Title: Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior

Summary: We develop a novel approach for deriving the time-varying functional connectivity of cortical networks and show this information encodes rapid fluctuations in behavioral state.

Authors: Hadas Benisty[1,2], Andrew H. Moberly[1], Sweyta Lohani[1], Daniel Barson[1], Ronald R. Coifman[2], Gal Mishne[3], Jessica A. Cardin[1], Michael J. Higley[1,*]

Affiliations: [1] Department of Neuroscience, Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA
[2] Program in Applied Mathematics, Yale University, New Haven, CT 06510, USA
[3] Halıcıoğlu Data Science Institute, University of California San Diego, La Jolla, CA, USA
[*] Corresponding Author

Lead Contact: m.higley@yale.edu

Abstract

Experimental work across a variety of species has demonstrated that spontaneously generated behaviors are robustly coupled to variation in neural activity within the cerebral cortex (*1-10*). Indeed, functional magnetic resonance imaging (fMRI) data suggest that functional connectivity in cortical networks varies across distinct behavioral states, providing for the dynamic reorganization of patterned activity (*5, 7, 11, 12*). However, these studies generally lack the temporal resolution to establish links between cortical signals and the continuously varying fluctuations in spontaneous behavior typically observed in awake animals. Here, we took advantage of recent developments in wide-field, mesoscopic calcium imaging (*13*) to monitor neural activity across the neocortex of awake mice. Applying a novel approach to quantifying rapidly time-varying functional connectivity, we developed a "graph of graphs" method to show that spontaneous behaviors are represented by fast changes in both the activity and correlational structure of cortical network activity. Both the approach and key results were generalizable to cellular-resolution data obtained via 2-photon imaging. Finally, dynamic functional connectivity of mesoscale signals revealed subnetworks that are not predicted by traditional anatomical atlas-based parcellation of the cortex. These results provide new insight into how behavioral information is represented across the mammalian neocortex and demonstrate an analytical framework for investigating time-varying functional connectivity in neural networks.

Variations in spontaneous motor output, such as locomotion and facial movements, are thought to reflect shifts in behavioral state that are coupled to changes in perceptual ability (*6, 14-18*) and co-occur with fluctuations in neuronal activity across the neocortex (*1-5*). Work using either fluorescence imaging or fMRI to monitor cortical dynamics broadly across multiple regions also suggests that the spatiotemporal correlations between neural signals in large-scale networks vary with transitions between states (*7-11, 19, 20*). However, these analyses typically rely on averaging activity within identified epochs (e.g., sleep versus wakefulness or quiescence versus arousal), ignoring the potential for rapid, continuous variation in both behavioral state and neural dynamics. Recent efforts using linear modeling suggested that uninstructed movements are represented in the activity recorded in mouse cortex (*2, 3*). Here, we sought to establish the ability of patterned cortical activity to accurately encode rapid, spontaneous changes in behavioral state.

Beginning with the assumption that behavior can be represented as some function of multi-dimensional neural activity, we explored how this function could be estimated using a standard multivariate Taylor expansion (see Methods). We considered a model of behavior, *b(t)*, derived from the first two terms of the Taylor expansion:

$$b(t) \approx \beta_0 + \boldsymbol{\beta}_1^T \boldsymbol{x}_t + \sum_i \sum_j \beta_2(i,j)\hat{C}_t(i,j) + \epsilon \qquad (1)$$

Here, the first-order term $\boldsymbol{x}_t$ is an N-dimensional vector corresponding to the time-varying neural activity across N cortical parcels at time t. Indeed, representing behavior as a linear combination of time-varying neural signals is a common approach (*14, 21-24*). The second-order term $\hat{C}_t(i,j)$ corresponds to the time-varying pairwise correlations between parcels i and j, reflecting the rapid dynamics of functional connectivity across cortical networks. We hypothesized that modelling behavior by combining both a linear term in activity $\boldsymbol{x}_t$ and correlations $\hat{C}_t$, which are a nonlinear second-order function of $\boldsymbol{x}_t$, would provide significantly improve decoding accuracy, suggesting distinct roles for variations in both signal activity and functional connectivity in cortical function.

To address this question, we carried out widefield, mesoscopic calcium imaging (*13*) in awake, head-fixed mice expressing the red fluorescent indicator jRCaMP1b (*25*) (Figure 1a). Indicator expression was mediated by neonatal injection of wild-type mice with AAV9-Syn-jRCaMP1b (see Methods) (*26, 27*). We simultaneously monitored cortical activity and spontaneous behavioral metrics including pupil diameter, facial movements, and locomotion (Figure 1b-f, see Methods) (*3, 4*). Previous studies have often relied on categorical definitions of behavioral state, averaging within epochs according to thresholding of motor signals (*4, 14*). However, analysis of our data indicates that, rather than falling into discrete clusters, these metrics are continuously distributed across a range of rapidly varying values (Calinski-Harabasz Index values vs. # of clusters for K-means clustering using 2-6 clusters, p=0.99, ANOVA, Figure 1b-c). We also find that these signals are only modestly correlated with each other (Figure 1d, Supplemental Figure S1), suggesting that they represent underlying latent variables corresponding to distinct behavioral dynamics and states.

After normalization and hemodynamic correction of imaging data (see Methods) (*1, 28*), we segmented the cortex into functional parcels using a graph theory-based approach that relies on spatiotemporal co-activity between pixels (LSSC, Figure 1e-h, Supplemental Figure S1)(*29*). This approach yields significantly smaller reconstruction errors for a similar number of parcels in comparison to the Allen Institute CCFv3 anatomy-based atlas (*30*) (Supplemental Figure S1). Moreover, LSSC performs similarly to principal component analysis but with the distinct advantage of generating discrete, disjointed parcels that can provide straightforward links between function and structure (Supplemental Figure S1). We next extracted $\boldsymbol{x}_t$, the time-varying fluctuations in the fluorescence signal associated with each parcel (Figure 1e, see Methods), corresponding to the first-order terms in the Taylor polynomial approximation of behavior (eq. 1). As expected, variation in activity appeared to be coupled to changes in behavioral metrics over rapid (sub-second) time scales (Figure 1e-h). We then calculated the time-varying, pair-wise correlations $\hat{C}_t$ between LSSC parcels using a sliding 3-second window (0.1 second step-size, Figure 1f-i, see Methods), corresponding to the second-order terms in the Taylor polynomial. On average, correlations across the cortex were high (r=0.6±0.03, n=6 mice), but their moment-to-moment values also appeared to co-vary with rapid behavioral changes (Figure 1f-i).

Dynamic functional connectivity encodes spontaneous behavior

Coordinated patterns of neural activity across the cortex are inherently high-dimensional but may represent low dimensional latent variables (*31-33*). To explore the relationship between behavior and dynamic

functional connectivity, we developed a novel strategy to calculate correlational dynamics with high temporal resolution and extract their low dimensional representation, ϕ_t. Fitting a linear regression model for behavioral dynamics whose predictors are time-varying cortical activity and functional connectivity, following the model in eq. (1), leads to poor predictive power (Supplemental Figure 2) due to over-fitting caused by the high-dimensionality of pairwise correlations $\hat{C}_t$ (~10^3 pairs per animal). Therefore, we developed a novel strategy to extract a lower dimensional representation, ϕ_t, capturing the intrinsic dynamics of the correlational signals using Riemannian geometry and diffusion embedding (*34*). Each correlation matrix over a short temporal window (e.g., 3 seconds) can be viewed as a spatial graph whose N^2 nodes are cortical parcels connected by weighted edges equal to the instantaneous pairwise correlation coefficients between parcels. Sliding the window over time (Δt=0.1 seconds) produces a series of correlation matrices, which can also be viewed as a time-varying graph (Figure 2a). We then built a "graph of graphs", where each node is now a time-point represented by the correlation matrix at that time (see Methods). A distance measure between correlation matrices is necessary to set the edge weights of this temporal graph. Since correlation matrices lie on a non-Euclidean Riemannian manifold called the Semi-Positive Definite (SPD) cone, Euclidean distances do not properly represent similarity in this space (Figure 2b). Therefore, we used Riemannian geometry to calculate pairwise geodesic distances between correlation matrices. We applied diffusion embedding to the graph of graphs and extracted the low dimensional representation, ϕ_t , of the temporal dynamics of functional connectivity (Figure 2c).

We constructed a cross-validated linear regression model combining the cortical activity for all parcels ($\boldsymbol{x}_t$, ranging from 48-53 parcels per animal across 6 mice) and the first 20 leading components of the embedded correlations, denoted by $\phi_t^{(20)}$ to predict the continuously varying behavioral signals for pupil diameter, facial movement, or locomotion (Figure 2d-e). We found that behavior can be robustly decoded by this joint model (Pupil: R^2=0.52±0.04; Face: R^2=0.59±0.04; Wheel: R^2=0.45±0.06; n=6 mice, Figure 2e). These values were not significantly different from decoding using the windowed (smoothed) measure of activity (Figure S2). As expected, predictive accuracy was significantly impaired when using either raw correlations or Euclidean rather than Riemannian distances for the diffusion embedding (Supplemental Figure S2). Moreover, predictive performance was robust to changes in model parameters, did not improve with the inclusion of more than 20 embedding components, and did not vary appreciably for window lengths of 3-10 seconds (Supplemental Figure S2).

To investigate the relative contributions of activity versus connectivity dynamics in modeling behavior, we recreated a joint model while temporally shuffling one of the two predictors. Shuffling either term significantly impaired prediction accuracy relative to the unshuffled full model (Pupil: R^2=0.2±0.04, p=0.001 for shuffling $\phi_t^{(20)}$, R^2=0.48±0.04, p=0.02 for shuffling $\boldsymbol{x}_t$; Face: R^2=0.38±0.05, p=0.002 for shuffling $\phi_t^{(20)}$, R^2=0.49±0.06, p=0.006 for shuffling $\boldsymbol{x}_t$; Wheel: R^2=0.31±0.07, p=0.01 for shuffling $\phi_t^{(20)}$, R^2=0.32±0.05, p=0.003 for shuffling $\boldsymbol{x}_t$; Paired t-test, Figure 2e). Surprisingly, models in which correlational data was preserved were similar or better at decoding behavior than activity-preserved models, reaching significance for variations in pupil diameter (Pupil: p=0.002; Face: p=0.06; Wheel: p=0.37; Paired t-test, Figure 2d-e). To further examine the ability of activity or connectivity signals to independently predict behavioral signals, we generated single predictor models which produced similar results, (Figure 2e). Finally, we note that, while the time-averaged activity and pairwise correlations significantly differ for high versus low behavioral state epochs (see Methods), the temporal dynamics of spatially averaged cortical activity and correlation signals are poorly predictive of the rapidly time-varying fluctuations in behavior (Supplemental Figure S2). Altogether, these findings indicate that inclusion of dynamic functional connectivity significantly improves decoding power for modeling of behavioral state, suggesting that cortical network function relies not only on the absolute amount of activity but also on the coordination of activity across widespread areas.

Network connectivity does not encode sensory information

Spontaneous cortical activity likely reflects latent signals corresponding to internally generated brain processes. Thus, we asked whether extrinsic sensory information was similarly represented by large-scale networks. We presented the mouse with a series of visual stimuli (drifting sinusoidal gratings, see Methods) and quantified evoked activity using mesoscopic calcium imaging. Contrast-dependent responses were largest in visual areas but were also observed broadly across other cortical regions. However, evoked responses had minimal impact on the correlational structure of activity across the cortex (Supplemental Figure S3). Linear modeling showed that the stimulus could be robustly decoded using activity in visual cortex, with prediction accuracy exhibiting contrast-dependence. However, correlational structure was only very weakly predictive of

the stimulus, differing significantly from the activity-based model (F= 57.1, p=5.5e^{-10}, ANOVA for combined stimulus contrasts >50%, n=6 mice; Supplemental Figure S3). These results suggest that brief sensory inputs drive large fluctuations in cortical activity without substantial alteration of functional network connectivity.

Dynamic correlations of cellular networks also predict behavior

To determine the generalizability of our approach and also examine encoding by neural correlations at a different spatial scale, we monitored local circuit activity using 2-photon calcium imaging of single GCaMP6s-expressing neurons (*35*) in the primary visual cortex (see Methods) simultaneously with measurements of pupil diameter, facial movement, and locomotion (Figure 3a-f). As above, we looked at both time-varying activity and embedded pair-wise correlations for identified neurons. Unlike large-scale network signals, time-averaged correlations between neurons were broadly distributed around zero (R^2=-0.001±0.006, n=6 mice).

We then generated a cross-validated linear model combining activity and embedded correlation dynamics across cells and attempted to predict rapid fluctuations in behavior. As with mesoscopic imaging, cellular data also robustly predicted behavior (Pupil: R^2=0.59±0.04; Face: R^2=0.44±0.08; Wheel: R^2=0.39±0.1; n=6 mice, Figure 3). Again, modeling performance was poorer using raw correlations and Euclidean distances for embedding but was insensitive to using smoothed activity signals (Supplemental Figure S4). Furthermore, modeling was again robust to changes in the number of embedding components and window length (Supplemental Figure S4).

To calculate the relative contributions of activity versus correlations in the joint model, we similarly shuffled one of the two predictors. As above, shuffling either variable significantly impaired prediction accuracy relative to the unshuffled model (Pupil: R^2=0.53±0.05, p=0.04 for shuffling $\phi_t^{(20)}$, R^2=0.42±0.05, p=0.0002 for $\boldsymbol{x}_t$; Face: R^2=0.39±0.09, p=0.004 for shuffling $\phi_t^{(20)}$, R^2=0.32±0.05, p=0.03 for shuffling $\boldsymbol{x}_t$; Wheel: R^2=0.32±0.1, p=0.04 for $\phi_t^{(20)}$, R^2=0.26±0.07, p=0.02 for $\boldsymbol{x}_t$; Paired t-test, Figure 3g). Models preserving either the activity or correlational data gave similar accuracy, with activity-based analysis showing modestly better performance for pupil fluctuations (p=0.025 for Pupil, p=0.14 for Face, p=0.15 for Wheel, Paired t-test, Figure 3g). Single-predictor models again produced similar results (Figure 3g). In summary, applying our novel approach for quantifying time-varying correlations in neural data to cellular resolution imaging, we again find that including dynamic functional connectivity significantly enhances prediction accuracy in models linking neural signals to fluctuations in behavioral state.

Dynamic functional connectivity suggests distinct cortical subnetworks

The improved accuracy of behavioral prediction using embedding of mesoscopic correlation matrices suggests they may reflect underlying principles of structural organization in large-scale cortical networks. We therefore examined the spatial interpretation of ϕ_t by asking how the time-varying correlation for each pair of parcels is represented by the overall embedding. This approach allows us to determine whether the embedding is primarily capturing spatially organized subsets of pairwise correlations while remaining agnostic to variation in behavioral state per se. We quantified the goodness-of-fit using $\phi_t^{(20)}$ to model the correlation time series between a target parcel and each of the other parcels across the cortex (Figure 4a-b, see Methods). Averaging these goodness-of-fit matrices across all animals (n=6 mice) revealed substantial spatial heterogeneity that was conserved across different individuals. Embedding primarily represented correlations between each target parcel and one or both of a posterior and anterolateral subdivision of the cortex (Figure 4c, Supplemental Figure S5). This spatial pattern was clearly evident after making a grand average across all parcels and animals (Figure 4c). Intuitively, this result indicates that independent of behavior, dynamic large-scale correlations of cortical areas are dominated by the interrelationship of each cortical parcel with one or both of these two subnetworks. Surprisingly, this functional organization is highly distinct from regional boundaries defined by traditional, anatomy-based atlases such as the CCFv3 (Figure 4e).

To further examine whether coordinated activity across this anterolateral/posterior partition does correspond to spontaneous behavioral variation, we calculated a time-varying participation coefficient that measures the functional connectivity of parcels within versus between two groupings (*36*) defined by a line bisecting the cortex (Figure 4e, see Methods). We quantified the difference in average behavioral metrics for time points corresponding to the upper versus lower deciles of the participation coefficients and repeated this analysis for different angles of the bisecting line. Our results showed a significant variation across angles for all behaviors (F=9.5, p=1.6e-5 for pupil; F=5.9, p=0.0006 for Face; F=6.3, p=0.0004 for Wheel; ANOVA, n=6 mice, Figure

4f, Supplemental Figure S5) with the peak value corresponding to the observed boundary of anterolateral and posterior networks. This result supports the conclusion that these anatomical subdivisions reflect behaviorally relevant functional cortical architecture.

Finally, we examined whether embedding of correlations between single neurons also reflects underlying spatial organization in the local circuit. Similar to our analysis for mesoscopic data, we quantified the goodness-of-fit using $\phi_t^{(20)}$ to model the correlations for each pair of neurons in a field of view and plotted this value versus the inter-somatic distance (Figure 4g). These results revealed no appreciable relationship, indicating that behaviorally relevant coupling between cells is independent of their relative physical location within the local network.

Overall, we developed a novel strategy for analyzing the continuously time-varying functional connectivity in cortical networks, viewing them as a graph of graphs in a non-Euclidean space. Our method provides an analytical framework for extracting the intrinsic dynamics of short-term correlations and uses Riemannian geometry to correctly evaluate distances between correlation matrices extracted at different time points. These distances are then used to set the weights of a graph-of-graphs, allowing us to extract a low-dimensional representation for the manifold of the correlations and capture their underlying dynamics. Using this approach and including both first-order (activity) and nonlinear second-order (embedded functional connectivity) terms for modeling behavior enabled us to significantly improve decoding power. The trained models accurately predict rapid (sub-second) dynamics of pupil diameter, facial movement, and locomotion, metrics commonly used to assess variation in arousal state (*1, 4, 6, 14*). Our method was generalizable for two different data sets, mesoscopic and 2-photon imaging, and yielded the surprising finding that higher-order statistics (i.e., correlational signals) produce similar or better predictive accuracy than time-varying changes in activity. Since the activity and correlation elements in the model are linearly independent, our results indicate that dynamic functional connectivity between neural elements represents a unique aspect of the data that is coupled to the animal's behavior.

The representation of information by time-varying cortical signals has been a focus of recent studies using diverse approaches to monitor human brain activity (*5, 7, 11, 37*). For example, shifts in wakefulness correspond to changes in average resting state connectivity (*38*), resting state fluctuations predict somatosensory perception (*15*), and working memory-based task performance corresponds to spatially heterogeneous variation in timescales of patterned activity (*39, 40*). However, in part due to relatively low temporal resolution of fMRI, these efforts have relied on averaging activity within sustained, categorical behavioral states. In contrast, we find that variations in pupil diameter, facial movement, and locomotion do not appear to cluster into distinct regimes, a result more consistent with continuously and rapidly varying states. Indeed, these behavioral dynamics are matched by sub-second fluctuations in functional connectivity that are similar to synaptic integration windows for single cells. Indeed, recent studies have combined mesoscopic imaging with either fMRI (*41*) or single-cell monitoring (*26, 42-44*) finding dynamic connectivity across spatial scales that supports the hypothesis that neurons may be sensitive to convergent synaptic input driven by correlated large-scale activity. These multi-modal approaches combined with novel analytical methods such as those presented here will likely drive rapid discoveries into the functional organization of brain networks in diverse systems.

With the rapidly expanding efforts to link large-scale cortical signals with sensory representations, movement, and task performance (*2, 8, 9, 26, 43, 45*), several distinct strategies have been developed to analyze the spatiotemporal organization of network activity, including singular value decomposition and non-negative matrix factorization (*8, 46*). Here, we show that functional parcellation of cortical regions (*29*) followed by embedding of time-varying correlations based on Riemannian geometry, provides a robust means to quantify dynamic functional connectivity that accurately decodes spontaneous fluctuations in behavior. With the increasing interest in analysis of neural manifolds, our results also highlight the necessity of considering the geometry of the manifold on which the data lie to accurately reveal their intrinsic representation. Notably, we did not find similarly strong representation of sensory information in mesoscopic correlations, suggesting that external inputs modulate activity without altering functional connectivity. However, this result may also reflect a lack of behavioral relevance for the stimuli as presented here. We and others have shown that training can modify the sensory and motor representations by single cortical neurons (*14, 47-49*), and future studies must determine if development, experience, or learning produce functional reorganization of large-scale networks.

Finally, our results suggest that the cortex can be spatially segmented into two broad subnetworks, an anterolateral and posterior division, a functional division that emerges from analysis of spontaneous activity but also reflects variation in behavioral state metrics. We previously suggested such a division based on correlations between single cell activity and mesoscopic cortical signals (*26*). Interestingly, this organization does not map readily onto standard anatomical segmentation of the cortex, such as the CCFv3 (*30*). We propose that the dynamic modulation and plasticity of synaptic strength may support the translation between such structural and functional views of connectivity in cortical networks, a hypothesis that awaits experimental validation.

Author Contributions

HB, RRC, GM, JAC, and MJH designed the study. HB, DB, RRC, GM, and MJH developed the analytical approach. HB carried out all analyses. AHM and SL collected experimental data. HB and MJH wrote the manuscript.

Acknowledgements

The authors thank members of the Higley and Cardin laboratories for helpful input throughout all stages of this study. We thank Rima Pant for generation of AAV vectors. We thank the GENIE Project for jRCaMP1b plasmids. This work was supported by funding from the NIH (MH099045 and MH121841 to MJH, EY022951 to JAC, MH113852 to MJH and JAC, EY031133 to AHM, EY026878 to the Yale Vision Core, EB026936 to GM and RRC), an award from the Yale Kavli Institute of Neuroscience (to MJH and RRC), an award from the Swartz Foundation (to HB), an award from Aligning Science Across Parkinson's (to MJH), an award from the Simons Foundation SFARI (to MJH), and a BBRF Young Investigator Grant (to SL).

Conflicts of Interest

The authors declare no conflicts of interest exist.

Data Availability Statement

The full datasets generated and analyzed in this study are available from the corresponding authors on reasonable request.

Code Availability Statement

Custom written MATLAB scripts used in this study are available on github (https://github.com/cardin-higley-lab).

Materials and Methods

All animal handling and experiments were performed according to the ethical guidelines of the Institutional Animal Care and Use Committee of the Yale University School of Medicine. Mesoscopic imaging data were collected as part of a previous study (*1*), with experimental details provided below for clarity. Analysis results presented here represent wholly new findings and have not appeared elsewhere.

Animals

Male and female C57BL/6 mice were kept on a 12h light/dark cycle, provided with food and water ad libitum, and housed individually following headpost implants. Imaging experiments were performed during the light phase of the cycle. For mesoscopic imaging, brain-wide expression of jRCaMP1b (*25*) was achieved via postnatal sinus injection as described previously (*26, 27*). Briefly, P0-P1 litters were removed from their home cage and placed on a heating pad. Pups were kept on ice for 5 min to induce anesthesia via hypothermia and then maintained on a metal plate surrounded by ice for the duration of the injection. Pups were injected bilaterally with 4 ul of AAV9-hSyn-NES-jRCaMP1b (2.5x10^13 gc/ml, Addgene). Mice also received an injection of AAV9-hSyn-$GRAB_{ACh3.0}$ to express the genetically encoded cholinergic sensor $GRAB_{ACh3.0}$ (*50*), although these data were not used in the present study. Once the entire litter was injected, pups were returned to their home cage. For two-photon imaging experiments, a similar procedure was used to drive cortex-wide expression of GCaMP6s (*35*).

Surgical procedures

All surgical implant procedures were performed on adult mice (>P50). Mice were anesthetized using 1-2% isoflurane and maintained at 37°C for the duration of the surgery. For mesoscopic imaging, the skin and fascia above the skull were removed from the nasal bone to the posterior of the intraparietal bone and laterally between the temporal muscles. The surface of the skull was thoroughly cleaned with saline and the edges of the incision secured to the skull with Vetbond. A custom titanium headpost was secured to the posterior of the nasal bone with transparent dental cement (Metabond, Parkell), and a thin layer of dental cement was applied to the entire dorsal surface of the skull. Next, a layer of cyanoacrylate (Maxi-Cure, Bob Smith Industries) was used to cover the skull and left to cure ~30 min at room temperature to provide a smooth surface for transcranial imaging. A similar procedure was used to prepare mice for two-photon imaging, with the addition of a dual-layer glass window implanted into a small (~4 mm square) craniotomy placed over the left primary visual cortex. The edges of the window were then sealed to the skull with dental cement.

Mesoscopic imaging

Widefield mesoscopic calcium imaging was performed using a Zeiss Axiozoom with a 1x, 0.25 NA objective with a 56 mm working distance (Zeiss). Epifluorescent excitation was provided by an LED bank (Spectra X Light Engine, Lumencor) using two output wavelengths: 395/25 (isosbestic for $GRAB_{ACh3.0}$) and 575/25 nm (jRCaMP1b). Emitted light passed through a dual camera image splitter (TwinCam, Cairn Research) then through either a 525/50 ($GRAB_{ACh3.0}$) or 630/75 (jRCaMP1b) emission filter (Chroma) before it reached two sCMOS cameras (Orca-Flash V3, Hamamatsu). Images were acquired at 512x512 resolution after 4x pixel binning, and each channel was acquired at 10 Hz with 20 ms exposure using HCImage software (Hamamatsu).

Two-photon imaging

Two-photon imaging was performed using a MOM microscope (Sutter Instruments) coupled to a 16x, 0.8 NA objective (Nikon). Excitation was driven by a Titanium-Sappire Laser (Mai-Tai eHP, DeepSee, Spectra-Physics) tuned to 920 nm. Emitted light was collected through a 525/50 filter and a gallium arsenide phosphide photomultiplier tube (Hamamatsu). Images were acquired at 512x512 resolution at 30 Hz using a galvo-resonant scan system controlled by ScanImage software (Vidrio).

Behavioral monitoring

All imaging was performed in awake, behaving mice that were head-fixed so that they could freely run on a cylindrical wheel. A magnetic angle sensor (Digikey) attached to the wheel continuously monitored wheel motion. Mice received at least three wheel-training habituation sessions before imaging to ensure consistent running bouts. During widefield imaging sessions, the face (including the pupil and whiskers) was illuminated with an IR LED bank and imaged with a miniature CMOS camera (Blackfly s-USB3, Flir) with a frame rate of 10 Hz using FlyCam2 software (Flir).

Visual stimulation

For visual stimulation experiments, sinusoidal drifting gratings (2 Hz, 0.04 cycles/degree) with varied contrast were generated using custom-written functions based on Psychtoolbox in Matlab and presented on an LCD monitor at a distance of 20 cm from the right eye. Stimuli were presented for 2 seconds with a 5 second inter-stimulus interval.

Data Analysis

All analyses were conducted using custom-written scripts in MATLAB (Mathworks). SVM classifiers were trained using publicly available software (*51*).

Preprocessing of behavior data

Pupil diameter and facial movements were extracted from face videography using FaceMap (*3*). For subsequent analysis, facial movement is defined as the the first component of FaceMap-based decomposition. Singular value decomposition (SVD) was applied to the face movie to extract the principal components (PCs) explaining the distinct movements apparent on the mouse's face. Wheel position was obtained from a linear angle detector attached to the wheel axle by unwrapping the temporal phase and then computing the traveled distance (cm). Locomotion speed was computed as the differential of the smoothed distance (cm/sec) using a 0.5 second window. Epochs of sustained locomotion and quiescence were extracted using change-point detection as described (*1*). High/low Pupil and Face epochs were extracted from within quiescence segments where z-score normalized values exceeded high/low thresholds of 60%/40% quantiles.

Preprocessing of mesoscopic imaging data

Imaging frames for green and red collection paths were grouped and down-sampled from 512X512 to 256X256 followed by an automatic 'rigid' transformation (imregtform, Matlab). In some cases, registration points were manually selected and a 'similarity' geometric transformation was applied. Detrending was applied using a low pass filter ($N = 100$, $f_{\text{cutoff}} = 0.001Hz$). Time traces were obtained using $(\Delta F/F)_i = \left(F_i - F_{i,o}\right)/F_{i,o}$ where F_i is the fluorescence of pixel i and $F_{i,o}$ is the corresponding low-pass filtered signal.

Hemodynamics correction

Hemodynamic artifacts were removed using a linear regression accounting for spatiotemporal dependencies between neighboring pixels (*1*). We used the isosbestic excitation of $\text{GRAB}_{\text{ACh3.0}}$ (395 nm) co-expressed in these mice as a means of measuring activity-independent fluctuations in fluorescence associated with hemodynamic signals. Briefly, given two $p \times 1$ random signals y_1 and y_2 corresponding to $\Delta F/F$ of p pixels for two excitation wavelengths "green" and "UV", we consider the following linear model:

$$y_1 = x + z + \eta,$$
$$y_2 = Az + \xi,$$

where x and z are mutually uncorrelated $p \times 1$ random signals corresponding to p pixels of the neuronal and hemodynamic signals, respectively. η and ξ are white Gaussian $p \times 1$ noise signals and A is an unknown $p \times p$ real invertible matrix. We estimate the neuronal signal as the optimal linear estimator for x (in the sense of Minimum Mean Squared Error):

$$\hat{x} = H\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad H = \Sigma_{xy}\Sigma_y^{-1},$$

where $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is given by stacking y_1 on top of y_2, $\Sigma_y = \mathbb{E}[yy^T]$ is the autocorrelation matrix of y and $\Sigma_{xy} = \mathbb{E}[xy^T]$ is the cross-correlation matrix between x and y. The matrix Σ_y is estimated directly from the observations, and the matrix Σ_{xy} is estimated by(*1*):

$$\Sigma_{xy} = \left(\Sigma_{y_1} - \sigma_\eta^2 I - \left(\Sigma_{y_1 y_2}\left(\Sigma_{y_2} - \sigma_\xi^2 I\right)^{-1}\Sigma_{y_2}^{-1}\Sigma_{y_1 y_2}^T\right)^T \quad 0\right),$$

where σ_η^2 and σ_ξ^2 are the noise variances of η and ξ, respectively, and I is the $p \times p$ identity matrix. The noise variances σ_η^2 and σ_ξ^2 are evaluated according to the median of the singular values of the corresponding correlation matrices Σ_{y_1} and Σ_{y_2} (*52*). This analysis is usually performed in patches where the size of the patch, p, is

determined by the amount of time samples available and estimated parameters. In the present study, we used a patch size of $p = 9$. The final activity traces were obtained by z-scoring the corrected $\Delta F/F$ signals per pixel.

Parcellation of mesoscopic data using LSSC

Functional parcellation of mesoscopic data was performed using Local Selective Spectral Clustering (LSSC)(*29*). Briefly, this method identifies areas of co-activity by building a graph where nodes are pixels and edge weights are determined by pairwise similarities between activity traces of pixels obtained by the following kernel:

$$\boldsymbol{K}(i,j) = \exp\left\{-\left\|(\Delta F/F)_i - (\Delta F/F)_j\right\|^2/\sigma^2\right\}$$

where σ is a parameter expressing a similarity radius. A row-stochastic matrix $\boldsymbol{P}$ is obtained by normalizing the rows such that $\boldsymbol{P} = \boldsymbol{D}^{-1}\boldsymbol{K}$, where $\boldsymbol{D}(i,i) = \sum_j \boldsymbol{K}(i,j)$. The matrix $\boldsymbol{P}$ can be viewed as a transition matrix of a Markov chain of the graph where $\boldsymbol{P}(i,j)$ is the probability to jump from node (pixel) i to node (pixel) j. We obtain a non-linear embedding of pixels by calculating the d right eigenvectors with the largest eigenvalues of $\boldsymbol{P}$:

$$(\Delta F/F)_i \rightarrow \psi^{(n)}(i) = \begin{pmatrix} \psi^1(i) \\ \vdots \\ \psi^n(i) \end{pmatrix}$$

Overall, by taking n to be significantly smaller than the number of time samples, every pixel is represented by a lower dimensional embedding $\psi^{(n)}$.

We evaluate the embedded representation $\psi^{(n)}$ and calculate the spectral embedding norm (*53*) of every pixel $s_i = \left\|\psi^{(n)}(i)\right\|$. LSSC uses an iterative approach for parcellating the brain where the inputs are the embedded representation of all pixels $\psi^{(n)}$ and their corresponding norms, $s_i, i = 1, \dots, p$, and lastly, a list of all pixels sorted by decreasing order of the embedding norm denoted by l. On each iteration the following operations are performed until coverage of at least ϑ percent of the mask brain pixels is assigned to parcels:

1. Select the first item on the list l (the pixel having the maximal norm, noted by i^*)
2. Select the axes in which i^* has the largest values, i.e., the subset: $L_{i*} = \left\{\ell_1, \ell_2, \dots, \ell_{d_i}\right\}$ such that $\left|\psi^{\ell_1}(i)\right| \geq \left|\psi^{\ell_2}(i)\right| \geq \left|\psi^{\ell_3}(i)\right| \geq \cdots$
3. Obtain the pixels whose embeddings are closer to $\psi^{(n)}(i)$ than to the origin based on the axes L_{i*} and assign them to the cluster k, i.e.:
$$C_k = \{j \mid \|\psi^{L_i}(i) - \psi^{L_i}(j)\|^2 < \|\psi^{L_i}(j)\|^2\}$$
4. Remove the set C_k from the list l: $l \leftarrow l \setminus C_k$
5. $k \leftarrow k + 1$
6. If at least ϑ percent of the mask of the brain is assigned to a specific parcel, then break.

The output is therefore a set of clusters $\{C_k\}$ where each clusters contains the pixels in that cluster. To increase robustness, we divided every session into 10 disjoint segments (folds), extracted the embedding on every fold and evaluated the embedding norm as the maximal value across all 10 folds. We refined the brain parcellation by merging parcels whose time traces are correlated more than a given threshold. Overlapping pixels were assigned to the parcel with closest centroid (in the embedding space). Additionally, unassigned isolated pixels (if any) were assigned to the (spatially) closest parcel. Isolated pixels within the borders of more than one parcel were assigned to the closest cluster (in the embedding space). Each animal and session was parcellated to reach a 95% coverage of the mask of the brain where clusters were merged based on a threshold of 0.99, resulting in ~45 parcels per hemisphere. Time series for parcels were extracted by averaging values for all pixels within the parcel (see preprocessing of mesoscopic data above).

ROI extraction by LSSC

We also used LSSC to identify cell bodies from the two-photon imaging data. The overall approach is similar to the parcellation process except for the stopping condition, where iterations continue until a maximal

number of cells is reached. In the refinement stage, identified cells that smaller than 15 pixels were discarded and overlapping regions were resolved by de-mixing (*29*).

Taylor expansion for estimating behavior as a function of neuronal activity

We formulate the link between temporal dynamics of neuronal activity $\boldsymbol{x}_t \in \mathbb{R}^N$ and an observed behavior b_t as:

$$b_t = f(\boldsymbol{x_t})$$

where f is an unknown function. Assuming that $f(\boldsymbol{x}_t)$ is 2 times differentiable, we can write its second-order Taylor's expansion as:

$$b_t = f(\boldsymbol{x}_t) \approx f(\overline{\boldsymbol{x}}) + \sum_i \left.\frac{\partial f}{\partial x_t(i)}\right|_{\overline{\boldsymbol{x}}} \left(x_t(i) - \bar{x}(i)\right) + \frac{1}{2}\sum_i \sum_j \left.\frac{\partial^2 f}{\partial x_t(i)\partial x_t(j)}\right|_{\overline{\boldsymbol{x}}} \left(x_t(i) - \bar{x}(i)\right)\left(x_t(j) - \bar{x}(j)\right) + \epsilon$$

where $\overline{\boldsymbol{x}}$ is the average neuronal activity (across time), and ϵ is the error of neglecting higher orders of $\boldsymbol{x}_t$. Simplifying this equation leads to:

$$b_t \approx \beta_0 + \sum_i \beta_1(i) x_t(i) + \sum_i \sum_j \beta_2(i,j) C_t(i,j) + \epsilon \qquad (3)$$

where $C_t(i,j) = \left(x_t(i) - \bar{x}(i)\right)\left(x_t(j) - \bar{x}(j)\right)$ is the time trace of the instantaneous interaction between brain region i and brain region j and $\beta_n, n = 0,1,2$ are the model parameters. Overall eq. (3) proposes a linear model for behavior based on two temporal signals - the activity $x_t(i)$ and the pairwise interaction $C_t(i,j)$, which is a nonlinear second-order function of elements of $\boldsymbol{x}_t$. Since the elements $x_t(i)$ and $C_t(i,j)$ are linearly independent for all $i,j \in \{1,N\}$, can measure the decoding power of each of these two components $\boldsymbol{x}_t$ and $\boldsymbol{C}_t$ independently.

In eq. (3) the instantaneous interactions $C_t(i,j)$ are evaluated based on a single time point. In practice, estimating all pairwise interactions at a single point is highly sensitive to noise. Thus, we evaluate the interactions over a short sliding time window to obtain the sample covariance $\hat{C}_t(i,j)$ as a smoothed and more robust estimation for the temporal evolution of $C_t(i,j)$:

$$\hat{C}_t(i,j) = \frac{\sum_{\tau=t-N_t/2}^{t+N_t/2}\left(x_\tau(i)-\overline{x}_t(i)\right)\left(x_\tau(j)-\overline{x}_t(j)\right)}{\sqrt{\sum_{\tau=t-N_t/2}^{t+N_t/2}\left(x_\tau(i)-\overline{x}_t(i)\right)^2}\sqrt{\sum_{\tau=t-N_t/2}^{t+N_t/2}\left(x_\tau(j)-\overline{x}_t(j)\right)^2}} \qquad (4)$$

where $\overline{x}_t(j)$ is the smoothed averaged activity:

$$\overline{x}_t(i) = \frac{1}{N_t}\sum_{\tau=t-N/2}^{t+N_t/2} x_\tau(i)$$

Inserting $\hat{C}_t(i,j)$ into (3) leads to eqn. (1). Overall, $\widehat{\boldsymbol{C}}$ is a three-dimensional tensor of parcels by parcels by time, where each element $\hat{C}_t(i,j)$ is a time trace of the instantaneous correlation coefficient between parcel i and parcel j. For most analyses, N_t was 30 (corresponding to a 3 second moving window). In all cases, the time-step was set to be 1 frame (0.1 second).

Riemannian projection of correlation matrices

Correlation matrices are Symmetric and Positive Definite (SPD, i.e., symmetric and full rank) and whose underlying geometry is a manifold shaped like a cone with a Riemannian metric (Supplemental Figure 2) (*54*). The distances between two correlation matrices on this cone is defined by the geodesic distance, the length of the arc connecting these matrices, whereas the Euclidean distance is not an accurate measure for this geodesic distance. To accurately capture distances between SPD matrices, Riemannian geometry is often used to project them onto a tangent Euclidean space where the geodesic length is approximated by the Euclidian distances between the corresponding projections. This evaluation becomes more accurate if the plane is tangent to the cone at a point that is relatively close to all relevant matrices, usually taken as their Riemannian mean.

Briefly, let $\{\boldsymbol{C}_k\}$ be a set of K SPD matrices. Denote $\overline{\boldsymbol{C}}$ as the Riemannian mean of the set and $\overline{\boldsymbol{S}}$ as its equivalent in the tangent plane. $\overline{\boldsymbol{C}}$ and $\overline{\boldsymbol{S}}$ are calculated using the following iterative equations:

$$\overline{\boldsymbol{S}}_n = \frac{1}{K}\sum_{k=1}^{K} \overline{\boldsymbol{C}}_n^{1/2} \log\left(\overline{\boldsymbol{C}}_n^{-1/2} \boldsymbol{C}_k \overline{\boldsymbol{C}}_n^{-1/2}\right) \overline{\boldsymbol{C}}_n^{1/2}$$

$$\overline{\boldsymbol{C}}_{n+1} = \overline{\boldsymbol{C}}_n^{1/2} \exp\left(\overline{\boldsymbol{C}}_n^{-1/2} \overline{\boldsymbol{S}}_n \overline{\boldsymbol{C}}_n^{-1/2}\right) \overline{\boldsymbol{C}}_n^{1/2}$$

Where log (·) and exp (·) are the matrix logarithm and matrix exponential, respectively, and where the Euclidean mean is used to initialize: $\overline{\boldsymbol{C}}_0 = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{C}_k$. Convergence is obtained when the Frobenius norm of $\overline{\boldsymbol{S}}_n$ is smaller than a pre-set parameter ε: $|\overline{\boldsymbol{S}}_n|_F < \varepsilon$.

The projections of $\{\boldsymbol{C}_k\}$ onto the tangent plane to the cone at the Riemannian mean are given by:

$$\boldsymbol{S}_k = \overline{\boldsymbol{C}}^{1/2} \log\left(\overline{\boldsymbol{C}}^{-1/2} \boldsymbol{C}_k \overline{\boldsymbol{C}}^{-1/2}\right) \overline{\boldsymbol{C}}^{1/2}, \qquad k = 1, \dots, K$$

As presented previously (*55*), the pairwise distances, $d_R^2(\boldsymbol{C}_k, \boldsymbol{C}_l)$ on the cone between correlation matrices $\{\boldsymbol{C}_k\}$ can be approximated by the Euclidean distances between their corresponding projections $\{\boldsymbol{S}_k\}$**:**

$$d_R^2(\boldsymbol{C}_k, \boldsymbol{C}_l) \cong \left\|\tilde{\boldsymbol{S}}_k - \tilde{\boldsymbol{S}}_l\right\|^2,$$

where $\tilde{\boldsymbol{S}}_{\boldsymbol{k}} = \log\left(\overline{\boldsymbol{C}}^{-1/2} \boldsymbol{C}_k \overline{\boldsymbol{C}}^{-1/2}\right)$. This method requires that all matrices $\{\boldsymbol{C}_k\}$ would be full rank (*56, 57*). In practice this is not always the case if the number of time points for evaluation of the correlation matrices is smaller than p, i.e. $N_t < p$. Therefore, we add a regularization term, $\lambda\boldsymbol{I}$**,** to each correlation matrix $\boldsymbol{C}_k$ (*31*) where λ is set to the median of the singular values of $\boldsymbol{x}_t$ (*52*).

Dimensionality reduction by diffusion embedding

The series of matrices $\boldsymbol{C}_t$ are symmetrical and therefore the dimension of their projections, $\{\boldsymbol{S}_t\}$, is equal to $\binom{p}{2}$ resulting in a high dimensional signal. To analyze the dynamics of this signal, we used diffusion geometry to obtain a low dimensional representation, capturing the dynamical properties of the correlation traces. Unlike LSSC where we reduce the dimension across time samples, in this case we reduce the dimension of parcels; we evaluated the $N_t \times N_t$ kernel matrix of $\{\boldsymbol{S}_t\}$:

$$\boldsymbol{A}(i,j) = \exp\left\{-\left\|\boldsymbol{S}_i - \boldsymbol{S}_j\right\|^2/\sigma^2\right\} \qquad (3)$$

where σ, which is a scale parameter evaluated as the median of pairwise distances between each projected matrix and its k-nearest neighbors where k=20. Note that our results are highly robust to variation in this parameter (Supplemental Figure 2). We then normalized the kernel $\boldsymbol{A}$ to be row-stochastic and obtained the low dimensional representation for the correlation traces as the right eigenvectors of the normalized kernel matrix:

$$\boldsymbol{S}_t \to \phi_t^{(n)} = \begin{pmatrix} \phi_t^1 \\ \vdots \\ \phi_t^n \end{pmatrix}$$

For comparison, we carried out similar diffusion embedding based on Euclidean distances (Supplemental Figure S2).

Dimensionality reduction by principal component analysis

As a comparison to LSSC, we also used principal component analysis to reduce the dimensionality of widefield data (Supplemental Figure S2). Principal components were derived using the 'pca' function in Matlab.

Visual response analysis

Visual responses were evaluated, per parcel, as the difference between peak response during stimulus presentation and the average activity during the preceding two seconds. The responses were averaged per contrast value and normalized by the response to 100% contrast. To quantify the accuracy with which visual responses are encoded by visual activity, or embedded network activity/correlations, we trained a binary classifier (linear SVM, libsvm) to separate the visual response and the two seconds prior to stimulus onset. We used 10-fold cross validation to estimate the classification accuracy for every contrast value based on each predictor.

Modeling Behavior

Behavioral variables (pupil, facial movements, running speed) were modeled using linear ridge regression with 10-fold cross validation. Each session was divided into 10-disjoint continuous segments, where on each fold one segment was set aside for testing and the other segments were used for training. We assessed the predictive power of neuronal activity using the following predictors: raw activity and smoothed activity (using a 3 second moving window). For pairwise correlations, we used: raw correlation traces, diffusion embedding of correlation traces using either Euclidean or Riemannian distances. To directly compare the predictive power of activity versus embedded correlations, we combined these predictors and evaluated the goodness of fit of the joint model. We then shuffled either activity or embedded correlations through time and trained the resulting model to assess $R^2_{\text{shuffled activity}}$ and $R^2_{\text{shuffled } \phi_t}$.

Reconstruction Error

Reconstruction error of diffusion embedding of functional connectivity was evaluated by:

$$\varepsilon^2 = \frac{\sum_{i,j}\sum_t \left(\hat{C}_t(i,j) - \hat{C}_t^n(i,j)\right)^2}{\sum_{i,j}\sum_t \left(\hat{C}_t(i,j) - \bar{C}(i,j)\right)^2}$$

where $\widehat{\boldsymbol{C}}(t)$ is the pairwise functional connectivity of brain parcels, $\overline{\boldsymbol{C}}$ is the temporal average and where $\widehat{\boldsymbol{C}}_t^n$ is the reconstruction of $\widehat{\boldsymbol{C}}_t$ based on n leading components of its embedded trace $\phi_t^{(n)}$:

$$\hat{C}_t^n(i,j) = \sum_{i=1}^{n} \left(\phi_{t'}^i \cdot \hat{C}_t(i,j)\right) \varphi_t^i$$

where φ_t^i are the left eigenvectors of the normalized kernel matrix.

Modeling correlations data by embedding

To quantify the relationship between the embedding of functional connectivity across the cortex ϕ_C and the time-varying correlation between specific pairs of parcels, we used linear regression (10-fold cross validation) and obtained an R^2 value for every pair-wise correlation trace. To match LSSC parcels across animals, we identified the LSSC parcels whose center of mass were closest to each Allen Atlas brain parcel (23 parcels overall in a single hemisphere) and extracted a 23×23 matrix of R^2 values per session. We averaged these matrices across animals and extracted the rows corresponding to individual parcels. Each row was then represented as a separate brain map image, color-coded by the R^2 value corresponding to the correlation between the target (specific to that image) and each of the other parcels.

Evaluating Integrated Network Configuration

We evaluated a time trace of the average participation coefficient (across parcels), based on the correlation time traces (3 seconds) and an arbitrary partition of the brain using a line bisecting the neocortex. The participation coefficient is calculated per parcel in a given time window t as:

$$P_t(i) = 1 - \sum_{s=1}^{N_M} \left(\frac{\kappa_t(i,s)}{k_t(i)}\right)^2$$

where $\kappa_t(i,s)$ is the sum of the correlations (absolute values) between parcel i and all the parcels in sub-module s and $k_t(i)$ is the sum of all correlations (absolute values) between parcel i and all other parcels. Therefore, a segregated network would be characterized by high connectivity between parcels related to the same module and low connectivity between sub-modules. In this case the average participation coefficient (across parcels) would approach zero. The opposite would happen in an integrated network where parcels communicate outside their

sub-modules just as much as they do within their sub-modules. In this case the average participation coefficient would approach 1.

We then evaluated the difference of the behavior variables (pupil size, facial movement, and locomotion speed) between time points where the network was in an extreme integrated state (top 10% participation coefficient) and extreme segregated state (bottom 10% participation coefficient) (*36*). By rotating the line in 30, 60, 90, 120, 150 degrees we measured the delta of behavior variables on integrated and segregated configurations based on different ways for partition of the brain into sub-networks.

References

1. S. Lohani *et al.*, Dual color mesoscopic imaging reveals spatiotemporally heterogeneous coordination of cholinergic and neocortical activity. *BioRXiv*, (2020).
2. S. Musall, M. T. Kaufman, A. L. Juavinett, S. Gluf, A. K. Churchland, Single-trial neural dynamics are dominated by richly varied movements. *Nat Neurosci* **22**, 1677-1686 (2019).
3. C. Stringer *et al.*, Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, 255 (2019).
4. M. Vinck, R. Batista-Brito, U. Knoblich, J. A. Cardin, Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron* **86**, 740-754 (2015).
5. M. D. Fox, M. E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* **8**, 700-711 (2007).
6. M. J. McGinley, S. V. David, D. A. McCormick, Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87**, 179-192 (2015).
7. D. J. Lurie *et al.*, Questions and controversies in the study of time-varying functional connectivity in resting fMRI. *Netw Neurosci* **4**, 30-69 (2020).
8. C. J. MacDowell, T. J. Buschman, Low-Dimensional Spatiotemporal Dynamics Underlie Cortex-wide Neural Activity. *Curr Biol* **30**, 2665-2680 e2668 (2020).
9. M. P. Vanni, A. W. Chan, M. Balbi, G. Silasi, T. H. Murphy, Mesoscale Mapping of Mouse Cortex Reveals Frequency-Dependent Cycling between Distinct Macroscale Functional Modules. *J Neurosci* **37**, 7513-7533 (2017).
10. M. H. Mohajerani *et al.*, Spontaneous cortical activity alternates between motifs defined by regional axonal projections. *Nat Neurosci* **16**, 1426-1435 (2013).
11. M. Breakspear, Dynamic models of large-scale brain activity. *Nat Neurosci* **20**, 340-352 (2017).
12. S. Gao, G. Mishne, D. Scheinost, Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics. *Human brain mapping* **42**, 4510-4524 (2021).
13. J. A. Cardin, M. C. Crair, M. J. Higley, Mesoscopic Imaging: Shining a Wide Light on Large-Scale Neural Dynamics. *Neuron* **108**, 33-43 (2020).
14. L. Tang, M. J. Higley, Layer 5 Circuits in V1 Differentially Control Visuomotor Behavior. *Neuron* **105**, 346-354 e345 (2020).
15. M. Boly *et al.*, Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc Natl Acad Sci U S A* **104**, 12187-12192 (2007).
16. J. M. Palva, S. Palva, Roles of multiscale brain activity fluctuations in shaping the variability and dynamics of psychophysical performance. *Progress in brain research* **193**, 335-350 (2011).
17. J. W. de Gee *et al.*, Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains. *Elife* **9**, (2020).
18. E. A. K. Jacobs, N. A. Steinmetz, A. J. Peters, M. Carandini, K. D. Harris, Cortical State Fluctuations during Sensory Decision Making. *Curr Biol* **30**, 4944-4955 e4947 (2020).
19. G. G. Gregoriou, S. J. Gotts, H. Zhou, R. Desimone, High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207-1210 (2009).
20. T. Ito *et al.*, Task-evoked activity quenches neural correlations and variability across cortical areas. *PLoS computational biology* **16**, e1007983 (2020).
21. K. C. Wood, C. F. Angeloni, K. Oxman, C. Clopath, M. N. Geffen, Neuronal activity in sensory cortex predicts the specificity of learning in mice. *Nature communications* **13**, 1167 (2022).
22. L. N. Driscoll, N. L. Pettit, M. Minderer, S. N. Chettih, C. D. Harvey, Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell* **170**, 986-999 e916 (2017).
23. K. M. Hallinen *et al.*, Decoding locomotion from population neural activity in moving C. elegans. *Elife* **10**, (2021).
24. Y. Livneh *et al.*, Estimation of Current and Future Physiological States in Insular Cortex. *Neuron* **105**, 1094-1111 e1010 (2020).
25. H. Dana *et al.*, Sensitive red protein calcium indicators for imaging neural activity. *Elife* **5**, (2016).
26. D. Barson *et al.*, Simultaneous mesoscopic and two-photon imaging of neuronal activity in cortical circuits. *Nature methods* **17**, 107-113 (2020).
27. A. S. Hamodi, A. Martinez Sabino, N. D. Fitzgerald, D. Moschou, M. C. Crair, Transverse sinus injections drive robust whole-brain expression of transgenes. *Elife* **9**, (2020).
28. Y. Ma *et al.*, Wide-field optical mapping of neural activity and brain haemodynamics: considerations and novel approaches. *Philos Trans R Soc Lond B Biol Sci* **371**, (2016).

29. G. Mishne, R. R. Coifman, M. Lavzin, J. Schiller, Automated cellular structure extraction in biological images with applications to calcium imaging data. *BioRXiv*, (2018).
30. Q. Wang *et al.*, The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936-953 e920 (2020).
31. K. Abbas *et al.*, Geodesic Distance on Optimally Regularized Functional Connectomes Uncovers Individual Fingerprints. *Brain connectivity* **11**, 333-348 (2021).
32. S. Gao, G. Mishne, D. Scheinost, in *Medical image computing and computer assisted intervention,* A. L. e. a. Martel, Ed. (Springer, 2020).
33. M. Hauser, A. Ray, Principles of Riemannian Geometry in Neural Networks. *Advances in Neural Information Processing Systems* **30**, (2017).
34. S. Lafon, Y. Keller, R. R. Coifman, Data fusion and multicue data matching by diffusion maps. *IEEE Trans Pattern Anal Mach Intell* **28**, 1784-1797 (2006).
35. T. W. Chen *et al.*, Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295-300 (2013).
36. M. Fukushima *et al.*, Structure-function relationships during segregated and integrated network states of human brain functional connectivity. *Brain Struct Funct* **223**, 1091-1106 (2018).
37. V. D. Calhoun, R. Miller, G. Pearlson, T. Adali, The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* **84**, 262-274 (2014).
38. E. Tagliazucchi, H. Laufs, Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. *Neuron* **82**, 695-708 (2014).
39. R. Gao, R. L. van den Brink, T. Pfeffer, B. Voytek, Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *Elife* **9**, (2020).
40. R. V. Raut, A. Z. Snyder, M. E. Raichle, Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proc Natl Acad Sci U S A* **117**, 20890-20897 (2020).
41. E. M. R. Lake *et al.*, Simultaneous cortex-wide fluorescence Ca(2+) imaging and whole-brain fMRI. *Nature methods* **17**, 1262-1271 (2020).
42. K. B. Clancy, I. Orsolic, T. D. Mrsic-Flogel, Locomotion-dependent remapping of distributed cortical networks. *Nat Neurosci* **22**, 778-786 (2019).
43. A. J. Peters, J. M. J. Fabre, N. A. Steinmetz, K. D. Harris, M. Carandini, Striatal activity topographically reflects cortical activity. *Nature* **591**, 420-425 (2021).
44. X. Liu *et al.*, Multimodal neural recordings with Neuro-FITM uncover diverse patterns of cortical-hippocampal interactions. *Nat Neurosci* **24**, 886-896 (2021).
45. Y. Gallero-Salas *et al.*, Sensory and Behavioral Components of Neocortical Signal Flow in Discrimination Tasks with Short-Term Memory. *Neuron* **109**, 135-148 e136 (2021).
46. S. Saxena *et al.*, Localized semi-nonnegative matrix factorization (LocaNMF) of widefield calcium imaging data. *PLoS computational biology* **16**, e1007791 (2020).
47. A. Puscian, H. Benisty, M. J. Higley, NMDAR-dependent emergence of behavioral representation in primary visual cortex. *Cell Reports* **in press**, (2020).
48. J. Poort *et al.*, Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron* **86**, 1478-1490 (2015).
49. H. Makino, T. Komiyama, Learning enhances the relative impact of top-down processing in the visual cortex. *Nat Neurosci* **18**, 1116-1122 (2015).
50. M. Jing *et al.*, An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. *Nature methods* **17**, 1139-1146 (2020).
51. C. C. Chang, C. J. Lin, LIBSVM: A Library for Support Vector Machines. *Acm T Intel Syst Tec* **2**, (2011).
52. M. Gavish, D. L. Donoho, The Optimal Hard Threshold for Singular Values is 4/root 3. *Ieee T Inform Theory* **60**, 5040-5053 (2014).
53. X. Cheng, G. Mishne, Spectral Embedding Norm: Looking Deep into the Spectrum of the Graph Laplacian. *SIAM J Imaging Sci* **13**, 1015-1048 (2020).
54. S. Diamond, S. Boyd, CVXPY: A Python-Embedded Modeling Language for Convex Optimization. *J Mach Learn Res* **17**, (2016).
55. O. Tuzel, F. Porikli, P. Meer, Pedestrian detection via classification on Riemannian manifolds. *Ieee T Pattern Anal* **30**, 1713-1727 (2008).
56. A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing* **112**, 172-178 (2013).
57. O. Yair, M. Ben-Chen, R. Talmon, Parallel Transport on the Cone Manifold of SPD Matrices for Domain Adaptation. *Ieee T Signal Proces* **67**, 1797-1811 (2019).

Figures and Figure Legends

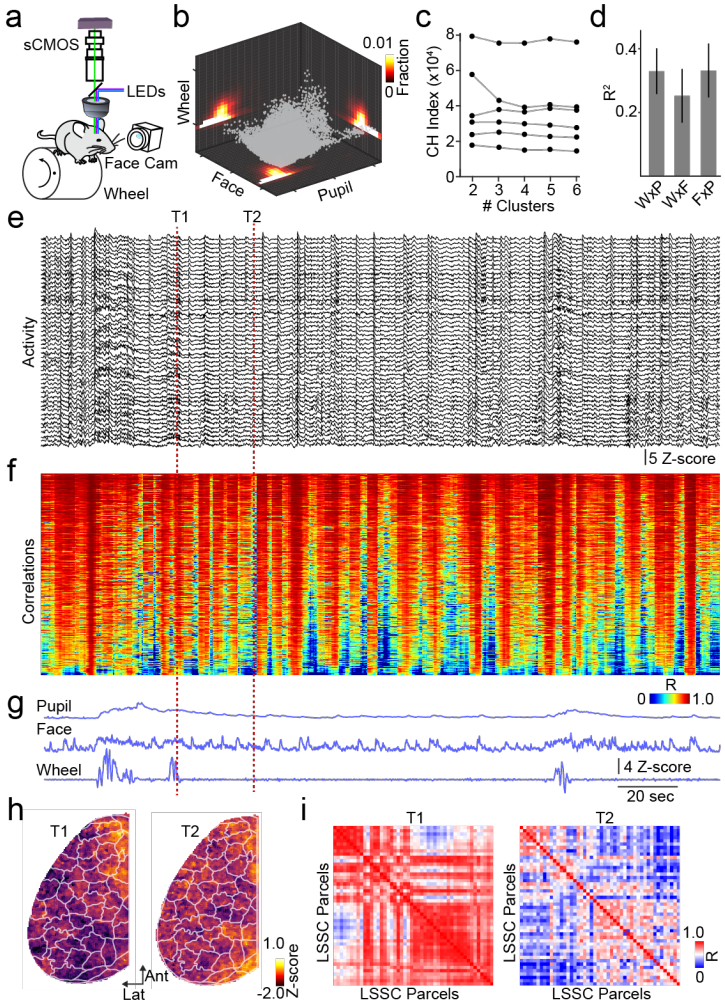


Figure 1. Mesoscopic imaging of cortical activity and functional connectivity. **a**, Schematic illustrating the setup for simultaneous behavioral monitoring and mesoscopic calcium imaging. **b**, Scatter plot illustrating the distribution of Z-scored behavioral metric values (locomotion, facial movement, and pupil size) collapsed over time and across all subjects (n=6 mice). **c**, Population data showing Calinski-Harabasz index values for K-means clustering of behavioral metrics for all subjects. **d**, Population data showing average (±SEM) Pearson's R^2 values for the relationships between wheel (W), pupil (P), and facial movements (F) for all subjects. **e**, Example time series from one animal showing cortical activity across the cortex. Each trace corresponds to one LSSC-based parcel. **f**, Heat map illustrating the time-series of pairwise correlations between each parcel from (D). Data are sorted by increasing standard deviation. **g**, Time series of behavioral metrics corresponding to the data shown in (D) and (E). **h**, Example LSSC-based functional parcellation of the neocortex for the data shown above. Left and right images are for the timepoints indicated by vertical red lines. **i**, Example pairwise correlation matrices for the data in (E) at the time points indicated.

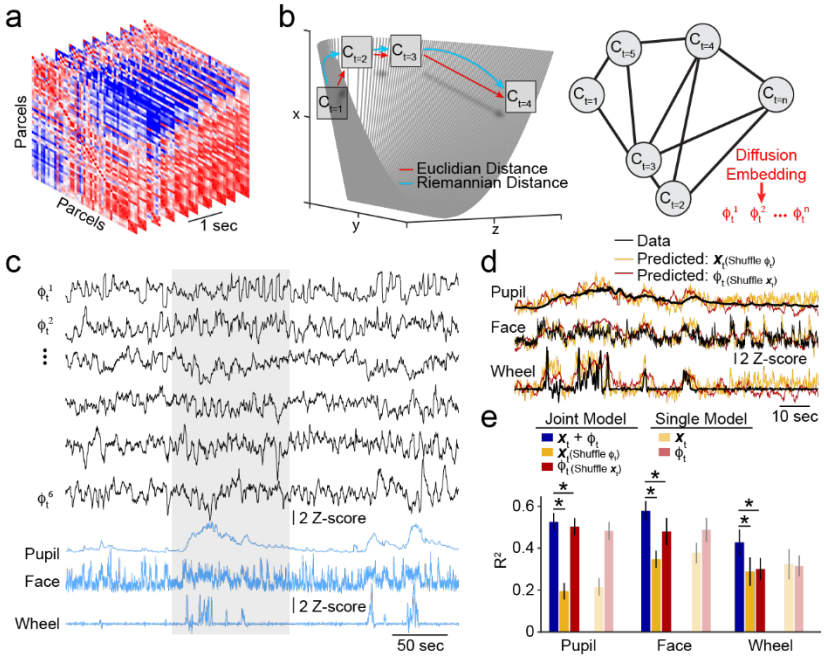


Figure 2. Dynamic functional connectivity encodes rapid behavioral variations.
a, Example sequential pairwise, parcel-based correlation matrices, derived from a sliding window applied to neural activity across the cortex. **b**, Left, Schematic illustrating the cone-shaped Riemannian manifold (based on 2×2 SPD matrices where $x, z > 0$ and $y^2 < xz$, see Methods) used to calculate distances between correlation matrices. The Riemannian measurement reflects geodesic distance that is ignored when using Euclidean distance. Right, Illustration of a "graph of graphs", whose nodes are SPD matrices and edges are weighted by the length of the geodesic arc along the Riemannian cone, that is used to extract diffusion embedding (ϕ_t) components. **c**, Example diffusion embedding components capturing dynamics of functional connectivity ϕ_t. Simultaneous time series for behavioral metrics are shown below in blue. **d**, Example behavioral data (black traces) showing fluctuations in pupil diameter, facial movement, and locomotion superimposed on predicted behavior estimated using a joint model in which either time-varying activity (red) or embedded correlations (yellow) have been shuffled for the region of data highlighted in (C). **e**, Population data showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using a joint model of activity and embedded correlations (blue), joint model with shuffled ϕ_t (yellow), joint model with shuffled activity (red), single predictor model using activity (pale yellow), and single predictor model using ϕ_t (pale red). * indicates $p<0.05$ (see main text).

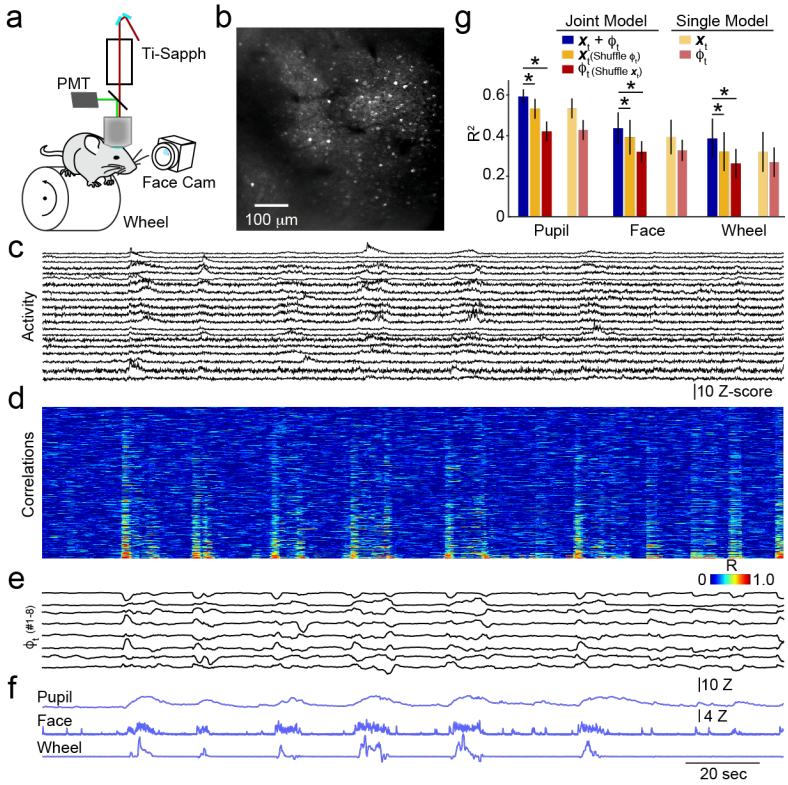


Figure 3. Local circuit dynamics encode spontaneous behavioral variation.
a, Schematic illustrating the setup for simultaneous behavioral monitoring and 2-photon calcium imaging. **b**, Example field of view showing individual GCaMP6s-expressing neurons in visual cortex. **c**, Example time series showing neuronal activity for all neurons in the field of view. **d**, Heat map illustrating the time-series of pairwise correlations between each neuron from (C). Data are sorted by increasing standard deviation. **e**, Example of the first six diffusion embedding components based on data in (D). **f**, Time-series for behavioral metrics corresponding to data in (C-E). **g**, Population data showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using a joint model of activity and embedded correlations (blue), joint model with shuffled ϕ_t (yellow), joint model with shuffled activity (red), single predictor model using activity (pale yellow), and single predictor model using ϕ_t (pale red). * indicates $p<0.05$ (see main text).

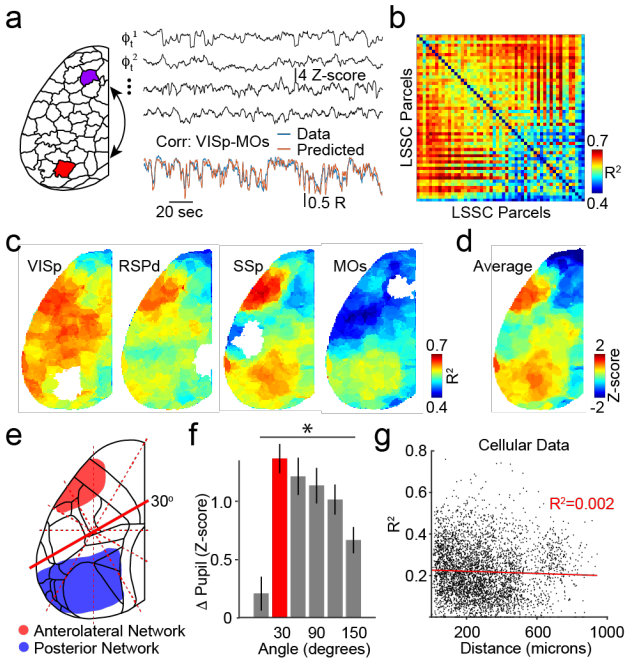
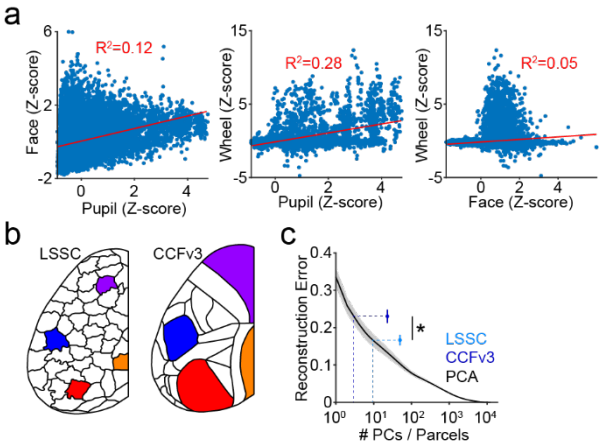
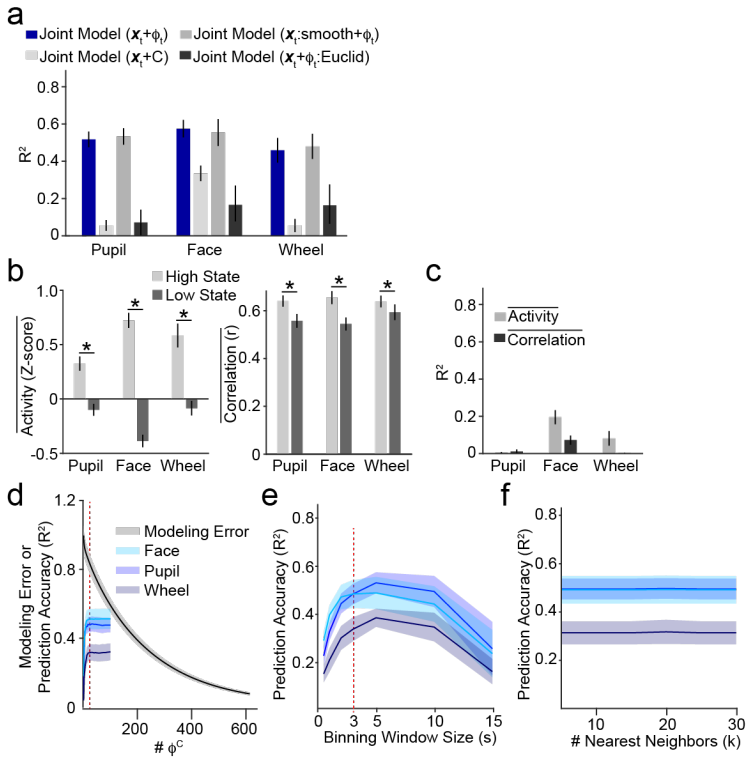


Figure 4. Dynamic functional connectivity reveals distinct cortical subnetworks.
a, Left illustration of LSSC-based parcellation, highlighting two parcels corresponding approximately to supplemental motor cortex (MOs) and primary visual cortex (VISp) based on CCFv3. Right, example components of correlation embedding for one animal (black), pairwise time-varying correlation between VISp and MOs (blue), and the predicted VISp-MOs correlation based on embedding. **b**, Example matrix from one animal showing the goodness of fit (R^2) for modeling the time-varying correlations between each pair of parcels using $\phi_t^{(20)}$. **c**, Average (n=6 mice) maps showing mean R^2 values for modeling the pairwise correlations of each cortical parcel with the indicated target parcel (shown in white). **d**, Grand average map showing R^2 values as in (C) collapsed across all animals (n=6) and all cortical parcels. **e**, Schematic illustrating anterolateral (red) and posterior (blue) subnetworks derived from (D), superimposed on the CCFv3 parcellation (black lines). **f**, Population data showing the average (±SEM) difference in pupil size for epochs corresponding to maximal and minimal network segregation versus a line angle bisecting LSSC parcels. 30° (red) corresponds to the division between anterolateral and posterior networks indicated in (E). * indicates $p<0.05$ (see main text). **g**, Goodness of fit (R^2) for modeling dynamic pairwise correlations between single neurons versus their inter-somatic distance.

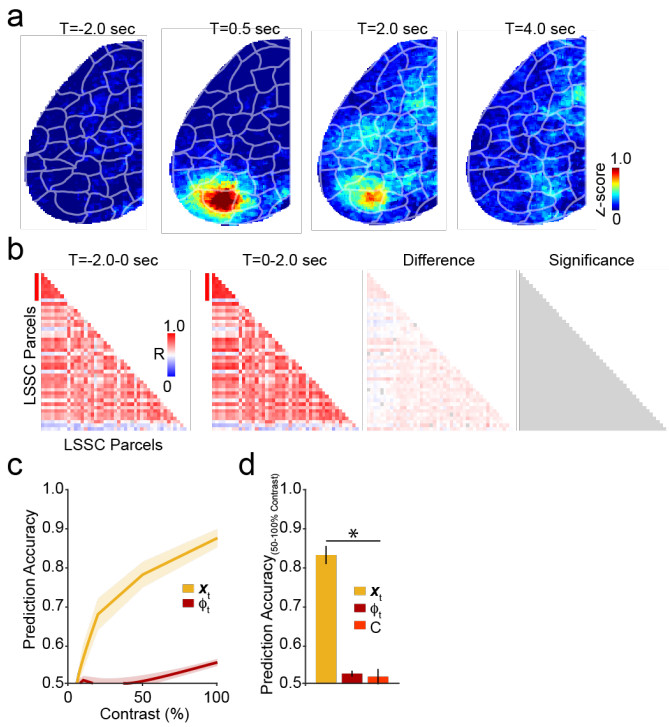
Supplemental Figures



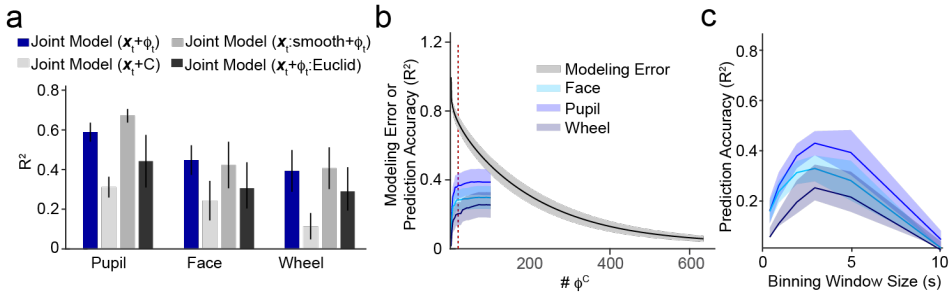
Supplemental Figure 1. Comparison of parcellation methods and average neural signaling across behavioral state.
a, Example scatter plots illustrating the relationship between Z-scored facial movement (Face), Pupil diameter (Pupil), and Locomotion (Wheel) for a representative subject. Pearson's correlation line and goodness of fit (R^2) are shown in red. **b**, Example LSSC-based functional parcellation (left) and CCFv3-based anatomical parcellation (right) for the data in Figure 1. Coloring indicates approximately matching areas (MO_s purple, SS_P blue, VIS_P red, RS orange). **c**, Reconstruction error for raw activity using either LSSC- or CCFv3-based parcellation or principal component analysis (PCA). PCA indicates error versus number of components included in the model. LSSC and CCFv3 points indicate the number of parcels obtained for each method.



Supplemental Figure 2. Comparison of behavioral prediction across different models. **a**, Widefield imaging population data of showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using a joint model of activity and embedded correlations (replicated from Figure 2, blue), joint model with raw correlations (light gray), joint model with smoothed (windowed) activity (medium gray), and joint model using Euclidean embedding (dark gray). **b**, Population data showing average (± SEM) of cortical activity (left) and correlations (right) across all LSSC-based parcels, comparing low versus high divisions of the indicated behavioral state. * indicates $p<0.05$ (see Main Text). **c**, Prediction accuracy for modeling behavioral variables using the average activity or correlations across all cortical parcels. **d**, Modeling reconstruction error of embedded correlations (black) or prediction error of behavioral metrics (shades of blue) versus the number of ϕ_C components. Shaded areas indicate ±SEM. **e**, Prediction accuracy (R^2) for behavioral metrics using $\phi_t^{(20)}$ versus sliding temporal window size for calculating correlations. **f**, Prediction accuracy (R^2) of behavioral metrics using $\phi_t^{(20)}$ versus the number of neighbors used for evaluating the scale factor of the diffusion kernel. * indicates $p<0.05$.

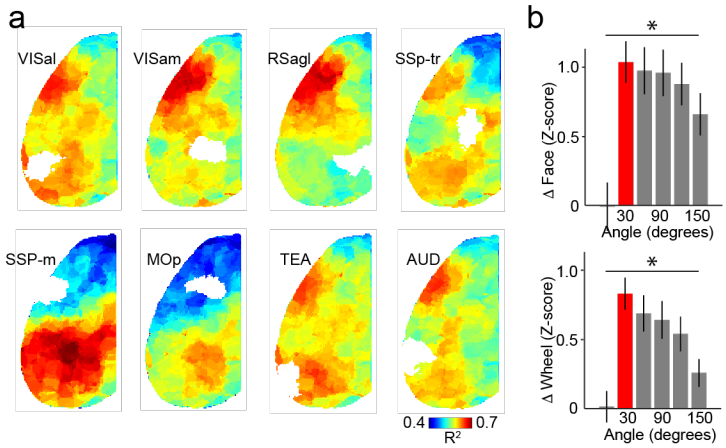


Supplementary Figure 3. Network representation of visually-evoked activity.
a, Example average (n=50 stimuli, one animal) mesoscopic imaging frames showing evoked visual responses (100% contrast) relative to stimulus onset. **b**, Example average correlation matrices corresponding to data from (a) showing functional connectivity of parcels before (first panel) and after (second panel) visual stimulation (100% contrast). The differences in correlation for each parcel (third panel) and the significant parcels (fourth panel) are shown. All cells are gray, indicating no significant differences ($p<0.05$, permutation test, Benjamini-Hotchberg multiple comparisons correction) for any parcel pair. **c**, Prediction accuracy for detecting the presentation of a visual stimulus for varying contrasts using activity in visual cortex (yellow) or ϕ_t (dark red). Shaded areas indicate SEM (n=6 mice). **d**, Population averages (n=6 mice) showing prediction accuracy of visual responses (averaging across trials with ≥50% stimulus contrast) for visual cortex activity (yellow), ϕ_t (red), and raw correlations (red). * indicates $p<0.05$ (ANOVA, Tukey's post-hoc tests, see main text).



Supplemental Figure 4. Diffusion embedding of cellular correlations is robust to choice of modeling parameters.

a, Cellular imaging population data of showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using a joint model of activity and embedded correlations (replicated from Figure 3, blue), joint model with raw correlations (light gray), joint model with smoothed (windowed) activity (medium gray), and joint model using Euclidean embedding (dark gray). **b**, Modeling reconstruction error of correlations (black) or prediction error of behavioral metrics (shades of blue) versus the number of ϕ_C components. Shaded areas indicate ±SEM. **c**, Prediction accuracy (R^2) of behavioral metrics using $\phi_t^{(20)}$ versus sliding temporal window size for calculating correlations.



Supplemental Figure 5. Dynamic connectivity reveals consistent spatial organization of cortical subnetworks.

a, Average (n=6 mice) maps showing mean R^2 values for modeling the pairwise correlations of each cortical parcel with the indicated target parcel (showing in white). **b**, Population data showing the average (±SEM) difference in facial movement (top) and wheel speed (bottom) for epochs corresponding to maximal and minimal network segregation versus a line angle bisecting LSSC parcels (as in Figure 4). * indicates $p<0.05$ (ANOVA).