

Structured Sparsification of Signal-Transmission Networks Enhances Visual Information Coding

Jiaji Zhu^{1,2,3} and Xiaoxuan Jia^{1,2,3,*}

¹Center for Life Sciences and Artificial Intelligence, School of Life Sciences, Tsinghua University, Beijing, China

²IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing, China

³Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing, China

*Correspondence: Xiaoxuan Jia (jxiaoxuan@gmail.com)

ABSTRACT

Sensory coding depends on the architecture of cortical neural networks, yet the principles linking network organization to coding performance remain poorly understood. To address this question, we inferred spiking signal-transmission networks at single-neuron resolution from simultaneous Neuropixels recordings across six visual areas and related their organization to the fidelity and speed of sensory coding. Improved coding was associated with structured sparsification of these networks, characterized by fewer, more local, modular, feature-specific, and feedforward connections. To identify the mechanisms through which structured sparsification enhances coding, we used rate-based network models within a linear Fisher information framework. Our results show that structured sparsification improves coding fidelity by reducing shared variability, decreasing signal–noise alignment, and sharpening neuronal selectivity. It also accelerates coding by promoting more feedforward transmission of sensory signals. Together, these findings provide a mechanistic link between the architecture of cortical signal-transmission networks and the fidelity and speed of population codes.

1 INTRODUCTION

Sensory perception relies on the cortical network that encode and transmit information. The organization of these network across scales, from local connectivity (Recanatesi et al., 2019; Dahmen et al., 2023) to global topology (Papadopoulos et al., 2026; Wyrick and Mazzucato, 2021; Litwin-Kumar and Doiron, 2012), determines how information is routed and how neuronal activity is propagated, which shapes the precision and speed of sensory coding (Yuste, 2015; Hopfield, 1982). The network structure is therefore central to understanding sensory information processing. However, the mechanisms by which network structure shapes the encoding and transmission of external sensory information remain poorly understood.

Theoretical work has identified several ways in which network structure can shape population codes. Recurrent circuit models show that connectivity can organize shared trial-to-trial variability (Rosenbaum and Doiron, 2014; Rosenbaum et al., 2017; Huang et al., 2019), which, when aligned with stimulus-sensitive population directions, limits Fisher information (Moreno-Bote et al., 2014; Kanitscheider et al., 2015b; Huang et al., 2022). Structured connectivity can also constrain the dimensionality and dynamics of population activity. For example, low-rank connectivity links connectivity geometry to low-dimensional dynamics and computational capacity (Mastrogiuseppe and Ostojic, 2018), whereas local motifs and recurrent coupling regulate the dimensionality of network activity (Recanatesi et al., 2019). Clustered network architectures organize ensemble dynamics and trial-to-trial variability (Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015), modulate metastable dynamics, thereby affecting representational dimensionality, discriminability, and coding speed (Mazzucato et al., 2016, 2019; Papadopoulos et al., 2026). Together, these studies show that network structure can influence coding by changing neuronal variability, response dimensionality, and dynamics. However, these theoretical links are typically established by imposing selected architectural features in model circuits, and remain insufficiently grounded in experimental measurements. Whether analogous structure–coding relationships appear in networks inferred

from experimentally observed cortical activity, and whether reconfiguration of these networks can account for changes in the capacity and speed of sensory coding, remain unresolved.

Locomotion provides a natural experimental setting to address these questions because it modulates visual coding and may concurrently reconfigure network structure. In the mouse visual cortex, locomotion enhances visually evoked responses and improves the speed and accuracy of visual coding (Niell and Stryker, 2010; Vinck et al., 2015; Dadarlat and Stryker, 2017; Christensen and Pillow, 2022; Ayaz et al., 2013). It also changes the coupling between neurons (Dadarlat and Stryker, 2017; Clancy et al., 2019; Horrocks et al., 2024). Because correlations and functional coupling are shaped in part by the recurrent circuit architecture (Doiron et al., 2016), these findings raise the possibility that locomotion improves visual coding not only by modulating response gain or variability, but also by reorganizing cortical neuronal interactions. We therefore used locomotion as a behavioral lever to ask whether single-neuron-resolution networks are reconfigured during different behavioral states and whether this reconfiguration can account for changes in the capacity and speed of sensory coding.

To address these questions, we analyzed Neuropixels recordings from the Allen Brain Observatory (Siegle et al., 2021) and inferred single-neuron-resolution signal-transmission networks from cross-correlograms between simultaneously recorded spike trains. Candidate edges were identified by narrow, asymmetric CCG peaks at short, biologically plausible latencies (within 13 ms). We applied spike-time jitter correction to remove correlations from stimulus locking, global state fluctuations, and other slower co-modulations, thereby isolating temporally precise interactions that could not be explained by broad activity covariation alone (Smith and Kohn, 2008; Jia et al., 2022; Siegle et al., 2021). These inferred connections therefore differ from conventional functional-connectivity estimates such as correlation: they capture directed, millisecond-timescale interactions consistent with putative monosynaptic or multi-synaptic transmissions (Ren et al., 2020; English et al., 2017; Fujisawa et al., 2008). Nevertheless, they remain statistical inferences and do not constitute anatomical proof of synaptic connectivity. For cautious, we refer to these networks as signal-transmission networks.

We found that improved coding was associated with a structured sparsification of the single-neuron resolution signal-transmission networks: overall connection density was reduced and network organization was shifted toward more localized, feature-specific, and feedforward architectures, with more input projections to weakly selective neurons. To determine the mechanism through which this structured sparsification affected neural coding, we combined rate-based network models with population coding analysis. Structured sparsification improved coding by reshaping the signal-noise geometry that determines linear Fisher information. Each network property corresponded to a distinct aspect of this geometry: reduced connection density decreased shared variability; more localized connectivity weakened the information-limiting effects of noise correlations; and feature-specific connectivity, together with more input projections to weakly selective neurons, enhanced stimulus selectivity. In addition, more feedforward architecture also straightened neural trajectories and accelerated coding. By linking the architecture of signal-transmission networks to the accuracy and speed of population codes, these analyses provide a principled framework for understanding how cortical networks shape sensory information processing.

2 RESULTS

Consider a neural population with signal-transmission matrix W , driven by sensory inputs $I(\theta)$ and internal noise ξ (Stringer et al., 2019; Musall et al., 2019), where θ indicates the stimulus parameter. The structure of W shapes the population responses and therefore modulates information-coding capacity and dynamics. How locomotion reconfigures W and how this network reconfiguration modulates the coding capacity and speed are the central questions we address in this study (Fig. 1a).

We analyzed the Allen Brain Observatory Neuropixels dataset (Siegle et al., 2021), which contains simultaneous recordings of spiking activity from six cortical visual areas (Fig. 1b; V1, LM, RL, AL, PM, AM) and two thalamic nuclei (LGN, LP) in mice passively viewing a variety of artificial and natural visual stimuli, while monitoring running speed and pupil size (Fig. 1b, c). We classified trials as locomotion or stationary based on mean running speed (threshold: 1 cm/s) to determine different behavioral states. In each behavioral state, we inferred pairwise connections using jitter-corrected cross-correlograms (CCGs) (Jia et al., 2013; Siegle et al., 2021; Smith and Kohn, 2008). Connections were inferred between neuron pairs whose jitter-corrected CCGs exhibited sharp peaks with short-latency: positive peaks reflect increased firing probability of the target neuron after the source neuron fires, whereas negative peaks reflect decreased firing probability (Fig. 1d). Since these statistical couplings have been suggested to reflect signal transmission efficacy (Smith and Kohn, 2008; Kohn et al., 2016; Jia et al., 2013; Siegle et al., 2021), potentially including monosynaptic, polysynaptic and common-input-mediated interactions, we refer to the resulting networks as signal-transmission networks (Fig. 1e).

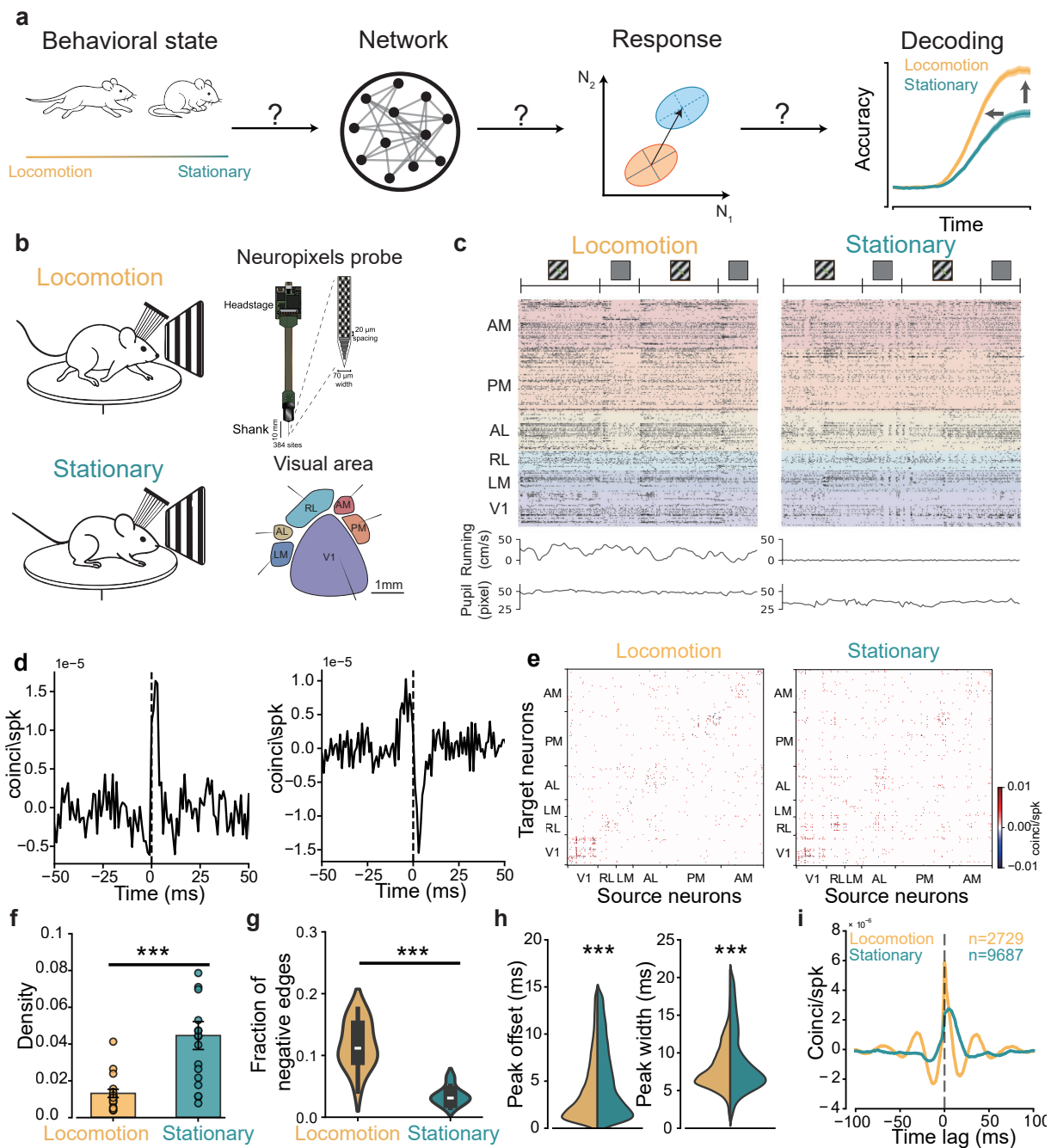


Fig. 1. Locomotion reduces the network density. **a**, Schematic illustration of the key questions addressed in this study. We investigated how behavioral state altered network organization in the visual cortex, how network reconfiguration reshaped neural responses, and how changes in neural responses affected information decoding. **b**, Schematic of data collection using Neuropixels probes inserted across six visual cortical areas (AM, PM, AL, RL, LM, and V1). **c**, Example raster plot of simultaneously recorded units from six areas during drifting grating stimulation, shown for both locomotion and stationary states, with the running-speed and pupil size trace below. **d**, Example positive CCG trace (left) and negative CCG trace (right). **e**, Example adjacency matrix during locomotion (left) and stationary states (right). **f**, Network density during locomotion and stationary states in response to drifting grating stimuli. Error bars indicate SEM; points indicate individual mice (Wilcoxon signed-rank test, $n = 17$ mice, $p = 1.52 \times 10^{-5}$). Density was defined as the fraction of significant edges relative to all possible edges. **g**, Fraction of negative edges during locomotion and stationary states in response to drifting grating stimuli. (Wilcoxon signed-rank test, $n = 17$ mice, $p = 1.52 \times 10^{-5}$). Fraction of negative edges was defined as the fraction of significant negative edges among all significant edges. Violin plots show the distribution; white lines indicate medians and dark boxes the interquartile range. **h**, CCG peak offset (Wilcoxon signed-rank test, $n = 17$

mice, $p = 1.52 \times 10^{-5}$) and peak width (Wilcoxon signed-rank test, $n = 17$ mice, $p = 0.0006$). **i**, Average CCG trace for significant pairs during drifting gratings stimuli.

Locomotion reorganizes the signal-transmission network through structured sparsification

Locomotion has been shown to reduce correlated variability in the visual cortex (Dadarlat and Stryker, 2017). Because correlated variability and functional coupling are both shaped in part by circuit architecture (Kohn et al., 2016), this raised the question of whether locomotion influences the signal-transmission network similarly. To address this, we quantified network density d as the fraction of inferred connections among all possible directed connections, $d = \frac{E}{N(N-1)}$, where E denotes the number of inferred connections and N is the number of neurons. We found that the density was consistently lower during locomotion than during stationary periods for drifting gratings (Fig. 1f, Wilcoxon signed-rank test, $p = 1.52 \times 10^{-5}$), static gratings, natural scenes, and natural movies (Extended Data Fig. 1b). This decrease was robust across visual areas (Fig. 2c). These results indicated that locomotion sparsified the signal-transmission network, consistent with previous reports that locomotion decorrelates neural activity (Dadarlat and Stryker, 2017; Erisken et al., 2014; Horrocks et al., 2024). Our findings further extended these observations beyond V1 to multiple visual areas and moved beyond noise correlations to directed, spike-timing-based signal transmission. Notably, the sparsification was sign-biased: the reduction in density was restricted to positive connections (Extended Data Fig. 2a); Negative-connection density was unchanged for drifting and static gratings, and was even higher during locomotion for natural scenes and natural movies (Extended Data Fig. 2b). Therefore the fraction of negative connections was higher during locomotion (Fig. 1g, Wilcoxon signed-rank test, $p = 1.52 \times 10^{-5}$). Beyond this network sparsification, locomotion accelerated and sharpened signal transmission, as reflected by shorter offsets and narrower widths of CCG peaks (Fig. 1h, i; peak offset: Wilcoxon signed-rank test, $n = 17$ mice, $p = 1.52 \times 10^{-5}$; peak width: Wilcoxon signed-rank test, $n = 17$ mice, $p = 6.0 \times 10^{-4}$; Extended Data Fig. 3a, b).

Naive cross-correlograms are sensitive to firing rate (Stark and Abeles, 2009), which is strongly modulated by behavioral state (Niell and Stryker, 2010). However, our jitter-corrected CCGs were normalized by the geometric mean of the two neurons' firing rates and further corrected by jittering (Methods) to minimize the influence of firing rates fluctuations beyond the jitter window. To confirm that the density decrease was not driven by firing rate modulation, we performed a rate-matching control in which we randomly subsampled spikes to equate firing rates across states while preserving spike-timing structure (Extended Data Fig. 4a). Networks constructed from rate-matched spike trains still showed reduced density during locomotion (Extended Data Fig. 4c), indicating that this effect cannot be explained solely by firing-rate changes.

Taken together, locomotion systematically sparsified the signal-transmission network W , preferentially eliminating positive over negative connections. This pattern suggested that the sparsification was structured rather than random. In the following sections, we characterize which features of W undergo structured sparsification and reveal their effects on information coding and dynamics.

Locomotion shifts the network towards local and modular organization

Anatomical organization constrains neuronal connectivity, such that the connection probability between pairs of neurons is generally higher when the neurons are in the same area and decreases with anatomical distance (Markov et al., 2013, 2014). This spatial constraint shapes noise covariance (Rosenbaum et al., 2017; Huang et al., 2019), determines the stability of neural dynamics (Rosenbaum and Doiron, 2014; Huang et al., 2019), modulates the intrinsic timescale (Zeraati et al., 2023), and influences information-coding capacity (Huang et al., 2019). We asked whether locomotion-induced sparsification was spatially structured by testing whether its magnitude differed between within-area and between-area connections and varied with anatomical distance. To quantify the magnitude of sparsification, we defined a relative change of density as $(d_{stationary} - d_{locomotion}) / (d_{stationary} + d_{locomotion})$, where $d_{locomotion}$ and $d_{stationary}$ denote density during locomotion and stationary states, respectively. We found that locomotion sparsified connection across all six visual areas, yet between area connections were sparsified more drastically (Fig. 2a–c): the relative change was significantly higher for between-area than for within-area connections (Fig. 2d, e, Extended Data Fig. 5a, b) and the fraction of within-area connections increased during locomotion (Fig. 2f, paired t-test, $p = 2.46 \times 10^{-10}$). Similarly, during locomotion, density declined more steeply with distance between recorded units (Fig. 2g–i, Extended Data Fig. 5c, d). These results revealed the spatial structure of locomotion-induced sparsification: it preferentially removed long-range connections, shifting the signal-transmission network toward a more localized organization.

We examined whether this localization was accompanied by higher network modularity (Deco et al., 2015; Cohen and D'Esposito, 2016; Sporns, 2013; Baum et al., 2017). We quantified modularity using the Louvain community-detection

algorithm (Blondel et al., 2008), which captures the extent to which a network is partitioned into modules with stronger within-module than between-module connectivity (Fig. 2j, k). The resolution parameter of the detection method was chosen to maximize the deviation from a degree- and density-matched null model (Methods; Extended Data Fig. 7a). Modularity was significantly higher during locomotion (Fig. 2l, Wilcoxon signed-rank test, $p = 6.10 \times 10^{-5}$), and this increase was not solely due to differences in density. A null-model-corrected modularity score, obtained by subtracting the modularity of a pair-preserving null model, remained significantly higher during locomotion (Extended Data Fig. 7c).

Similar structured sparsification was observed across different significant-CCG-peak thresholds (Extended Data Fig. 6a–c). These results indicate that sparsification was spatially structured, biasing the network towards a more localized organization. Between-area and long-range connections were preferentially eliminated while short-range and within-area connections were comparatively preserved, shifting the network from a globally integrated architecture towards a more localized and modular organization.

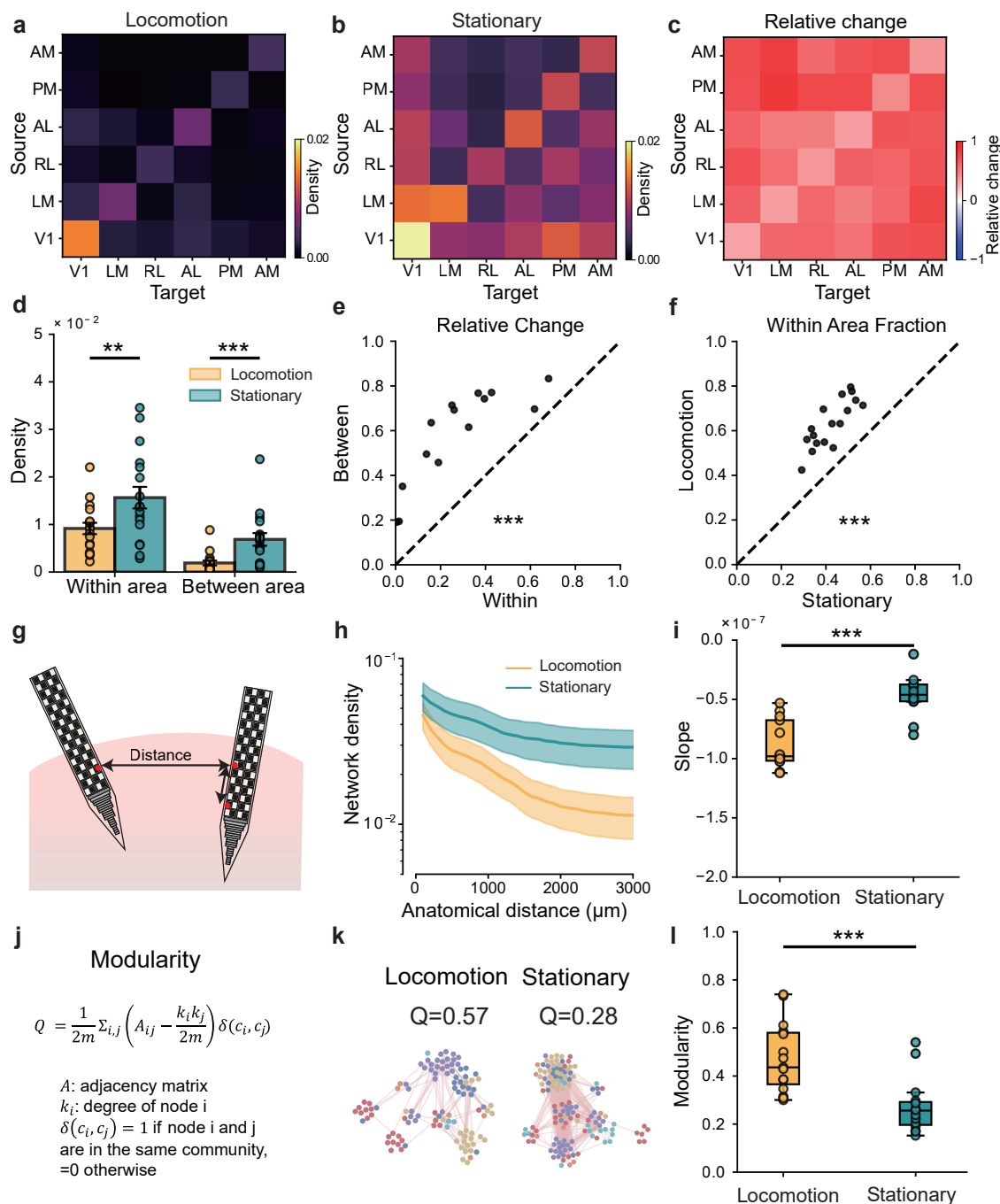


Fig. 2. Locomotion localizes connectivity and increases network modularity. **a**, Network density of connections among six visual areas during locomotion. **b**, Same as in **a**, but during the stationary state. **c**, Relative change in network density across the six visual areas. Relative change in density was defined as $(d_{stationary} - d_{locomotion}) / (d_{stationary} + d_{locomotion})$, where $d_{stationary}$ and $d_{locomotion}$ denote network density during stationary and locomotion states. **d**, Network density of within-area and between-area connections during drifting grating stimuli. Error bars indicate SEM, and points indicate individual mice (Wilcoxon signed-rank test, $n = 17$ mice, within-area: $p = 1.52 \times 10^{-3}$; between-area: $p = 1.52 \times 10^{-5}$). **e**, Relative changes in within-area density and between-area density. Points represent individual mice (paired t -test, $n = 17$ mice, $p = 1.81 \times 10^{-10}$). **f**, Within-area fraction of connections, defined as the number of connections between neurons in the same area divided by the total number of connections (paired t -test, $n = 17$ mice, $p = 2.46 \times 10^{-10}$). **g**, Schematic illustration of anatomical distance between neurons. **h**, Network density as a function of anatomical distance. Shading denotes $\pm$ SEM. ($n = 17$ mice). **i**, Slopes of curves in **h** for each mouse. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times$ IQR; points represent mice (paired t -test, $n = 17$ mice, $p = 1.43 \times 10^{-8}$). **j**, Schematic illustration of network modularity. **k**, Example network visualization from a representative mouse during locomotion (left) and stationary (right) states. **l**, Network modularity across conditions. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times$ IQR; points represent mice (Wilcoxon signed-rank test, $n = 17$ mice, $p = 6.10 \times 10^{-5}$).

Localized networks improve coding by reducing information-limiting correlations

To test whether and how sign-biased and spatially localized sparsification can improve population coding, we built a two-layer rate-based network with independently tunable connection density, spatial range, and edge sign fraction (Fig. 3a–c). Each layer contained 200 neurons, with a 4:1 ratio of excitatory to inhibitory units. To approximate hierarchical visual processing (Siegle et al., 2021), we used a two-layer hierarchical model in which stimuli entered the lower layer and propagated to the higher layer, with spatially structured connectivity defined by distance-dependent connection probabilities. We computed linear Fisher information (Fig. 3d) to quantify the amount of stimulus information accessible to an optimal linear decoder (Moreno-Bote et al., 2014):

$$I = f'^T \Sigma^{-1} f' \quad (1)$$

where Σ is the noise covariance matrix and $f' = df(\theta)/d\theta$ denotes the derivative of the mean population response f with respect to the stimulus parameter θ . Larger linear Fisher information corresponds to higher population decoding accuracy. Fisher information depends on three components: the magnitude of noise covariance, signal strength and signal-noise alignment (Methods). We examined how structured sparsification changes each of these components.

As previously reported (Dadarlat and Stryker, 2017), locomotion reduced the shared variability of neural activity. We asked whether sparsification and sign-biased reconfiguration of networks contribute to this effect. To test this, we varied the total density and measured the noise covariance (Fig. 3e). Decreasing density reduced the eigenvalues of the noise covariance matrix (Fig. 3f), suggesting that sparsification reduced shared variability. We further tested the effect of an increased negative fraction. At fixed density ($p = 0.1$), increasing the negative fraction also reduced eigenvalues (Fig. 3g). These results suggest that sparsification and an increased negative fraction reduce the magnitude of noise covariance, thereby increasing Fisher information (Fig. 3h).

Linear Fisher information depends not only on the magnitude of shared noise, but also on its alignment with the stimulus-related signal (Moreno-Bote et al., 2014). Noise becomes information-limiting when it lies along the signal direction, which we quantified as the cosine similarity between the signal derivative (f') and the dominant noise direction (u_1). Feedforward projections are a source of information-limiting correlation, because signal and noise propagate through the same pathways and are therefore difficult for downstream populations to disentangle (Kanitscheider et al., 2015b). We hypothesized that more spatially restricted feedforward projections would reduce signal–noise alignment by reducing shared feedforward circuitry. We also hypothesized that increasing the within-area fraction of connections would weaken information-limiting correlation by reducing the relative contribution of feedforward input.

To test these predictions, we varied the spatial scale of feedforward connectivity σ_{ffwd} (Fig. 3i). Decreasing σ_{ffwd} reduced the alignment between signal and dominant shared noise (quantified by cosine similarity) (Fig. 3j), indicating weaker signal–noise alignment. Increasing the within-area fraction produced a similar effect (Fig. 3k). These observations were robust to a range of parameters (Extended Data Fig. 8a, b). Because locomotion shifts the network towards more within-area and spatially localized connectivity structure (Fig. 2), these results suggest that locomotion reduces the information-limiting effect of noise correlations (Fig. 3l). Consistently, alignment between f' and the leading eigenvectors of the noise covariance was lower during locomotion in the experimental data (Fig. 3s).

Unlike the spatial scale of feedforward connections, the spatial scale of recurrent connections primarily shapes network

integration (Cohen and D’Esposito, 2016; Deco et al., 2015; Rosenbaum et al., 2017). We hypothesized that broader recurrent connectivity would promote population-wide shared dynamics, concentrating variability into a few dominant modes and reducing the effective dimensionality of population activity. To test this, we varied the recurrent spatial scale σ_{rec} (Extended Data Fig. 8c, Fig. 3m). Narrower recurrent connectivity substantially reduced the first few dominant eigenvalues of the noise covariance, while leaving the smaller eigenvalues largely unchanged (Fig. 3n). As a result, intrinsic dimensionality increased (Fig. 3o).

Consistent with the simulations, noise correlation decreased during locomotion (Fig. 3r) and the leading eigenvalue decreased during locomotion in the mouse visual cortex (Fig. 3q) and intrinsic dimensionality was higher during locomotion (Fig. 3t). Increased intrinsic dimensionality does not by itself imply higher Fisher information, because if f' were orthogonal to the dominant noise modes, changes in their eigenvalues would have little effect on coding. In the mouse visual cortex, however, the dominant eigenmodes of the noise covariance were also most aligned with f' , as indicated by the highest cosine similarity (Fig. 3s). The increase in dimensionality (Fig. 3t) thus reflects selective reduction of noise along signal-relevant dimensions thereby also increasing Fisher information (Fig. 3p).

These results demonstrate that sign-biased and spatially ordered sparsification reshapes noise covariance to increase Fisher information through multiple mechanisms. Reduced density and an increased negative fraction suppress shared variability, reducing the magnitude of noise covariance. More within-area and more localized feedforward connectivity reduces signal–noise alignment. Narrower recurrent connectivity redistributes noise away from signal-relevant dimensions, reducing the information-limiting effect of noise correlations.

Locomotion reorganizes the network to enhance feature selectivity

Linear Fisher information is proportional to signal magnitude $\|f'\|$, which is large when neurons respond selectively (Moreno-Bote et al., 2014). Neurons with similar tuning are more likely to be connected (Cossell et al., 2015; Ding et al., 2025; Ko et al., 2011), and such feature-specific connectivity can enhance neuronal selectivity (Huang et al., 2022). We therefore asked whether sparsification was structured along tuning similarity. We computed tuning similarity as the difference in preferred direction (drifting gratings) or orientation (static gratings) and then quantified how density changed with tuning similarity. This density was normalized by the overall network density to account for differences in overall density across states, producing a relative density measurement. In both states, relative density generally decreased with increasing tuning difference (Fig. 4a, b), consistent with a feature-specific connectivity pattern. Neurons with identical tuning preferences exhibited higher relative density during locomotion for drifting gratings and static gratings (Fig. 4a, b). The same effect was observed when tuning similarity was quantified using signal correlations (Extended Data Fig. 9a, b). These results indicated that sparsification was structured along tuning similarity, producing a more feature-specific network during locomotion.

Previous work has shown that feature-specific connectivity can enhance selectivity when orientation preference is organized in a pinwheel map and internal noise is negligible (Huang et al., 2022). However, in mouse visual cortex, neither assumption holds: orientation preference follows a salt-and-pepper organization (Ohki and Reid, 2007; Ringach et al., 2016; Ohki et al., 2005), and non-visual fluctuations contribute substantially to response variability (Stringer et al., 2019; Musall et al., 2019). To test whether feature-specific connectivity still improves selectivity in this regime, we defined connection probability to decrease as tuning similarity decreases, with F_{fs} controlling the degree of feature specificity (Fig. 4c, d). Increasing F_{fs} systematically sharpened population tuning, increasing both the orientation selectivity index and the signal-derivative magnitude ($\|f'\|$) (Fig. 4e, f, Extended Data Fig. 10a). Responses were normalized before computing f' , ensuring that the increase in $\|f'\|$ reflected sharper tuning rather than a trivial consequence of elevated firing rates.

Locomotion also increased the portion of inputs to low-selectivity neurons while decreased the portion of inputs to highly selective neurons. To quantify this effect, we used the normalized in-degree, defined as the in-degree of neurons divided by the mean of in-degree across the population. Neurons were divided into five subpopulations according to their selectivity (Fig. 4g), spanning the lowest selectivity quintile (0–20th percentile) to the highest (80–100th percentile). During locomotion, more selective neurons showed lower normalized in-degree (Fig. 4h), indicating that locomotion preferentially reduced input projections to selective neurons. To simulate this effect in the model, we introduced a parameter β that modulates the probability of input projections to selective neurons in our simulations. This parameter determines how fast in-degree decays with increasing selectivity. A high value of β indicated fast decay, corresponding to fewer input projections to selective neurons and more input projections to less selective neurons (Fig. 4i and Methods). We found that, when connections are feature specific ($F_{fs} > 0$), orientation selectivity increased monotonically with increasing β (Fig. 4j, Extended Data Fig. 10b).

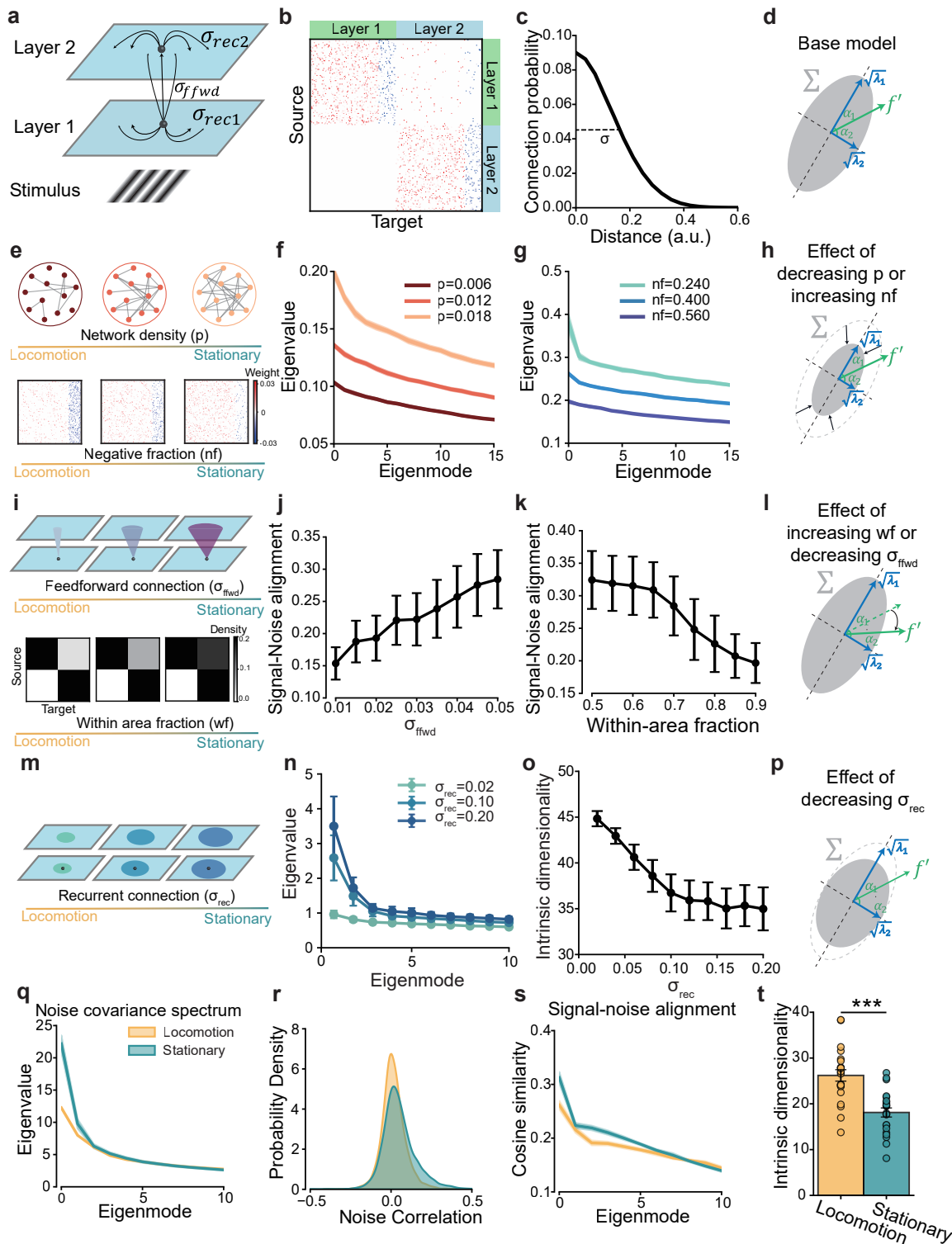


Fig. 3. Localized networks reshape noise covariance to reduce information-limiting correlation. **a**, Schematic of a network model. Neurons in each layer are arranged in a 2D plane. The network consists of two layers. **b**, Example adjacency matrix of the network. Red and blue entries indicate positive and negative values, respectively. **c**, Distance dependence of connection probability. **d**, Fisher information was decomposed into the magnitude of the signal ($\|f'\|$), the magnitude of noise covariance ($\sqrt{\lambda_i}$), and the alignment between signal and dominant noise directions (α_1). **e**, Top, increasing network density from sparse to dense. Bottom, decreasing the fraction of negative connections. **f**, Eigenvalue spectrum of the noise covariance matrix at different network densities. Shading denotes SEM across 20 independently generated network realizations and is too small to be visible. **g**, Eigenvalue spectrum of the noise covariance matrix. **h**, Effect of decreasing p or increasing nf . **i**, Feedforward connection (σ_{fwd}). **j**, Signal-noise alignment. **k**, Signal-noise alignment. **l**, Effect of increasing wf or decreasing σ_{fwd} . **m**, Recurrent connection (σ_{rec}). **n**, Eigenvalue spectrum of the noise covariance matrix at different recurrent connection strengths. **o**, Intrinsic dimensionality. **p**, Effect of decreasing σ_{rec} . **q**, Noise covariance spectrum. **r**, Probability density of noise correlation. **s**, Signal-noise alignment. **t**, Intrinsic dimensionality.

at different fractions of negative connections (nf). Shading denotes SEM across 20 independently generated network realizations and is too small to be visible. **h**, Reduced density or a higher fraction of negative connections reduced noise covariance. **i**, Top, increasing the spatial scale of feedforward connections from narrow to broad. Bottom, decreasing the within-area fraction of connections. **j**, Effect of feedforward connection spatial scale on the alignment between signal direction (f') and dominant noise direction (the dominant eigenvector of the noise covariance matrix). Alignment is defined as the cosine similarity, which is higher when signal and noise are more similarly oriented. Error bars indicate SEM. ($n = 20$ repeats). **k**, Effect of within-area fraction on the alignment between signal direction (f') and dominant noise direction. Error bars indicate SEM. ($n = 20$ repeats). **l**, Within-area, localized networks reduced signal–noise alignment. **m**, Increasing the spatial scale of recurrent connections from narrow to broad. **n**, Eigenvalue spectrum of the noise covariance matrix at different spatial scales of recurrent connectivity. Error bars indicate SEM. ($n = 20$ repeats). **o**, Intrinsic dimensionality decreased as the spatial scale of recurrent connections increased. Error bars indicate SEM. ($n = 20$ repeats). **p**, Localized networks increased the intrinsic dimensionality of noise. **q**, Eigenvalue spectrum of the noise covariance matrix in the experimental data. Shaded areas indicate SEM. ($n = 17$ mice). **r**, Distribution of noise correlations across states ($n = 17$ mice). **s**, Signal–noise alignment, quantified as the cosine similarity between f' and covariance eigenvectors, in the experimental data. Shaded areas indicate SEM ($n = 17$ mice). **t**, Intrinsic dimensionality of neural responses. Points represent individual mice. Error bars indicate SEM. (Paired t-test, $p = 8.86 \times 10^{-7}$, $n = 17$ mice).

This is because an input projection enhances a target neuron’s selectivity only when it originates from a presynaptic neuron that is more selective and shares similar orientation preference (Methods). Highly selective neurons have few such presynaptic source neurons, so most of their inputs reduce selectivity; decreasing their input number therefore increases selectivity. In contrast, low-selectivity neurons have many more-selective presynaptic source neurons, so increasing their input number raises the contribution of selectivity-enhancing inputs and thereby increases selectivity (Extended Data Fig. 11a–c).

Consistent with previous reports (Dadlart and Stryker, 2017; Christensen and Pillow, 2022) and our simulations, locomotion increased single-neuron selectivity in the experimental data: both the direction selectivity index (DSI, drifting gratings) and orientation selectivity index (OSI, static gratings) were significantly higher during locomotion (Fig. 4k, l). This sharpening was also reflected in an increased magnitude of the normalized signal derivative $\|f'\|$ (Fig. 4m, paired t-test, $p = 7.1 \times 10^{-7}$).

These results revealed that sparsification was structured along two axes: it preferentially removed connections between dissimilarly tuned neurons and removed input projections to selective neurons. Both forms of structured sparsification sharpened tuning, increased signal magnitude, and improved coding capacity (Fig. 4n).

Locomotion promotes feedforward organization and accelerates decoding

Beyond enhancing information coding, locomotion also accelerates information transmission (Niell and Stryker, 2010). Fast visual processing has been linked to feedforward network architecture. A feedforward network matched human performance in a rapid visual discrimination task (VanRullen, 2007), whereas models with greater feedback processing better predicted neural responses when categorization was slowed by challenging images (Kar et al., 2019). We asked whether structured sparsification could reconfigure feedforward and feedback organization. We quantified connection directionality between pairs of visual cortical areas using a directionality score measuring the relative prevalence of source-to-target versus target-to-source connections (Methods). During both states, the directionality scores from anatomically lower areas to higher area were generally positive (Fig. 5a), suggesting a feedforward network architecture. While during locomotion, the directionality scores (Methods) for lower-to-higher area pairs became more positive (Fig. 5b, c, paired t-test, $p = 0.002$), indicating strengthened feedforward structure during locomotion. This result was robust across alternative directionality measures, including reachability and canonical correlation analysis (Extended Data Fig. 12a–d).

To test whether a more feedforward network can influence the speed of information coding, we constructed a three-layer hierarchical network model with both feedforward and feedback projections between areas (Fig. 5d, e). We defined feedforwardness as $d_{ff}/(d_{ff} + d_{fb})$, where d_{ff} and d_{fb} are the feedforward and feedback densities, respectively. Feedforwardness ranged from 0 to 1, with 0.5 indicating equal feedforward and feedback densities and 1 indicating exclusively feedforward connectivity. Total inter-areal connection density was held constant at 0.1 while feedforwardness was systematically varied.

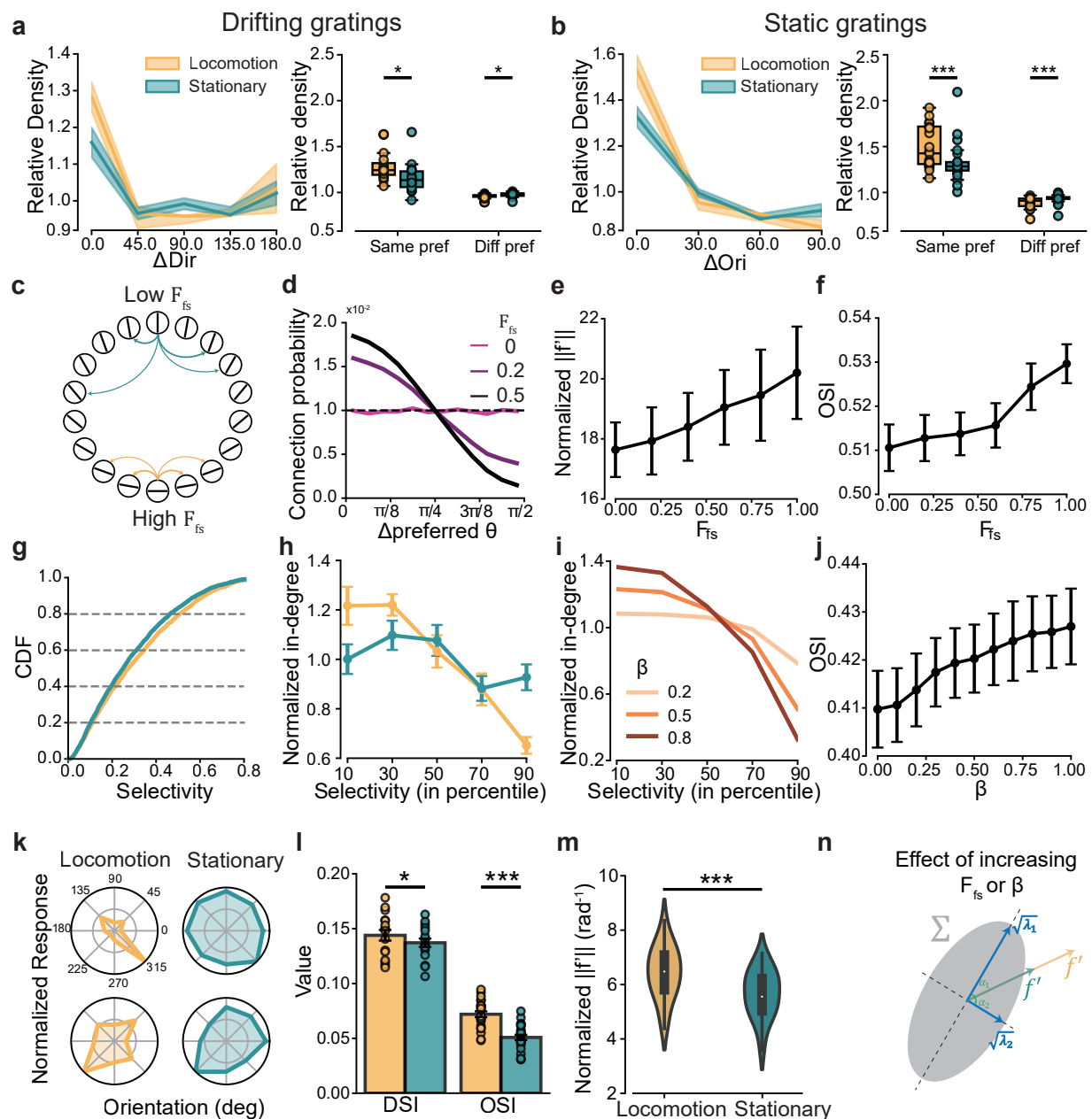


Fig. 4. Locomotion increases feature selectivity through network reconfiguration. **a**, Left: Relative connection density between neuron pairs with different direction preferences during drifting grating stimulation. Shaded regions indicate mean $\pm$ SEM ($n = 17$ mice). Right: Relative density of connections between neurons with the same direction preference ($\Delta Dir = 0$) versus different direction preferences ($\Delta Dir \neq 0$). (same pref; paired t -test, $n = 17$ mice, $p = 0.017$; diff pref; paired t -test, $n = 17$ mice, $p = 0.02$). **b**, Same as in **a**, but based on orientation preference instead of direction preference ($n = 24$ mice). (same pref; paired t -test, $n = 24$ mice, $p = 6.3 \times 10^{-5}$; diff pref; paired t -test, $n = 24$ mice, $p = 5.6 \times 10^{-5}$). **c**, Illustration of feature specificity. Each circle on the ring represents a neuron, and the bar inside each circle indicates its preferred orientation. **d**, Relationship between tuning similarity and connection probability under different levels of feature specificity. **e**, Normalized $\|f'\|$ as a function of feature specificity in simulations. Error bars indicate SEM ($n = 20$ repeats). **f**, In simulations, the orientation selectivity index (OSI) increased with feature specificity. Error bars indicate SEM ($n = 20$ repeats). **g**, Cumulative distribution functions of the selectivity index in stationary and locomotion states. **h**, Normalized in-degree of neurons with different selectivity levels (percentile). Error bars indicate SEM ($n = 17$ mice). **i**, Normalized in-degree of neurons with different selectivity levels under different values of β . Shaded areas indicate SEM ($n = 20$ repeats), which is too small to be visible. **j**, Relationship between β and OSI when $F_{fs} = 0.2$. **k**, Example tuning curves across states for two neurons in the experimental data. Responses were normalized to the maximum stimulus-evoked

response. **l**, Direction selectivity index (DSI) during drifting grating stimulation (paired t -test, $n = 17$ mice, $p = 0.029$) and orientation selectivity index (OSI) during static grating stimulation (paired t -test, $n = 24$ mice, $p = 4.4 \times 10^{-11}$). **m**, Normalized magnitude of f' in the experimental data during drifting gratings stimuli. Neural responses were normalized by z -scoring (paired t -test, $n = 17$ mice, $p = 7.1 \times 10^{-7}$). **n**, Higher feature specificity F_{fs} and larger β increased signal magnitude.

We presented gratings spanning nine orientations to the model (40 repeats each; 200 ms duration) and decoded stimulus orientation from the modeled neural population activity. Decoding latency was defined as the earliest time at which accuracy exceeded the 95th percentile of a shuffled control (threshold=0.19). Increasing feedforwardness monotonically reduced decoding latency (Fig. 5f, g, Extended Data Fig. 13a). Consistent with simulations, we found that decoding latencies were shorter during locomotion in mouse visual cortex (Fig. 5h, i). While faster decoding during locomotion has previously been attributed to gain modulation (Wyrick and Mazzucato, 2021), our results identify a complementary network-level mechanism, namely, a shift toward more feedforward inter-areal organization.

Increasing feedforwardness also produced more direct and less tangled population dynamics, which are associated with smoother, more noise-robust latent dynamics (Russo et al., 2018; Horrocks et al., 2024). We quantified this using two metrics (Fig. 5j and Methods): the distance ratio and the tangling index (Russo et al., 2018; Horrocks et al., 2024). In the model, increasing feedforwardness reduced both metrics (Fig. 5k, l, Extended Data Fig. 13b, c). Experimentally, we observed reductions in both metrics during locomotion (Fig. 5m, n), consistent with the model's predictions and prior work (Horrocks et al., 2024).

These findings suggest that structured sparsification preferentially removed feedback connections, shifting visual cortex toward a more feedforward configuration. This feedforward shift enables more efficient signal propagation to higher-order areas, producing earlier decodable stimulus representations and more direct population dynamics. Together, these effects provide a network-level explanation for the faster information processing observed during locomotion. This reconfiguration may help animals respond more rapidly to visual cues during movement, supporting navigation, predator avoidance, and other survival-critical behaviors.

3 DISCUSSION

We investigated how cortical networks shape sensory coding by examining how reconfiguration of signal-transmission networks was associated with enhanced population coding during locomotion. Locomotion induced a structured sparsification of the single-neuron-resolution signal-transmission networks (Fig. 6a, b): inferred interactions became sparser overall and were reorganized toward a more local and within-area architectures, with greater feature specificity, more inputs to low-selectivity neurons, and increased feedforwardness. Rate-based circuit models linked this structured sparsification to enhanced coding capacity (Fig. 6c): structured sparsification reshaped the signal-noise geometry and temporal dynamics of neural responses to improve the accuracy and speed of coding. Specifically, reduced network density and an increased fraction of negative interactions suppressed shared variability; localized and within-area connectivity reduced signal-noise alignment; feature-specific connectivity and more inputs to low-selectivity neurons increased signal magnitude by sharpening tuning. A more feedforward network organization shortened decoding latency and untangled population trajectories, increasing the speed of coding. These results link the architecture of cortical signal-transmission networks to the geometry and timing of population codes, and suggest that these networks can be flexibly reconfigured under different behavioral and internal states to improve the capacity and speed of information coding.

Running increases visually evoked activity in V1 and improves the reliability of sensory encoding across cortical areas (Niell and Stryker, 2010; Polack et al., 2013; Vinck et al., 2015; Dadarlat and Stryker, 2017; Christensen and Pillow, 2022). These effects have been linked to membrane-potential changes (Polack et al., 2013), neuromodulatory state (McGinley et al., 2015), VIP-SST disinhibitory circuits (Fu et al., 2014), altered interactions between excitatory and inhibitory cell types (Dipoppa et al., 2018), and intrinsic-timescale changes that regulate cortical processing speed (Wyrick and Mazzucato, 2021). Our results extend prior findings by demonstrating that locomotion can change the signal-transmission network between neurons. Locomotion accelerated the signal transmission and sparsified the signal-transmission network (Fig. 1). Furthermore, we revealed that the sparsification is structured, shifting the network towards localized, modularized, feature specific, and feedforward architecture (Fig. 2, Fig. 4 and Fig. 5).

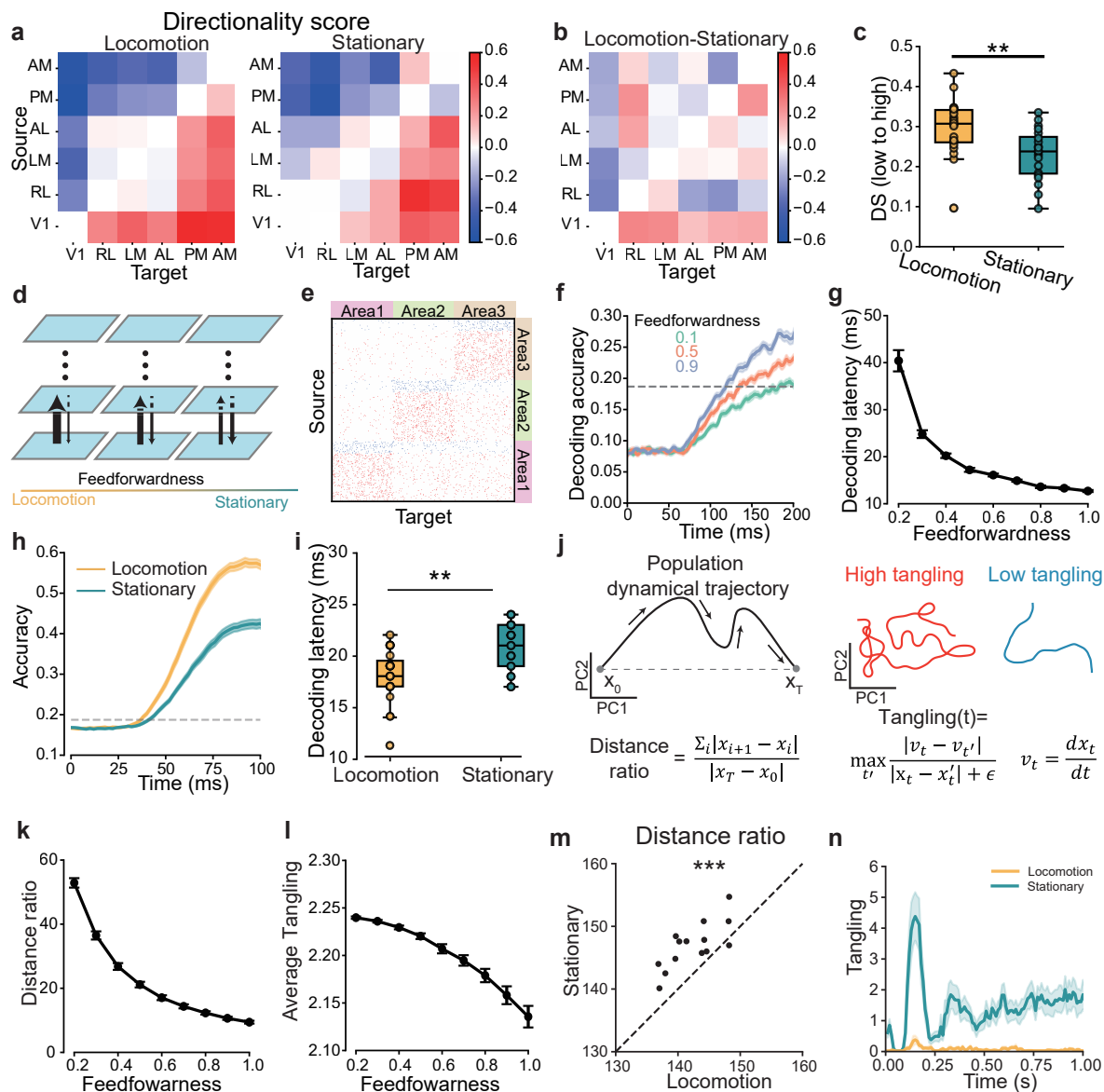


Fig. 5. Feedforward reorganization supports faster visual decoding during locomotion. **a**, Directionality score calculated from the distribution of connection counts for each pair of cortical areas during locomotion (left) and stationary (right) states. Rows indicate source, columns indicate target. **b**, Difference in directionality score between the locomotion and stationary states. **c**, Directionality score (DS) for connections from anatomically lower to higher visual areas. The error distribution was estimated by bootstrap resampling over mice, with 20% of mice sampled in each repeat (paired t -test, $n = 20$ repeats, $p = 0.002$). **d**, Schematic of a multi-area network model. Feedforwardness gradually decreases from left to right. **e**, Example adjacency matrix at feedforwardness = 0.5. **f**, Time course of decoding accuracy in simulations under different values of feedforwardness ($n = 20$ repeats). Shading indicates SEM. The dashed line indicates the 95th percentile of decoding accuracy under random guessing. **g**, Relationship between decoding latency and feedforwardness. Decoding latency was defined as the first time point at which decoding accuracy exceeded the 95th percentile of decoding accuracy under random guessing. ($n = 20$ repeats). **h**, Time course of decoding accuracy in the experimental data during static grating presentation. Shading indicates SEM ($n=17$ mice). **i**, Decoding latency in the experimental data. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times \text{IQR}$; points represent mice (Wilcoxon signed-rank test, $n = 17$ mice, $p = 0.006$). **j**, Schematic illustration of the distance ratio and tangling index. **k**, In simulations, the distance ratio decreased as feedforwardness increased. Error bars indicate SEM ($n = 20$ repeats). **l**, In simulations, the tangling index decreased as feedforwardness increased. Error bars indicate SEM ($n = 20$ repeats). **m**, Distance ratio in the experimental data during locomotion and stationary states (paired t -test, $n = 17$ mice, $p = 3.5 \times 10^{-5}$). **n**, Tangling index of neural trajectories in the experimental data during locomotion and stationary states. Shaded areas indicate SEM ($n = 17$ mice).

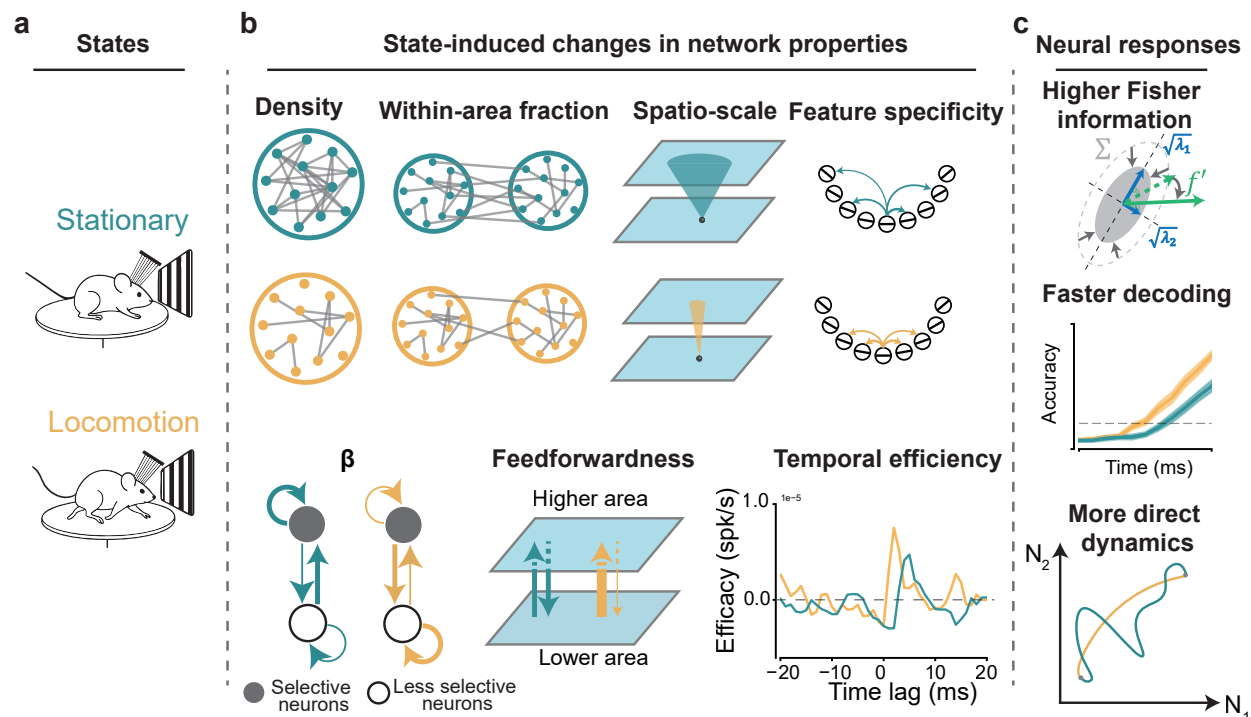


Fig. 6. Overview of locomotion-induced reconfiguration of signal-transmission networks and its effects on Fisher information and neural dynamics. **a**, Illustration of stationary and locomotion states. **b**, Behavioral state-induced changes in network properties, including decreased density, a higher within-area fraction, a narrower spatial scale, heightened feature specificity, reduced incoming projection to highly selective neurons, increased feedforwardness, and improved temporal efficiency of signal transmission during the locomotion state. **c**, Effects of network changes on coding. Top: changes in the network properties shown in **b** result in reduced noise covariance, larger signal magnitude, and signal–noise misalignment, leading to increased Fisher information. Middle: network changes lead to higher decoding accuracy and shorter decoding latency during locomotion. Bottom: network reconfiguration straightened neural dynamics.

Our results attribute the locomotion-induced improvement in coding to changes in network structure. Locomotion enhances coding in the visual cortex (Dadgarlat and Stryker, 2017; Christensen and Pillow, 2022), but the mechanisms underlying this improvement remain incompletely understood (Mineault et al., 2016; Christensen and Pillow, 2022; Horrocks et al., 2024). Previous studies have linked enhanced information coding to reduced shared variability (Dadgarlat and Stryker, 2017), reduced signal–noise alignment (Horrocks et al., 2024), and improved selectivity and response reliability (Dadgarlat and Stryker, 2017; Christensen and Pillow, 2022). Our results provide a network-level mechanism through which these changes can arise. Network sparsification and an increased fraction of negative interactions reduced noise correlations (Fig. 3), network localization reduced signal–noise alignment (Fig. 3), and feature-specific connectivity and reduced input projection to selective neurons increased selectivity (Fig. 4). Together, these findings suggest that prior explanations for improved coding during locomotion can be unified within a single framework: locomotion reconfigures the structure of neuronal interactions.

Our results suggest that the spatial scale of connectivity can be flexibly reconfigured to enhance information coding. Connection probability between neuronal pairs typically decays with anatomical distance, but this distance dependence is more than a physical constraint. The spatial scale of connectivity shapes the stability of balanced firing rates (Rosenbaum and Doiron, 2014), the spatial structure of noise correlations (Rosenbaum et al., 2017), the dimensionality of shared variability (Huang et al., 2019), dynamical patterns, and linear Fisher information (Huang et al., 2022). We demonstrated that locomotion changed the spatial scale of connectivity, shifting the network towards a more localized structure (Fig. 2). Furthermore, our simulations showed that such localization reduced signal–noise alignment and increased the dimensionality of shared variability, thereby reducing the information-limiting effects of noise correlations (Moreno-Bote et al., 2014). These findings indicate that spatial scale is not only a fixed architectural feature, but also a state-dependent network parameter that can regulate population coding (Rosenbaum and Doiron, 2014; Rosenbaum et al., 2017; Huang et al., 2019, 2022). Our simulations of how spatial scale influences Fisher information yield results that differ from previous work, which assumed a pinwheel orientation-preference map and negligible internal noise (Huang et al., 2022). By contrast, our model adopts the salt-and-pepper organization characteristic of mouse visual cortex (Ohki and Reid,

2007; Ringach et al., 2016; Ohki et al., 2005), and includes per-neuron noise to capture non-visual variability, including movement-related signals, which are substantial in mouse visual cortex (Musall et al., 2019; Akella et al., 2025; Stringer et al., 2019). Therefore, our results apply to a different set of assumptions from those considered in the previous study.

The feedforward component of this reorganization offers a network-level explanation of why locomotion accelerates visual processing. Previous work attributed this acceleration to gain modulation (Wyrick and Mazzucato, 2021). In our model, increasing feedforwardness reduced decoding latency and produced straighter, less tangled neural trajectories. This does not imply that feedforward processing is generally superior to recurrent processing. Feedforward propagation supports rapid sensory discrimination, whereas recurrent processing is important when stimuli are ambiguous or occluded, or when they require contextual integration (VanRullen, 2007; Kar et al., 2019). Locomotion may therefore shift the visual system toward a regime suited to fast extraction of salient sensory input, potentially supporting faster responses to threatening stimuli, while other states favor broader recurrent integration with contextual or internally generated activity.

More generally, structured sparsification and increased feedforward organization suggest that cortical networks adapt their communication architecture to behavioral context and internal state. Recent work has shown that visual cortical variability is not best explained by a fixed stimulus code, but rather by state-dependent changes in the relative contributions of visual input, behavior, and internal brain dynamics (Akella et al., 2025), and theoretical explanations of interareal communication have proposed that information is routed through selected population subspaces or dynamically regulated channels (Semedo et al., 2019; Palmigiano et al., 2017). Our results connect these two views: state-dependent encoding can be reflected in a selective reconfiguration of the interaction graph at single-neuron resolution, which in turn determines which population subspaces are accessible to downstream circuits. The structured sparsification observed during locomotion (Fig. 1, Fig. 2, Fig. 4, and Fig. 5) may reflect a broader network mechanism for improving cortical information coding. Because enhanced coding is also associated with attention (Cohen and Maunsell, 2009; Harris and Thiele, 2011; Saproo and Serences, 2010), arousal (Vinck et al., 2015; Papadopoulos et al., 2026), learning (Yan et al., 2014; Jeanne et al., 2013; Sanayei et al., 2018), and development (Smith et al., 2015; Chang and Fitzpatrick, 2022), our results suggest that structured sparsification could provide a common network-level principle linking these diverse behavioral and developmental states. Consistent with this possibility, attention, learning, and development are often accompanied by decorrelation of neural activity (Mitchell et al., 2009; Cohen and Maunsell, 2009; Gu et al., 2011; Jeanne et al., 2013; Sanayei et al., 2018; Smith et al., 2015; Chang and Fitzpatrick, 2022). Thus, the network reconfiguration identified here offers a testable framework for determining whether enhanced information processing in these contexts arises from structured sparsification.

The present study has several limitations. First, the inferred networks should be viewed as signal-transmission networks rather than anatomical networks or effective networks. Jitter-corrected cross-correlograms isolate fast-timescale spike-timing relationships and reduce slow shared fluctuations, but significant short-lag interactions can still arise from monosynaptic coupling, polysynaptic pathways, common input, unobserved intermediate neurons, or state-dependent coordination among distributed populations (Smith and Kohn, 2008; Stark and Abeles, 2009; Jia et al., 2013, 2022); by the same logic, negative CCG peaks should not be equated directly with inhibitory synapses. Our conclusions therefore concern state-dependent reconfiguration of fast-timescale communication, not changes in anatomical wiring or direct causal links.

Second, locomotion is a coarse label for a richer behavioral and physiological state. Running covaries with arousal, neuromodulatory tone, eye movements, whisking, facial movements, and other uninstructed behaviors (Vinck et al., 2015; Reimer et al., 2014, 2016; Stringer et al., 2019; Musall et al., 2019; Akella et al., 2025). Our binary locomotion-stationary classification captures a robust and ethologically meaningful transition, but it does not identify which behavioral or neuromodulatory variables drive structured sparsification.

These constraints motivate several complementary experimental directions. First, causal tests would combine large-scale recordings with optogenetic perturbations to determine whether effective networks are reconfigured according to the same structured-sparsification framework. Second, future behavioral-state analyses could infer signal-transmission networks across a more continuous state manifold, incorporating running, pupil size, eye position, facial motion, body movement, and neuromodulatory signals, to clarify which dimensions of state drive the reconfiguration. Future studies could test whether structured sparsification is cell-type specific. Because locomotion modulates firing rates differently across cell types (Dipoppa et al., 2018; Polack et al., 2013), it may also reshape functional connectivity in a cell-type-dependent manner. Identifying the neuronal populations that drive these changes would clarify whether structured sparsification reflects a global network-state shift or selective reconfiguration of certain excitatory or inhibitory circuits. Moreover, the biological mechanisms underlying this structured sparsification remain unclear. Locomotion-induced changes in neural activity have been attributed to neuromodulatory systems (McGinley et al., 2015) and cell-type-specific modulation (Polack et al., 2013; Dipoppa et al., 2018). Whether these mechanisms can account for the observed structure of sparsification, and through what cellular or circuit-level processes, remains an open question. Addressing it will require careful cell type and neuromodulation monitoring, which is an important direction for future work.

In summary, structured sparsification is a candidate mechanism by which behavioral state reconfigures the architecture of cortical communication to improve coding. By linking experimentally inferred signal-transmission networks to signal-noise geometry and neural dynamics, our results provide a framework for understanding how network architecture shapes the capacity and timing of population codes.

4 METHODS

Dataset

The data analyzed and discussed in this paper are part of the publicly released Allen Institute Brain Observatory Neuropixels dataset (Siegle et al., 2021). The Neuropixels project uses high-density extracellular probes to record neuronal spiking activity simultaneously from multiple regions of the mouse brain. In this work, we analyzed recordings obtained from six Neuropixels probes targeting six visual cortical areas (V1, LM, RL, AL, PM, AM) and two subcortical areas (LGN, LP). During data acquisition, mice were head-fixed but allowed to run freely on a wheel equipped with an encoder that continuously tracked running speed and distance. While recording, mice passively viewed six categories of visual stimuli: gray screens, full-field flashes, drifting gratings, static gratings, natural scenes, and natural movies. Trials were classified as locomotion if the mean running speed during stimulus presentation exceeded 1 cm/s, and as stationary otherwise. Sessions were excluded if trials in either state accounted for less than 20% of all trials. To balance behavioral conditions, we subsampled trials such that the number of trials in the two states was equal within each session. Only neurons with mean firing rates above 2 Hz during stimulus presentation in both states were included. Following these criteria, an average of 228 ± 61 units per mouse were retained for analysis.

Statistics and data analyses

We used Python as the primary programming language for data analyses, together with analytical tools including SciPy (Virtanen et al., 2020). For pairwise comparisons between locomotion and stationary states measured across mice ($n = 17$ mice for drifting gratings, $n = 24$ mice for static gratings), the Wilcoxon signed-rank test was used for non-normally distributed metrics (e.g., network density, peak offset, peak width, modularity) and paired t -tests were used for metrics whose paired differences were approximately normally distributed (e.g., relative density changes between within-area and between-area connections, selectivity indices, distance ratio). Error bars, unless otherwise specified, represent standard error of the mean (SEM). For the feedforwardness analysis, error distributions were estimated by bootstrap resampling over mice ($n = 20$ repeats, 20% of mice per repeat), and significance was assessed using a paired t -test. Decoding performance was evaluated using multinomial logistic regression with repeated stratified 5-fold cross-validation (4 repeats). For modularity analysis, 50 surrogate networks per session were generated using a pair-preserving null model, and only modules containing at least four neurons were retained.

Firing rate-matching control

To test whether locomotion-induced firing-rate modulation accounted for differences in network density, we performed a firing-rate-matching control. For each neuron, spikes were randomly subsampled from the behavioral state with the higher firing rate until its firing rate matched that in the lower-rate state. Thus, for most neurons, spikes were subsampled from locomotion trials to match the stationary firing rate, whereas for the small fraction of neurons with higher stationary firing rates, spikes were subsampled from stationary trials to match the locomotion firing rate. We then inferred signal-transmission networks from the rate-matched spike trains using jitter-corrected CCGs and compared network densities across behavioral states.

Connectivity

We analyzed interactions between pairs of simultaneously recorded neurons by calculating the spike-train cross-correlogram (CCG) (Jia et al., 2022; Siegle et al., 2021; Smith and Kohn, 2008; Tang et al., 2024). For a pair of neurons with spike trains X and Y , the CCG at lag τ was defined as

$$\text{CCG}_{XY}(\tau) = \frac{\frac{1}{K} \sum_{k=1}^K \sum_{t \in \mathcal{T}_\tau} X_{t,k} Y_{t+\tau,k}}{\Theta(\tau) \sqrt{\lambda_X \lambda_Y}}, \quad (2)$$

where K is the number of trials, T is the number of time bins in each trial, $X_{t,k}$ and $Y_{t,k}$ are the spike counts of the two neurons at time bin t on trial k , and $\mathcal{T}_\tau = \{t : 1 \leq t \leq T, 1 \leq t + \tau \leq T\}$ denotes the set of valid time bins for lag τ . The

parameters λ_X and λ_Y are the mean firing rates of the two neurons. The factor $\Theta(\tau)$ is the triangular correction for the number of overlapping time bins at lag τ .

The jitter-corrected CCG was computed by subtracting the expected CCG produced by spike trains with spike times randomly jittered within a jitter window (Harrison and Geman, 2009; Smith and Kohn, 2008). Jitter-corrected CCGs can be computed analytically (Harrison and Geman, 2009). For spike train X , let W_s index be the windows of length L , and let

$$X_{s,k}^L = \sum_{t \in s} X_{t,k} \quad (3)$$

be the spike count in jitter window s on trial k . The across-trial mean count in that jitter window was

$$\bar{X}_s^L = \frac{1}{K} \sum_{k=1}^K X_{s,k}^L. \quad (4)$$

We normalized the binned spike counts as

$$\tilde{X}_{s,k}^L = X_{s,k}^L / \bar{X}_s^L. \quad (5)$$

After expanding $\tilde{X}_{s,k}^L$ back to the original time resolution by assigning the same value to each time bin within window s , the PSTH-jitter mean for spike train X was

$$E(X_{t,k}^L | X) = \bar{X}_t \tilde{X}_{s(t),k}^L, \quad (6)$$

where $s(t)$ denotes the jitter window containing time bin t , and $\bar{X}_t = \frac{1}{K} \sum_{k=1}^K X_{t,k}$. The same procedure was applied to spike train Y .

The jittered CCG was then computed as

$$\text{CCG}_{XY}^{\text{jittered}}(\tau) = \frac{\frac{1}{K} \sum_{k=1}^K \sum_{t \in \mathcal{T}_\tau} E(X_{t,k}^L | X) E(Y_{t+\tau,k}^L | Y)}{\Theta(\tau) \sqrt{\lambda_X \lambda_Y}}. \quad (7)$$

Finally, the jitter-corrected CCG was

$$\text{CCG}_{XY}^{\text{jitter-corrected}}(\tau) = \text{CCG}_{XY}(\tau) - \text{CCG}_{XY}^{\text{jittered}}(\tau). \quad (8)$$

Consistent with our previous studies, we chose a jitter window of $L = 25$ ms. A positive sharp peak was deemed significant if the maximum of the jitter-corrected CCG amplitude within the 0–13 ms time window exceeded $\mu + 4\sigma$. A negative sharp peak was deemed significant if the minimum of the jitter-corrected CCG amplitude within the same time window was lower than $\mu - 4\sigma$. Here, μ and σ indicate the mean and standard deviation of jitter-corrected CCG values during $[-100 \text{ ms}, -50 \text{ ms}] \cup [50 \text{ ms}, 100 \text{ ms}]$.

Directionality scores

The directionality score was computed following Siegle et al. (2021). For each pair of visual areas, we denoted the number of significant CCG connections from source area (area A) to target area (area B) as C_{BA} and the number of connections from target area (area B) to source area (area A) as C_{AB} . The directionality score from area A to area B was defined as

$$DS_{BA} = \frac{C_{BA} - C_{AB}}{C_{BA} + C_{AB}} \quad (9)$$

Thus, $DS_{BA} > 0$ indicates that there are more connections from area A to area B, while $DS_{BA} < 0$ indicates otherwise. We also computed a directionality score based on reachability. A target neuron was considered reachable from a source neuron if there was a directed path, consisting of one or more directed connections, from the source to the target. The number of reachable pairs from area A to area B was R_{BA} , and the reachability-based directionality score was

$$DS_{BA}^R = \frac{R_{BA} - R_{AB}}{R_{BA} + R_{AB}} \quad (10)$$

Canonical correlation analysis (CCA) was also used to quantify directional population interactions between brain areas, following Semedo et al. (2022). For each ordered pair of visual areas, we computed the canonical correlation between population activity vector in area A $x_A(t)$ and the time-lagged population activity vector in area B $x_B(t + \tau)$. We used the first canonical correlation as the interaction strength between the two areas. For each direction, we summed the first canonical correlations across positive lags to obtain the total lagged interaction strength from area A to area B,

$$S_{BA} = \sum_{\tau>0} \rho_{BA}(\tau), \quad (11)$$

where $\rho_{BA}(\tau)$ is the first canonical correlation between activity in area A at time t and activity in area B at time $t + \tau$. Similarly, S_{AB} denotes the summed lagged interaction strength in the opposite direction. The CCA-based directionality score was defined as

$$DS_{BA}^{CCA} = \frac{S_{BA} - S_{AB}}{S_{BA} + S_{AB}}. \quad (12)$$

We computed this score separately for locomotion and stationary trials, using matched trial counts across behavioral states. Spike counts were extracted in 10-ms windows with 5-ms steps, and lags from 5 to 45 ms were evaluated for each ordered pair of visual areas.

Network model

The network consisted of multiple layers representing different areas along the visual hierarchy. Neurons in each layer were arranged on a uniform grid covering a unit square $\Gamma = [0, 1] \times [0, 1]$. Each layer contained N neurons, with 80% excitatory and 20% inhibitory. To simplify the representational landscape of image space, each neuron was tuned to a single angular feature representing the orientation of static gratings, randomly and uniformly sampled in $[0, \pi]$. The preferred orientation of neuron i was denoted by $\hat{\theta}_i$. We assumed that the input from LGN can be modeled as

$$I_i(\theta) = I_0 + H_i \exp(\kappa_i \cos(2(\theta - \hat{\theta}_i))), \quad (13)$$

where κ_i controls the sharpness of this LGN input tuning. Unless otherwise stated, we set $I_0 = 1.0$, $H_i = 1$, and $\kappa_i = 1.5$. Neuronal activity followed a rate equation:

$$\tau \frac{dr_i}{dt} = -r_i + \sum_j J_{ij} \phi(r_j) + I_i(\theta) + \xi_i \quad (14)$$

where $\tau = 10$ is the membrane time constant, J_{ij} is the connection from neuron j to neuron i , and ξ_i is an independent Gaussian noise with $\xi_i \sim \mathcal{N}(0, \sigma_{\text{noise}}^2)$ and $\sigma_{\text{noise}} = 1.5$. The transfer function was $\phi(x) = \max(0, x)$. Synaptic strengths were set to $J_{ee} = 0.06$, $J_{ie} = 0.06$, $J_{ei} = -0.24$, and $J_{ii} = -0.28$.

Distance-dependent connectivity. The total connection density within layer n is d_{nn} . The cross-layer connection density from source layer m to target layer n is denoted d_{nm} . Within each layer, the connection probability from neuron j to neuron i is distance-dependent:

$$p_{ij} \propto d_{nn} \exp(-\|x_i - x_j\|^2 / (2\sigma_{\text{rec}}^2)) \quad (15)$$

where x_i and x_j denote the spatial coordinates of neurons i and j , respectively, and σ_{rec} controls the spatial scale of recurrent connectivity. Feedforward connections from layer 1 to layer 2 also followed the same distance-dependent rule, but with a different spatial scale parameter σ_{ffwd} . Following the convention above, for a connection from neuron j in source layer m to neuron i in target layer n , the connection probability is

$$p_{ij} \propto d_{nm} \exp(-\|x_i - x_j\|^2 / (2\sigma_{\text{ffwd}}^2)). \quad (16)$$

Feature-specific connectivity. Feature-specific connectivity was achieved by making the connection probability depend on feature similarity:

$$p_{ij} \propto 1 + F_{fs} \cos(2(\hat{\theta}_i - \hat{\theta}_j)) \quad (17)$$

with F_{fs} controlling the strength of feature specificity ($0 \leq F_{fs} \leq 1$). $\hat{\theta}_i$ and $\hat{\theta}_j$ are the preferred orientation of neuron i and neuron j , respectively.

Taken together, the connection probability from neuron j in source layer m to neuron i in target layer n is

$$p_{ij} = d_{nm} \exp(-\|x_i - x_j\|^2 / (2\sigma_a^2)) (1 + F_{fs} \cos(2(\hat{\theta}_i - \hat{\theta}_j))) / C \quad (18)$$

where C is the normalization term.

$$C = \frac{1}{N_n N_m} \sum_{i \in n} \sum_{j \in m} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) (1 + F_{fs} \cos(2(\hat{\theta}_i - \hat{\theta}_j))) \quad (19)$$

N_n and N_m denote the numbers of neurons in layers n and m , respectively, both of which equal N . $\sigma_a = \sigma_{\text{rec}}$ when $n = m$, and $\sigma_a = \sigma_{\text{ffwd}}$ otherwise.

Increased number of input projections to weakly selective neurons. To model the observation that highly selective neurons receive fewer recurrent inputs whereas non-selective neurons receive more recurrent inputs, we introduced a hyperparameter β . Specifically, each neuron's κ_i (see equation (13)) was independently sampled from a gamma distribution,

$$\kappa_i \sim \text{Gamma}(\alpha, \theta_\kappa), \quad (20)$$

where we used shape $\alpha = 0.7$ and scale $\theta_\kappa = 1.0$.

Let P_{ij}^{base} denote the baseline probability of a recurrent connection from presynaptic neuron j to postsynaptic neuron i , after incorporating the spatial and feature-specific connectivity terms. We defined a postsynaptic targeting factor based on the κ_i value:

$$\rho_i = \exp(-\beta \kappa_i), \quad (21)$$

$\beta = 0$ indicate that all neurons have the same number of input projections, while higher β corresponds to increased the number of inputs to weakly selective neurons and reduced number of inputs to highly selective neurons. The unnormalized biased connection probability from neuron j to neuron i is

$$\tilde{P}_{ij} = P_{ij}^{\text{base}} \rho_i. \quad (22)$$

For each presynaptic neuron j , we then normalized the probabilities across all valid postsynaptic targets to preserve the original expected out-degree of neuron j :

$$P_{ij} = \tilde{P}_{ij} \frac{\langle P_{ij}^{\text{base}} \rangle_{i \in \mathcal{V}_j}}{\langle \tilde{P}_{ij} \rangle_{i \in \mathcal{V}_j}}, \quad (23)$$

where $\mathcal{V}_j$ denotes the valid postsynaptic targets for presynaptic neuron j , excluding self-connections for recurrent connectivity and $\langle \rangle_{i \in \mathcal{V}_j}$ denotes the average over the valid postsynaptic targets of neuron j .

Thus, $\beta = 0$ leaves the baseline connection probabilities unchanged. For $\beta > 0$, neurons with larger κ_i have smaller targeting factors ρ_i and therefore receive fewer recurrent inputs on average, whereas neurons with smaller κ_i receive more recurrent inputs. The targeting factor modulates only the probability that a postsynaptic neuron is targeted; it does not directly modify synaptic weights. After normalization, probabilities were clipped to the valid range $[0, 1]$ when necessary.

Feedforwardness of inter-area connectivity. We introduced a feedforwardness parameter f that controls the the fraction of feedforward connections.

$$f = \frac{d_{\text{ff}}}{d_{\text{ff}} + d_{\text{fb}}}, \quad (24)$$

where d_{ff} and d_{fb} denote the densities of feedforward and feedback inter-area connections, respectively. Thus, $f = 1$ corresponds to a purely feedforward architecture, $f = 0.5$ to balanced feedforward and feedback connectivity, and $f = 0$ to a purely feedback architecture. Feedforward and feedback projections used the same spatial and feature-dependent connectivity rules; f controls only projection direction.

Simulations. All simulations were performed by numerically integrating the rate equations with a forward Euler method. For each parameter setting, results were averaged across 20 independently generated network realizations. We used nine stimulus orientations ranging uniformly from 0 to π . Forty trials were simulated per stimulus orientation with independent Gaussian noise on each trial and time step. Activity was initialized at zero and simulated for 200 time steps; the average response over this period was used as the trial response. For time-resolved analyses of trajectory geometry and decoding latency, simulations additionally included a 50-time-step prestimulus baseline and a 250-time-step poststimulus period.

Intrinsic dimensionality

Intrinsic dimensionality was computed as the participation ratio of eigenvalues of the noise covariance matrix. Mathematically, let λ_i be the i th eigenvalue of the noise covariance matrix Σ . Intrinsic dimensionality is

$$PR = (\sum_i \lambda_i)^2 / \sum_i \lambda_i^2 \quad (25)$$

Linear Fisher information

To quantify the stimulus information available to an optimal *linear* decoder, we computed the linear Fisher information around a reference stimulus value θ_0 as

$$I(\theta_0) = f'(\theta_0)^\top \Sigma(\theta_0)^{-1} f'(\theta_0), \quad (26)$$

where $f(\theta) = \mathbb{E}[r | \theta] \in \mathbb{R}^N$ denotes the mean population response (tuning curve) and $f'(\theta_0) = \frac{df}{d\theta}|_{\theta_0}$ is the vector of tuning-curve derivatives with respect to the stimulus parameter. $\Sigma(\theta_0) \in \mathbb{R}^{N \times N}$ is the noise covariance of the trial-to-trial response variability around θ_0 .

In practice, we computed the generalized linear Fisher information (Kafashan et al., 2021). We estimated $f'(\theta_0)$ and $\Sigma(\theta_0)$ by symmetric finite differences using two nearby stimulus values $\theta_{1,2} = \theta_0 \pm \delta\theta/2$:

$$f'(\theta_0) \approx \frac{f(\theta_1) - f(\theta_2)}{\delta\theta}, \quad \Sigma(\theta_0) \approx \frac{\text{cov}(r | \theta_1) + \text{cov}(r | \theta_2)}{2}. \quad (27)$$

This approximation treats the covariance as locally constant over the interval $\delta\theta$ and reduces bias relative to a one-sided difference. We then bias-corrected the estimated linear Fisher information (Kanitscheider et al., 2015a):

$$\hat{I}_{bc}(\theta_0) = \hat{f}'(\theta_0)^\top \hat{\Sigma}(\theta_0)^{-1} \hat{f}'(\theta_0) \cdot \frac{2T - N - 3}{2T - 2} - \frac{2N}{T \delta\theta^2}, \quad (28)$$

where T is the number of trials and N is the number of neurons.

Eigenmode and geometric interpretation of linear Fisher information. To provide a geometrically interpretable decomposition, we expressed linear Fisher information in the eigenbasis of the covariance. For the noise covariance matrix $\Sigma(\theta_0)$, let

$$\Sigma(\theta_0) = U \Lambda U^\top, \quad \Sigma(\theta_0)^{-1} = U \Lambda^{-1} U^\top, \quad (29)$$

where $U = [u_1, \dots, u_N]$ contains the eigenvectors and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$ contains the eigenvalues (noise variances along the principal axes). Expressing the signal direction in this eigenbasis gives

$$f'(\theta) = \sum_i a_i u_i, \quad a_i = u_i^\top f'(\theta), \quad (30)$$

so that Fisher information can be written as

$$I(\theta) = \sum_i \frac{a_i^2}{\lambda_i} = \|f'(\theta)\|^2 \sum_i \frac{\cos^2 \alpha_i}{\lambda_i}, \quad (31)$$

where α_i is the angle between $f'(\theta)$ and the i th covariance eigenvector. Thus, linear Fisher information is determined by three geometric factors: the magnitude of the signal, $\|f'(\theta)\|$; the magnitude of noise along each covariance axis, λ_i ; and the alignment of the signal with high- or low-variance noise directions, quantified by $\cos^2 \alpha_i$.

Modularity

The modularity for directed networks (Leicht and Newman, 2008) is defined as

$$\hat{Q} = \frac{1}{m} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i^{\text{in}} k_j^{\text{out}}}{m} \right) \delta(\sigma_i, \sigma_j), \quad (32)$$

where A is the adjacency matrix of the network, m is the number of edges, and k_i^{in} and k_i^{out} represent the in-degree and out-degree of node i , respectively. δ is the Kronecker delta function, and σ_i denotes the community assignment of node i . $\delta(\sigma_i, \sigma_j) = 1$ if and only if nodes i and j are assigned to the same community. γ is a resolution parameter that controls the size of modules. We ignored connection strength as well as the sign of connections by only considering the existence of edges. The Louvain method is a community-detection algorithm for partitioning networks into groups of nodes with dense connections within groups and sparse connections between groups (Blondel et al., 2008). The algorithm uses the original modularity $\hat{Q}$ as a quality function to optimize the partitioning of the network. The Louvain method operates through a series of iterative steps that merge neighboring modules to maximize modularity gain until a locally optimal partition is reached. The algorithm follows a bottom-up approach, starting from single-node modules and iteratively merging modules to form larger ones. The Louvain method was used to compute modularity.

To determine the resolution parameter for module analysis, we obtained a modularity-difference curve by varying γ and computing the difference between the modularity of empirical and surrogate networks generated by the pair-preserving model, and then selected the value of γ that maximized this difference (Harris et al., 2019; Tang et al., 2024). The pair-preserving model randomizes the network while preserving network size, density, degree distribution, and neuron pair distribution (Milo et al., 2002; Tang et al., 2024). This procedure yielded the module partition that deviated most from the reference model. We used the Z-score of modularity to quantify the strength of modular organization relative to a reference model. The Z-score of modularity is defined as

$$Z_Q = \frac{\hat{Q} - \langle q \rangle}{\sqrt{\langle q^2 \rangle - \langle q \rangle^2}}. \quad (33)$$

Here, q denotes the modularity obtained from one realization of the reference model. We randomly generated 50 surrogate networks for each session. Only modules with a size of at least four neurons were included in subsequent analyses to eliminate noise contributions from isolated single neurons, pairs, and triplets.

Selectivity index

In total, eight different directions of drifting gratings, evenly spaced in $[0, 2\pi)$, and six different orientations of static gratings, evenly spaced in $[0, \pi)$, were presented to the mice. We calculated each neuron's direction selectivity index (DSI) from its drifting-grating tuning curve and its orientation selectivity index (OSI) from its static-grating tuning curve as follows:

$$\text{DSI} = \left| \frac{\sum r(\theta) e^{i\theta}}{\sum r(\theta)} \right|, \quad (34)$$

$$\text{OSI} = \left| \frac{\sum r(\theta) e^{2i\theta}}{\sum r(\theta)} \right|, \quad (35)$$

where θ is the stimulus direction or orientation and $r(\theta)$ is the mean neuronal response for a given stimulus parameter θ .

Selectivity-enhancing input projections

Consider a population of neurons responding to external stimuli parameterized by their orientation, θ . The response of neuron m to a stimulus with orientation θ is denoted by $r_m(\theta)$. When neuron m receives an input projection from neuron n , its modified response is modeled as

$$r'_m(\theta) = r_m(\theta) + j r_n(\theta), \quad (36)$$

where j denotes the effective coupling strength. We define the response-weighted coupling strength as

$$\alpha = j \frac{R_n}{R_m}, \quad (37)$$

where R_m and R_n denote the response amplitudes of neurons m and n , respectively. We also define the difference between their preferred orientations as

$$\Delta\theta = \theta_n - \theta_m, \quad (38)$$

where θ_m and θ_n are the preferred orientations of neurons m and n , respectively. The modified orientation selectivity of neuron m is then given by

$$s'_m = \frac{|s_m + \alpha s_n \exp(2i\Delta\theta)|}{1 + \alpha}. \quad (39)$$

Equivalently, the exact post-input selectivity is

$$s'_m = \frac{\sqrt{s_m^2 + \alpha^2 s_n^2 + 2\alpha s_m s_n \cos(2\Delta\theta)}}{1 + \alpha}. \quad (40)$$

For weak coupling,

$$|\alpha| \ll 1, \quad (41)$$

a first-order expansion of Eq. (40) in α gives

$$s'_m \simeq s_m + \alpha [s_n \cos(2\Delta\theta) - s_m]. \quad (42)$$

Therefore, the first-order change in the selectivity of neuron m is

$$\Delta s_m \equiv s'_m - s_m \simeq \alpha [s_n \cos(2\Delta\theta) - s_m]. \quad (43)$$

The excitatory input projection enhances the selectivity of neuron m when

$$\Delta s_m > 0, \quad (44)$$

or, equivalently,

$$s_n \cos(2\Delta\theta) > s_m. \quad (45)$$

This condition can be satisfied only if the presynaptic neuron is more selective than the postsynaptic neuron:

$$s_n > s_m. \quad (46)$$

When $s_n > s_m$, their preferred orientations must also be sufficiently aligned:

$$|\theta_n - \theta_m| < \frac{1}{2} \cos^{-1} \left(\frac{s_m}{s_n} \right). \quad (47)$$

We refer to input projections that satisfy these selectivity and alignment conditions as selectivity-enhancing projections.

An inhibitory input projection enhances selectivity when

$$s_n \cos(2\Delta\theta) < s_m. \quad (48)$$

Distance ratio

The distance ratio was computed as previously reported (Horrocks et al., 2024). We denote the neural response at time t as r_t , which is an $N \times 1$ vector, with N being the number of neurons. $R = [r_1, \dots, r_T]$. We applied principal component analysis to R and retained the principal dimensions that explain 90% of total variance, yielding a lower-dimensional representation matrix X of size $m \times T$, where m is the number of dimensions. x_t is the position of the neural state in m -dimensional neural space at time t . To quantify how direct the trial-averaged population trajectory was as it moved between stimulus onset and stimulus offset, we computed a distance ratio. This ratio is the cumulative path length traveled by the trajectory between the two endpoints, divided by the Euclidean distance between those same endpoints.

$$\begin{aligned} \text{Cumulative distance} &= \sum_{t=T_{\text{start}}+1}^{T_{\text{end}}} \|x_t - x_{t-1}\| \\ \text{Direct distance} &= \|x_{T_{\text{end}}} - x_{T_{\text{start}}}\| \\ \text{Distance Ratio} &= \frac{\text{Cumulative distance}}{\text{Direct distance}} \end{aligned} \quad (49)$$

where T_{start} is the time of stimulus onset and T_{end} is the time of stimulus offset.

Tangling index

We calculated population tangling using previously described methods (Horrocks et al., 2024; Russo et al., 2018). Trajectory tangling was quantified as

$$Q_{tangle}(t) = \max_{t'} \frac{\|\dot{x}_t - \dot{x}_{t'}\|^2}{\|x_t - x_{t'}\|^2 + \epsilon} \quad (50)$$

where x_t is the PCA-decomposed neural state in m -dimensional neural space at time t (as described in the Distance ratio section), which is an $m \times 1$ vector, with m being the number of dimensions. $\dot{x}_t$ is the temporal derivative of x_t . We computed the derivative of x_t as $\dot{x}_t = (x_t - x_{t-\Delta t})/\Delta t$, with $\Delta t = 1$ ms. $\|\cdot\|$ is the Euclidean norm, and ϵ is a small constant that prevents division by zero. Neural tangling $Q_{tangle}(t)$ is large when the neural population state at time t closely matches a state observed at another time t' , yet the associated state velocities (time derivatives) differ substantially. To obtain a temporally local estimate, we constrained the search for t' to within ± 100 ms of t . This avoids spuriously large tangling driven by temporally distant epochs.

Peak width

To characterize the temporal precision of pairwise coupling between neuronal signals, we quantified the peak width of the cross-correlogram (CCG) using a smoothed half-amplitude approach. For each significant sharp CCG peak, we first baseline-corrected the trace by subtracting the mean CCG value computed over the baseline period (time lags from -100 to -50 ms and 50 to 100 ms). The baseline-corrected traces were then smoothed using a Savitzky–Golay filter (Gorry, 1991) (window length = 11 ms, polynomial order = 3) to suppress high-frequency noise while preserving the overall peak shape. Positive and negative peaks were treated separately. For positive connections, we identified the time point of the maximum correlation and used one-half of the positive peak amplitude as the width threshold. For negative peaks, we identified the time point of the minimum correlation and used one-half of the absolute negative peak amplitude as the width threshold. The left and right boundaries were defined as the nearest time points on either side of the peak extremum where the smoothed trace crossed the corresponding half-amplitude threshold. The peak width was then quantified as the temporal distance (in time bins) between these two crossing points.

Population decoding analysis

To quantify stimulus information across behavioral states, we trained multinomial logistic regression classifiers on neuronal firing rates during stimulus presentation. Class labels corresponded to unique stimulus orientation and spatial frequency combinations for static gratings. And labels corresponded to direction and temporal frequency combinations for drifting gratings. Firing rates were z-scored across trials, and decoding performance was evaluated separately for each behavioral state using repeated stratified 5-fold cross-validation with four repeats. For decoding-latency analyses, firing rates were computed in 20-ms sliding windows with 2-ms steps. Decoding latency was defined as the first time window in which accuracy exceeded the 95th percentile of the chance distribution.

DATA AVAILABILITY

Data is freely accessible via the Allen SDK (Siegle et al., 2021) (Details at <https://portal.brainmap.org/circuits-behavior/visual-coding-neuropixels>). Source data for all main figures are provided with this paper. Source data are provided with this paper.

CODE AVAILABILITY

Code will be available upon publication.

REFERENCES

- Shailaja Akella, Peter Ledochowitsch, Joshua H. Siegle, Hannah Belski, Daniel D. Denman, Michael A. Buice, Severine Durand, Christof Koch, Shawn R. Olsen, and Xiaoxuan Jia. Deciphering neuronal variability across states reveals dynamic sensory encoding. *Nature Communications*, 16(1):1768, 2025. doi: 10.1038/s41467-025-56733-w.
- Ashl Ayaz, Aman B. Saleem, Marieke L. Schölvinck, and Matteo Carandini. Locomotion controls spatial integration in mouse visual cortex. *Current Biology*, 23(10):890–894, May 2013. doi: 10.1016/j.cub.2013.04.012.

- Graham L. Baum, Rastko Ciric, David R. Roalf, Richard F. Betzel, Tyler M. Moore, Russell T. Shinohara, Ari E. Kahn, Simon N. Vandekar, Petra E. Rupert, Megan Quarmley, Philip A. Cook, Mark A. Elliott, Kosha Ruparel, Raquel E. Gur, Ruben C. Gur, Danielle S. Bassett, and Theodore D. Satterthwaite. Modular Segregation of Structural Brain Networks Supports the Development of Executive Function in Youth. *Current biology: CB*, 27(11):1561–1572.e8, 2017. doi: 10.1016/j.cub.2017.04.051.
- Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008. doi: 10.1088/1742-5468/2008/10/P10008.
- Jeremy T. Chang and David Fitzpatrick. Development of visual response selectivity in cortical GABAergic interneurons. *Nature Communications*, 13(1):3791, 2022. doi: 10.1038/s41467-022-31284-6.
- Amelia J. Christensen and Jonathan W. Pillow. Reduced neural activity but improved coding in rodent higher-order visual cortex during locomotion. *Nature Communications*, 13(1):1676, 2022. doi: 10.1038/s41467-022-29200-z.
- Kelly B. Clancy, Ivana Orsolich, and Thomas D. Mrsic-Flogel. Locomotion-dependent remapping of distributed cortical networks. *Nature Neuroscience*, 22(5):778–786, 2019. doi: 10.1038/s41593-019-0357-8.
- Jessica R. Cohen and Mark D’Esposito. The Segregation and Integration of Distinct Brain Networks and Their Relationship to Cognition. *Journal of Neuroscience*, 36(48):12083–12094, 2016. doi: 10.1523/JNEUROSCI.2965-15.2016.
- Marlene R. Cohen and John H. R. Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, 12(12):1594–1600, 2009. doi: 10.1038/nn.2439.
- Lee Cossell, Maria Florencia Iacaruso, Dylan R. Muir, Rachael Houlton, Elie N. Sader, Ho Ko, Sonja B. Hofer, and Thomas D. Mrsic-Flogel. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518(7539):399–403, 2015. doi: 10.1038/nature14182.
- Maria C. Dadarlat and Michael P. Stryker. Locomotion Enhances Neural Encoding of Visual Stimuli in Mouse V1. *Journal of Neuroscience*, 37(14):3764–3775, 2017. doi: 10.1523/JNEUROSCI.2728-16.2017.
- David Dahmen, Stefano Recanatesi, Xiaoxuan Jia, Gabriel K. Ocker, Luke Campagnola, Stephanie Seeman, Tim Jarsky, Moritz Helias, and Eric Shea-Brown. Strong and localized recurrence controls dimensionality of neural activity across brain areas, 2023. Preprint at bioRxiv.
- Gustavo Deco, Giulio Tononi, Melanie Boly, and Morten L. Kringelbach. Rethinking segregation and integration: contributions of whole-brain modelling. *Nature Reviews Neuroscience*, 16(7):430–439, 2015. doi: 10.1038/nrn3963.
- Zhuokun Ding, Paul G. Fahey, Stelios Papadopoulos, Eric Y. Wang, Brendan Celii, Christos Papadopoulos, Andersen Chang, Alexander B. Kunin, Dat Tran, Jiakun Fu, Zhiwei Ding, Saumil Patel, Lydia Ntanavara, Rachel Froebe, Kayla Ponder, Taliah Muhammad, J. Alexander Bae, Agnes L. Bodor, Derrick Brittain, JoAnn Buchanan, Daniel J. Bumbarger, Manuel A. Castro, Erick Cobos, Sven Dorkenwald, Leila Elabbady, Akhilesh Halageri, Zhen Jia, Chris Jordan, Dan Kapner, Nico Kemnitz, Sam Kinn, Kisuk Lee, Kai Li, Ran Lu, Thomas Macrina, Gayathri Mahalingam, Eric Mitchell, Shanka Subhra Mondal, Shang Mu, Barak Nehoran, Sergiy Popovych, Casey M. Schneider-Mizell, William Silversmith, Marc Takeno, Russel Torres, Nicholas L. Turner, William Wong, Jingpeng Wu, Wenjing Yin, Szi-chieh Yu, Dimitri Yatsenko, Emmanouil Froudarakis, Fabian Sinz, Krešimir Josić, Robert Rosenbaum, H. Sebastian Seung, Forrest Collman, Nuno Maçarico Da Costa, R. Clay Reid, Edgar Y. Walker, Xaq Pitkow, Jacob Reimer, and Andreas S. Tolias. Functional connectomics reveals general wiring rule in mouse visual cortex. *Nature*, 640(8058):459–469, 2025. doi: 10.1038/s41586-025-08840-3.
- Mario Dipoppa, Adam Ranson, Michael Krumin, Marius Pachitariu, Matteo Carandini, and Kenneth D. Harris. Vision and locomotion shape the interactions between neuron types in mouse visual cortex. *Neuron*, 98(3):602–615.e8, 2018. doi: 10.1016/j.neuron.2018.03.037.
- Brent Doiron, Ashok Litwin-Kumar, Robert Rosenbaum, Gabriel K Ocker, and Krešimir Josić. The mechanics of state-dependent neural correlations. *Nature Neuroscience*, 19(3):383–393, 2016. doi: 10.1038/nn.4242.

- Daniel F. English, Sam McKenzie, Talfan Evans, Kanghwan Kim, Euisik Yoon, and György Buzsáki. Pyramidal cell–interneuron circuit architecture and dynamics in hippocampal networks. *Neuron*, 96(2):505–520.e7, 2017. doi: 10.1016/j.neuron.2017.09.033.
- Sinem Erisken, Agne Vaiceliunaite, Ovidiu Jurjut, Matilde Fiorini, Steffen Katzner, and Laura Busse. Effects of Locomotion Extend throughout the Mouse Early Visual System. *Current Biology*, 24(24):2899–2907, 2014. doi: 10.1016/j.cub.2014.10.045.
- Yu Fu, Jason M. Tucciarone, J. Sebastian Espinosa, Nengyin Sheng, Daniel P. Darcy, Roger A. Nicoll, Z. Josh Huang, and Michael P. Stryker. A cortical circuit for gain control by behavioral state. *Cell*, 156(6):1139–1152, 2014. doi: 10.1016/j.cell.2014.01.050.
- Shigeyoshi Fujisawa, Asohan Amarasingham, Matthew T. Harrison, and György Buzsáki. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neuroscience*, 11(7):823–833, 2008. doi: 10.1038/nn.2134.
- Peter A. Gorry. General least-squares smoothing and differentiation of nonuniformly spaced data by the convolution method. *Analytical Chemistry*, 63(5):534–536, 1991. doi: 10.1021/ac00005a031.
- Yong Gu, Sheng Liu, Christopher R. Fetsch, Yun Yang, Sam Fok, Adhira Sunkara, Gregory C. DeAngelis, and Dora E. Angelaki. Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron*, 71(4):750–761, 2011. doi: 10.1016/j.neuron.2011.06.015.
- Julie A. Harris, Stefan Mihalas, Karla E. Hirokawa, Jennifer D. Whitesell, Hannah Choi, Amy Bernard, Phillip Bohn, Shiella Caldejon, Linzy Casal, Andrew Cho, Aaron Feiner, David Feng, Nathalie Gaudreault, Charles R. Gerfen, Nile Graddis, Peter A. Groblewski, Alex M. Henry, Anh Ho, Robert Howard, Joseph E. Knox, Leonard Kuan, Xiuli Kuang, Jerome Lecoq, Phil Lesnar, Yaoyao Li, Jennifer Luviano, Stephen McConoughey, Marty T. Mortrud, Maitham Naeemi, Lydia Ng, Seung Wook Oh, Benjamin Ouellette, Elise Shen, Staci A. Sorensen, Wayne Wakeman, Quanxin Wang, Yun Wang, Ali Williford, John W. Phillips, Allan R. Jones, Christof Koch, and Hongkui Zeng. Hierarchical organization of cortical and thalamic connectivity. *Nature*, 575(7781):195–202, 2019. doi: 10.1038/s41586-019-1716-z.
- Kenneth D. Harris and Alexander Thiele. Cortical state and attention. *Nature Reviews Neuroscience*, 12(9):509–523, 2011. doi: 10.1038/nrn3084.
- Matthew T. Harrison and Stuart Geman. A Rate and History-Preserving Resampling Algorithm for Neural Spike Trains. *Neural Computation*, 21(5):1244–1258, 2009. doi: 10.1162/neco.2008.03-08-730.
- J J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. doi: 10.1073/pnas.79.8.2554.
- Edward A. B. Horrocks, Fabio R. Rodrigues, and Aman B. Saleem. Flexible neural population dynamics govern the speed and stability of sensory encoding in mouse visual cortex. *Nature Communications*, 15(1):6415, 2024. doi: 10.1038/s41467-024-50563-y.
- Chengcheng Huang, Douglas A. Ruff, Ryan Pyle, Robert Rosenbaum, Marlene R. Cohen, and Brent Doiron. Circuit Models of Low-Dimensional Shared Variability in Cortical Networks. *Neuron*, 101(2):337–348.e4, 2019. doi: 10.1016/j.neuron.2018.11.034.
- Chengcheng Huang, Alexandre Pouget, and Brent Doiron. Internally generated population activity in cortical networks hinders information transmission. *Science Advances*, 8(22):eabg5244, 2022. doi: 10.1126/sciadv.abg5244.
- James M. Jeanne, Tatyana O. Sharpee, and Timothy Q. Gentner. Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, 78(2):352–363, 2013. doi: 10.1016/j.neuron.2013.02.023.
- Xiaoxuan Jia, Seiji Tanabe, and Adam Kohn. Gamma and the Coordination of Spiking Activity in Early Visual Cortex. *Neuron*, 77(4):762–774, 2013. doi: 10.1016/j.neuron.2012.12.036.
- Xiaoxuan Jia, Joshua H. Siegle, Séverine Durand, Gregory Heller, Tamina K. Ramirez, Christof Koch, and Shawn R. Olsen. Multi-regional module-based signal transmission in mouse visual cortex. *Neuron*, 110(9):1585–1598.e9, 2022. doi: 10.1016/j.neuron.2022.01.027.

- Mohammad Mehdi Kafashan, Anna W. Jaffe, Selmaan N. Chettih, Ramon Nogueira, Iñigo Arandia-Romero, Christopher D. Harvey, Rubén Moreno-Bote, and Jan Drugowitsch. Scaling of sensory information in large neural populations shows signatures of information-limiting correlations. *Nature Communications*, 12(1):473, 2021. doi: 10.1038/s41467-020-20722-y.
- Ingmar Kanitscheider, Ruben Coen-Cagli, Adam Kohn, and Alexandre Pouget. Measuring fisher information accurately in correlated neural populations. *PLOS Computational Biology*, 11(6):e1004218, 2015a. doi: 10.1371/journal.pcbi.1004218.
- Ingmar Kanitscheider, Ruben Coen-Cagli, and Alexandre Pouget. Origin of information-limiting noise correlations. *Proceedings of the National Academy of Sciences*, 112(50), 2015b. doi: 10.1073/pnas.1508738112.
- Kohitij Kar, Jonas Kubilius, Kailyn Schmidt, Elias B. Issa, and James J. DiCarlo. Evidence that recurrent circuits are critical to the ventral stream’s execution of core object recognition behavior. *Nature Neuroscience*, 22(6):974–983, 2019. doi: 10.1038/s41593-019-0392-5.
- Ho Ko, Sonja B. Hofer, Bruno Pichler, Katherine A. Buchanan, P. Jesper Sjöström, and Thomas D. Mrsic-Flogel. Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473(7345):87–91, 2011. doi: 10.1038/nature09880.
- Adam Kohn, Ruben Coen-Cagli, Ingmar Kanitscheider, and Alexandre Pouget. Correlations and Neuronal Population Information. *Annual review of neuroscience*, 39:237–256, 2016. doi: 10.1146/annurev-neuro-070815-013851.
- E. A. Leicht and M. E. J. Newman. Community Structure in Directed Networks. *Physical Review Letters*, 100(11):118703, 2008. doi: 10.1103/PhysRevLett.100.118703.
- Ashok Litwin-Kumar and Brent Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience*, 15(11):1498–1505, 2012. doi: 10.1038/nn.3220.
- N. T. Markov, M. M. Ercsey-Ravasz, A. R. Ribeiro Gomes, C. Lamy, L. Magrou, J. Vezoli, P. Misery, A. Falchier, R. Quilodran, M. A. Gariel, J. Sallet, R. Gamanut, C. Huissoud, S. Clavagnier, P. Giroud, D. Sappey-Mariniér, P. Barone, C. Dehay, Z. Toroczkai, K. Knoblauch, D. C. Van Essen, and H. Kennedy. A Weighted and Directed Interareal Connectivity Matrix for Macaque Cerebral Cortex. *Cerebral Cortex*, 24(1):17–36, 2014. doi: 10.1093/cercor/bhs270.
- Nikola T. Markov, Mária Ercsey-Ravasz, David C. Van Essen, Kenneth Knoblauch, Zoltán Toroczkai, and Henry Kennedy. Cortical High-Density Counterstream Architectures. *Science*, 342(6158):1238406, 2013. doi: 10.1126/science.1238406.
- Francesca Mastrogiuseppe and Srdjan Ostojic. Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. *Neuron*, 99(3):609–623.e29, 2018. doi: 10.1016/j.neuron.2018.07.003.
- L. Mazzucato, G. La Camera, and A. Fontanini. Expectation-induced modulation of metastable activity underlies faster coding of sensory stimuli. *Nature Neuroscience*, 22(5):787+, 2019. doi: 10.1038/s41593-019-0364-9.
- Luca Mazzucato, Alfredo Fontanini, and Giancarlo La Camera. Dynamics of multistable states during ongoing and evoked cortical activity. *Journal of Neuroscience*, 35(21):8214–8231, 2015. doi: 10.1523/JNEUROSCI.4819-14.2015.
- Luca Mazzucato, Alfredo Fontanini, and Giancarlo La Camera. Stimuli reduce the dimensionality of cortical activity. *Frontiers in Systems Neuroscience*, 10:11, 2016. doi: 10.3389/fnsys.2016.00011.
- Matthew J. McGinley, Stephen V. David, and David A. McCormick. Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron*, 87(1):179–192, 2015. doi: 10.1016/j.neuron.2015.05.038.
- R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298(5594):824–827, 2002. doi: 10.1126/science.298.5594.824.
- Patrick J. Mineault, Elaine Tring, Joshua T. Trachtenberg, and Dario L. Ringach. Enhanced Spatial Resolution During Locomotion and Heightened Attention in Mouse Primary Visual Cortex. *The Journal of Neuroscience*, 36(24):6382–6392, 2016. doi: 10.1523/JNEUROSCI.0430-16.2016.

- Jude F. Mitchell, Kristy A. Sundberg, and John H. Reynolds. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, 63(6):879–888, 2009. doi: 10.1016/j.neuron.2009.09.013.
- Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. *Nature Neuroscience*, 17(10):1410–1417, 2014. doi: 10.1038/nn.3807.
- Simon Musall, Matthew T. Kaufman, Ashley L. Juavinett, Steven Gluf, and Anne K. Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10):1677–1686, 2019. doi: 10.1038/s41593-019-0502-4.
- Cristopher M. Niell and Michael P. Stryker. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, 65(4):472–479, 2010. doi: 10.1016/j.neuron.2010.01.033.
- Kenichi Ohki and R Clay Reid. Specificity and randomness in the visual cortex. *Current Opinion in Neurobiology*, 17(4): 401–407, 2007. doi: 10.1016/j.conb.2007.07.007.
- Kenichi Ohki, Sooyoung Chung, Yeang H. Ch’ng, Prakash Kara, and R. Clay Reid. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature*, 433(7026):597–603, 2005. doi: 10.1038/nature03274.
- Agostina Palmigiano, Theo Geisel, Fred Wolf, and Demian Battaglia. Flexible information routing by transient synchrony. *Nature Neuroscience*, 20(7):1014–1022, 2017. doi: 10.1038/nn.4569.
- Lia Papadopoulos, Suhyun Jo, Kevin Zumwalt, Michael Wehr, Santiago Jaramillo, David A. McCormick, and Luca Mazzucato. Modulation of metastable ensemble dynamics explains the inverted-u relationship between tone discriminability and arousal in auditory cortex. *Neuron*, 114(4), 2026. doi: 10.1016/j.neuron.2025.11.011.
- Pierre-Olivier Polack, Jonathan Friedman, and Peyman Golshani. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nature Neuroscience*, 16(9):1331–1339, 2013. doi: 10.1038/nn.3464.
- Stefano Recanatesi, Gabriel Koch Ocker, Michael A. Buice, and Eric Shea-Brown. Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity. *PLOS Computational Biology*, 15(7):e1006446, 2019. doi: 10.1371/journal.pcbi.1006446.
- Jacob Reimer, Emmanouil Froudarakis, Cathryn R. Cadwell, Dimitri Yatsenko, George H. Denfield, and Andreas S. Tolias. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, 84(2):355–362, 2014. doi: 10.1016/j.neuron.2014.09.033.
- Jacob Reimer, Matthew J. McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A. McCormick, and Andreas S. Tolias. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, 7:13289, 2016. doi: 10.1038/ncomms13289.
- Naixin Ren, Shinya Ito, Hadi Hafizi, John M. Beggs, and Ian H. Stevenson. Model-based detection of putative synaptic connections from spike recordings with latency and type constraints. *Journal of Neurophysiology*, 124(6):1588–1604, 2020. doi: 10.1152/jn.00066.2020.
- Dario L. Ringach, Patrick J. Mineault, Elaine Tring, Nicholas D. Olivas, Pablo Garcia-Junco-Clemente, and Joshua T. Trachtenberg. Spatial clustering of tuning in mouse primary visual cortex. *Nature Communications*, 7(1):12270, 2016. doi: 10.1038/ncomms12270.
- Robert Rosenbaum and Brent Doiron. Balanced Networks of Spiking Neurons with Spatially Dependent Recurrent Connections. *Physical Review X*, 4(2):021039, 2014. doi: 10.1103/PhysRevX.4.021039.
- Robert Rosenbaum, Matthew A. Smith, Adam Kohn, Jonathan E. Rubin, and Brent Doiron. The spatial structure of correlated neuronal variability. *Nature Neuroscience*, 20(1):107–114, 2017. doi: 10.1038/nn.4433.
- Abigail A. Russo, Sean R. Bittner, Sean M. Perkins, Jeffrey S. Seely, Brian M. London, Antonio H. Lara, Andrew Miri, Najja J. Marshall, Adam Kohn, Thomas M. Jessell, Laurence F. Abbott, John P. Cunningham, and Mark M. Churchland. Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. *Neuron*, 97(4):953–966.e8, 2018. doi: 10.1016/j.neuron.2018.01.004.

- Mehdi Sanayei, Xing Chen, Daniel Chicharro, Claudia Distler, Stefano Panzeri, and Alexander Thiele. Perceptual learning of fine contrast discrimination changes neuronal tuning and population coding in macaque V4. *Nature Communications*, 9(1):4238, 2018. doi: 10.1038/s41467-018-06698-w.
- Sameer Saproo and John T. Serences. Spatial attention improves the quality of population codes in human visual cortex. *Journal of Neurophysiology*, 104(2):885–895, 2010. doi: 10.1152/jn.00369.2010.
- João D. Semedo, Anna I. Jasper, Amin Zandvakili, Aravind Krishna, Amir Aschner, Christian K. Machens, Adam Kohn, and Byron M. Yu. Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, 13(1):1099, 2022. doi: 10.1038/s41467-022-28552-w.
- João D. Semedo, Amin Zandvakili, Christian K. Machens, Byron M. Yu, and Adam Kohn. Cortical areas interact through a communication subspace. *Neuron*, 102(1):249–259.e4, 2019. doi: 10.1016/j.neuron.2019.01.026.
- Joshua H. Siegle, Xiaoxuan Jia, Séverine Durand, Sam Gale, Corbett Bennett, Nile Graddis, Gregory Heller, Tamina K. Ramirez, Hannah Choi, Jennifer A. Luviano, Peter A. Groblewski, Ruweida Ahmed, Anton Arkhipov, Amy Bernard, Yazan N. Billeh, Dillan Brown, Michael A. Buice, Nicolas Cain, Shiella Caldejon, Linzy Casal, Andrew Cho, Maggie Chvilicek, Timothy C. Cox, Kael Dai, Daniel J. Denman, Saskia E. J. De Vries, Roald Dietzman, Luke Esposito, Colin Farrell, David Feng, John Galbraith, Marina Garrett, Emily C. Gelfand, Nicole Hancock, Julie A. Harris, Robert Howard, Brian Hu, Ross Hytten, Ramakrishnan Iyer, Erika Jessett, Katelyn Johnson, India Kato, Justin Kiggins, Sophie Lambert, Jerome Lecoq, Peter Ledochowitsch, Jung Hoon Lee, Arielle Leon, Yang Li, Elizabeth Liang, Fuhui Long, Kyla Mace, Jose Melchior, Daniel Millman, Tyler Mollenkopf, Chelsea Nayan, Lydia Ng, Kiet Ngo, Thuyahn Nguyen, Philip R. Nicovich, Kat North, Gabriel Koch Ocker, Doug Ollerenshaw, Michael Oliver, Marius Pachitariu, Jed Perkins, Melissa Reding, David Reid, Miranda Robertson, Kara Ronellenfitch, Sam Seid, Cliff Slaughterbeck, Michelle Stoecklin, David Sullivan, Ben Sutton, Jackie Swapp, Carol Thompson, Kristen Turner, Wayne Wakeman, Jennifer D. Whitesell, Derric Williams, Ali Williford, Rob Young, Hongkui Zeng, Sarah Naylor, John W. Phillips, R. Clay Reid, Stefan Mihalas, Shawn R. Olsen, and Christof Koch. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*, 592(7852):86–92, 2021. doi: 10.1038/s41586-020-03171-x.
- Gordon B. Smith, Audrey Sederberg, Yishai M. Elyada, Stephen D. Van Hooser, Matthias Kaschube, and David Fitzpatrick. The development of cortical circuits for motion discrimination. *Nature Neuroscience*, 18(2):252–261, 2015. doi: 10.1038/nn.3921.
- Matthew A. Smith and Adam Kohn. Spatial and Temporal Scales of Neuronal Correlation in Primary Visual Cortex. *Journal of Neuroscience*, 28(48):12591–12603, 2008. doi: 10.1523/JNEUROSCI.2929-08.2008.
- Olaf Sporns. Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, 23(2):162–171, 2013. doi: 10.1016/j.conb.2012.11.015.
- Eran Stark and Moshe Abeles. Unbiased estimation of precise temporal correlations between spike trains. *Journal of Neuroscience Methods*, 179(1):90–100, 2009. doi: 10.1016/j.jneumeth.2008.12.029.
- Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, 2019. doi: 10.1126/science.aav7893.
- Disheng Tang, Joel Zylberberg, Xiaoxuan Jia, and Hannah Choi. Stimulus type shapes the topology of cellular functional networks in mouse visual cortex. *Nature Communications*, 15:5753, 2024. doi: 10.1038/s41467-024-49704-0.
- Rufin VanRullen. The power of the feed-forward sweep. *Advances in Cognitive Psychology*, 3(1):167–176, 2007. doi: 10.2478/v10053-008-0022-3.
- Martin Vinck, Renata Batista-Brito, Ulf Knoblich, and Jessica A. Cardin. Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron*, 86(3):740–754, 2015. doi: 10.1016/j.neuron.2015.03.028.
- Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod

- Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, and Paul van Mulbregt. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020. doi: 10.1038/s41592-019-0686-2.
- David Wyrick and Luca Mazzucato. State-Dependent Regulation of Cortical Processing Speed via Gain Modulation. *Journal of Neuroscience*, 41(18):3988–4005, 2021. doi: 10.1523/JNEUROSCI.1895-20.2021.
- Yin Yan, Malte J. Rasch, Minggui Chen, Xiaoping Xiang, Mei Huang, Si Wu, and Wu Li. Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nature Neuroscience*, 17(10):1380–1387, 2014. doi: 10.1038/nn.3805.
- Rafael Yuste. From the neuron doctrine to neural networks. *Nature Reviews Neuroscience*, 16(8):487–497, 2015. doi: 10.1038/nrn3962.
- Roxana Zeraati, Yan-Liang Shi, Nicholas A. Steinmetz, Marc A. Gieselmann, Alexander Thiele, Tirin Moore, Anna Levina, and Tatiana A. Engel. Intrinsic timescales in the visual cortex change with selective attention and reflect spatial connectivity. *Nature Communications*, 14(1):1858, 2023. doi: 10.1038/s41467-023-37613-7.

ACKNOWLEDGEMENTS

This work was supported by funding from the Tsinghua-Peking Center for Life Sciences, National Natural Science Foundation of China (92370116), the Beijing Major Science and Technology Project (No. Z251100008125058 to X.J.), Tsinghua University Initiative Scientific Research Program, and Tsinghua University Dushi Program. We thank Disheng Tang, Zimoren Zhang, Ying Zhou and Fan Ye for helpful discussions, and Wenhao Zhang and Gabriel Koch Ocker for helpful feedback on the manuscript.

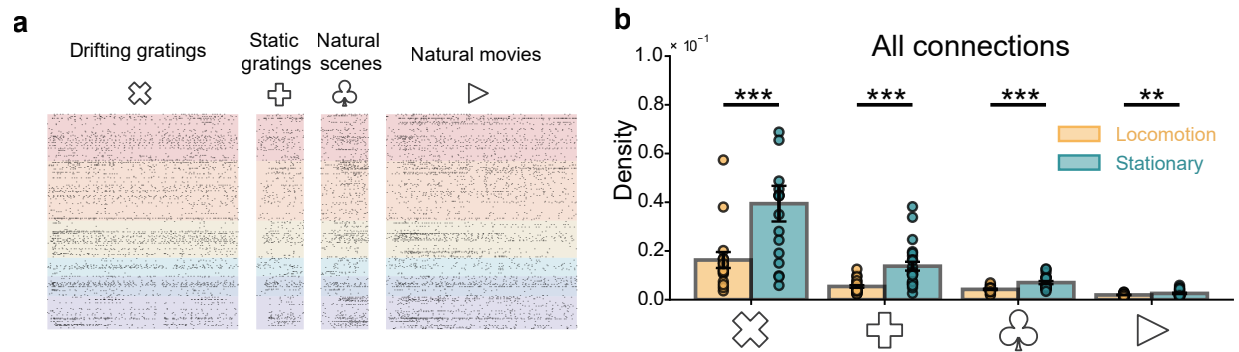
AUTHOR CONTRIBUTIONS

J.Z. and X.J. conceived the study and designed the analyses. J.Z. performed the data analysis, developed the computational models, generated the figures, and drafted the manuscript. X.J. supervised the project, acquired funding, and provided conceptual and methodological guidance. J.Z. and X.J. interpreted the results, revised the manuscript, and approved the final manuscript.

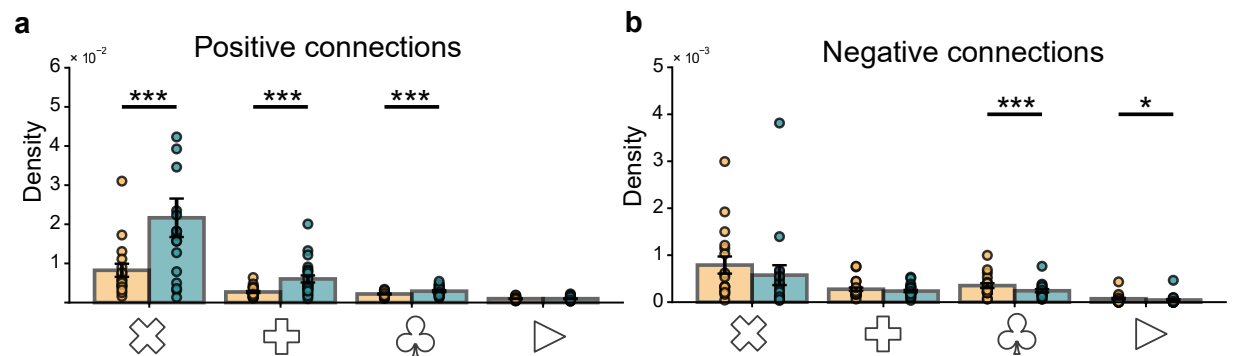
COMPETING INTERESTS

The authors declare no competing interests.

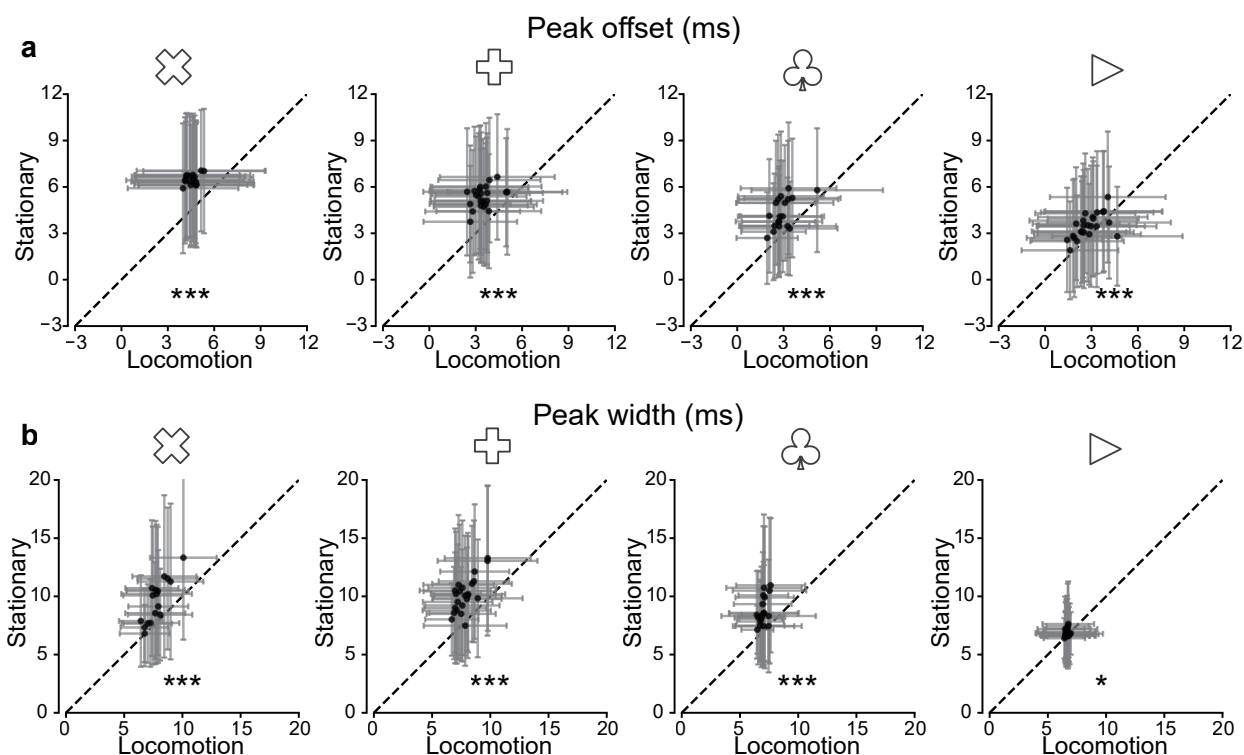
EXTENDED DATA FIGURES



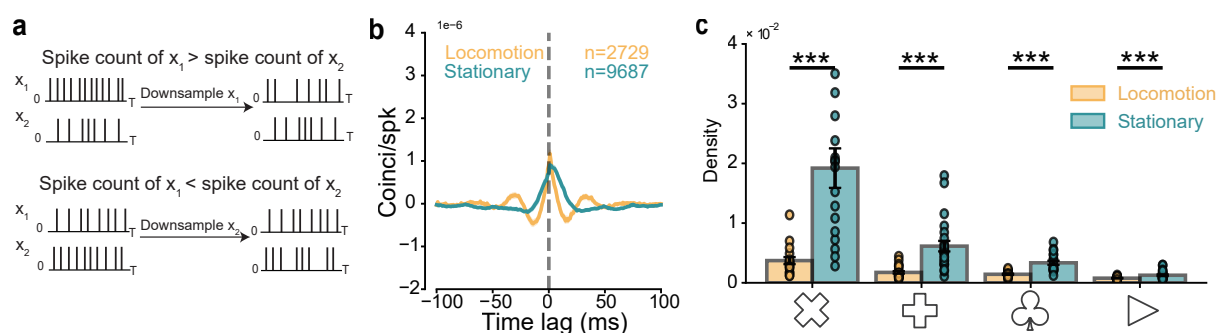
Extended Data Fig. 1. Locomotion-induced sparsification is robust to stimulus type. **a**, Example raster plots from mouse visual cortex during presentation of drifting gratings, static gratings, natural scenes, and natural movies. **b**, Network density across four visual stimulus types at a confidence threshold of 4. Points represent individual mice (Wilcoxon signed-rank test, from left to right $p = 3.05 \times 10^{-5}$, 4.77×10^{-7} , 2.67×10^{-6} , 2.64×10^{-3} ; $n = 17, 24, 19, 24$ mice).



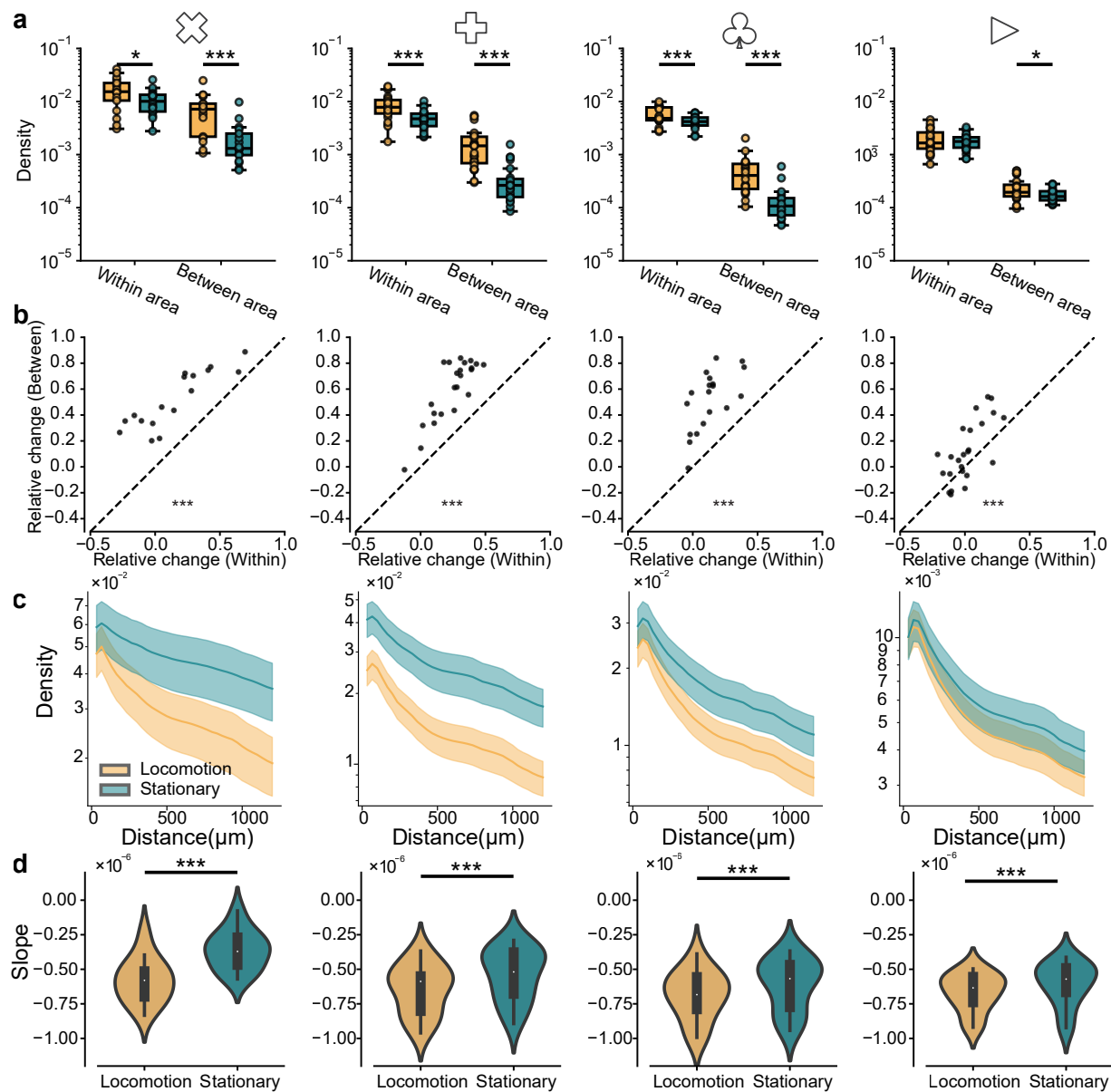
Extended Data Fig. 2. Sparsification in positive and negative connections. **a**, Density of positive connections across four visual stimulus types. Points represent individual mice (Wilcoxon signed-rank test, from left to right $p = 3.05 \times 10^{-5}$, 4.76×10^{-7} , 3.81×10^{-6} , 1.35×10^{-1} ; $n = 17, 24, 19, 24$ mice). **b**, Density of negative connections across four visual stimulus types (Wilcoxon signed-rank test, from left to right $p = 6.00 \times 10^{-2}$, 3.34×10^{-1} , 3.77×10^{-6} , 1.47×10^{-2} ; $n = 17, 24, 19, 24$ mice). Density of negative connections was not significantly different across states for drifting gratings and static gratings, and was higher during locomotion for natural scenes and natural movies.



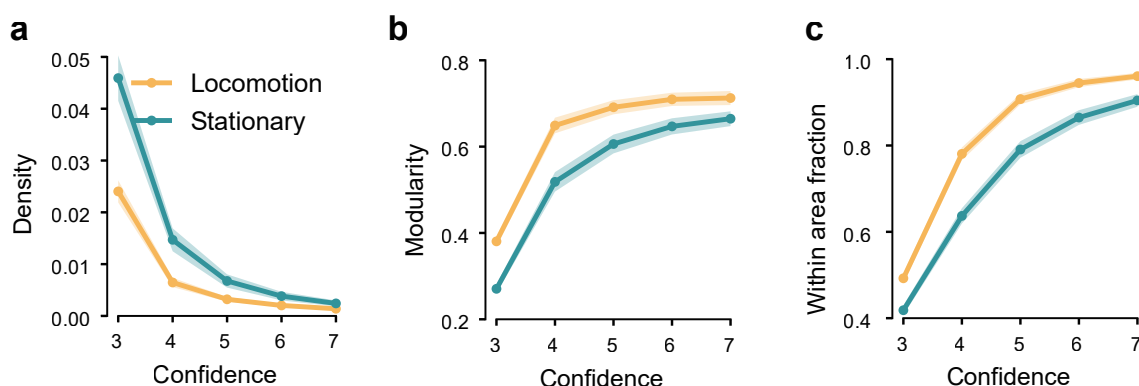
Extended Data Fig. 3. Offset and width of significant CCG peaks. **a**, Peak offset across four visual stimulus types. Points represent individual mice. Error bars indicate STD. (Wilcoxon signed-rank test, from left to right $p = 1.5 \times 10^{-5}$, 2.4×10^{-7} , 7.6×10^{-6} , 2.1×10^{-4} ; $n = 17, 24, 19, 24$ mice). **b**, Peak width across four visual stimulus types. Points represent individual mice. Error bars indicate STD. (Wilcoxon signed-rank test, from left to right $p = 6.0 \times 10^{-4}$, 4.8×10^{-7} , 7.61×10^{-6} , 1.2×10^{-2} ; $n = 17, 24, 19, 24$ mice).



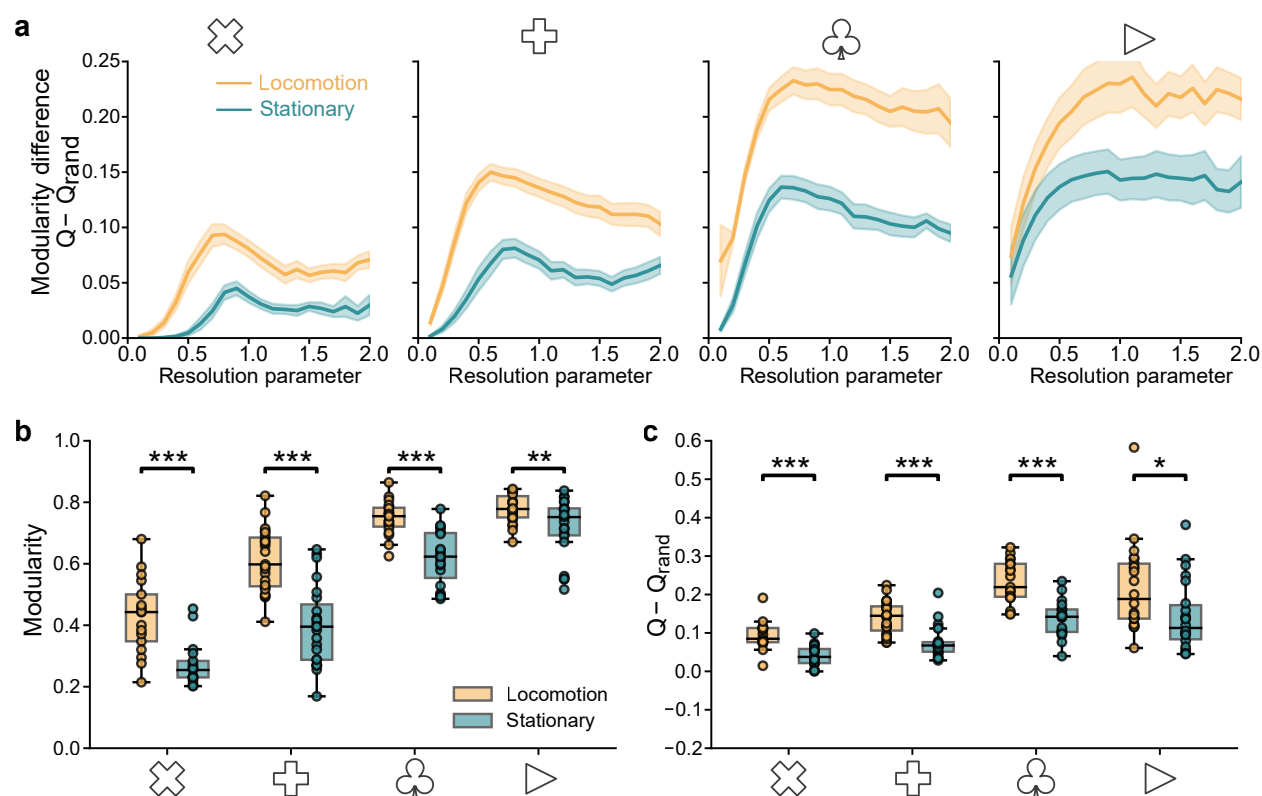
Extended Data Fig. 4. Controlling for firing rate does not eliminate the density reduction. **a**, Schematic illustration of spike downsampling. **b**, CCG trace of downsampled spikes with the same pairs in Fig. 1i. **c**, Network density inferred from downsampled spikes across four visual stimulus types. Points represent individual mice. Error bars indicate SEM. (Wilcoxon signed-rank test, from left to right $p = 1.52 \times 10^{-5}$, 2.38×10^{-7} , 3.81×10^{-6} , 3.65×10^{-5} ; $n = 17, 24, 19, 24$ mice). Across stimulus types, firing rates before rate matching during locomotion and stationary: drifting gratings, 8.511 ± 0.919 vs. 5.795 ± 0.623 ; static gratings, 9.098 ± 0.788 vs. 6.951 ± 0.594 ; natural scenes, 9.743 ± 0.711 vs. 7.358 ± 0.655 ; and natural movies, 8.784 ± 0.911 vs. 6.721 ± 0.885 . After rate matching, firing rates were 5.470 ± 0.593 , 6.511 ± 0.610 , 6.938 ± 0.619 , and 6.131 ± 0.749 for drifting gratings, static gratings, natural scenes, and natural movies, respectively.



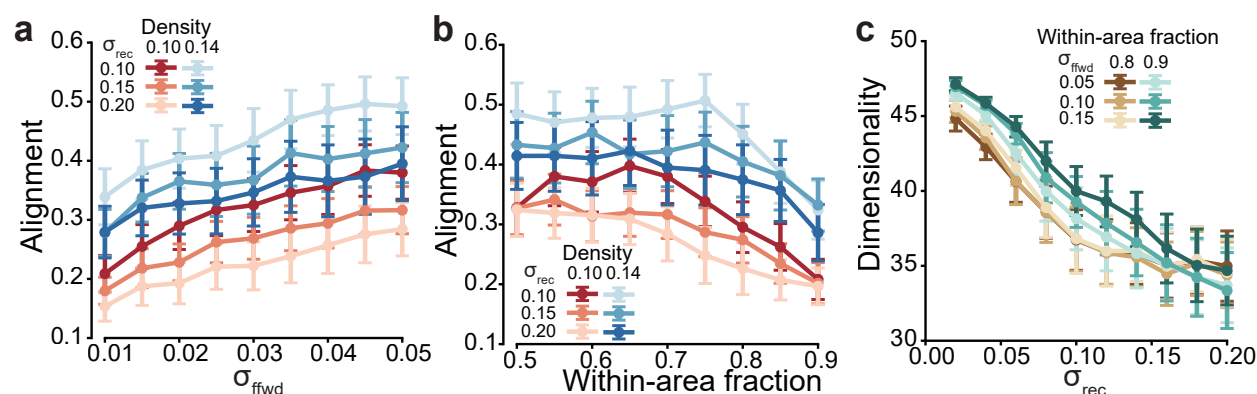
Extended Data Fig. 5. The anatomical pattern of sparsification is consistent across stimulus types. **a**, Within-area and between-area network density across four visual stimulus types. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times \text{IQR}$. Points represent individual mice. (Wilcoxon signed-rank test, from left to right, within area: $p = 1.74 \times 10^{-2}$, 1.19×10^{-6} , 6.45×10^{-4} , 0.56; between area: $p = 1.52 \times 10^{-5}$, 4.76×10^{-7} , 7.62×10^{-6} , 0.03; $n = 17, 24, 19, 24$ mice). **b**, Relative change in network density across four visual stimulus types. Points represent individual mice. (Wilcoxon signed-rank test, from left to right, $p = 1.52 \times 10^{-5}$, 2.38×10^{-7} , 3.81×10^{-6} , 8.71×10^{-4} ; $n = 17, 24, 19, 24$ mice). **c**, Spatial decay of network density across four visual stimulus types. **d**, Slopes of the spatial decay of network density across four visual stimulus types. (Paired t-test, from left to right, $p = 5.19 \times 10^{-7}$, 2.31×10^{-7} , 4.88×10^{-4} , 7.70×10^{-4} ; $n = 17, 24, 19, 24$ mice).



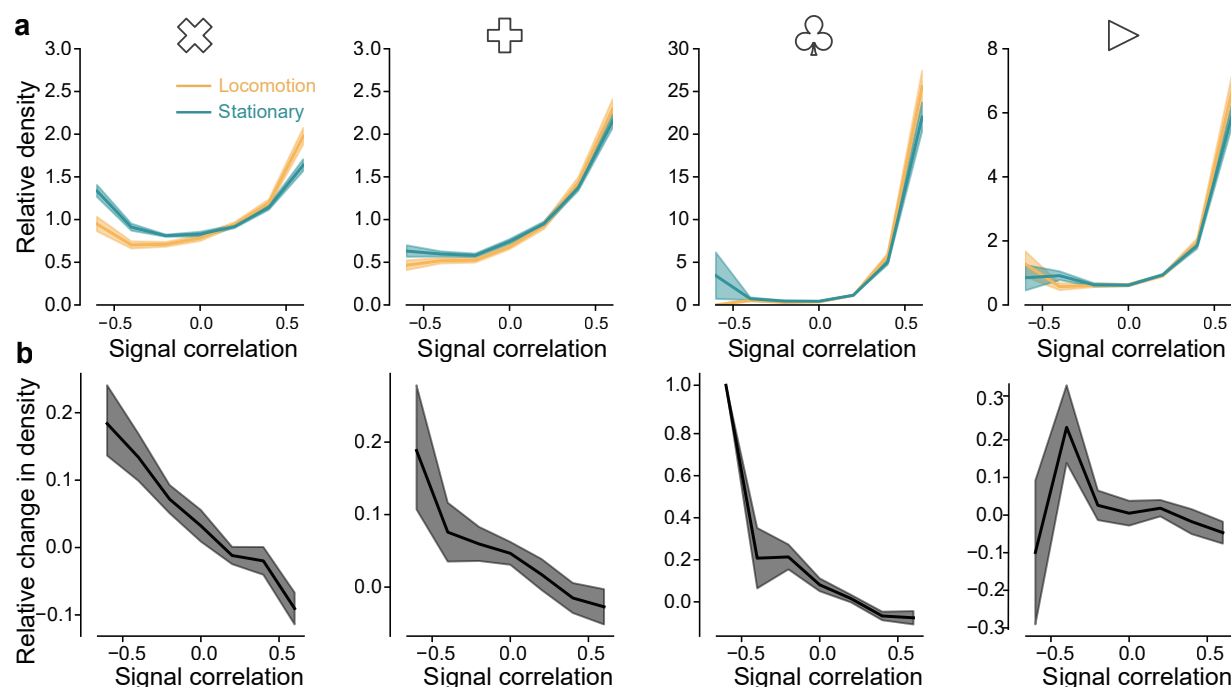
Extended Data Fig. 6. Structured sparsification across significant-peak thresholds. a–c, Density (a), modularity (b), and within-area fraction (c) across significant-peak thresholds. Shaded areas indicate SEM ($n = 84$ mice).



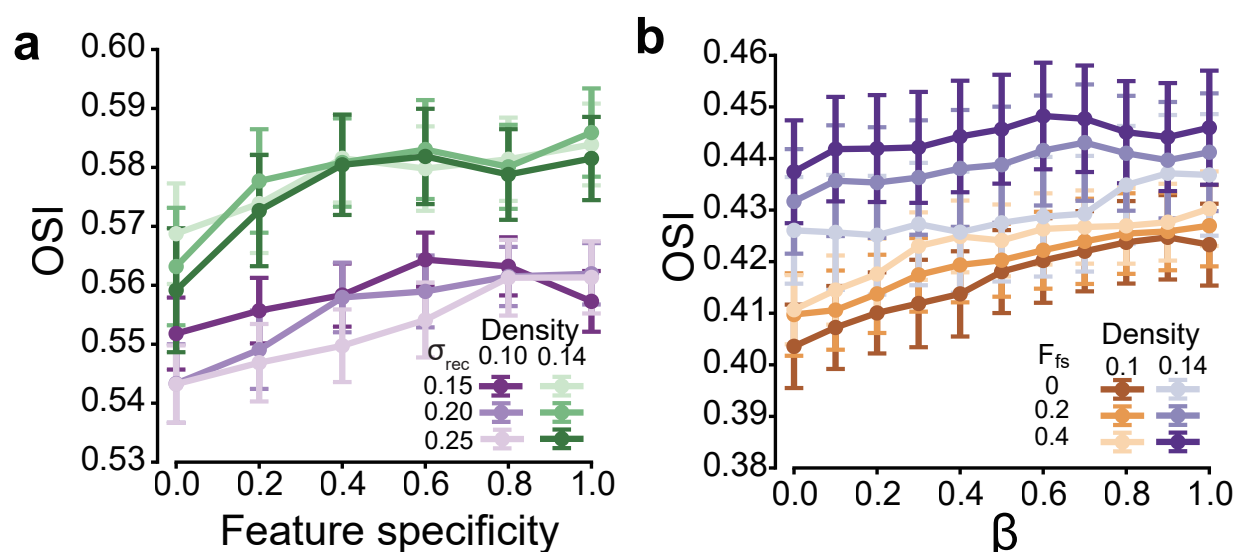
Extended Data Fig. 7. Network modularity across stimulus types. a, Relationship between the modularity difference and the resolution parameter. Shaded areas indicate SEM. ($n = 17, 24, 19, 24$ mice). b, Network modularity across four visual stimulus types. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times$ IQR. (Paired t-test, from left to right, $p = 1.50 \times 10^{-5}, 4.15 \times 10^{-10}, 2.86 \times 10^{-6}, 7.45 \times 10^{-3}$; $n = 17, 24, 19, 24$ mice). c, Network modularity difference ($Q - Q_{rand}$) across four visual stimulus types. Box plots show the median (center line), interquartile range (box), and whiskers extending to $1.5 \times$ IQR. (Paired t-test, from left to right, $p = 1.42 \times 10^{-6}, 1.05 \times 10^{-9}, 8.64 \times 10^{-6}, 1.22 \times 10^{-2}$; $n = 17, 24, 19, 24$ mice).



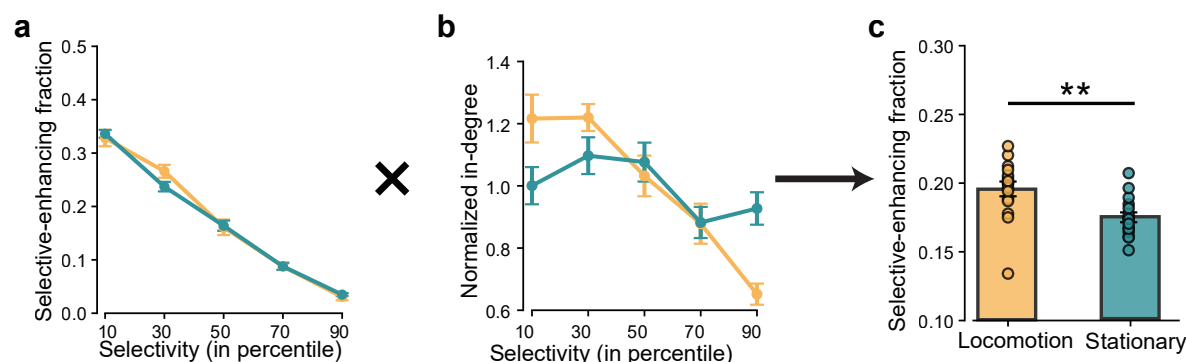
Extended Data Fig. 8. Spatial scale of connectivity shapes noise geometry. Network simulations showed that the spatial organization of recurrent and feedforward connectivity systematically reshaped the geometry of population variability across different choices of parameters. **a**, Increasing the spatial scale of feedforward connectivity, σ_{ffwd} , increased signal–noise alignment, indicating stronger overlap between stimulus-driven signal directions and dominant noise-covariance modes. This effect was observed across different connection densities and recurrent spatial scales, σ_{rec} . **b**, Increasing the fraction of within-area connections reduced signal–noise alignment across different densities and recurrent spatial scales. **c**, Increasing the recurrent spatial scale, σ_{rec} , reduced intrinsic dimensionality, consistent with broader recurrent coupling concentrating variability into fewer dominant modes. Points and error bars indicate mean $\pm$ SEM across 20 simulation repeats.



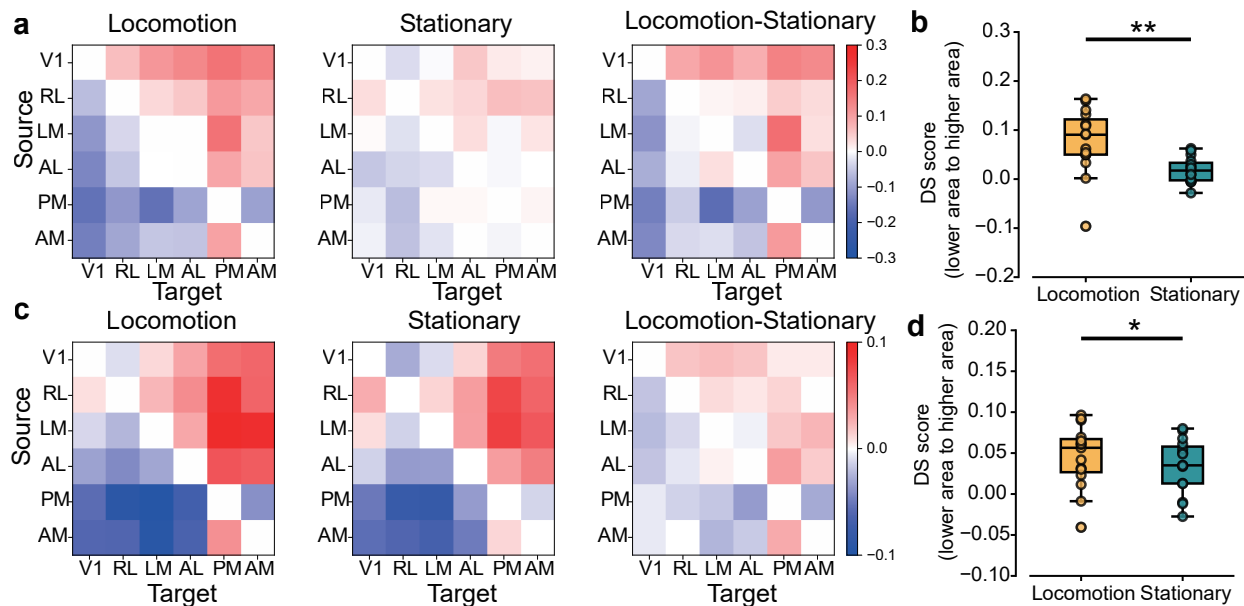
Extended Data Fig. 9. Feature-specific connectivity using signal correlation as a measure of feature similarity. **a**, Relationship between signal correlation and relative density across four visual stimulus types. **b**, Relationship between signal correlation and the relative change in density across four visual stimulus types. Shaded areas indicate SEM. (From left to right, $n = 17, 24, 19, 24$ mice).



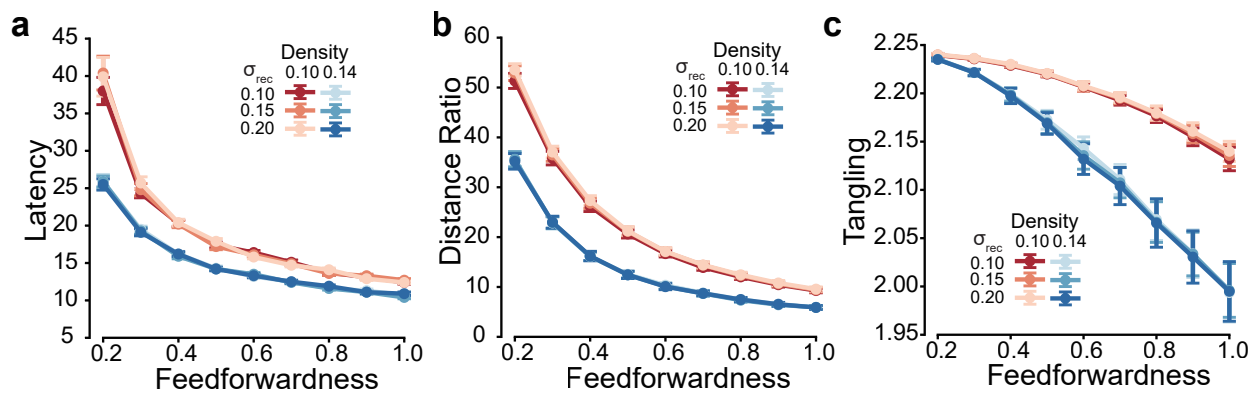
Extended Data Fig. 10. Structured sparsification increases selectivity. **a**, Selectivity increases as feature specificity increases, under different densities and spatial scales of recurrent connections. **b**, Selectivity increases as β increases, under different densities and feature specificity values. Points and error bars indicate mean $\pm$ SEM across 20 simulation repeats.



Extended Data Fig. 11. Fraction of selective-enhancing input projections. **a**, Fraction of selective-enhancing input projections across neuron populations with different selectivity. **b**, Normalized in-degree of neurons with different selectivity levels (percentile). Error bars indicate SEM ($n = 17$ mice). **c**, Overall fraction of selective-enhancing input projections. The overall selective-enhancing fraction was calculated by weighting each subpopulation's selective-enhancing fraction by its in-degree. Error bars indicate SEM (Wilcoxon signed-rank test, $p = 0.0021$, $n = 17$ mice).



Extended Data Fig. 12. Feedforwardness quantified using alternative measures. **a**, Directionality score measured using reachability during locomotion and stationary states, and for the locomotion-minus-stationary difference. **b**, Directionality score from anatomically lower to higher areas during locomotion and stationary states. (Paired t-test, $n = 15$ pairs, $p = 0.007$). **c**, Same as in **a**, but with directionality score measured using canonical correlation analysis. **d**, Same as in **b**, but with directionality score measured using canonical correlation analysis. (Paired t-test $n = 15$ pairs, $p = 0.013$).



Extended Data Fig. 13. Feedforward architecture accelerates decoding and straightens neural dynamics. **a**, Decoding latency decreases as feedforwardness increases, under different densities and spatial scales of recurrent connectivity. **b**, Distance ratio decreases as feedforwardness increases, under different densities and spatial scales of recurrent connectivity. **c**, Tangling index decreases as feedforwardness increases, under different densities and spatial scales of recurrent connectivity. Points and error bars indicate mean $\pm$ SEM across 20 simulation repeats.