Title

Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior

Permalink

https://escholarship.org/uc/item/6kg7x8r0

Journal

Nature Neuroscience, 27(1)

ISSN

1097-6256

Authors

Benisty, Hadas
Barson, Daniel
Moberly, Andrew H
et al.

Publication Date

2024

DOI

10.1038/s41593-023-01498-y

Published in final edited form as:
Nat Neurosci. 2024 January ; 27(1): 148–158. doi:10.1038/s41593-023-01498-y.

Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior

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

[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

Abstract

Experimental work across species has demonstrated that spontaneously generated behaviors are robustly coupled to variation in neural activity within the cerebral cortex. Functional MRI data suggest that temporal correlations in cortical networks vary across distinct behavioral states, providing for the dynamic reorganization of patterned activity. However, these studies generally lack the temporal resolution to establish links between cortical signals and the continuously varying fluctuations in spontaneous behavior observed in awake animals. Here, we used wide-field, mesoscopic calcium imaging to monitor cortical dynamics in awake mice and developed an approach to quantify rapidly time-varying functional connectivity. We show that spontaneous behaviors are represented by fast changes in both the magnitude and correlational structure of cortical network activity. Combining mesoscopic imaging with simultaneous cellular resolution 2-photon microscopy also demonstrated that correlations among neighboring neurons and between local and large-scale networks also encode behavior. Finally, the 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 neocortex and demonstrate an analytical framework for investigating time-varying functional connectivity in neural networks.

[†] Corresponding Author.
[*] These authors contributed equally to this manuscript.

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

Competing Interests Statement
The authors declare no conflicts of interest exist.

Code Availability
Custom written MATLAB scripts used in this study are available on github (https://github.com/cardin-higley-lab/Benisty_Higley_2023).

Introduction

Cognitive functions such as perception and attention require the dynamic activity of neuronal networks defined by synaptic connectivity over local and long-range spatial scales[1–3]. Moreover, animals cycle through behavioral states categorized by a variety of physical markers, including pupil dilation, facial movement, and locomotion, that are associated with changes in cognitive performance[4–9]. Such variation in motor behaviors and arousal are themselves strongly coupled to fluctuations in neural activity, and several recent studies have demonstrated clear modulation of spontaneous and sensory-evoked firing rates associated with transitions between behavioral states[10–14]. Work in both human and non-human subjects suggests that the spatiotemporal correlations between neural signals in large-scale networks spanning multiple brain regions also co-vary with state[1,14–20]. These correlations are often viewed as functional connectivity between nodes in a network, reflecting either true structural (synaptic) connections or common inputs[21]. From these results, behavioral state fluctuations might be viewed as shifts between distinct network configurations optimized for contextually-relevant cognitive functions.

While many analyses of state-dependent correlations in cortical networks rely on binning data within identified epochs (e.g., sleep versus wakefulness or quiescence versus arousal), others have explored continuously time-varying functional connectivity[15,22]. However, methodological challenges to monitoring large-scale brain dynamics at high spatial and temporal resolution have obscured the investigation of rapid co-variation in behavioral state and network connectivity. As neurons can be exquisitely sensitive to patterned or synchronized input[23], it seems reasonable to hypothesize that fast changes in network correlations are closely linked to the integrative function of single cells. Furthermore, the short-term correlation between two neuronal signals is a non-linear, second-order function of those signals that expresses the dynamics of their connectivity. Thus, temporal dynamics of activity and short-term correlations might differentially represent behaviorally relevant information. Nevertheless, the relative contributions of dynamic neural activity versus dynamic inter-node correlations to decoding behavior is largely unknown.

To explore this question, we used widefield mesoscopic and cellular resolution 2-photon calcium (Ca2+) imaging, both independently and simultaneously, to monitor neural activity in the awake, head-fixed mouse. We developed a methodological strategy that views the dynamic correlations between individual cortical regions and/or neurons as a “graph of graphs”, extracting their time-varying latent variables through non-Euclidean diffusion embedding[24]. We then asked how accurately these signals could be used to decode spontaneous fluctuations in behavior (measured by pupil dilation, facial movements, and locomotion), and sensory input. Our results show that time-varying network correlations carry significant information about behavioral metrics that can outperform analyses based on the dynamics of activity alone. In addition, the dynamic multimodal correlations between local cellular and large-scale networks are also significantly predictive of behavior. We then show that both cortical parcels and single neurons are dynamically correlated to one of two broad mesoscale networks that are distinct from traditional anatomical boundaries. Overall, these findings demonstrate that rapid fluctuations in functional connectivity across spatial scales provide a robust representation of spontaneous behavior.

Results

Monitoring spontaneous behavior and cortical dynamics

We carried out widefield, mesoscopic calcium imaging[3] 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–g, see Methods)[13,28]. Pupil diameter and locomotion were extracted from video recording of the eye and running wheel-speed, respectively. Facial movements were derived from video data with FaceMap[28], and the first extracted component was used for most subsequent analyses. Although some studies have used categorical definitions of behavioral state according to thresholding of motor signals[9,13,14,29], our data suggest that 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). These results suggest pupil, locomotion, and facial metrics may reflect at least partially distinct, highly dynamic latent variables, such as different neuromodulatory influences[14]. Here, we focused on the general ability of functional connectivity to encode behavior and therefore separately compared the relationships of time-varying network activity and correlations to each of these observable variables.

After normalization and hemodynamic correction of imaging data (see Methods)[14,30], we segmented the cortex into functional parcels using a graph theory-based approach that relies on spatiotemporal co-activity between pixels (LSSC, Figure 1h, Supplemental Figure S2)[31]. This approach yields comparable reconstruction errors for a similar number of parcels in comparison to either the Allen Institute CCFv3 anatomy-based atlas[32] or a uniform grid-based segmentation (see Methods). Principal component analysis yielded similar reconstruction errors as LSSC, and localized semi-nonnegative matrix factorization (LocalNMF)[33] gave near-zero errors, though both approaches generate non-disjoint parcels that may be more difficult to interpret (Supplemental Figure S2). We next extracted $\boldsymbol{x}_t$, the time-varying fluctuations in the fluorescence signal associated with each parcel (Figure 1e, see Methods). 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 1 f–i, see Methods). On average, correlations across the cortex were high but with large fluctuations over time (average r=0.6±0.03, average CV_r=0.4±0.1, n=6 mice). Indeed, the moment-to-moment values also appeared to co-vary with rapid behavioral changes (Figure 1f–i).

Dynamic functional connectivity encodes spontaneous behavior

We next began with a basic assumption that behavior can be represented as some function of multi-dimensional, time-varying neural activity. That is:

$$b_t = f(\boldsymbol{x}_t)$$

(1)

Here, $\boldsymbol{x}_t$ is an N-dimensional vector corresponding to the time-varying neural activity across N cortical parcels at time t. This relationship can be approximated using the first two terms of a standard Taylor expansion (see Methods):

$$b_t \approx \beta_0 + \boldsymbol{\beta}_1^T \boldsymbol{x}_t + \sum_i \sum_j \beta_2(i,j)\hat{C}_t(i,j) + \epsilon \tag{2}$$

Again, $\boldsymbol{x}_t$ is the time-varying neural activity and the second-order term $\hat{C}_t(i,j)$ corresponds to the time-varying pairwise sample correlations between parcels i and j. Representing behavior as a linear combination of time-varying neural signals is a common approach[9,34–37], and we hypothesized that 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 significantly improve decoding accuracy.

In an initial effort, fitting a linear ridge regression model for behavioral dynamics whose predictors are time-varying cortical activity and functional connectivity led to poor predictive power (Supplemental Figure S2) 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[24]. Each correlation matrix over a short temporal window (e.g., 3 seconds) can be viewed as a graph whose N 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, 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 LSSC 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–f). Prediction accuracy fell off rapidly for higher-order components of facial movement (Supplemental Figure S2). Additionally, accuracy was impaired when using either raw correlations, Euclidean rather than Riemannian distance for the diffusion embedding, or PCA for dimensionality reduction (Supplemental Figure S2). Smoothing or diffusion embedding of activity did not significantly impact the analysis (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). We also compared model prediction accuracy for different methods of cortical segmentation. Analyses with LSSC was similar to grid-based and anatomical (CCFv3) parcels as well as non-disjoint segmentations derived from either PCA or LocalNMF (Supplemental Figure S2). Finally, to further quantify the potential for decoding behavior from our neural data, we considered a fully connected artificial neural network (see Methods). However, this non-linear approach showed no improvement over our joint linear model (Supplemental Figure S2).

To investigate the relative contributions of activity versus connectivity dynamics in modeling behavior, we recreated the joint model while temporally (circularly) 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 2f). Surprisingly, models in which correlational data were preserved performed similarly 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=o.37; Paired t-test, Figure 2f). 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 2f).

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 averaged cortical activity (across all parcels) or correlations (across all pairs of parcels) are poorly predictive of the rapid fluctuations in behavior (Supplemental Figure S2). Altogether, these findings indicate that inclusion of rapidly time-varying 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 dynamic coordination of activity across widespread areas.

Sensory inputs modestly influence network connectivity

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. First, we looked at responses to visual stimulation outside of any learning or task-based conditions. We presented the mouse with a series of 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 strong contrast-dependence (Spearman’s R=0.9±0.1, p=0.0001). However, prediction accuracy using mesoscopic correlational structure was significantly lower (p=5.6e-5, paired T-test) and not coupled to stimulus contrast (Spearman’s R=0.1+0.2, p=0.25, Supplemental Figure S3).

We then asked whether network correlational structure was more affected by sensory inputs with learned association to behavior. We adopted a visually cued eyeblink task in which the drifting grating stimulus is paired with an aversive corneal air puff (see Methods). Our previous work showed that task performance is contrast-dependent and requires V1 activity[9]. As with untrained animals, the visual stimulus evoked a large response in V1 and surrounding areas but did not produce significant changes in pairwise correlations between areas. However, linear modeling based on either V1 activity or mesoscopic correlations could robustly predict trial occurrence in a contrast-dependent manner (Spearman’s R=0.91±0.05, p=2.0e-5 for activity and 0.59+0.17, p=0.01 for embedded correlations), although the former again significantly outperformed the latter (p=5.8e-5, paired t-test, Supplemental Figure S3). These results suggest that brief sensory inputs drive large fluctuations in cortical activity, with their impact on network correlations exhibiting sensitivity to the behavioral relevance of the stimulus.

Dynamic correlations of cellular networks 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 cellular resolution 2-photon calcium imaging of GCaMP6s-expressing neurons[38] in the primary visual cortex (see Methods, Figure 3a–b). As above, we looked at both time-varying activity $\boldsymbol{x}_t$ and embedded pair-wise correlations ϕ_t for identified neurons simultaneously with measurements of pupil diameter, facial movement, and locomotion (Figure 3c–f). Unlike large-scale network signals, correlations between neurons were broadly distributed around zero (average r=−0.001±0.006, average CV_r=6.1±1.47, 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 robust to changes in model parameters (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 shuffling $\phi_t^{(20)}$, R^2=0.26±0.07, p=0.02 for shuffling $\boldsymbol{x}_t$; 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 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 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 subsets of pairwise correlations. 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 within-subject spatial heterogeneity that was conserved across different individuals. The overall 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 4d). Intuitively, this result indicates that independent of behavior, the 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 distinct from the cortical segmentation defined by traditional, anatomy-based atlases such as the CCFv3 (Figure 4d). In particular, the anterolateral region includes rostral representations in primary motor cortex as well as upper limb, mouth, and nose representations in primary somatosensory cortex. The posterior region includes visual, auditory, and parietal association areas.

To further examine whether coordinated activity across this anterolateral/posterior partition preferentially encodes spontaneous behavioral variation, we systematically bisected the cortex with a line rotated about the midpoint (Figure 4d). We then modeled behavior using either (1) the activity in these two parcels or (2) the correlations between the two parcels and quantified the prediction accuracy as a function of the angle of division. Accuracy using time-varying activity was independent of angle, suggesting limited spatial heterogeneity in how behavioral metrics are encoded by fluctuations in the magnitude of cortical signals (Figure 4e, Supplemental Figure S5). In contrast, accuracy using the time-varying correlations between the two parcels was significantly higher for the angle matching the anterolateral-posterior division noted above (Figure 4d). This result was relatively insensitive to the window length used for embedding (Supplemental Figure S5). Overall, these findings suggests that encoding of spontaneous behaviors by large-scale correlations is spatially organized into distinct subnetworks in the neocortex.

Dynamic connectivity across spatial scales encodes behavior

Our analyses revealed that spontaneous behaviors can be accurately decoded from the temporal dynamics of correlations between neural signals at both local circuit and mesoscopic spatial scales. However single neurons are embedded in large-scale networks by virtue of long-range synaptic connections. Thus, we wondered whether the functional connectivity across these scales was similarly dynamic. To this end, we carried out simultaneous wide-field and cellular 2-photon imaging (Figure 5a–b, see Methods)[26]. We first analyzed these multimodal data sets separately as described above and quantified the accuracy with which linear models based on time-varying activity $\boldsymbol{x}_t$ or embedded correlations $\phi_t^{(20)}$ could predict behavior. Activity was predictive for mesoscopic (Pupil: R^2=0.13±0.03; Face: R^2=0.14±0.04; Wheel: R^2=0.12±0.04; n=7 mice) and cellular (Pupil: R^2=0.31±0.04; Face: R^2=0.4±0.07; Wheel: R^2=0.5±0.08) data. Similarly, embedded correlations were also predictive for mesoscopic (Pupil: R^2=0.35±0.03; Face: R^2=0.38±0.06; Wheel: R^2=0.51±0.06) and cellular (Pupil: R^2=0.14±0.03; Face: R^2=0.12±0.05; Wheel: R^2=0.15±0.06) data (Figure 5c). As above, correlations-based prediction accuracy was greater for mesoscopic data (Pupil: p=0.002; Face: p=0.004; Wheel: p=0.0002; paired t-test, n=7 mice) and activity-based prediction accuracy was greater for cellular data (Pupil: p=0.0001; Face: p=0.001; Wheel: p=0.0004; paired t-test, n=7 mice).

To explore functional connectivity across spatial scales, we developed a strategy to calculate the time-varying correlations between cells and parcels for a sliding 3-second window followed by diffusion embedding (Figure 5d–e, see Methods). This analysis revealed considerable heterogeneity in the degree of multimodal correlation dynamics exhibited by different cells, measured as the standard deviation of correlation fluctuations averaged for a single cell across all brain parcels across the imaging session (Figure 5e, Supplemental Figure S6, see Methods). The variation in both the correlations and the embedding components appeared to track with behavioral metrics (Figure 5e–g). Indeed, the embedding of the correlations for the dual imaging data could predict fluctuations in pupil diameter, facial movement, and locomotion (Figure 5h), with results robust across a range of model parameters (Pupil: R^2=0.29±0.06; Face: R^2=0.16±0.03; Wheel: R^2=0.22±0.06; n=7 mice, Supplemental Figure S6). Shuffling any one of the correlation signals (mesoscopic, cellular, or dual) did not significantly reduce model accuracy, suggesting a substantial amount of overlap in the information represented across spatial scales.

We explored the spatial interpretation of the dual mesoscopic-cellular embedding by quantifying the accuracy with which the correlations between a single neuron and the mesoscopic parcels are represented by the overall embedding. In general, cells with the most dynamic correlations (largest standard deviation) exhibited the strongest spatial heterogeneity in their modeling accuracy (Figure 5i, Supplemental Figure S6). However, the spatial pattern was generally conserved for all cells and was clearly evident after averaging across the population (Figure 5j), again showing a division of the cortex into anterolateral and posterior subnetworks, with cells in visual cortex being dynamically correlated most strongly with the latter. These results provide an additional independent demonstration that these subdivisions reflect a fundamental organizing principle in the cortex.

Discussion

Our results show that functional connectivity in cortical networks is highly dynamic, varying on a sub-second time-scale that tracks with continuous metrics of spontaneous behaviors. Including these dynamic correlations in a linear model predicted behavioral fluctuations with high accuracy. This result was true for both large-scale networks monitored with mesoscopic calcium imaging and local networks monitored at cellular resolution with 2-photon microscopy. Moreover, combining these modalities revealed that behavior was also accurately represented by the dynamic correlations between local and cortex-wide networks. The spatial organization of dynamic correlations between either parcels or neurons and cortex-wide activity revealed two distinct subnetworks not obviously predicted from standard anatomical divisions.

The representation of behavioral information by time-varying cortical signals has been a focus of recent studies using diverse approaches to monitor brain activity[1,2,10,15,22,39]. In rodent models, variations in behavioral state or arousal are coupled to changes in firing rate, pairwise correlations, and neuromodulatory signaling[11–14,40,41]. In particular, arousal- and motor-related variables (e.g., pupil diameter, locomotion, whisking) are represented at the cellular and network scale[11,12,14]. Several groups have also demonstrated that the spatial patterns of large-scale activity in the neocortex are markedly different when comparing across state[14,16,20]. For example, spontaneous cortical activity can be decomposed using various methods into repeating spatiotemporal motifs that may correspond to sensory or motor signals[20]. Similarly in human subjects, shifts in wakefulness correspond to changes in average resting state connectivity[42], resting state fluctuations predict somatosensory perception[4], and working memory-based task performance corresponds to spatially heterogeneous variation in timescales of patterned activity[43,44]. Time-varying functional connectivity in fMRI studies has received considerable recent focus, with ongoing debates as to the mechanisms and behavioral relevance of coordinated signals over large-scale networks[15]. Indeed, a recent study also demonstrated the utility of Euclidean manifold learning to reduce the dimensionality of connectivity matrices[22], suggesting the generality of this approach.

While some prior efforts to characterize spontaneous behaviors have relied on categorical definitions of state, our data indicate 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. The exact mapping of these observed motor signals onto latent neural variables (and whether they are independent) is not well understood. For example, all three behavioral metrics are linked to variation in cholinergic and adrenergic activity[41,45,46]. We also recently found that acetylcholine release was more robustly linked to facial movement (e.g., whisking) than locomotion[14]. Both Stringer et al.[12] and Musall et al.[11] demonstrated that rapid dynamics of behavior variables could accurately encode neural activity across the cortex. Here, we further explored the dynamics of functional connectivity expressed as the correlation between cortical parcels or neurons. Intuitively, time-varying activity and pairwise correlations can be viewed as first- and second-order terms in a Taylor expansion of a function relating behavior to neural signals. Thus, correlations cannot be linearly derived from the underlying activity and represent a potential mechanism to encode an independent component of behavioral dynamics. This conclusion is strongly supported by our results, where shuffling either activity or correlations significantly reduces modeling accuracy for both mesoscopic and 2-photon data.

The temporal scale of the neural and behavioral dynamics is similar to synaptic integration windows for single cells, suggesting the hypothesis that neurons may be sensitive to convergent synaptic input driven by correlated large-scale activity. Our results combining mesoscopic and 2-photon imaging demonstrate that the functional connectivity across these divergent spatial scales is also dynamic and accurately predictive of fluctuations in behavior. Thus, we propose that network activity within and across spatial scales in the neocortex is coordinated as a function of spontaneous behaviors. In the near future, ongoing development of multi-modal approaches, combining fluorescence imaging, fMRI, and electrophysiology[26,47–49], will likely drive additional discoveries into the functional organization of brain networks in diverse systems.

Our method for viewing time-varying functional connectivity in cortical networks as a graph of graphs provides a conceptual and 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 (we note that similar analyses using Euclidean geometry yielded substantially poorer prediction accuracy). 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. A similar strategy has also been applied to fMRI data[22]. 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. Our method was generalizable for three different data sets (mesoscopic and 2-photon imaging alone and in combination) and yielded the surprising finding that higher-order statistics (i.e., correlational signals) can produce similar or better predictive accuracy than time-varying changes in activity.

Several distinct strategies have been developed to analyze the spatiotemporal organization of network activity, including singular value decomposition and non-negative matrix factorization[16,33]. Here, we show that parcellation of cortical regions[31] 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. Notably, our overall findings were similar for a number of either structural or functional approaches for segmenting the cortex. We did find that LSSC-based parcellation gave robust results with a good balance of reconstruction error and number of disjoint parcels. With the increasing interest in analysis of neural manifolds, our results also highlight the necessity of considering the geometry of the manifold (i.e., Riemannian versus Euclidean) on which the data lie to accurately reveal their intrinsic representation. Finally, our approach modestly out-performed a neural network-based model. This finding suggests that much of the information linking neural signals and spontaneous behavior is present in the dynamics of the activity and correlations.

Surprisingly, the representation of task-independent sensory information in mesoscopic correlations was relatively weak. However, the representation of conditioned sensorimotor event was more robust, indicating that functional connectivity may be altered by external inputs with learned behavioral relevance. These finding are consistent with recent work suggesting spontaneous behavior and external stimuli can be represented in orthogonal dimensions[12], but task-relevant behaviors are broadly present in large-scale patterns of network activity[11,29,50]. Additionally, we and others have shown that training can modify the sensory and motor representations by single cortical neurons[9,51–53], and future studies must continue to explore how development, experience, and learning shape the functional organization of large-scale networks.

Finally, our results show 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 demonstrated such a division based on correlations between single cell activity and mesoscopic cortical signals[26], a result that is also present in our dual mesoscopic and 2-photon data here. Work from other labs using widefield imaging also provided evidence for the existence of spatially similar state-dependent subnetworks[17,48]. Intriguingly, large-scale network activity may also be cell type-specific, as recent findings found distinct patterns for different populations of layer 5 projection neurons[29,50]. Interestingly, our functional subnetworks do not map readily onto standard anatomical segmentation of the cortex, such as the CCFv3[32]. For example, the anterolateral division distinctly omits more caudal motor and somatosensory representations, while the posterior division broadly encompasses visual, auditory, and parietal regions while omitting retrosplenial areas. The mechanisms driving these partitions are unclear, but may reflect poorly mapped intracortical connections, heterogeneous neuromodulatory signaling, or indirect connections through subcortical hubs such as the thalamus[14,54]. Indeed, recent work using brain-wide individual animal connectome sequencing (BRICseq) found anatomical evidence for inter-region connectivity that can support such disjoint subnetworks[55]. Finally, we hypothesize 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.

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. Some of the mesoscopic imaging data were collected as part of a previous study[14], with experimental details provided below for clarity. Analysis results presented here represent wholly new findings and have not appeared elsewhere.

Animals

Adult (P60–100) male and female c57/Bl6 mice (n=19 mice in total) were kept on a 12h light/dark cycle, provided with food and water ad libitum, and housed individually following headpost implants. The animal housing facility was kept between 22–24 degrees C and 60–70% humidity. Imaging experiments were performed during the light phase of the cycle. For most mesoscopic imaging experiments, brain-wide expression of jRCaMP1b[25] was achieved via postnatal sinus injection as described previously[26,27]. Briefly, Po-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 (2.5×10^13gc/ml, Addgene). Mice also received an injection of AAV9-hSyn-GRAB$_{ACh3.0}$ to express the genetically encoded cholinergic sensor GRAB$_{ACh3.0}$[56], 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 and eyeblink conditioning experiments, a similar procedure was used to drive cortex-wide expression of GCaMP6s[38]. For dual mesoscopic and two-photon imaging experiments, adult (P60–70) mice transgenically expressing GCaMP6s in cortical excitatory neurons (CaMK2a-tTA; tetO-GCaMP6s; VIP-Cre)[57] were used. These animals were also injected with AAV driving Cre-dependent GCaMP6s and Cre-dependent tdTomato, though all red fluorescent cells were excluded from the present analysis. For eyeblink conditioning experiments, we used adult (P6o-70) CaMK2a-tTA; tetO-GCaMP6s.

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 skull 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. For dual mesoscopic and two-photon imaging, a 2mm glass microprism (Tower Optical) was placed on top of a dual-layer glass window implanted over the right primary visual cortex[26].

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 GCaMP6s) or 630/75 (jRCaMP1b) emission filter (Chroma) before it reached two sCMOS cameras (Orca-Flash V3, Hamamatsu). Images were acquired at 512×512 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, o.8 NA objective (Nikon). Excitation was driven by a Titanium-Sapphire 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 512×512 resolution at 30 Hz using a galvo-resonant scan system controlled by ScanImage software (Vidrio).

Dual mesoscopic and two-photon imaging

Dual imaging was carried out using a custom microscope combining a Zeiss Axiozoom (as above) and a Sutter MOM (as above), as described previously[26]. To image through the implanted prism, a long-working distance objective (20x, o.4 NA, Mitutoyo) was used. Frame acquisitions were interleaved with an overall rate of 9.15 Hz, with each cycle alternating sequentially between a 920nm two-photon acquisition (512×512 resolution), a 395/25nm widefield excitation acquisition, and a 470/20nm widefield excitation acquisition. Widefield data were collected through a 525/50nm filter into a sCMOS camera (Orca Fusion, Hamamatsu) at 576×576 resolution after 45x pixel binning with 20ms exposure.

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 Spinview software (Flir).

Visual stimulation

For visual stimulation experiments, sinusoidal drifting gratings (2 Hz, 0.04 cycles/degree, 20 degrees of visual field) 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.

Eyeblink Conditioning

Conditioning was carried out as previously published[9]. Briefly, mice were acclimated to head-fixation on the running wheel for several days. For training, each trial started with the onset of a 500 ms visual stimulus comprising a sinusoidal drifting grating (2 Hz, 0.04 cycles/degree, 20 degrees of visual field) that co-terminated with a 50 ms air puff directed to the cornea. Each training day comprised 60 pairings at 100% contrast. After reaching stable performance for 2–3 consecutive days, mice were moved to the next stage where contrast was varied randomly across trials. For all sessions, trials were separated by an exponentially distributed inter-stimulus interval ranging from 18–33 seconds.

Data analysis

All analyses were conducted using custom-written scripts in MATLAB (Mathworks). SVM classifiers were trained using publicly available software[58]. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications[14,26]. Data distributions were assumed to be normal, but this was not formally tested. All animals were selected randomly from our colony for inclusion in the study, and there were no comparisons across subject groups. No blinding was used in the study given the lack of experimental groups. No data points were excluded from the study.

Preprocessing of behavior data—Pupil diameter and facial movements were extracted from face videography using FaceMap[28]. 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 0.4 second window. Epochs of sustained locomotion and quiescence were extracted using change-point detection as described[14]. 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 $(\Delta F/F)_i = (F_i - F_{i,o})/F_{i,o}$ where F_i is the fluorescence of pixel i and $F_{i,o}$ is the corresponding low-pass filtered signal.

Hemodynamic correction—Hemodynamic artifacts were removed using a linear regression accounting for spatiotemporal dependencies between neighboring pixels[14]. We used the approximate isosbestic excitation of GCaMP6s or GRAB$_{\text{ACh3.0}}$ (395 nm) 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““U””, we consider the following linear model:

$$\begin{aligned} y_1 &= x + z + \eta, \\ y_2 &= Az + \xi, \end{aligned}$$

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 $y = \binom{y_1}{y_2}$ is given by stacking y_1 on top of y_2, $\Sigma_y = \mathbb{E}\left[yy^T\right]$ is the correlation matrix between y and $\Sigma_{xy} = \mathbb{E}\left[xy^T\right]$ is the correlation matrix between x and y. The matrix Σ_y is estimated directly from the observations, and the matrix Σ_{xy} is estimated by[14]:

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

where σ_η^2 and σ_ξ^2 are the noise variances of η and ξ, respectively, and $\boldsymbol{I}$ 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}[59]. 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 primarily using Local Selective Spectral Clustering (LSSC)[31]. 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:

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

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 $\boldsymbol{P}$:

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

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[60] of every pixel $s_i = \| \psi^{(n)}\left(i\right) \|$. 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, $s_i, i = 1, \ldots, p$, 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: $L_{i*} = \{ \ell_1, \ell_2, \ldots, \ell_{d_i} \}$ 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.:

$$C_k = \left\{ j \mid \| \psi^{L_i}(i) - \psi^{L_i}(j) \|^2 < \| \psi^{L_i}(j) \|^2 \right\}$$

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 o.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).

As a comparison to LSSC we used PCA to reduce the dimension of the widefield signal and used Localized semi-nonnegative matrix factorization (LocaNMF)[33] for parcellation of mesoscopic data. As an additional comparison, we spatially decimated the mesoscopic data by factor of 8, which is equivalent to using a fixed grid dividing the brain to patches of 8×8 pixels, used as “Grid” parcels.

Cellular 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[31].

Taylor expansion for estimating behavior as a function of neuronal activity—We formulate the link between temporal dynamics of neuronal activity $\mathbf{x}_t \in \mathbb{R}^N$ 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:

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

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:

$$b_t \approx \beta_0 + \sum_i \beta_1\left(i\right)x_t\left(i\right) + \sum_i \sum_j \beta_2\left(i,j\right)C_t\left(i,j\right) + \epsilon \tag{3}$$

where $C_t(i,j) = (x_t(i) - \overline{x}(i))(x_t(j) - \overline{x}(j))$ 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 $\hat{C}_t\left(i,j\right)$ as a smoothed and more robust estimation for the temporal evolution of $C_t(i,j)$:

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

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

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

Inserting $\hat{C}_t\left(i,j\right)$ into (3) leads to eqn. (1). Overall, $\widehat{\boldsymbol{C}}$ is a three-dimensional tensor of parcels by parcels by time, where each element $\hat{C}_t\left(i,j\right)$ 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)[61,62]. 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 $\bar{\boldsymbol{C}}$ as the Riemannian mean of the set and $\bar{\boldsymbol{S}}$ as its equivalent in the tangent plane. $\bar{\boldsymbol{C}}$ and $\bar{\boldsymbol{S}}$ are calculated using the following iterative equations:

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

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

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

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

$$\boldsymbol{S}_k = \bar{\boldsymbol{C}}^{1/2}\log\left(\bar{\boldsymbol{C}}^{-1/2}\boldsymbol{C}_k\bar{\boldsymbol{C}}^{-1/2}\right)\bar{\boldsymbol{C}}^{1/2}, \quad k = 1, \ldots, K$$

As presented previously[63], 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 $\{\boldsymbol{S}_k\}$:

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

where $\widetilde{\boldsymbol{S}}_{\boldsymbol{k}} = \log\left(\bar{\boldsymbol{C}}^{-1/2}\boldsymbol{C}_k\bar{\boldsymbol{C}}^{-1/2}\right)$. This method requires that all matrices $\{\boldsymbol{C}_k\}$ would be full rank[64,65]. 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 $\boldsymbol{C}_k$[66] where λ is set to the median of the singular values of $\boldsymbol{x}_t$[59].

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 $\{\boldsymbol{S}_t\}$:

$$\boldsymbol{A}\left(i, j\right) = \exp\left\{-\| \boldsymbol{S}_i - \boldsymbol{S}_j \|^2 / \sigma^2\right\} \tag{5}$$

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).

Normalizing the kernel $\boldsymbol{A}$ to be row-stochastic and taking the right eigenvectors leads to the low dimensional representation for the correlation traces:

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

The size of the kernel matrix $\boldsymbol{A}$ is determined by the available time points recorded on each experiment ($N_T \approx 10^4$ for all sessions). To reduce computational complexity involving eigenvalue decomposition of large matrices we used Nyström out-of-sample extension[67] as follows. We randomly choose a smaller subset of time points $\{t_i\}_{i=1}^N$, where $N < N_T$. We obtain the normalized kernel using these time points and extract its eigen-decomposition $\lambda_k, \phi_{t_i}^k, k = 1, \ldots, n$. We then extend this low dimensional representation to all time points:

$$\phi_t^k = \frac{1}{\lambda_k} \sum_{i=1}^{N} \exp\left\{-\| \boldsymbol{S}_t - \boldsymbol{S}_{t_i} \|^2 / \sigma^2\right\} \phi_{t_i}^k$$

For comparison to using Riemannian geometry for calculating distances in the kernel, 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 for unconditioned and conditioned experiments 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 (unless otherwise mentioned) 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) and diffusion embedding based on Euclidean distances. For pairwise correlations, we used: raw correlation traces, PCA with number of components selected to account for >95% variance, and 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 circularly shuffled either activity or embedded correlations through time and trained the resulting model to assess $R^2_{\text{shuffled activity}}$ and $R^2_{\text{shuffled } \phi_t}$.

As a comparison to linear modeling we trained three fully connected Artificial Neural Networks for prediction of pupil/FaceMap/running speed based on neuronal activity traces as predictors. We performed a grid-search for the optimal configuration for prediction of behavior, based on 10-fold cross validation where on each fold 80% was used for training, 10% for parameters tunning and 10% for testing. Overall, the optimal performance was achieved with two inner layers of 64 and 4 units, learning rate of 0.05 and Relu activation.

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 predictive dynamics of cortical subnetworks—We evaluated the average activity (across parcels) of two sub-cortical networks using an arbitrary partition of the brain using a line bisecting the neocortex. We then used these two time traces to predict pupil size, facial movement, and locomotion speed. By rotating the line in 30, 60, 90, 120, 150 degrees we measured the R^2 values for prediction of behavior variables based on different ways for partition of the brain into sub-networks. We then evaluated the time trace of correlation between these two sub-networks using a sliding window of 3sec. This time we used the correlation trace to predict each behavior and evaluate the R^2 values as a function of the angle. For the correlations we also evaluated the R^2 as a function of the analysis window length (varying from 0.5sec to 15sec) for angle of 30degrees.

Instantaneous Multimodal Connectivity—Denote $\boldsymbol{x}_t^{(1)} \in \mathbb{R}^{N_1}$ as the neuronal activity of N_1 brain parcels at time t and denote $\boldsymbol{x}_t^{(2)} \in \mathbb{R}^{N_2}$ as the neuronal activity of N_2 cells at time t. We define $\Sigma_t(i, j)$ as a 3-dimensional tensor of N_1 over N_2 over time, expressing the dynamics of multimodal connectivity between cells and parcels:

$$\Sigma_t(i, j) = \frac{\sum_{\tau = t - \frac{N_t}{2}}^{t + \frac{N_t}{2}} \left(x_\tau^1(i) - \overline{x}_t^1(i)\right)\left(x_\tau^2(j) - \overline{x}_t^2(j)\right)}{\sqrt{\sum_{\tau = t - \frac{N_t}{2}}^{t + \frac{N_t}{2}} \left(x_\tau^1(i) - \overline{x}_t^1(i)\right)^2}\sqrt{\sum_{\tau = t - \frac{N_t}{2}}^{t + \frac{N_t}{2}} \left(x_\tau^2(j) - \overline{x}_t^2(j)\right)^2}}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2, \quad t = 1, \ldots, N_T$$

where N_T is the numer of time points. For a given time point t, the matrix Σ_t, is non-symmetric and therefore not bound to the Riemannian cone (as opposed to the sample covariance matrices $\widehat{\boldsymbol{C}}_t$, which are SPD). Therefore, we used Euclidean distances to evaluate the diffusion kernel between correlation matrices related to different time points:

$$\widetilde{\boldsymbol{A}}(t, t') = \exp\left\{- \| \Sigma_t - \Sigma_{t'} \|^2 / \sigma^2\right\} \tag{6}$$

where σ is a scale parameter evaluated as the median of pairwise distances between each correlation matrix and its k-nearest neighbors, where $k = 200$. Note that as for embedding of correlation matrices, our results here are also highly robust to variation in k (Supplemental Figure 6). Here we again reduce computational complexity using Nyström extension and evaluate the kernel $\widetilde{\boldsymbol{A}}$ based on a smaller randomly selected sub-set of time points, $t' = 1, \ldots, N$, where $N < N_T$. We normalize the kernel matrix to be row-stochastic and obtain its eigen-decomposition:

$$\Sigma_{t'} \rightarrow \varphi_{t'}^{(n)} = \begin{pmatrix} \phi_{t'}^1 \\ \vdots \\ \phi_{t'}^n \end{pmatrix}$$
$$\lambda_k, \qquad k = 1, \ldots, n$$

We then extend this representation to all time points $t = 1, \ldots, N_T$:

$$\varphi_t^k = \frac{1}{\lambda_k} \sum_{t'=1}^{N} \exp\left\{ - \| \Sigma_t - \Sigma_{t'} \|^2 / \sigma^2 \right\} \varphi_{t'}^k$$

Standard Deviation of Multimodal Connectivity—We estimate the average variability of correlations between a given cell and the mesoscopic cortical network as:

$$\sigma_j = \frac{1}{N_1} \frac{1}{N_t} \sum_{i=1}^{N_1} \sum_{t=1}^{N_t} \left(\Sigma_t(i,j) - \overline{\Sigma}(i,j) \right)^2, \qquad j = 1, \ldots, N_2$$

where $\overline{\Sigma}\left(i,j\right) = \frac{1}{N_T} \sum_{t=1}^{N_T} \Sigma_t\left(i,j\right)$.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank members of the Higley and Cardin laboratories for helpful input throughout all stages of this study. We thank Rima Pant and Estella Murillo for generation of AAV vectors. We thank the GENIE Project for jRCaMP1b plasmids. This work was supported by funding from the NIH (MH099045, MH121841, and EY033975 to MJH, EY022951 to JAC, MH113852 to MJH and JAC, EY029581 and GM007205 to DB, EY031133 to AHM, EY026878 to the Yale Vision Core, EB026936 to GM and RRC), the NSF (CCF-2217058 to GM), an award from the Yale Kavli Institute of Neuroscience (to MJH and RRC), an award from the Swartz Foundation (to HB), an award from the Simons Foundation SFARI (to MJH), an award from the Smith-Magenis Syndrome Research Foundation (to MJH and JAC), and a BBRF Young Investigator Grant (to SL).

Data Availability

The full datasets generated and analyzed in this study are available from the corresponding authors on reasonable request. Data for mesoscopic imaging experiments with CCFv 3-based parcellation have been deposited on the figshare archive: https://figshare.com/projects/Benisty_Higley_2023/175317.

References

1. Breakspear M Dynamic models of large-scale brain activity. Nat Neurosci 20, 340–352 (2017). [PubMed: 28230845]
2. Calhoun VD, Miller R, Pearlson G & Adali T The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron 84, 262–274 (2014). [PubMed: 25374354]
3. Cardin JA, Crair MC & Higley MJ Mesoscopic Imaging: Shining a Wide Light on Large-Scale Neural Dynamics. Neuron 108, 33–43 (2020). [PubMed: 33058764]
4. Boly M, et al. Baseline brain activity fluctuations predict somatosensory perception in humans. Proc Natl Acad Sci U S A 104, 12187–12192 (2007). [PubMed: 17616583]
5. de Gee JW, et al. Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains. Elife 9(2020).
6. Jacobs EAK, Steinmetz NA, Peters AJ, Carandini M & Harris KD Cortical State Fluctuations during Sensory Decision Making. Curr Biol 30, 4944–4955 e4947 (2020). [PubMed: 33096037]

7. McGinley MJ, David SV & McCormick DA Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron 87, 179–192 (2015). [PubMed: 26074005]

8. Palva JM & Palva S Roles of multiscale brain activity fluctuations in shaping the variability and dynamics of psychophysical performance. Progress in brain research 193, 335–350 (2011). [PubMed: 21854973]

9. Tang L & Higley MJ Layer 5 Circuits in V1 Differentially Control Visuomotor Behavior. Neuron 105, 346–354 e345 (2020). [PubMed: 31757603]

10. Fox MD & Raichle ME Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8, 700–711 (2007). [PubMed: 17704812]

11. Musall S, Kaufman MT, Juavinett AL, Gluf S & Churchland AK Single-trial neural dynamics are dominated by richly varied movements. Nat Neurosci 22, 1677–1686 (2019). [PubMed: 31551604]

12. Stringer C, et al. Spontaneous behaviors drive multidimensional, brainwide activity. Science 364, 255 (2019). [PubMed: 31000656]

13. Vinck M, Batista-Brito R, Knoblich U & Cardin JA Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. Neuron 86, 740–754 (2015). [PubMed: 25892300]

14. Lohani S, et al. Spatiotemporally heterogeneous coordination of cholinergic and neocortical activity. Nat Neurosci 25, 1706–1713 (2022). [PubMed: 36443609]

15. Lurie DJ, et al. Questions and controversies in the study of time-varying functional connectivity in resting fMRI. Netw Neurosci 4, 30–69 (2020). [PubMed: 32043043]

16. MacDowell CJ & Buschman TJ Low-Dimensional Spatiotemporal Dynamics Underlie Cortex-wide Neural Activity. Curr Biol 30, 2665–2680 e2668 (2020). [PubMed: 32470366]

17. Vanni MP, Chan AW, Balbi M, Silasi G & Murphy TH Mesoscale Mapping of Mouse Cortex Reveals Frequency-Dependent Cycling between Distinct Macroscale Functional Modules. J Neurosci 37, 7513–7533 (2017). [PubMed: 28674167]

18. Gregoriou GG, Gotts SJ, Zhou H & Desimone R High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science 324, 1207–1210 (2009). [PubMed: 19478185]

19. Ito T, et al. Task-evoked activity quenches neural correlations and variability across cortical areas. PLoS computational biology 16, e1007983 (2020). [PubMed: 32745096]

20. Mohajerani MH, et al. Spontaneous cortical activity alternates between motifs defined by regional axonal projections. Nat Neurosci 16, 1426–1435 (2013). [PubMed: 23974708]

21. Cohen MR & Kohn A Measuring and interpreting neuronal correlations. Nat Neurosci 14, 811–819 (2011). [PubMed: 21709677]

22. Gonzalez-Castillo J, et al. Manifold Learning for fMRI time-varying FC. bioRxiv (2023).

23. Spruston N Pyramidal neurons: dendritic structure and synaptic integration. Nat Rev Neurosci 9, 206–221 (2008). [PubMed: 18270515]

24. Lafon S, Keller Y & Coifman RR Data fusion and multicue data matching by diffusion maps. IEEE Trans Pattern Anal Mach Intell 28, 1784–1797 (2006). [PubMed: 17063683]

25. Dana H, et al. Sensitive red protein calcium indicators for imaging neural activity. Elife 5(2016)

26. Barson D, et al. Simultaneous mesoscopic and two-photon imaging of neuronal activity in cortical circuits. Nature methods 17, 107–113 (2020). [PubMed: 31686040]

27. Hamodi AS, Martinez Sabino A, Fitzgerald ND, Moschou D & Crair MC Transverse sinus injections drive robust whole-brain expression of transgenes. Elife 9(2020).

28. Syeda A, et al. Facemap: a framework for modeling neural activity based on orofacial tracking. BioRXiv (2022).

29. Mohan H, et al. Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. Nat Neurosci 26, 481–494 (2023). [PubMed: 36690901]

30. Ma Y, 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).

31. Mishne G, Coifman RR, Lavzin M & Schiller J Automated cellular structure extraction in biological images with applications to calcium imaging data. BioRXiv (2018).

32. Wang Q, et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell 181, 936–953 e920 (2020). [PubMed: 32386544]

33. Saxena S, et al. Localized semi-nonnegative matrix factorization (LocaNMF) of widefield calcium imaging data. PLoS computational biology 16, e1007791 (2020). [PubMed: 32282806]

34. Wood KC, Angeloni CF, Oxman K, Clopath C & Geffen MN Neuronal activity in sensory cortex predicts the specificity of learning in mice. Nature communications 13, 1167 (2022).

35. Driscoll LN, Pettit NL, Minderer M, Chettih SN & Harvey CD Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. Cell 170, 986–999 e916 (2017). [PubMed: 28823559]

36. Hallinen KM, et al. Decoding locomotion from population neural activity in moving C. elegans. Elife 10(2021).

37. Livneh Y, et al. Estimation of Current and Future Physiological States in Insular Cortex. Neuron 105, 1094–1111 e1010 (2020). [PubMed: 31955944]

38. Chen TW, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300 (2013). [PubMed: 23868258]

39. Gonzalez-Castillo J, et al. Imaging the spontaneous flow of thought: Distinct periods of cognition contribute to dynamic functional connectivity during rest. NeuroImage 202, 116129 (2019). [PubMed: 31461679]

40. Constantinople CM & Bruno RM Effects and mechanisms of wakefulness on local cortical networks. Neuron 69, 1061–1068 (2011). [PubMed: 21435553]

41. Polack PO, Friedman J & Golshani P Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat Neurosci 16, 1331–1339 (2013). [PubMed: 23872595]

42. Tagliazucchi E & Laufs H Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. Neuron 82, 695–708 (2014). [PubMed: 24811386]

43. Gao R, van den Brink RL, Pfeffer T & Voytek B Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. Elife 9(2020).

44. Raut RV, Snyder AZ & Raichle ME Hierarchical dynamics as a macroscopic organizing principle of the human brain. Proc Natl Acad Sci U S A 117, 20890–20897 (2020). [PubMed: 32817467]

45. Reimer J, et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nature communications 7, 13289 (2016).

46. Joshi S, Li Y, Kalwani RM & Gold JI Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron 89, 221–234 (2016). [PubMed: 26711118]

47. Lake EMR, et al. Simultaneous cortex-wide fluorescence Ca(2+) imaging and whole-brain fMRI. Nature methods 17, 1262–1271 (2020). [PubMed: 33139894]

48. Clancy KB, Orsolic I & Mrsic-Flogel TD Locomotion-dependent remapping of distributed cortical networks. Nat Neurosci 22, 778–786 (2019). [PubMed: 30858604]

49. Peters AJ, Fabre JMJ, Steinmetz NA, Harris KD & Carandini M Striatal activity topographically reflects cortical activity. Nature 591, 420–425 (2021). [PubMed: 33473213]

50. Musall S, et al. Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. Nat Neurosci 26, 495–505 (2023). [PubMed: 36690900]

51. Puscian A, Benisty H & Higley MJ NMDAR-Dependent Emergence of Behavioral Representation in Primary Visual Cortex. Cell Rep 32, 107970 (2020). [PubMed: 32726633]

52. Poort J, et al. Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. Neuron 86, 1478–1490 (2015). [PubMed: 26051421]

53. Makino H & Komiyama T Learning enhances the relative impact of top-down processing in the visual cortex. Nat Neurosci 18, 1116–1122 (2015). [PubMed: 26167904]

54. Miller-Hansen AJ & Sherman SM Conserved patterns of functional organization between cortex and thalamus in mice. Proc Natl Acad Sci U S A 119, e2201481119 (2022). [PubMed: 35588455]

55. Huang L, et al. BRICseq Bridges Brain-wide Interregional Connectivity to Neural Activity and Gene Expression in Single Animals. Cell 182, 177–188 e127 (2020). [PubMed: 32619423]

56. Jing M, et al. An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. Nature methods 17, 1139–1146 (2020). [PubMed: 32989318]

57. Wekselblatt JB, Flister ED, Piscopo DM & Niell CM Large-scale imaging of cortical dynamics during sensory perception and behavior. J Neurophysiol 115, 2852–2866 (2016). [PubMed: 26912600]

58. Chang CC & Lin CJ LIBSVM: A Library for Support Vector Machines. Acm T Intel Syst Tec 2(2011).

59. Gavish M & Donoho DL The Optimal Hard Threshold for Singular Values is 4/root 3. Ieee T Inform Theory 60, 5040–5053 (2014).

60. Cheng X & Mishne G Spectral Embedding Norm: Looking Deep into the Spectrum of the Graph Laplacian. SIAM J Imaging Sci 13, 1015–1048 (2020). [PubMed: 34136062]

61. Diamond S & Boyd S CVXPY: A Python-Embedded Modeling Language for Convex Optimization. J Mach Learn Res 17(2016).

62. Venkatesh M, Jaja J & Pessoa L Comparing functional connectivity matrices: A geometry-aware approach applied to participant identification. NeuroImage 207, 116398 (2020). [PubMed: 31783117]

63. Tuzel O, Porikli F & Meer P Pedestrian detection via classification on Riemannian manifolds. Ieee T Pattern Anal 30, 1713–1727 (2008).

64. Barachant A, Bonnet S, Congedo M & Jutten C Classification of covariance matrices using a Riemannian-based kernel for BCI applications. Neurocomputing 112, 172–178 (2013).

65. Yair O, Ben-Chen M & Talmon R Parallel Transport on the Cone Manifold of SPD Matrices for Domain Adaptation. Ieee T Signal Proces 67, 1797–1811 (2019).

66. Abbas K, et al. Geodesic Distance on Optimally Regularized Functional Connectomes Uncovers Individual Fingerprints. Brain connectivity 11, 333–348 (2021). [PubMed: 33470164]

67. Fowlkes C, Belongie S, Chung F & Malik J Spectral grouping using the Nystrom method. Ieee T Pattern Anal 26, 214–225 (2004).

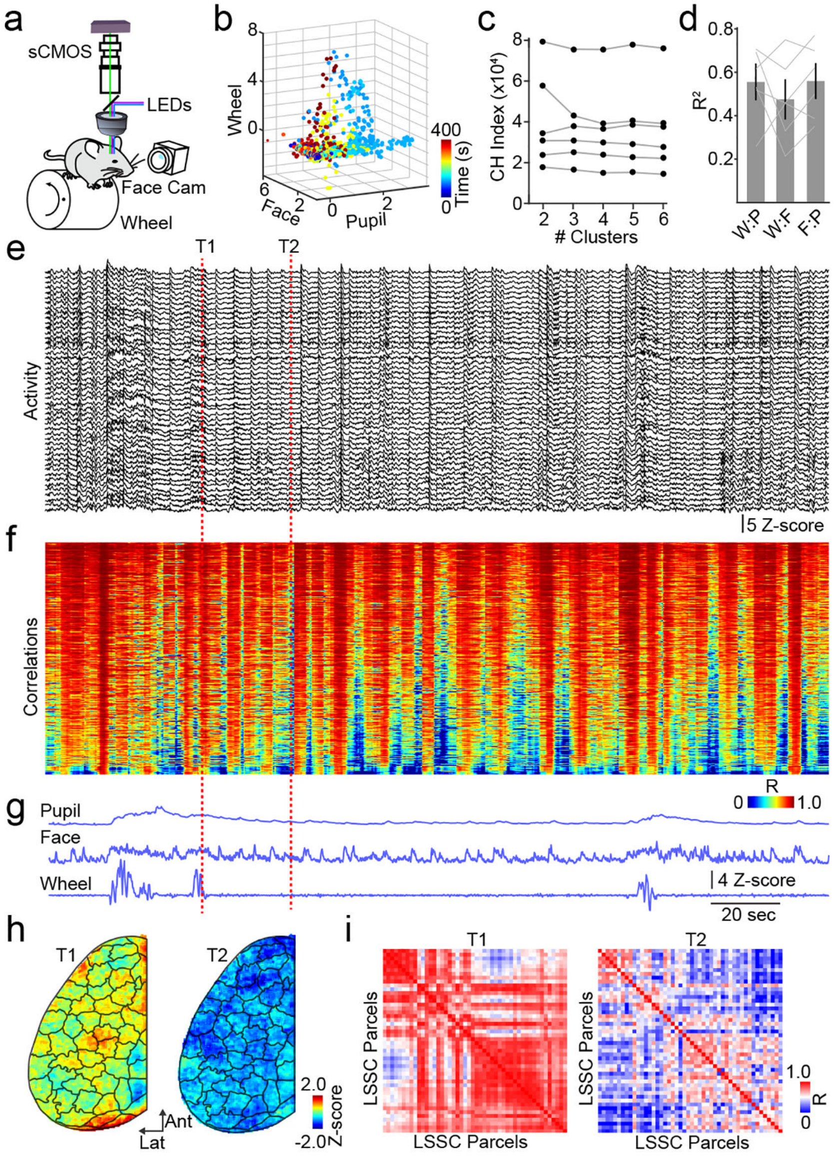

Figure 1. Mesoscopic imaging of cortical activity and functional connectivity.

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) over a 400 second window for the example mouse shown in (e-i). **c**, Population data showing Calinski-Harabasz index values for K-means clustering of behavioral metrics for all subjects. **d**, Population data (n=6 independent mice) 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 (e). Data are sorted by increasing standard deviation. **g**, Time series of behavioral metrics corresponding to the data shown in (e) and (f). **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.

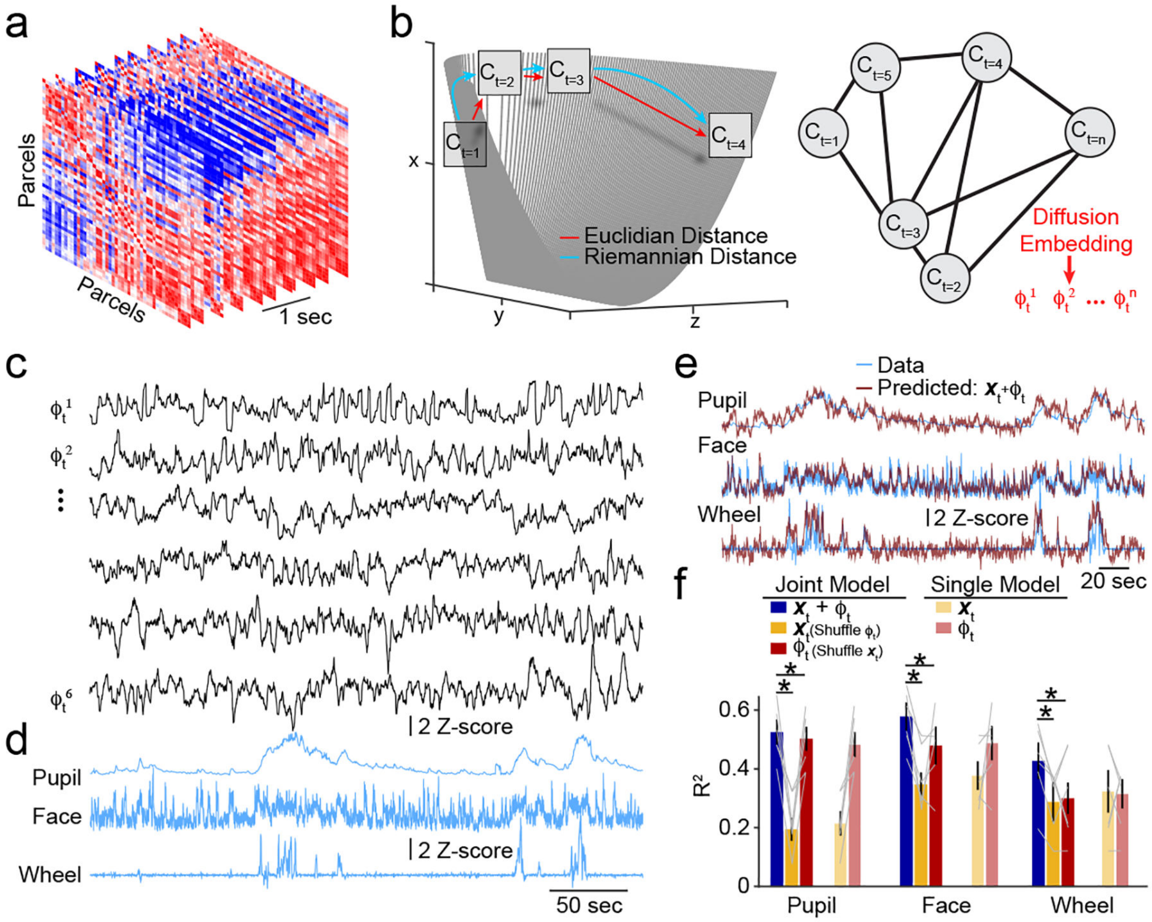

Figure 2. Dynamic functional connectivity encodes rapid behavioral variations.

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 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 individual 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. **d**, Time series for behavioral metrics corresponding to data in (c). **e**, Example behavioral data (blue traces) from (d) showing fluctuations in pupil diameter, facial movement, and locomotion superimposed on predicted behavior (red traces) estimated using a joint model based on time-varying activity and embedded correlations. **f**, Population data (n=6 independent mice) 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 for two sided paired t-test (see main text). Full model compared to shuffling $\phi_t^{(20)}$: p=0.001 for pupil, p=0.002 for face and p=0.01

for wheel. Full model compared to shuffling x_i: p=0.02 for pupil, p=0.006 for face, p=0.003 for wheel.

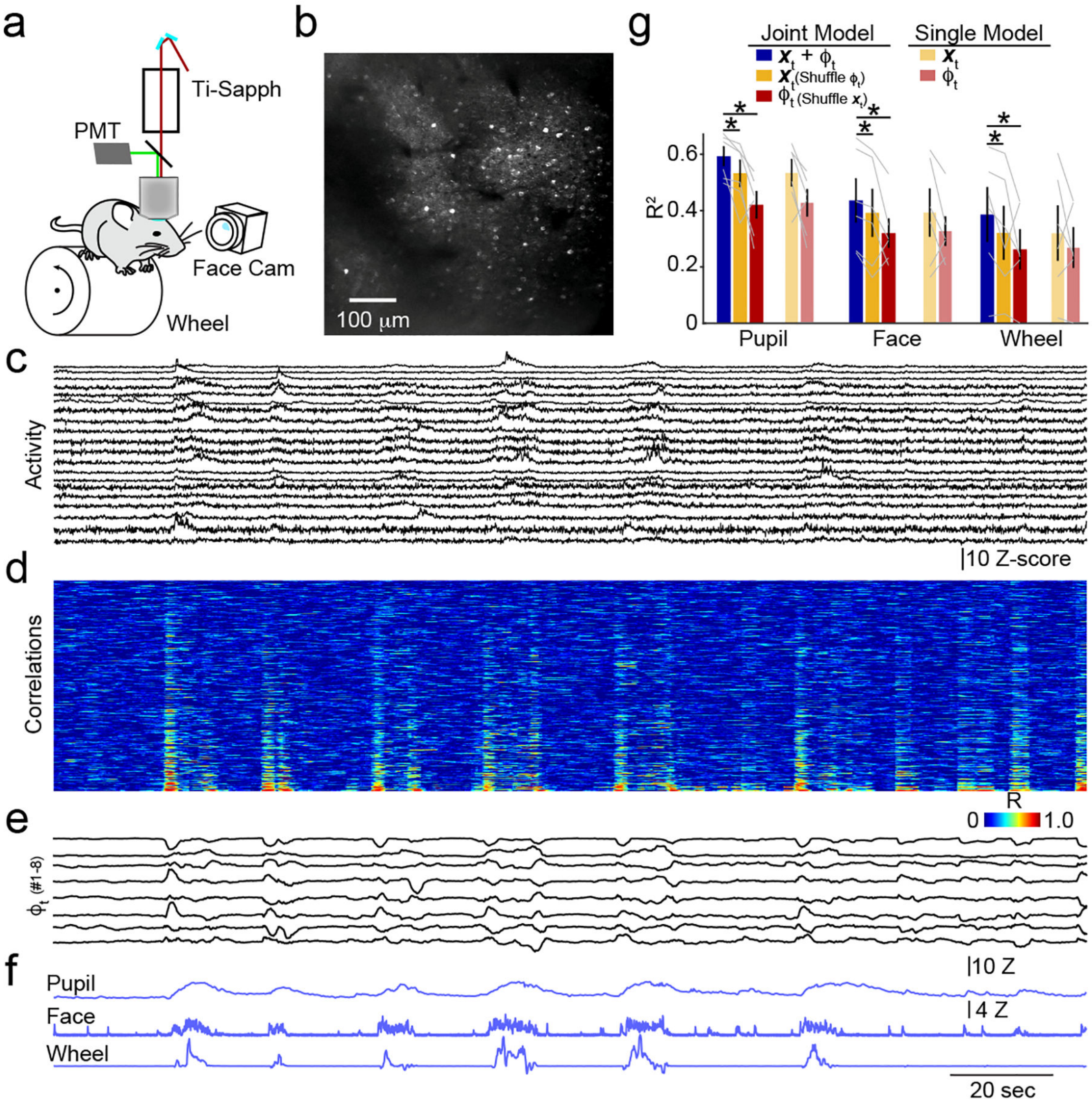

Figure 3. Local circuit dynamics encode spontaneous behavioral variation.

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. Similar results were obtained for each of 6 mice. **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-d). **g**, Population data (n=6 independent mice) 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 $p<0.05$ for two sided paired t-test (see main text). Full model compared to shuffling $\phi_t^{(20)}$: p=0.04 for pupil, p=0.03 for face and p=0.02 for wheel. Full model compared to shuffling $\boldsymbol{x}_t$: p=0.0002 for pupil, p=0.006 for face p=0.003 for wheel.

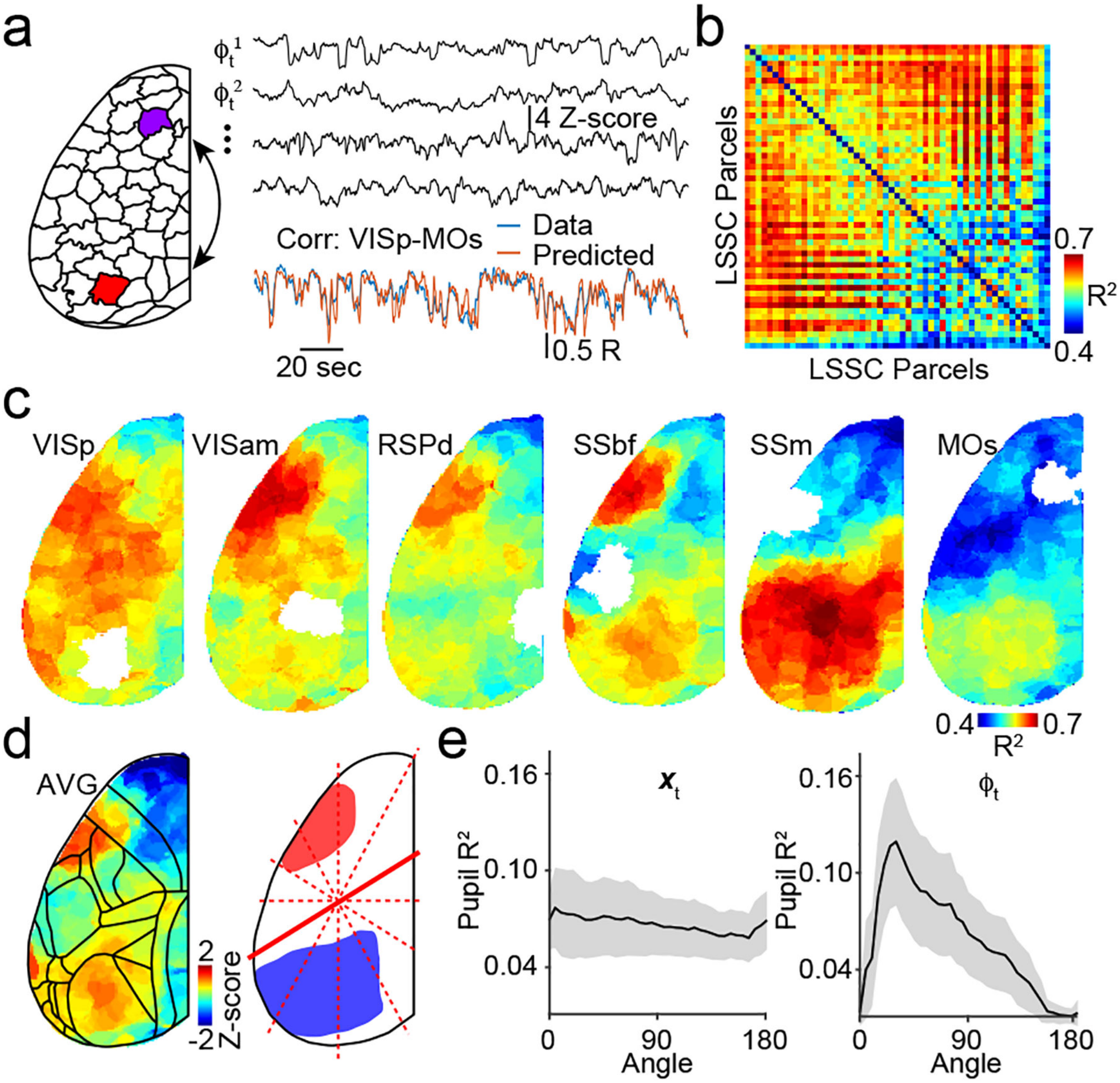

Figure 4. Dynamic functional connectivity reveals distinct cortical subnetworks.

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**, Left, grand average map showing R^2 values as in (c) collapsed across all animals (n=6) and all cortical parcels. Right, schematic illustrating the anterolateral (red) and posterior subnetworks (blue) derived from data in (c). Red dashed lines indicate angles for bisecting LSSC parcels into arbitrary subnetworks, with solid line (30°) corresponding to anterolateral/posterior division. e, Left, population data showing the average (±SEM) prediction accuracy for modeling pupil fluctuation based on time-varying activity in two subnetworks determined by bisecting lines in (d). Right, average (±SEM) prediction accuracy for pupil fluctuation based on time-varying correlation between two subnetworks determined by lines in (d).

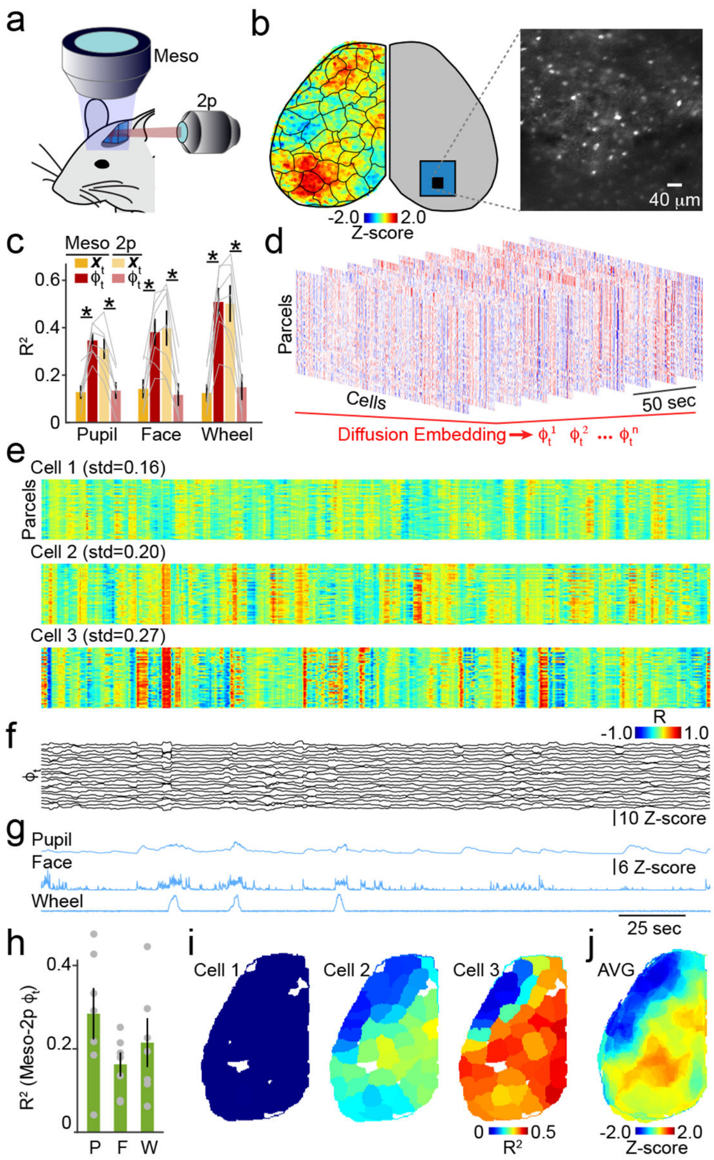

Figure 5. Functional connectivity across spatial scales encodes behavior.

a, Schematic illustrating the setup for simultaneous mesoscopic and 2-photon imaging. **b**, Left, example mesoscopic imaging frame and schematic of microprism placement in the contralateral hemisphere. Right, example 2-photon imaging frame collected through the prism. **c**, Population data (n=7 independent mice) showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using either activity (yellow) or ϕ_t (red) for mesoscopic or 2-photon data. * indicates p<0.05 for two sided paired t-test (see main text). Comparing $\phi_t^{(20)}$ to x_t: for mesoscopic data, pupil: p=0.002; face: p=0.004; wheel: p=0.0002. For cellular data, pupil: p=0.0001; face: p=0.001; wheel: p=0.0004. **d**, Example sequential multimodal correlation matrices, derived from a sliding window applied to neural activity from mesoscopic (parcels) and 2-photon (cells) imaging, used for diffusion embedding. **e**, Dynamic multimodal correlation time series for three example cells, where each row represents a mesoscopic parcel. The standard deviation of correlation values over time, averaged across all rows is indicated. **f**, Example of the first 20 diffusion embedding components from the same animal (n=243 cells, 47 parcels). **g**, Time series for behavioral metrics corresponding to data in (e-g). **h**, Population data (n=7 independent mice) showing average (±SEM) prediction accuracy (R^2) for modeling behavior variables using ϕ_t derived from the embedding of dual mesoscopic and 2-photon correlations. **i**, Example maps for the cells in (e) showing R^2 values for modeling the correlation of the cell with each parcel using the overall diffusion embedding. **j**, Grand average map showing R^2 values as in (i) collapsed across all animals (n=6) and all cells and cortical parcels.