves, which a downstream readout could not access directly. In practice, it would have to be estimated from single-trial population responses. Such estimates are available in principle. Probabilistic population codes represent the uncertainty of a stimulus estimate in the population response itself, from which it can be read out by a downstream area (Ma et al., 2006), and neural correlates of decision confidence have been identified in higher brain regions (Kiani & Shadlen, 2009; Kepecs et al., 2008). Finally, our account treats drift as a byproduct of synaptic noise rather than as serving a function, in contrast to proposals that drift supports continual learning or robustness to perturbation (Rule et al., 2019; Aitken et al., 2022). While the present results are agnostic to the functional role of drift, theories that focus on the functional role of drift might offer a complementary perspective to understand this phenomenon. More broadly, representational drift offers a unique window into neural computation. Exploring other candidate objectives could identify new population-level metrics that remain stable despite drift in single neurons, and reveal the computations performed by different brain circuits.

AI USE STATEMENT

In this work, we used generative AI tools for checking mathematical proofs, to provide critical ingredients for proving the mathematical result in appendix A.2, and assisting with model simulations and data analysis. It was not used to generate research ideas, to design or provide feedback on research methodology or experiments, to generate synthetic data sets, to help develop theoretical models or conceptual frameworks, to design the model and analyses methods, or to interpret results. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Kyle Aitken, Marina Garrett, Shawn Olsen, and Stefan Mihalas. The geometry of representational drift in natural and artificial neural networks. *PLOS Computational Biology*, 18(11):e1010716, November 2022. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1010716.

Dominik F. Aschauer, Jens-Bastian Eppler, Luke Ewig, Anna R. Chambers, Christoph Pokorny, Matthias Kaschube, and Simon Rumpel. Learning-induced biases in the ongoing dynamics of sensory representations predict stimulus generalization. *Cell Reports*, 38(6):110340, February 2022. ISSN 22111247. doi: 10.1016/j.celrep.2022.110340.

Baktash Babadi and Haim Sompolinsky. Sparseness and Expansion in Sensory Representations. *Neuron*, 83(5):1213–1226, September 2014. ISSN 08966273. doi: 10.1016/j.neuron.2014.07.035.

H. B. Barlow. Possible Principles Underlying the Transformations of Sensory Messages. In Walter A. Rosenblith (ed.), *Sensory Communication*, pp. 216–234. The MIT Press, 1961. ISBN 978-0-262-51842-0. doi: 10.7551/mitpress/9780262518420.003.0013.

Joel Bauer, Uwe Lewin, Elizabeth Herbert, Julijana Gjorgjieva, Carl E. Schoonover, Andrew J. P. Fink, Tobias Rose, Tobias Bonhoeffer, and Mark Hübener. Sensory experience steers representational drift in mouse visual cortex. *Nature Communications*, 15(1):9153, October 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-53326-x.

Andrea Benucci, Aman B Saleem, and Matteo Carandini. Adaptation maintains population homeostasis in primary visual cortex. *Nature Neuroscience*, 16(6):724–729, June 2013. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.3382.

Alice Berners-Lee, Elizabeth Shtrahman, Julien Grimaud, and Venkatesh N. Murthy. Experience-dependent evolution of odor mixture representations in piriform cortex. *PLOS Biology*, 21(4):e3002086, April 2023. ISSN 1545-7885. doi: 10.1371/journal.pbio.3002086.

Nicolas Brunel and Jean-Pierre Nadal. Mutual Information, Fisher Information, and Population Coding. *Neural Computation*, 10(7):1731–1757, October 1998. ISSN 0899-7667, 1530-888X. doi: 10.1162/089976698300017115.

Marina Chistiakova, Nicholas M. Bannon, Maxim Bazhenov, and Maxim Volgushev. Heterosynaptic Plasticity: Multiple Mechanisms and Multiple Roles. *The Neuroscientist*, 20(5):483–498, October 2014. ISSN 1073-8584. doi: 10.1177/1073858414529829.

Jason R. Climer, Heydar Davoudi, Jun Young Oh, and Daniel A. Dombeck. Hippocampal representations drift in stable multisensory environments. *Nature*, 645(8080):457–465, September 2025. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-025-09245-y.

Daniel Deitch, Alon Rubin, and Yaniv Ziv. Representational drift in the mouse visual cortex. *Current Biology*, 31(19):4327–4339.e6, October 2021. ISSN 09609822. doi: 10.1016/j.cub.2021.07.062.

Lucie A.L. Descamps, Wesley P. Clawson, Miguel M. Carvalho, Thomas Rogerson, Omer Hazon, Oscar M.T. Chadney, Mark J. Schnitzer, and Clifford Kentros. Heterogeneous single-cell dynamics support stable population codes for objects in the mouse anterior cingulate cortex. *Cell Reports*, 45(9):117890, September 2026. ISSN 22111247. doi: 10.1016/j.celrep.2026.117890.

Laura N. Driscoll, Noah L. Pettit, Matthias Minderer, Selmaan N. Chettih, and Christopher D. Harvey. Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell*, 170(5):986–999.e16, August 2017. ISSN 00928674. doi: 10.1016/j.cell.2017.07.021.

Laura N. Driscoll, Lea Duncker, and Christopher D. Harvey. Representational drift: Emerging theories for continual learning and experimental future directions. *Current Opinion in Neurobiology*, 76:102609, October 2022. ISSN 09594388. doi: 10.1016/j.conb.2022.102609.

Deep Ganguli and Eero P. Simoncelli. Efficient Sensory Encoding and Bayesian Inference with Heterogeneous Neural Populations. *Neural Computation*, 26(10):2103–2134, October 2014. ISSN 0899-7667, 1530-888X. doi: 10.1162/NECO_a_00638.

Anthony J.G.D. Holtmaat, Joshua T. Trachtenberg, Linda Wilbrecht, Gordon M. Shepherd, Xiaoqun Zhang, Graham W. Knott, and Karel Svoboda. Transient and Persistent Dendritic Spines in the Neocortex In Vivo. *Neuron*, 45(2):279–291, January 2005. ISSN 08966273. doi: 10.1016/j.neuron.2005.01.003.

Adam Kepecs, Naoshige Uchida, Hatim A. Zariwala, and Zachary F. Mainen. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–231, 2008. doi: 10.1038/nature07200.

Roozbeh Kiani and Michael N. Shadlen. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, 324(5928):759–764, 2009. doi: 10.1126/science.1169405.

Michael P. Kilgard and Michael M. Merzenich. Cortical Map Reorganization Enabled by Nucleus Basalis Activity. *Science*, 279(5357):1714–1718, March 1998. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.279.5357.1714.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, March 2017. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1611835114.

Bart Kosko. Differential Hebbian learning. In *AIP Conference Proceedings*, volume 151, pp. 277–282. AIP, 1986. doi: 10.1063/1.36225.

Simon Laughlin. A Simple Coding Procedure Enhances a Neuron's Information Capacity. *Zeitschrift für Naturforschung C*, 36(9-10):910–912, October 1981. ISSN 1865-7125, 0939-5075. doi: 10.1515/znc-1981-9-1040.

Y. Loewenstein, A. Kuras, and S. Rumpel. Multiplicative Dynamics Underlie the Emergence of the Log-Normal Distribution of Spine Sizes in the Neocortex In Vivo. *Journal of Neuroscience*, 31 (26):9481–9488, June 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.6130-10.2011.

Wei Ji Ma, Jeffrey M. Beck, Peter E. Latham, and Alexandre Pouget. Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11):1432–1438, 2006. doi: 10.1038/nn1790.

Tyler D. Marks and Michael J. Goard. Stimulus-dependent representational drift in primary visual cortex. *Nature Communications*, 12(1):5169, August 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-25436-3.

Cengiz Pehlevan, Anirvan M. Sengupta, and Dmitri B. Chklovskii. Why Do Similarity Matching Objectives Lead to Hebbian/Anti-Hebbian Networks? *Neural Computation*, 30(1):84–124, January 2018. ISSN 0899-7667, 1530-888X. doi: 10.1162/neco_a_01018.

Shanshan Qin, Shiva Farashahi, David Lipshutz, Anirvan M. Sengupta, Dmitri B. Chklovskii, and Cengiz Pehlevan. Coordinated drift of receptive fields in Hebbian/anti-Hebbian network models during noisy representation learning. *Nature Neuroscience*, 26(2):339–349, February 2023. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-022-01225-z.

Neil C. Rabinowitz, Ben D.B. Willmore, Jan W.H. Schnupp, and Andrew J. King. Contrast Gain Control in Auditory Cortex. *Neuron*, 70(6):1178–1191, June 2011. ISSN 08966273. doi: 10.1016/j.neuron.2011.04.030.

Aviv Ratzon, Dori Derdikman, and Omri Barak. Representational drift as a result of implicit regularization. *eLife*, 12:RP90069, May 2024. ISSN 2050-084X. doi: 10.7554/eLife.90069.

Patrick D. Roberts. Computational consequences of temporally asymmetric learning rules: I. differential Hebbian learning. *Journal of Computational Neuroscience*, 7(3):235–246, 1999. doi: 10.1023/A:1008910918445.

Michael E Rule, Timothy O'Leary, and Christopher D Harvey. Causes and consequences of representational drift. *Current Opinion in Neurobiology*, 58:141–147, October 2019. ISSN 09594388. doi: 10.1016/j.conb.2019.08.005.

Carl E. Schoonover, Sarah N. Ohashi, Richard Axel, and Andrew J. P. Fink. Representational drift in primary olfactory cortex. *Nature*, 594(7864):541–546, June 2021. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-021-03628-7.

H S Seung and H Sompolinsky. Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences*, 90(22):10749–10753, November 1993. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.90.22.10749.

Genki Shimizu and Taro Toyoizumi. Reshaping neural representation via associative, presynaptic short-term plasticity. *arXiv preprint arXiv:2601.10397*, 2026.

Yikai Si and Shanshan Qin. Continual-learning rules shape representational drift. (arXiv:2608.16141), August 2026. doi: 10.48550/arXiv.2608.16141.

Eero P Simoncelli and Bruno A Olshausen. Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience*, 24(1):1193–1216, March 2001. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev.neuro.24.1.1193.

Taro Toyoizumi, Kazuyuki Aihara, and Shun-ichi Amari. Fisher information for spike-based population decoding. *Physical Review Letters*, 97(9):098102, 2006. doi: 10.1103/PhysRevLett.97.098102.

Gina G. Turrigiano and Sacha B. Nelson. Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5(2):97–107, February 2004. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn1327.

Suzanne van der Veldt, Gido M. van de Ven, Sanne Moorman, and Guillaume Etter. Learning continually with representational drift. (arXiv:2512.22045), December 2025. doi: 10.48550/arXiv.2512.22045.

Xue-Xin Wei and Alan A Stocker. A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nature Neuroscience*, 18(10):1509–1517, October 2015. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4105.

Xue-Xin Wei and Alan A. Stocker. Mutual Information, Fisher Information, and Efficient Coding. *Neural Computation*, 28(2):305–326, February 2016. ISSN 0899-7667, 1530-888X. doi: 10.1162/NECO_a_00804.

Xiaohui Xie and H. Sebastian Seung. Spike-based learning rules and stabilization of persistent neural activity. In *Advances in Neural Information Processing Systems*, volume 12, pp. 199–205, 2000.

Angela J. Yu and Peter Dayan. Uncertainty, Neuromodulation, and Attention. *Neuron*, 46(4):681–692, May 2005. ISSN 08966273. doi: 10.1016/j.neuron.2005.04.026.

CiCi Xingyu Zheng, Bin Yu, Saket Navlakha, and Alexei Koulakov. How the olfactory bulb maintains stable odor manifolds amid adaptation and representational drift. 2026. doi: 10.64898/2026.01.23.701335.

Noam E. Ziv and Naama Brenner. Synaptic Tenacity or Lack Thereof: Spontaneous Remodeling of Synapses. *Trends in Neurosciences*, 41(2):89–99, February 2018. ISSN 01662236. doi: 10.1016/j.tins.2017.12.003.

Yaniv Ziv, Laurie D. Burns, Eric D. Cocker, Elizabeth O. Hamel, Kunal K. Ghosh, Lacey J. Kitch, Abbas El Gamal, and Mark J. Schnitzer. Long-term dynamics of CA1 hippocampal place codes. *Nature Neuroscience*, 16(3):264–266, March 2013. ISSN 1546-1726. doi: 10.1038/nn.3329.

A Appendix

To derive equation 4, we first solve the allocation problem for a general resource constraint, then show that the per-neuron weight budget constraint we imposed via equation 3 is a special case.

A.1 Resource allocation under a general constraint

Consider maximizing the loss $\mathcal{L}[J] = \int_{\mathcal{S}} p(s) \log J(s)\, ds$ over functions J, subject to $\int_{\mathcal{S}} J(s)^\alpha\, ds = C$ for some fixed $\alpha > 0$. Substituting $u = J^\alpha$ turns this into the problem of maximizing a concave objective with a linear constraint: $\alpha^{-1} \int_{\mathcal{S}} p(s) \log u(s)\, ds$ subject to $\int_{\mathcal{S}} u(s)\, ds = C$.

Setting the functional derivative to zero gives $p(s)/u(s) = \lambda\alpha$, hence $u^\star \propto p$ and

$$J^\star(s) \;=\; C\, p(s)^{1/\alpha}, \tag{8}$$

where the multiplier has been absorbed into C. The exponent of the optimal allocation is therefore determined by which resource is constrained. A constraint on total Fisher information ($\alpha = 1$) gives $J^\star \propto p$, whereas a constraint on $\int \sqrt{J}\, ds$ ($\alpha = 1/2$) gives $J^\star \propto p^2$, the allocation obtained under a constraint on the number of neurons (Ganguli & Simoncelli, 2014).

A.2 Cumulative weight constraint

We now identify which constraint equation 3 imposes. To do so we adopt the density parameterization of Ganguli & Simoncelli (2014), in which a heterogeneous population is described by the local density of preferred stimuli together with a gain function, with tuning widths varying inversely with that density. Concretely, we assume that the learned solutions will have the following properties. First, because $W \geq 0$ and $\sum_j W_{ij} = 1$ limit each neuron's total input, tuning curves take the form

$$\mu_i(s) \;=\; c\, \psi\Big(\frac{s - s_i}{w_i}\Big), \tag{9}$$

with ψ a fixed unimodal profile satisfying $\psi(0) = 1$, s_i the preferred stimulus of neuron i, w_i its tuning width with peak amplitude c. Second, letting $d(s)$ denote the local density of preferred stimuli with $\int_{\mathcal{S}} d(s)\, ds = M$, tuning widths vary inversely with that density, $w(s) \propto 1/d(s)$.

Each neuron contributes to J over a range of width w, with slope $\mu_i' \sim c/w$, so its contribution scales as c^2/w^2 (according to equation 1). Summing over the $d(s)\, w(s)$ neurons active near s,

$$J(s) \;\propto\; \frac{c^2}{\sigma^2}\, \frac{d(s)}{w(s)} \;\propto\; c^2 d(s)^2, \tag{10}$$

using $w \propto 1/d$ in the second step. Hence $\sqrt{J} \propto c\, d(s)$ and

$$\int_{\mathcal{S}} \sqrt{J(s)}\, ds \;\propto\; c \int_{\mathcal{S}} d(s)\, ds \;=\; c\, M,$$

which depends only on the peak amplitude, the number of neurons and the fixed tuning profile. The weight budget therefore constrains $\int \sqrt{J}\, ds$, and equation 8 with $\alpha = 1/2$ gives $J^\star \propto p^2$.

A.3 Nonlinear output neurons

Let the output population apply a pointwise nonlinearity ϕ such that:

$$\mathbf{a}(s) = W\mathbf{x}(s), \qquad \boldsymbol{\mu}(s) = \phi\big(\mathbf{a}(s)\big), \qquad \mathbf{y} = \boldsymbol{\mu}(s) + \boldsymbol{\xi}, \quad \boldsymbol{\xi} \sim \mathcal{N}(0, \sigma^2 I). \tag{11}$$

Writing $G(s) = \mathrm{diag}\big(\phi'(\mathbf{a}(s))\big)$ for the matrix of single-neuron gains, differentiating w.r.t. stimulus s gives $\boldsymbol{\mu}'(s) = G(s)W\mathbf{x}'(s)$, so that

$$J(s; W) \;=\; \frac{1}{\sigma^2}\, \mathbf{x}'(s)^\top W^\top G(s)^2 W\, \mathbf{x}'(s) \;=\; \frac{1}{\sigma^2} \sum_i \phi'(a_i)^2 \big(\mathbf{w}_i^\top \mathbf{x}'\big)^2, \tag{12}$$

which reduces to Eq. equation 1 when $\phi' \equiv 1$. To derive a learning rule for this case, we differentiate with respect to the weights and apply the product rule:

$$\frac{\partial \mu_i'}{\partial \mathbf{w}_i} = \phi''(a_i)\left(\mathbf{w}_i^\top \mathbf{x}'\right)\mathbf{x} + \phi'(a_i)\,\mathbf{x}' = \phi''(a_i)\,a_i'\,\mathbf{x} + \phi'(a_i)\,\mathbf{x}' = \frac{d}{ds}\Big(\phi'(a_i)\,\mathbf{x}\Big),$$

Since $\nabla_{\mathbf{w}_i} \log J = (2/\sigma^2 J)\,\mu_i'\,\partial\mu_i'/\partial\mathbf{w}_i$, gradient ascent on $\mathcal{L}(W)$ gives,

$$\Delta W_{ij} \;=\; \eta\, U(s)\,\mu_i'(s)\,\frac{d}{ds}\Big(\phi'\big(a_i(s)\big)\,x_j(s)\Big), \qquad U(s) = J(s)^{-1}, \tag{13}$$

Following the argument of Section 2.3, the corresponding temporal-derivative form is:

$$\Delta W_{ij} \;=\; \eta\, U_t\,\dot{y}_{i,t}\,\frac{d}{dt}\Big(\phi'\big(a_{i,t}\big)\,x_{j,t}\Big).$$

Note that the nonlinearity enters the update rule only by weighting the presynaptic activity by the neuron's gain, so the rule reduces to Eqs. (6) and (7) when the gain is constant. Note also that the rule remains local. For $\phi = \tanh$, the gain is $\phi'(a_i) = 1 - \mu_i^2$, a function of the postsynaptic activity alone. However, saturation effects reduce $\phi'(a_i)$, so neurons tuned to high-probability stimuli contribute less Fisher information than Eq. equation 9 presumes. The fitted exponent under $\phi = \tanh$ is therefore slightly less than 2 in simulations.

A.4 Supplementary figures

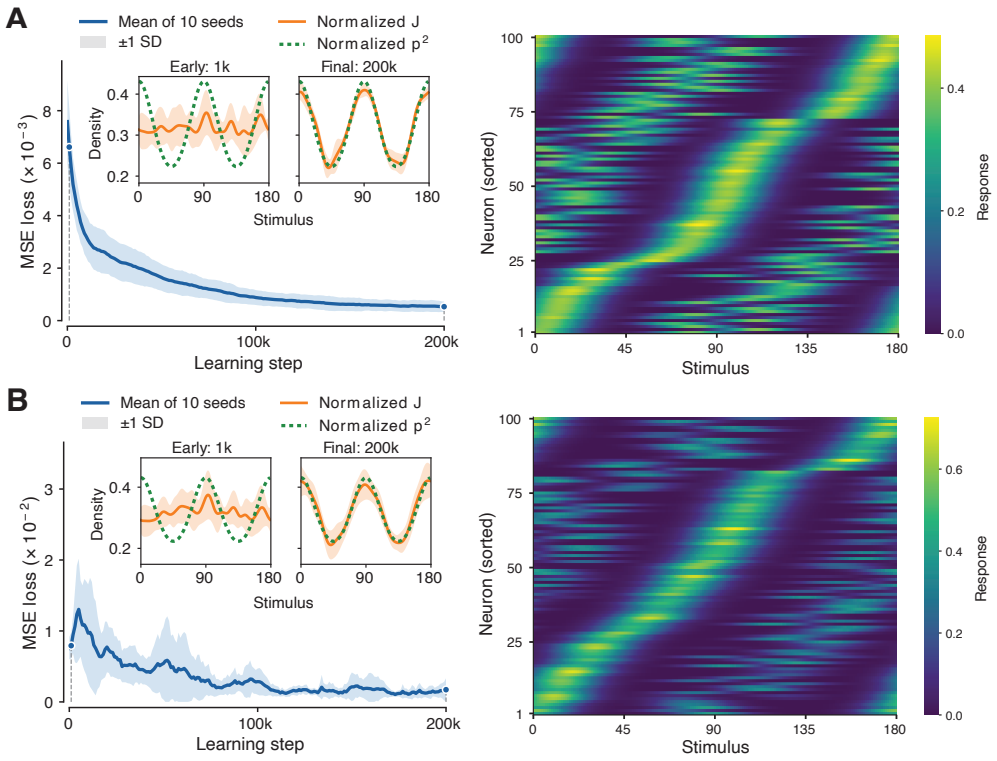

Figure A1: Fisher information profiles for alternative models. (A) Saturating output nonlinearity, $\boldsymbol{\mu}(s) = \tanh(W\mathbf{x}(s))$, trained with equation 13. (B) Stimulus derivatives replaced by temporal derivatives for stimuli following a Langevin trajectory, trained with equation 6. Left: mean squared error between the normalized Fisher information profile of the simulated outputs and theory over learning, with insets showing early and late time points. Right: Tuning curves after convergence, sorted by preferred stimulus.

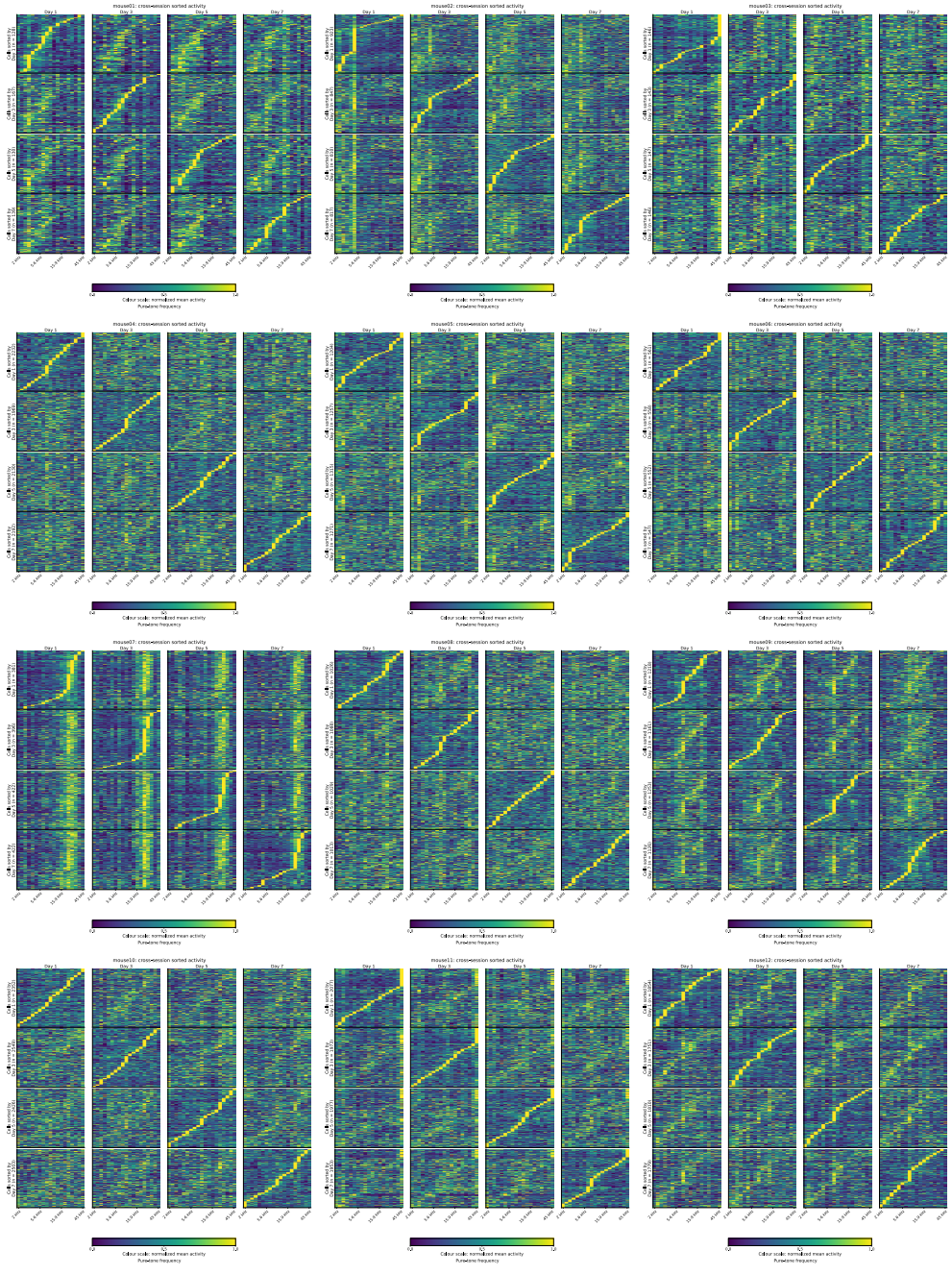

Figure A2: Single-neuron representational drift in auditory cortex (n=12 mice). Within each animal, columns are recording days and rows are the sort order taken from each day, so diagonal panels are sorted by the session they display.

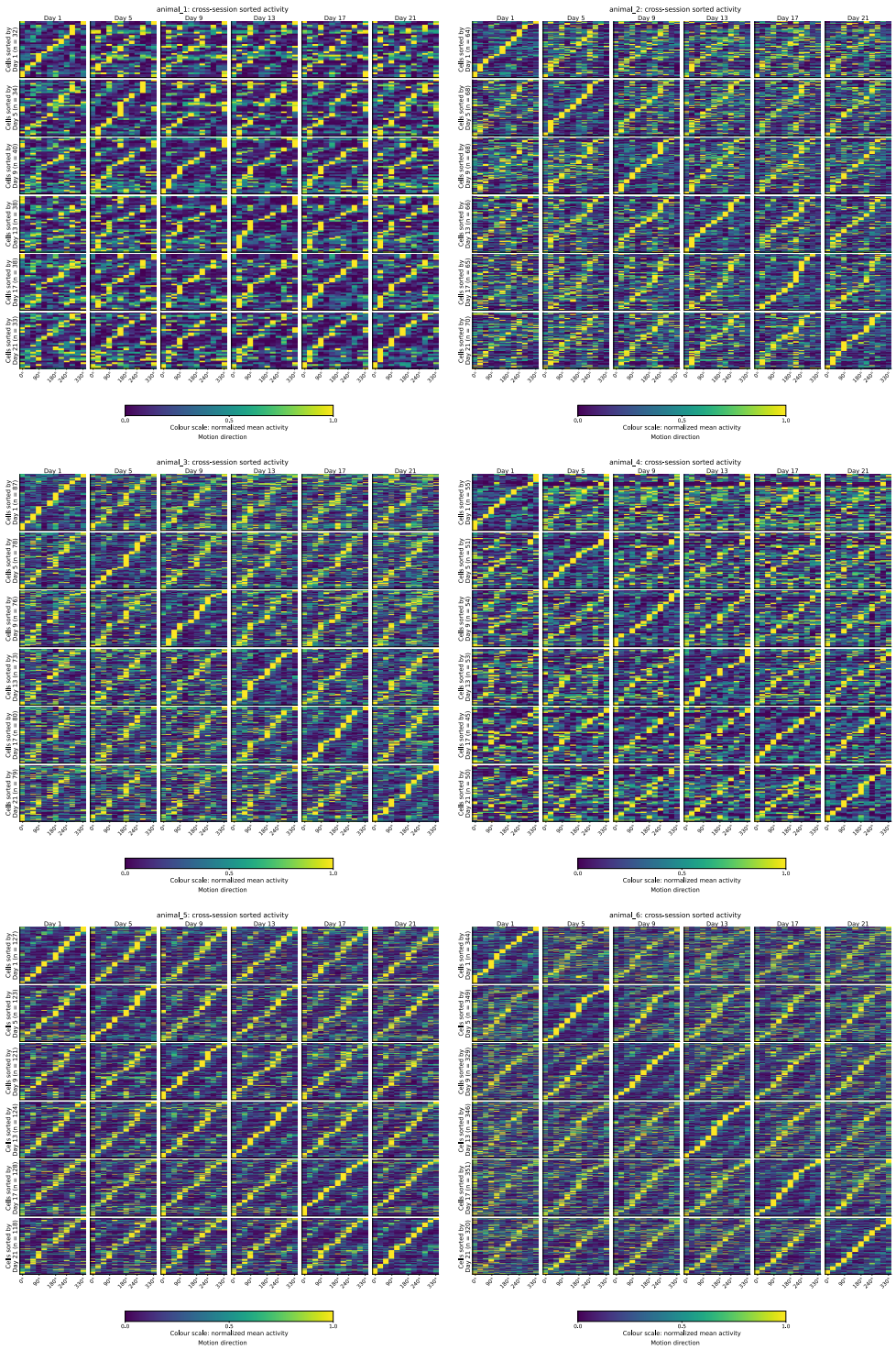

Figure A3: Single-neuron representational drift in visual cortex (n=6 mice). Conventions as in Figure A2.

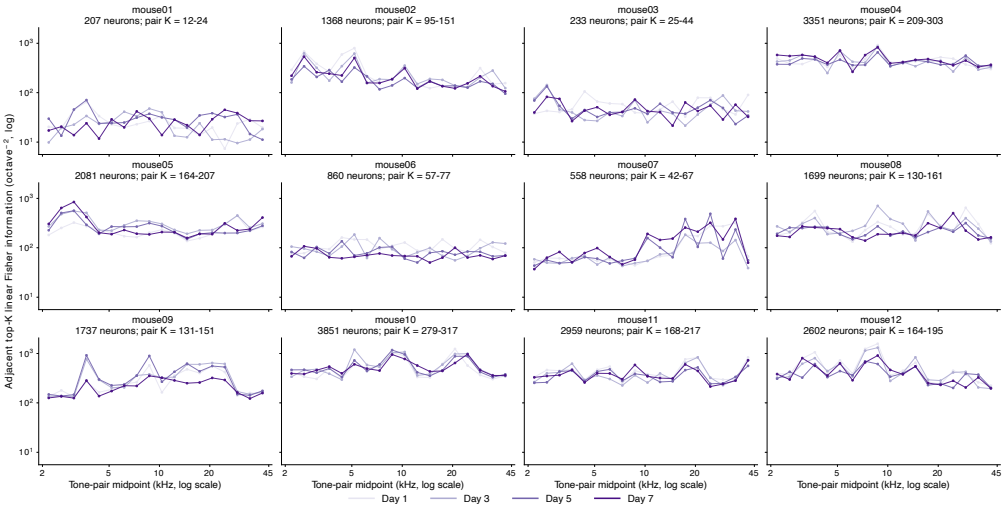

Figure A4: Fisher information profiles across tone frequency for each mouse in the auditory dataset, shown separately for each recording day. Neuron counts and the number of dimensions retained for covariance estimation are given above each panel.

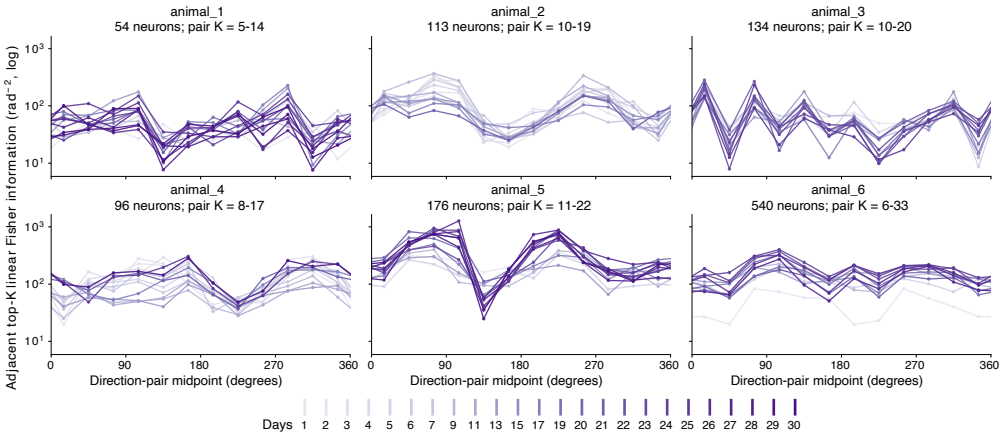

Figure A5: Fisher information profiles across motion direction for each animal in the visual dataset. Conventions as in Figure A4.

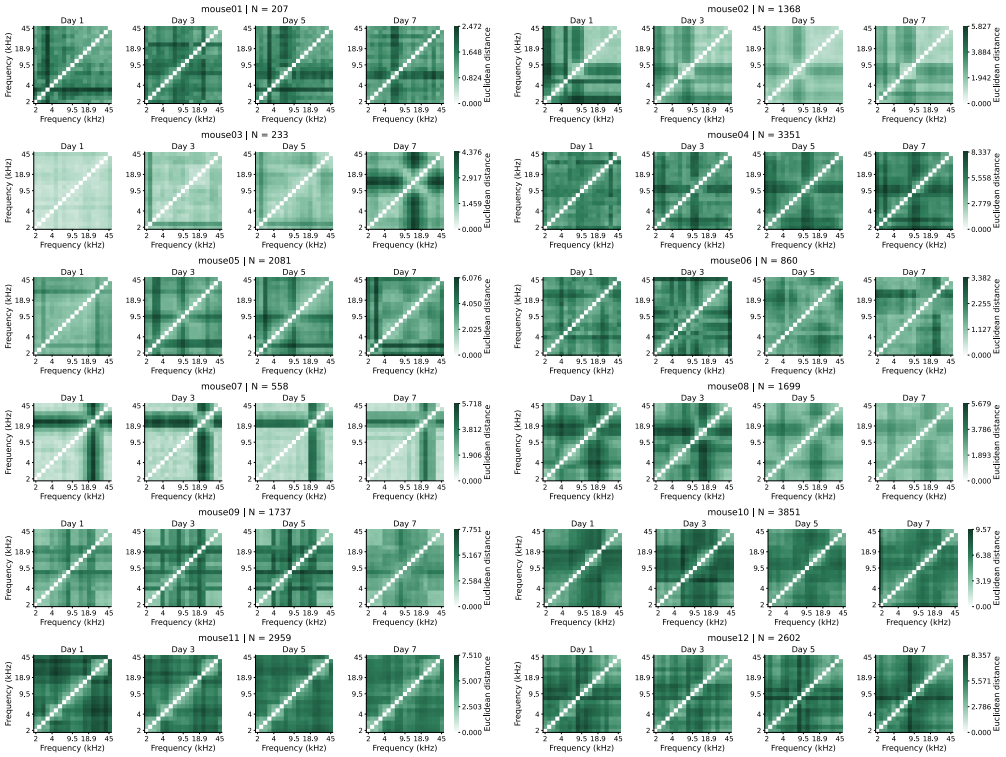

Figure A6: Representational dissimilarity matrices for each mouse in the auditory dataset, shown separately for each recording day. N denotes the number of neurons.

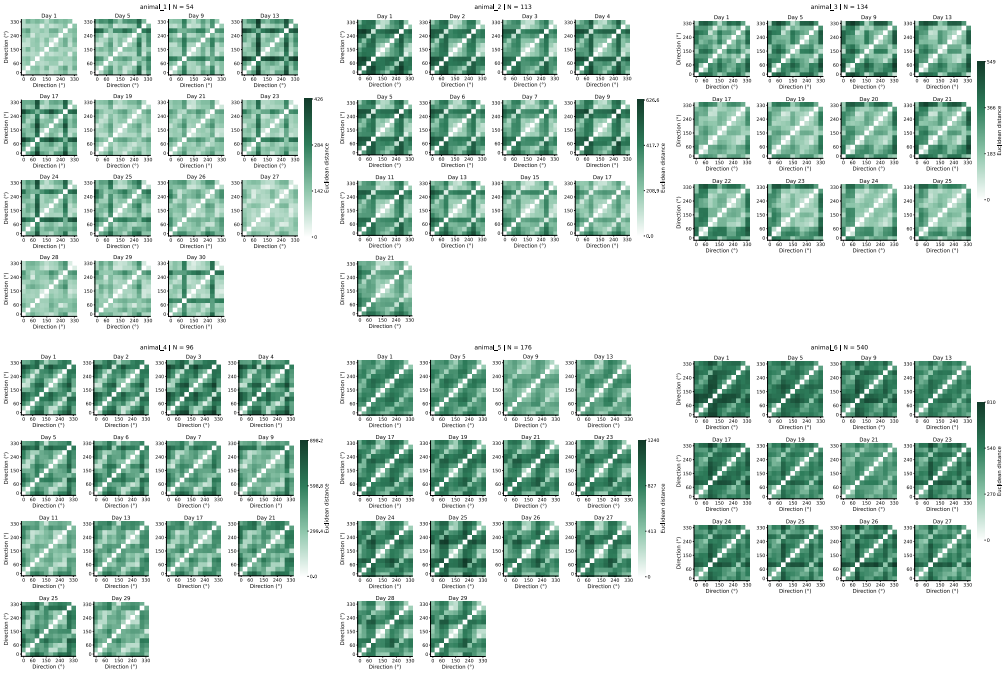

Figure A7: Representational dissimilarity matrices for each animal in the visual dataset. Conventions as in Figure A6.