Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Distributed cortical dynamics organize behavior across timescales

Pengfei Wei

pf.wei@seu.edu.cn

Southeast University

Yaning Han

Southern University of Science and Technology https://orcid.org/0000-0002-1650-2262

Quanying Liu

Southern University of Science and Technology https://orcid.org/0000-0002-2501-7656

Siyuan Liu

South China University of Technology

Heping Cheng

College of Future Technology Peking University https://orcid.org/0000-0002-9604-6702

Article

Keywords:

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11376554/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Distributed cortical dynamics organize behavior across timescales

Yaning Han[2, #], Quanying Liu[3, #], Siyuan Liu[4, *], Heping Cheng[5, 6, 7, 8, 9, 10, *], Pengfei Wei[1, 11, 12, *]

[1]School of Biological Science and Medical Engineering, State Key Laboratory of Digital Medicine, Southeast University, Nanjing 211189, China

[2]Department of Pharmacology, Joint Laboratory of Guangdong-Hong Kong Universities for Vascular Homeostasis and Diseases, SUSTech Homeostatic Medicine Institute, School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China

[3]Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen 518055, China

[4]South China University of Technology, Guangzhou 510641, China

[5]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[6]PKU-Nanjing Institute of Translational Medicine, Nanjing Raygen Health, Nanjing 211899, China

[7]College of Future Technology, Peking University, Beijing 100871, China

[8]Peking-Tsinghua Center for Life Sciences, College of Future Technology, Peking University, Beijing 100871, China

[9]National Biomedical Imaging Center, Beijing 100871, China

[10]State Key Laboratory of Membrane Biology, Beijing 100871, China

[11]Guangdong Provincial Key Laboratory of Brain Connectome and Behavior, the Brain Cognition and Brain Disease Institute, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[12]Jiangsu Brain-Computer Interface Research Institute, Nanjing 210046, China

[#]The authors contributed equally to this paper.

Correspondence to:

[*]Pengfei Wei, E-mail: pf.wei@seu.edu.cn

[*]Heping Cheng, E-mail: chengp@pku.edu.cn

[*]Siyuan Liu, E-mail: liu@scut.edu.cn

Abstract

The environment changes across a broad range of timescales, requiring animals to generate behavior that is organized over similarly diverse temporal scales. However, how the brain represents and coordinates behavior across these complex timescales remains unclear. Here we established a minimal sensory-decision-motor system in freely moving mice. We used miniature two-photon calcium imaging to record neural activity from primary somatosensory cortex (S1), dorsomedial prefrontal cortex (dmPFC), and primary motor cortex (M1), and combined this with deep-learning-based methods to quantify behavior at multiple timescales. We found that distributed neural dynamics spanning S1, dmPFC, and M1 were conserved across repeated behavioral episodes. The computational organization of these distributed neural dynamics was reflected in their dimensionality, as validated by all-optical interrogation. Low-dimensional dynamics encoded longer-timescale behavior, whereas high-dimensional dynamics encoded shorter-timescale behavior. Differences across cortical areas arose from distinct projection weights onto these dimensions. At longer behavioral timescales, coding differences between cortical areas became stronger in projection weights. Distributed neural dynamics were broadly present across behavioral episodes and evolved in an episode-specific manner, with reproducible temporal patterns within the same episode type and distinct patterns across different episode types. These findings show that behavioral timescales are encoded in distributed cortical dynamics, and their functional differences arise from how they project onto shared neural dimensions.

Introduction

Animal behavior is organized across multiple timescales, from short postures and movements to longer action sequences[1]. However, the neural circuit mechanisms that support this organization remain unclear. Classical circuit models propose that different timescales of behavior are encoded by different brain areas, such that lower areas represent sensory and motor details, whereas higher areas represent more abstract behavioral structure[2]. Recent large-scale recordings across the brain have challenged this view, showing that behavior is often represented by broad and distributed neural activity rather than by highly selective local circuits[3,4]. From the neural manifold view, this distributed activity may reflect dynamic neural ensembles that span multiple brain regions and jointly represent behavior. How behavioral timescales are encoded in the population dynamics of functionally distinct brain areas and whether a unified model can explain this process remain open questions.

A major challenge in studying neural coding across timescales is that behavioral features at different temporal scales overlap in time. Previous studies have linked motor cortical population activity to continuous movement trajectories, discrete action states, or longer self-initiated action sequences[5-9]. A key unresolved issue is that these features are not separated during natural behavior. Short kinematic details, movement motifs, and long behavioral events unfold together, and they must be coordinated by the same ongoing neural activity[10,11]. Recent behavioral frameworks, including our previous work on hierarchical behavioral decomposition[12-15], now allow these parallel behavioral timescales to be measured with high precision, from short posture dynamics to movement motifs and long action sequences[16-20]. These advances make it possible to ask how neural population activity encodes multiple timescales of natural behavior at the same time.

Here, we established a minimal sensory-decision-motor observation system in the animal brain to

study the neural coding of behavior across timescales. We used two-photon calcium imaging to record population activity from the primary somatosensory cortex (S1), medial prefrontal cortex (dmPFC), and primary motor cortex (M1) during freely moving behavior[21,22]. We found that neural trajectories were variable across the full behavioral session but became highly consistent within specific behavioral episodes. This episode-specific consistency was broadly distributed across sensory, prefrontal, and motor cortices. The organization of these distributed dynamics was structured across neural dimensions. High-dimensional dynamics encoded shorter-timescale behavioral kinematics, whereas low-dimensional dynamics encoded longer-timescale behavioral sequences. The dmPFC showed stronger projections onto low-dimensional dynamics, whereas M1 and S1 were more strongly weighted toward high-dimensional dynamics. All-optical interrogation further showed that perturbations reducing high-dimensional neural dynamics were accompanied by a selective loss of shorter-timescale behavioral details. This dimensional organization was broadly present throughout the full behavioral session, although its expression changed dynamically across behavioral contexts. Together, these findings show that behavioral timescales are encoded across brain regions through distinct dimensions of distributed neural dynamics, and that regional functional differences are characterized by differential weighting of these shared neural dimensions.

Results

Specific behavioral episodes evoke reproducible distributed cortical dynamics

To capture natural behavior across a broad range of timescales, we used the MouseVenue3D system developed in our laboratory (Fig. 1a), which combines synchronized multi-view behavioral recording with miniaturized two-photon calcium imaging (mTPM)[12,21,22]. We used a freely interacting two-mouse paradigm to generate continuous behavior spanning full behavioral sessions and shorter social interaction episodes. The mouse carrying the miniaturized two-photon microscope was defined as the subject mouse, whereas its unimplanted interaction partner was defined as the object mouse. We recorded population activity from the primary somatosensory cortex (S1), dorsomedial prefrontal cortex (dmPFC), and primary motor cortex (M1) as a minimal sensory-decision-motor system, yielding 6,125 neurons in total (Fig. 1b).

We first examined cortical population dynamics at the scale of the full behavioral session (Fig. 1c) and of individual social interaction episodes (Fig. 1d). During full sessions, low-dimensional neural trajectories continuously evolved through the embedding space, without converging onto an obvious recurrent structure (Fig. 1e). In contrast, trajectories extracted from individual social interaction episodes displayed similar geometries across repeated trials, despite occupying different absolute positions in the embedding space (Fig. 1f). Because individual interaction episodes varied in duration, we used Generalized Time Warping (GTW)[23] to normalize their temporal progression and quantify the similarity of their dynamical structure (Fig. 1g).

Following GTW alignment, neural trajectories from repeated social interaction episodes converged onto reproducible paths, revealing a common low-dimensional structure (Fig. 1h). Pairwise

trajectory distances were significantly reduced after alignment in S1, dmPFC, and M1, indicating that repeated interaction episodes were accompanied by conserved temporal dynamics (Fig. 1i). These aligned trajectories were highly reproducible across all three cortical regions, with normalized Procrustes similarities of approximately 0.7 (Fig. 1j). The degree of reproducibility did not differ significantly among S1, dmPFC, and M1, indicating that episode-specific dynamics were broadly distributed across the sensory-prefrontal-motor cortical axis rather than being restricted to a single region.

We next asked whether this recurrent geometry reflected behaviorally meaningful neural structure rather than coincidental similarity introduced by dimensionality reduction. Neural trajectory segments associated with matched behavioral actions were therefore compared with randomly paired segments (Fig. 1k). Trajectories associated with the same actions showed significantly greater similarity than random pairings, demonstrating that the recurrent neural geometry was systematically related to ongoing behavior.

These results show that distributed cortical dynamics exhibit distinct organizational properties across behavioral timescales. At the scale of the full behavioral session, neural trajectories evolved continuously over time. At the scale of specific social interaction episodes, however, reproducible dynamical structures emerged across repeated events. This episode-specific organization was distributed across sensory, prefrontal, and motor cortices, suggesting that behavioral timescale is a fundamental organizing dimension of distributed cortical dynamics. These findings establish a distributed and timescale-dependent organization of cortical dynamics during natural behavior.

Neural dimensions encode behavior across temporal scales

To examine how distributed cortical dynamics encode behavior across finer temporal scales, we first quantified behavior at multiple levels using hierarchical behavioral analysis developed in our laboratory. We reconstructed three-dimensional body poses of the subject and object mice from multi-view recordings (Fig. 2a). These behaviors were progressively quantified as low-level poses, mid-level movements, and high-level behavioral sequences (Fig. 2b), enabling us to examine the relationship between neural activity and behavior across timescales (Fig. 2c). Anti-Drift Pose Tracker (ADPT)[15] was used to reconstruct continuous three-dimensional skeletal configurations (Fig. 2d). These pose trajectories were further decomposed into discrete movement categories using Behavior Atlas (BeA)[14] (Fig. 2e), and temporally extended behavioral sequences were identified from the movement streams using Body Language-Bidirectional Encoder Representation from Transformers (BL-BERT)[13] (Fig. 2f).

We examined the temporal organization of behavior during social interaction episodes. Trial-aligned movement patterns revealed reproducible behavioral structure in both subject and object mice (Fig. 2g). In contrast, single-neuron calcium activity aligned to the same episodes was heterogeneous (Fig. 2h). This difference indicates that the temporal structure of behavior might be represented more clearly at the population level than at the level of individual neurons.

We therefore decomposed cortical population activity using principal component analysis (PCA). The leading principal components displayed reproducible temporal patterns aligned to the progression of social interaction episodes (Fig. 2i). PC1-PC3 showed clear changes associated with the interaction period, whereas this stereotyped temporal organization became progressively weaker in higher-order

PCs. This pattern was observed across S1, dmPFC, and M1, indicating that episode-related temporal structure was distributed across cortical population dynamics.

We next asked whether different neural dimensions preferentially encoded behavior at different temporal scales. Linear models were used to predict behavioral variables from neural PCs across cortical regions and behavioral levels (Fig. 2j). The first five PCs showed larger weights for movement and sequence variables, with prominent contributions from dmPFC-related features. In contrast, PCs beyond the first 100 dimensions showed larger weights for low-level pose variables, with stronger contributions from S1- and M1-related features. Thus, low-dimensional cortical dynamics preferentially encoded longer-timescale behavioral information, whereas higher-dimensional dynamics preferentially encoded shorter-timescale behavioral features.

These results reveal a dimension-dependent organization of distributed cortical dynamics across behavioral timescales. Low-dimensional dynamics preferentially represent movements and sequences over longer temporal scales, whereas high-dimensional dynamics emphasize fine-scale postural features. This organization is shared across cortical regions but differs in weighting, with dmPFC contributing more strongly to lower-dimensional, longer-timescale representations and S1 and M1 contributing more strongly to higher-dimensional, shorter-timescale representations.

High-dimensional dynamics causally shape behavioral details

The results above showed that distributed cortical dynamics encode behavior at different temporal scales through distinct neural dimensions. We next asked whether this dimensional organization has a causal role in behavioral control. To test this, we used all-optical interrogation in freely moving mice,

combining simultaneous calcium imaging of dmPFC neurons with global optogenetic stimulation to perturb the intrinsic dimensional structure of cortical dynamics (Fig. 3a). Max-projection images showed an overall reduction in neuronal fluorescence during optogenetic stimulation (Fig. 3b). Across full behavioral sessions, mean neural activity was significantly reduced, whereas responses during social interaction episodes were substantially more variable, suggesting that the effect of perturbation depended on behavioral context (Fig. 3c).

We next examined whether this neural perturbation altered behavior. During social interaction episodes, the non-locomotor speed of the subject mouse was significantly reduced under optogenetic stimulation, whereas no corresponding change was observed in the object mouse (Fig. 3d). This selective effect indicated that perturbing dmPFC activity altered the behavior of the stimulated animal rather than globally changing the interaction. The behavioral change was primarily expressed as a reduction in fine-scale, short-timescale behavioral details.

Given this timescale-specific behavioral effect, we next asked whether optogenetic perturbation selectively altered the dimensional organization of population dynamics. As in the unperturbed condition, single-neuron activity showed limited alignment with the temporal structure of social interaction episodes (Fig. 3e). In contrast, neural PCs remained aligned with interaction onset and termination, with this correspondence becoming progressively weaker at higher dimensions (Fig. 3f). We therefore compared the predictability of neural PCs before and during optogenetic perturbation. At higher-dimensional components accounting for progressively smaller fractions of population variance, PC predictability increased under perturbation, consistent with a reduction in the complexity of high-dimensional neural dynamics (Fig. 3g). Importantly, animals showing significant changes in these dimensions also exhibited a reduction in short-timescale behavioral details (Fig. 3h).

196 To directly characterize these dynamical changes, we modeled low- and high-dimensional neural
197 activity using recurrent switching linear dynamical systems (rSLDS) (Fig. 3i). Low-dimensional
198 trajectories showed no significant change in either radius or relative radial variability following
199 optogenetic perturbation. In contrast, both measures were significantly altered in high-dimensional
200 dynamics (Fig. 3j and k). These changes indicated a contraction of the high-dimensional neural state
201 space together with reduced dynamical variability, consistent with a loss of degrees of freedom in high-
202 dimensional population activity.

203 Together, these results establish a causal relationship between neural dimensionality and
204 behavioral timescale. Optogenetic perturbation selectively constrained high-dimensional dmPFC
205 dynamics and was accompanied by a reduction in short-timescale behavioral details. Thus, distinct
206 dimensions of distributed cortical dynamics not only encode behavior across timescales, but also
207 contribute to the control of behavioral features at corresponding temporal scales.

208

209 **Distributed cortical dynamics generalize across behavioral contexts**

210 The analyses above focused on social interaction episodes, in which distributed cortical dynamics
211 encoded behavior across temporal scales through distinct neural dimensions. We next asked whether
212 this dimensional organization generalized to broader behavioral contexts. Using social interaction
213 episodes as reference points, we extracted the intervening periods between successive interactions and
214 defined them as post-interaction episodes (Fig. 4a). Compared with social interaction episodes, these
215 periods contained more diverse and unconstrained behavioral patterns (Fig. 4b). Consistent with this

increased behavioral variability, single-neuron activity again showed no obvious stereotyped temporal structure (Fig. 4c).

We therefore examined whether population-level neural dynamics remained structured during post-interaction episodes. Principal-component trajectories showed reproducible temporal patterns, but these patterns appeared to separate into distinct groups across trials (Fig. 4d). Because post-interaction episodes varied substantially in duration, we used a Dynamic Time Alignment Kernel (DTAK)[18,24], which is sensitive to trajectory structure but relatively insensitive to temporal length, to cluster neural PC trajectories in an unsupervised manner (Fig. 4e). This analysis separated the trajectories into two groups. Across cortical regions, the mean PC1 trajectories of the two groups showed distinct temporal profiles: one group exhibited shorter, more rapidly varying dynamics, whereas the other showed longer-lasting and differently shaped dynamics (Fig. 4f). These results indicate that structured distributed cortical dynamics persist even under less constrained behavioral conditions.

We next asked whether these neural-dynamical groups were associated with distinct behavioral phenotypes. Using the unsupervised neural-PC cluster labels, we compared the corresponding movement compositions of the subject and object mice (Fig. 4g). In the subject mouse, the short-timescale cluster showed reduced immobility and increased running relative to the long-timescale cluster. In the object mouse, the short-timescale cluster showed increased walk–rise behavior. Representative behavioral examples are shown in Fig. 4h. These behavioral differences indicated that the short-timescale neural-dynamical group corresponded to rapid social re-engagement after an interaction, whereas the long-timescale group was associated with disengagement, during which the two animals returned to more independent behavior. Thus, the dimensional organization of distributed

cortical dynamics could distinguish broader behavioral contexts without explicitly using behavioral labels to define the neural states.

Together with the social-interaction results, these findings reveal a common dynamical framework spanning behavioral timescales from fine-scale poses to extended behavioral episodes (Fig. 4i). Within this framework, natural behavior was associated with three characteristic forms of distributed cortical dynamics: switching dynamics during social interaction, phasic dynamics during social re-engagement, and tonic dynamics during disengagement. Different neural dimensions further encoded behavioral information across temporal scales within each context. These results suggest that the dimensional organization of distributed cortical dynamics provides a general computational framework for representing behavior across timescales.

Discussion

Our findings reveal a dimensional principle by which distributed cortical dynamics organize behavior across timescales. Low-dimensional dynamics preferentially represented longer-timescale behavior, whereas high-dimensional dynamics captured shorter-timescale behavioral details. This organization was shared across S1, dmPFC, and M1, but each region expressed it with distinct weighting. The dmPFC was biased towards lower-dimensional, longer-timescale representations, whereas S1 and M1 were biased towards higher-dimensional, shorter-timescale features. This population-level organization is consistent with work showing that behavior and task variables are widely distributed across neural populations rather than being represented exclusively by anatomically isolated circuits[3,25-28]. Regional specialization and distributed coding may therefore emerge as complementary features of a common dynamical architecture.

This framework refines the conventional view of a cortical hierarchy of timescales. Intrinsic and representational timescales generally increase from sensory to association cortices, and cortical populations can encode information over markedly different temporal windows[2,29-31]. These observations have motivated the idea that cortical areas specialize in computations over distinct timescales. Our results preserve this regional specialization but suggest a different organizational basis. The same relationship between neural dimensionality and behavioral timescale was present across all three regions; what differed was each dimension's relative contribution. Cortical hierarchy may therefore be expressed as a gradient of dimensional weighting rather than a strict anatomical segregation of behavioral timescales.

The mapping between dimensionality and behavioral timescale also points to a functional division within population dynamics. Population-level dynamical descriptions have provided compact explanations for motor control, context-dependent computation, flexible timing, and latent cognitive variables[5,32-36]. Low-dimensional dynamics may provide stable coordinates for behavioral structure that must persist over longer intervals. Higher-dimensional dynamics, by contrast, can provide a richer set of degrees of freedom for rapidly varying details. This interpretation is consistent with the broader observation that lower-dimensional correlated representations can favour robustness, whereas high-dimensional representations increase representational capacity[37]. Such an organization would allow persistent behavioral structure and moment-to-moment flexibility to coexist within the same neural population.

Our perturbation experiments support a causal role for this dimensional organization. Optogenetic perturbation preferentially altered high-dimensional dmPFC dynamics, contracting the neural state space and reducing dynamical variability. These changes were accompanied by a selective loss of

short-timescale behavioral details, while lower-dimensional trajectory structure remained comparatively preserved. Previous work has shown that neural manifolds and behaviorally relevant dynamics can remain stable across behaviors and time[6,38,39], and that optical perturbations can causally probe population-level neural codes[40,41]. Our findings extend this view by indicating that high-dimensional activity is not simply residual variance around dominant low-dimensional dynamics, but contains behaviorally relevant degrees of freedom. Because our perturbation globally affected dmPFC activity, however, future experiments that directly perturb defined population dimensions will be required to establish dimension-specific causality.

An important feature of this framework is that its generality does not depend on a fixed neural trajectory. Natural and spontaneous behavior generates widespread, high-dimensional, and continuously varying brain activity[27,28,42], yet structured population dynamics can emerge when activity is examined relative to behaviorally meaningful epochs[5,6]. In our data, full-session trajectories evolved continuously, whereas specific interaction episodes revealed reproducible structure. Social interaction, re-engagement, and disengagement were associated with switching, phasic, and tonic dynamics, respectively. Despite these context-dependent trajectories, the dimensional relationship between neural activity and behavioral timescale remained preserved. Thus, what generalizes across behaviors may be the rule by which cortical dynamics are organized rather than any particular trajectory.

More broadly, this dimensional organization may offer a computational principle for artificial systems operating in dynamic environments. World models must preserve slowly varying latent structure for long-horizon prediction while retaining rapidly changing information required for immediate action[43]. Recurrent networks and state-space architectures similarly address the problem of

maintaining and transforming information over extended temporal ranges[44-47]. Our results raise the possibility that these temporal functions need not be assigned to separate modules. Instead, different timescales could coexist within a shared dynamical state space, with low-dimensional modes preserving longer-term structure and higher-dimensional modes retaining flexible detail. Functional specialization could then emerge through differential readout of these shared dimensions. Such a strategy provides a biologically grounded hypothesis for recurrent and world-model architectures that must coordinate information across timescales.

Methods

Animals. Forty male C57BL/6J mice (10-14 weeks old) were used in the experiment. Twenty mice without surgery were housed at 5 mice per cage, and the other fifteen mice with surgery were housed at 1 mouse per cage under a 12 h light-dark cycle at 22–25 °C with 40–70% humidity, and were allowed to access water and food ad libitum (15 in Shenzhen Institutes of Advanced Technology, Shenzhen, China, and another 5 in Nanjing Raygen Health, Nanjing, China). Fifteen mice that underwent surgery were anesthetized with isoflurane and placed in a stereotactic apparatus (RWD). The skull was exposed, and a craniotomy was performed for virus injection. To cover the entire primary somatosensory cortex (S1), dorsomedial prefrontal cortex (dmPFC), and primary motor cortex (M1), the mouse was microinjected at the following coordinates: AP, −0.60mm; ML, −2.40mm; DV, 2.00mm, AP, 2.00mm; ML, −0.30mm; DV, −1.80mm, and AP, 0.30mm; ML, −1.50mm; DV, −0.50mm. AAV9-CaMKII-GCaMP6s viruses were used for S1, and AAV2/9-hsyn-GCaMP6f viruses were used for dmPFC and M1. Five mice used for the optogenetic experiment followed the same surgery procedure as the above dmPFC mice, only the virus was different, which was AAV2/9-hsyn-GCaMP6f and AAV2/9-hsyn-ChRmine-mScarlet-Kv2.1. All animal experiments were conducted in accordance with the approval of the Animal Care and Use Committee at the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, and the Institutional Animal Care and Use Committee of PKU-Nanjing Institute of Translational Medicine (Approval ID: IACUC-2021-018).

Neural and behavioral data collection. The collection of neural and behavioral data is based on the MouseVenue3D system. Four Intel RealSense D435 cameras are mounted orthogonally on four supporting pillars made of stainless steel. The distance between the nearest cameras is 85 cm. The

cameras are adjusted to 45 cm off the ground to capture the whole view of the animal activities in the open field. Images were simultaneously recorded at 30 frames in 640×480 sizes by a PCI-E USB-3.0 data acquisition card and the PyRealsense2 Python camera interface package. The cameras are connected to a high-performance computer (i9 10900K, 128G RAM) equipped with a 512-gigabyte SSD and two 16-terabyte HDDs as an image acquisition platform. The computer also controls the camera calibration module.

The version of the miniature two-photon microscopy (mTPM) system is NUPERNOVA-600 (TRANSCEND VIVOSCOPE). The frame size is 512×512 and the frame rate is 4.84 Hz. The synchronization of 4 behavioral cameras and mTPM is based on the integrated synchronization module in mTPM. MouseVenue3D would send a TTL marker to mTPM every 30 frames. mTPM received the marker and output to the hardware time to a file. At the same time, mTPM writes the hardware time of every neural recording frame to another file. Later data alignment is based on these two timestamp files.

Two types of experiments were performed, including neural behavioral mapping and neural behavioral perturbation. In the neural behavioral mapping experiments, a subject mouse (implanted with an mTPM) and an object mouse (untreated) were allowed to interact freely in an open field for 15 minutes. During this period, we simultaneously recorded the subject's mTPM neural activity and multi-view behavior videos of both animals. In the neural behavioral perturbation experiments, the subject and object mice first engaged in 15 minutes of free social interaction without optogenetic stimulation, during which mTPM neural activity and multi-view behavior were recorded. Subsequently, the optogenetic laser was switched on, and a second 15-minute session of simultaneous mTPM imaging and behavioral recording was performed. Because the fluorescence intensity of the perturbation

experiment is weaker than that of mapping preparations, and ambient illumination can degrade imaging quality, we reduced the background lighting level during optogenetic experiments to ensure stable image quality. Within each experiment, the ambient illumination level was kept constant across pre-stimulation and stimulation sessions.

The quality control of neural data. The neural activities are extracted by suite2p in the 4.84 Hz frame rate. Other parameters of suite2p are kept default. The extracted neural activities are upsampled using step interpolation to 30 frames per second to align with the behavior data. In order to eliminate the frequency artifact caused by the upsample, the upsampled neural activities are filtered by an equiripple lowpass filter with 2.0 passband frequency, 2.2 stopband frequency, and 20 density factor. Further, the filtered neural data $F(t)$ are processed by the quality control step. We firstly extracted the running baseline $F_0(t)$, which was estimated as $F_0(t) = F_s(t) + m$. $F_s(t)$ is the smoothed fluorescence signal, calculated as the eighth percentile of $F(t)$ within a ±15s moving window centered at time t. m is a small constant value calculated by $m = \min(F_{std}(t)) + 0.1 * (\min(F_{std}(t)) + \max(F_{std}(t)))$. The neural activities finally expressed as the fractional change in the fluorescence with respect to baseline $\Delta F / F(t) = \frac{F(t) - F_0(t)}{F_0(t)}$. We manually checked that the baseline of $\Delta F / F$ is flatten but they existed noisy outliers. We set two hard thresholds $T_{\min} = -10$ and $T_{\max} = 20$ to exclude the neurons with high outliers.

Because optogenetic imaging exhibits a relatively low signal-to-noise ratio (SNR), we applied A fast blind zero-shot denoiser (Noise2Fast) to enhance image quality. Noise2Fast was used with its default parameters. To ensure a consistent noise baseline across conditions, every frame of both the mTPM

imaging and the corresponding perturbation sessions in the same mouse was denoised using Noise2Fast. The denoised image sequences were then processed by Suite2p for motion correction, ROI segmentation, and fluorescence extraction.

To further improve the SNR of the extracted neural signals, we implemented a one-dimensional (1D) Unet following the strategy of DeepCAD-RT. The 1D Unet consisted of an encoder and decoder, each composed of four layers of 1D convolutional blocks. The model was trained in a self-supervised manner using the odd-indexed frames as input and the even-indexed frames as targets (and vice versa). Training was performed with a mini-batch size of 16, a maximum of 20 epochs, and a learning rate of 1×10^{-4}. Because the predicted output sequence has half the temporal resolution of the original, we applied linear interpolation to restore the denoised signals to their original length. After that, the above upsampling, filtering, and baseline correction are applied to the denoised neural data.

Pose estimation using Anti-Drift Pose Tracker (ADPT). ADPT was trained on one NVIDIA GeForce RTX 3090 GPU. The definition of poses is the same as including "0: nose", "1: left_ear", "2: right_ear", "3: neck", "4: left_front_limb", "5: right_front_limb", "6: left_hind_limb", "7: right_hind_limb", "8: left_front_claw", "9: right_front_claw", "10: left_hind_claw", "11: right_hind_claw", "12: back", "13: root_tail", "14: mid_tail", and "15: tip_tail". The sleletons are definded including [0,1], [0,2], [1,2], [2,3], [1,3], [3,4], [3,5], [4,5], [4,6], [4,8], [5,7], [6,7], [5,9], [6,10], [7,11], [3,12], [4,12], [5,12], [6,12], [7,12], [13,12], [7,13], [6,13], [13,14], and [14,15]. Other parameters are set to the default.

We used the single animal pose estimation function of ADPT. We separately labeled 2615 frames of mice mounted FHIRM-TPM and 1852 frames of mice without neural recordings. Two ADPT models

are trained separately, and the predictions of the poses of each mouse are concatenated together for later 3D reconstruction.

The 3D reconstruction of two mice in the same open field has two steps. The first step is the 2D to 3D projection. The basic projection formula between 2D points and 3D space points is as follows.

$$s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Here, s represents the scaling factor, x and y are the points in the image, K is the camera internal reference, R is the rotation matrix, t is the translation matrix, and X, Y and Z represent the coordinates of the 3D points. The camera calibration step gave K, R, and t of each camera. Combing with the 2D points $[x, y, 1]^T$ frame by frame, the 3D points $[X, Y, Z, 1]^T$ can be soloved.

The second step of 3D reconstruction is ground rotation. The 3D skeletons from the first step exist in the relative coordinates to the calibration checkboard. Because the calibration checkboard is not always parallel to the ground, the 3D skeletons need to be rotated to the ground coordinates. We used the 3D points of two mice' hind paws to estimate the locations of the ground. We applied least square method to fit the plane of these 3D points as follows.

$$\arg\min \|Z - a_1 X - a_2 Y - a_3\|$$

Here, X, Y, and Z are the coordinates of 3D points. a_1 and a_2 are the weights of X and Y. a_3 is the intercede. So, the normal vector of the fitted plane is $[a_1, a_{2,}, a_3]$. The ground plane is defined to be parallel to the ground so its normal vector is $[0, 0, -1]$. Then the cross vector is $[-a_2, a_1, 0]$ and the rotation vector is $\frac{\arccos(-a_3)}{\sqrt{a_1^2 + a_2^2 + a_3^2}}[-a_2, a_1, 0]$. The rotated 3D skeletons $[X_r, Y_r, Z_r]$ can be calculated through Rodrigues formula $R(\cdot)$:

419 $$[X_r, Y_r, Z_r] = R\left(\frac{\arccos(-a_3)}{\sqrt{a_1^2 + a_2^2 + a_3^2}}[-a_2, a_1, 0]\right) \cdot [X, Y, Z]$$

420

421 **Movement decomposition and identification using Behavior Atlas (BeA).** BeA was trained on a
422 PC with an Intel i9-12900K CPU and 128 GB DDR4 RAM using the rotated 3D skeleton data. The
423 size of the smooth window of BeA is 500 ms, and the smoother is a median filter. The body points of
424 "12: back" and "13: root_tail" are used to align the direction of skeletons. The body sizes of each frame
425 are corrected to 25-pixel-centered distributions according to the length between "12: back" and "13:
426 root_tail". "14: mid_tail" and "15: tip_tail" are not selected for later movement decomposition. The
427 percentage of temporal dimensional reduction of the first layer is 5. The clustering number of the first
428 layer is 3. The kernel of the second layer is Gaussian. The kernel size of the second layer is 30. The
429 clustering number of the second layer is 24. The minimum and maximum time windows of the second
430 layer are 500 and 2000 ms. The initialization method of the second layer is spectral clustering. Other
431 BeA parameters are set to default. The movement decomposition is done one by one for each mouse
432 in each 15-minute video.
433 Further, the decomposed movements are reclusted for phenotype identification. Pairwise dynamic time
434 alignment kernel (DTAK) distances are calculated between all of the movement segmentations. The
435 kernel size is 30. The DTAK distance matrix is embedded using uniform manifold approximation and
436 projection (UMAP) with 199 n_neighbors and 0.05 min_dist. Hierarchical clustering with standard
437 Euclidean distance and Ward linkage is used to cluster the UMAP embedding. The clustering number
438 is 80. The recluster is separately done in the mice with or without FHIRM-TPM. In total, 160 clusters

of movements are manually identified. The clusters with the same name are merged together. Finally, 49 movement categories are defined.

Sequence extraction using Body Language-Bidirectional Encoder Representation from Transformers (BL-BERT). BL-BERT was trained on one NVIDIA GeForce RTX 3090 GPU. 400,000 sequences are sampled from the movements for training. The length of each sequence is 120 movement modules. The percentage of masks is 0.2. The training batch size is 50. The number of epochs is 20. The initial learning rate is 1.0. BL-BERT is trained using a warm-up of 3000 iterations. The resampling time of the mask for prediction is 100. The 35 cluster number is selected according to the silhouette coefficient, based on the Hamming distance matrix across each pair of the sequences.

The separation of social interaction episodes and post-interaction trials. According to our previous studies, the change point between close interaction and other behavior is about 5 cm of body distance, which can be converted to around 60 pixels using 0.88 cm/pixel[18]. Using 5 cm as the threshold between close interaction and other behavior, the full behavior data can be separated into dynamic segments. 5 s is chosen as the baseline of close interaction and the time after close interaction in the neural behavioral mapping data. The time length of the close interaction part is set to 2 s to 10 s to exclude the very short interaction and long-time stay close but no interaction. Using these criteria, 98 trials are separated from all 15 videos for further analysis. The trials are sorted according to their length for visualization.

In the perturbation experiments, we initially evaluated a 5 s baseline window. However, the neural principal component (PC) trajectories did not exhibit a clear or reproducible pattern under this baseline

length. After systematic inspection, we found that shortening the baseline window enhanced the visibility of consistent neural patterns. This limitation is likely related to the character of PCA. The baseline dynamics differ substantially from those during social interaction episodes, and the relatively low signal-to-noise ratio of the perturbation recordings makes the extracted PCs more susceptible to baseline-induced variance. We therefore tested a shorter 2 s baseline, which consistently revealed a robust and reproducible PC pattern in the perturbation data. Based on this empirical optimization, a 2 s baseline window was used for all subsequent perturbation analyses. Correspondingly, the time length of the close interaction part is set to 2 s to 4 s to match the shorter baseline.

The post-interaction trials are separated using the same 5 cm distance threshold. Considering that the concerned data are the after-interaction, we chose 1 s close interaction as the baseline to keep as much data as possible. The time length of the after-interaction trials is from 10 s to 180 s, which firstly covers the baseline and after-length of close interaction trials, and secondly contains as many trials as possible. Using these criteria, 219 after-interaction trials are separated from all 15 videos for further analysis. The trials are sorted according to their length for visualization.

Generalized time warping (GTW). Within each cortical area, social-episode trajectories were aligned using generalized time warping (GTW), which jointly performs spatial alignment by canonical correlation analysis (CCA) into a three-dimensional common latent space and temporal alignment by monotone time warping onto a common time axis[23]. We used a polynomial (order 3, weight 0.4) and tangent (order 3, weights 0.6 and 1) temporal basis, CCA dimensionality d=3, and the default optimization settings (GTW: 50 iterations; frame-to-time warping: 2 iterations; linear input with MATLAB quadratic programming solver).

Visualization of neural manifolds. Neural population activity was embedded into a three-dimensional space using UMAP, and the resulting manifold was visualized directly: the full-session embedding was plotted as a light-grey point cloud (subsampled every eighth time point), over which the far-close-far (FCF) social-episode trajectories were overlaid as colored curves, one color per cortical area. After cross-animal alignment, the GTW-aligned trajectories (the first three canonical components, denoted GTW 1-3) were rendered as thin individual traces together with their mean trajectory. To summarize the shared geometry of the aligned trajectories, a curved tubular surface was fitted to each cortical area's manifold as follows. Each trajectory was first resampled onto a common normalized GTW time axis with 600 time points using piecewise-cubic Hermite interpolation. The manifold centerline was taken as the running mean of the resampled trajectories (moving-average window of 21 points), and its unit tangent was computed as the normalized gradient along time. A parallel-transport frame (two orthonormal vectors spanning the plane normal to the tangent) was propagated along the centerline to avoid spurious rotation of consecutive cross-sections. At each time point, trial points were projected onto this normal plane, and a regularized 2×2 covariance matrix (with $10^{-6} \times$ trace shrinkage) was estimated; the cross-section radius was set to the 90th percentile of the resulting Mahalanobis distances, so that each elliptical cross-section enclosed 90% of trial points, and the lower-triangular Cholesky factor of the covariance defined the elliptical shape. The radius and shape parameters were smoothed along time with a moving-average window of 31 points, and the tube surface was reconstructed by sweeping the smoothed ellipse along the centerline (72 angular samples per cross-section). For display, the tube mesh and centerline were further smoothed with a Gaussian kernel (spans of 151 and 31 points, respectively); the centerline was colour-coded by normalized time

using a reversed Spectral colour map, an arrow indicated the direction of time, and the start and end of the centerline were marked with green and red dots. A single viewing angle, chosen by inspecting candidate azimuth–elevation combinations for each area, was applied to all cortical areas.

Reduction in pairwise Euclidean distance after GTW. For each cortical area, social-episode trajectories before and after GTW were resampled to a common length of 60 time points by piecewise-cubic Hermite interpolation (pchip), mean-centered and normalized to unit Frobenius norm. Pairwise Euclidean distances between all social episodes within each session were computed and averaged, yielding one value per session. Before-GTW and after-GTW distances from the same sessions were compared with a one-sided paired t-test (testing whether after-GTW distances were smaller); the resulting p-values were corrected for multiple comparisons across the three cortical areas using the Holm method. N = 5, 5, and 5 paired sessions for S1, dmPFC and M1, respectively.

Procrustes shape similarity across cortical areas. For GTW-aligned trajectories, we computed a normalized Procrustes similarity between pairs of trajectories as the sum of the singular values of their cross-covariance matrix, $\Sigma\sigma(x^{T}y)$, after each trajectory had been resampled to 60 time points, mean-centered, and normalized to unit norm. This quantity is invariant to rotation, reflection, translation, and scaling, and equals 1 when two manifolds have identical shape. Trial-level similarities were computed between every trajectory from one cortical area and every trajectory from another area, and were then averaged within each cross-area session pair ($5 \times 5 = 25$ session pairs per area pair). This analysis was descriptive; no statistical comparisons were performed among cortical-area pairs.

Cross-validated representational similarity analysis (RSA). Movement states were defined from behavior module labels (global movement identities that are comparable across sessions). Within each session, the UMAP embedding was standardized (z-scored per dimension), and bouts, contiguous segments of a single movement label, were retained if they were 15–300 frames long and each movement label had at least 6 (and at most 20, randomly subsampled) bouts. Bouts were resampled to 20 time points. For every cross-area session pair sharing at least 6 movement labels, the bouts of each label were split into two halves and averaged to form split-half templates, and a representational dissimilarity matrix (RDM) was built from the Euclidean distances between label templates, rank-transformed (Spearman) and z-scored. Cross-validated representational similarity was computed as the mean of (i) the correlation between the RDM from half 1 of area A and the RDM from half 2 of area B and (ii) the correlation between half 2 of area A and half 1 of area B. To test whether neural manifolds carrying matched movement labels corresponded more than expected by chance, movement-label identities were permuted 1,000 times within each session pair (preserving every trajectory and its temporal structure), and the cross-validated RSA was recomputed for each permutation. A one-sided p-value was obtained as (1 + number of permutations exceeding the observed value) / 1,001, and corrected across the three cortical-area pairs using the Holm method. N = 23, 25, and 23 cross-area session pairs for S1-M1, S1-dmPFC, and M1-dmPFC, respectively.

Linear mapping using the PCs of neural activities to hierarchical behaviors. Before linear mapping, the neural activities are transformed using principal component analysis (PCA). The PCA is used in each of the trials above. The coefficients of PCA are normalized and aligned according to their length, one PC by one PC. Because the behavior has three hierarchies and their continuities are

549 different, we applied two linear models to separately map neural PCs with continuous poses and
550 discrete movements and sequences.
551 The generalized linear model (GLM) is used for linear mapping between neural PCs and poses because
552 they are both continuous variables. The neural PCs of subject mice are separately mapped to the subject
553 and object poses trial by trial using GLM in close interaction data. The discriminant analysis is used
554 for linear mapping between neural PCs and movements and also between neural PCs and sequences
555 because the output movement and sequence labels are discrete variables. The neural PCs of subject
556 mice are separately mapped to the subject/object movements/sequences trial by trial. The mean
557 mapping accuracy of all of them is 99.81±0.63%. The high mapping accuracy has verified that GLM
558 and discriminant analysis can weight the combination of different neural PCs to the optimized fitting
559 of poses, movements, and sequences.
560
561 **The parameter settings of the recurrent switching linear dynamic system (rSLDS).** We used
562 rSLDS to describe the dynamics of the neural data in this paper. Before using rSLDS, the data are
563 intercepted from the beginning of close interactions to keep each data having the same initial baseline.
564 The threshold of distance to separate close interaction is 5 cm. The baseline of the beginning of close
565 interactions is 15 s. Each behavior together with neural recording data is cut to 600 s from the start of
566 the baseline to the end. Further, the data are separated into 3 groups for rSLDS modeling.
567 The first group is the close interaction group. This step is similar to the separation of close interaction
568 trials. The minimum time of close interaction is 2 s, the baseline is 2 s, and the time after interaction
569 is 2 s. The numbers of trials of S1, dmPFC, and M1 are separately 13, 17, and 25. The second group
570 is the after-interaction group. This step is also similar to the separation of after-interaction trials. The

minimum time of after-interaction is 5 s. The baseline is 1 s. The numbers of trials of S1, dmPFC, and M1 are separately 60, 80, and 79. The third group is the all-data group. Each trial is the 600 s data described before. After the separation of these 3 groups, their neural PCs are calculated trial by trial. The first 150 PCs of each trial are kept. The 150 neural PCs inside each brain area are concatenated for training rSLDS one by one brain area across 3 groups.

The rSLDS model consists of a discrete state $z_t \in \{1,...,K\}$, a continuous latent state $x_t \in \mathbb{R}^D$, and an observation $y_t \in \mathbb{R}^{D_{obs}}$. In this study, we set $K = 2$ (two latent dynamical regimes emerging from the concatenated PCA trajectories), $D = 2$ (for visualization and interpretability), and $D_{obs} = 150$ (the forst 150 PCs of neural activity). The generative process for each discrete state k is:

$$\text{(State transitions)}\ p(z_t = k \mid x_{t-1}) = \text{softmax}_k(R_k^{\bullet} x_{t-1} + r_k),$$

$$\text{(Continuous dynamics)}\ x_t \mid x_{t-1}, z_t = k \sim N(A_k x_{t-1} + b_k, Q_k),$$

$$\text{(Observations)}\ y_t \mid x_t \sim N(Cx_t + d, R), \text{ s.t. } C^{\bullet} C = I.$$

Recurrent transitions (transitions="recurrent_only") allow the switching probabilities to depend solely on the previous continuous latent state x_{t-1}. State-specific dynamics use diagonal linear Gaussian matrices A_k (dynamics="diagonal_gaussian") to reduce overfitting and enable vector field interpretation. Orthogonal Gaussian emissions (emissions="gaussian_orthog", single_subspace=True) ensure that the neural observation space projects into a stable, interpretable low-dimensional subspace. Model parameters directly correspond to the script variables: R is transitions.Rs, r is transitions.r, A_k is dynamics.As, and b_k is dynamics.bs.

We fit a separate rSLDS model for each cortical region using Laplace-EM variational inference (method="laplace_em") with a structured mean-field posterior (variational_posterior="structured_meanfield"). The initialization is random with a fixed seed for

reproducibility. The iteration time is 1000. The optimization target is the evidence lower bound. The posterior outputs are the continuous latent means $\hat{x}_{1:T}$ and the most likely discrete state sequence $\hat{z}_{1:T}$ were extracted using $\hat{z}_{1:T} = \arg\max p(z_{1:T} \mid \hat{x}_{1:T}, y_{1:T})$.

To visualize the state-dependent flow fields, we constructed a 2D grid over the latent space and computed the one-step drift $\Delta x = (A_k x + b_k) - x$. Each grid point was assigned to its most likely discrete state via the recurrent transition function $\arg\max_k (R_k^{\bullet} x + r_k)$. Vector fields show Δx arrows within each state region. Plot limits were adapted to $l_i = \max_t |\hat{x}_{t,i}| + 1$ to show the full fields.

Non-locomotion speed. Non-locomotion speed was used to quantify the animal's movement when no net body displacement occurred (i.e., excluding walking or running). This metric captures subtle kinematic events such as micromovements, sniffing, turning-in-place, and other posture changes that do not alter the animal's global position but reflect fine-scale motor dynamics. Let $[x_t, y_t, z_t]$ denote the 3D position of the animal at time t. The instantaneous speed is

$$v_t = \frac{\left\| [x_t, y_t, z_t] - [x_{t-1}, y_{t-1}, z_{t-1}] \right\|}{\Delta t},$$

where Δt is the interval time between two frames.

Non-locomotion speed is defined as

$$v_t^{nl} = \begin{cases} v_t, v_t < v_{thres} \\ 0, v_t \geq v_{thres} \end{cases}$$

where v_{thres} is the threshold to separate non-locomotion speed and locomotion speed. We applied 150 pixels/s as v_{thres}, empirically determined as approximately three times the mean baseline speed. Frames with instantaneous movement below this threshold were categorized as non-locomotion movements. The mean non-locomotion speed is then

$$\overline{v^{nl}} = \frac{\sum_{t=1}^{T} v_t^{nl}}{T},$$

where T is the frame length of the total time.

The mTPM illustration using max projection. To visualize the spatial activation of neurons in mTPM recordings, we generated maximum-intensity projection images from mTPM imaging stacks. For each imaging session, a mean reference frame was computed across the entire recording for each stack and subtracted from each frame to obtain a baseline-corrected signal. The resulting frames were normalized to the range [0,1] using linear rescaling. To suppress slow baseline drifts and emphasize transient activity, we applied a temporal rolling normalization using a sliding window of 100 frames. This rolling normalization enhanced local fluctuations while preserving spatial structure. Finally, for each stack, we computed a maximum-intensity projection across the temporal dimension:

$$I_{\max}(x, y) = \max_t I(x, y, t),$$

where $I(x, y, t)$ denotes the normalized, rolling-corrected fluorescence signal at pixel (x, y) and time t. The resulting max-projection images provide a compact representation of the spatial trace of neural activity over the entire recording and were used for visualization of mTPM activity patterns.

PC predictability. To quantify the dynamical predictability of each neural PC, we estimated its one-step temporal predictability using a minimal first-order autoregressive (AR(1)) model combined with K-fold cross-validation. For each PC time series $\{p_t\}_{t=1}^{T}$, we constructed paired samples (p_t, p_{t+1}) and fit a linear autoregressive model of the form:

$$p_{t+1} = ap_t + b,$$

where a and b were estimated by ordinary least squares.

To obtain a robust and unbiased estimate of predictability, we performed k-fold cross-validation (k=30). In each fold, the model was trained on the training subset and evaluated on held-out samples. Prediction errors were accumulated across all folds. We quantified predictability using Root Mean Squared Error (RMSE) metrics:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t}(p_{t+1} - \hat{p}_{t+1})^2}\ .$$

Lower RMSE indicates higher temporal predictability and stronger dynamical stability of the corresponding PC. This predictability metric captures the degree to which each neural PC follows a stable low-order dynamical rule, and was used to assess selective destabilization of high-dimensional components under optogenetic perturbation.

Directional alignment analysis between neural predictability and behavior. To assess whether optogenetic-induced changes in neural dynamics were behaviorally relevant, we examined the directional consistency between changes in PC predictability and changes in non-locomotion speed across individual mice. For each mouse i and each PC k, we computed:

$$\begin{aligned} \Delta P_{i,k} &= P_{i,k}^{\text{op}} - P_{i,k}^{\text{free}} \\ \Delta S_i &= S_i^{\text{op}} - S_i^{\text{free}} \end{aligned},$$

where P denotes PC predictability metric and S denotes the mean value of non-locomotion speed of each body point, under optogenetic (op) and control (free) conditions.

We focused on the directional relationship between neural destabilization and behavioral change caused by optogenetic perturbation. For each PC k, we counted across mice the number of cases in which the sign of predictability change was opposite to the sign of speed change:

$$\mathrm{match}_{i,k} = \mathrm{I}\left[\mathrm{sign}(\Delta P_{i,k}) = -\mathrm{sign}(\Delta S_i)\right],$$

which reflects the hypothesis that increased neural unpredictability is associated with reduced non-locomotion speed, and vice versa. To avoid spurious contributions from near-zero fluctuations, mice with $|\Delta P_{i,k}| < \varepsilon_P$ or $|\Delta S_i| < \varepsilon_S$ were excluded ($\varepsilon = 10^{-6}$).

For each PC, we performed a one-sided exact binomial test against the null hypothesis of random directional alignment (probability $p = 0.5$). The observed number of directionally consistent mice m_{obs} was compared to the binomial distribution:

$$p_{\mathrm{dir}} = \Pr(M \geq m_{\mathrm{obs}} \mid N, p = 0.5),$$

where N is the number of valid mice.

This test evaluates whether optogenetic-induced neural destabilization in a given PC is systematically aligned with no-locomotion speed changes across animals. PCs showing significant directional alignment indicate dimensions in which neural dynamical perturbations are kinematically relevant.

Radius of rSLDS latent trajectories. To quantify the geometric dispersion of latent trajectories inferred by rSLDS, we measured the radial extent of each trajectory in the two-dimensional latent space. For each latent trajectory $\{x_t\}_{t=1}^{T}$, we first computed its centroid:

$$c = \frac{1}{T}\sum_{t=1}^{T} x_t \ .$$

We then computed the instantaneous radius at each time point as the Euclidean distance from the centroid:

$$r_t = \left\| x_t - c \right\|_2 ,$$

and reported the mean trajectory radius as:

$$r_{\text{mean}} = \frac{1}{T} \sum_{t=1}^{T} r_t .$$

This mean radius summarizes how widely the latent trajectory disperses around its center; larger values indicate more spatially extended latent dynamics, whereas smaller values indicate more compact dynamics.

Relative radial variability. Based on r_t, the relative radial variability is calculated by

$$R_v = \frac{r_{\text{std}}}{r_{\text{mean}} + \varepsilon} ,$$

where r_{std} is the standard deviation. This dimensionless ratio measures the relative fluctuation of trajectory spread, independent of its absolute scale. Small R_v values indicate that the trajectory maintains a nearly constant radius, corresponding to geometrically stable and confined dynamics, whereas large R_v values indicate strong radial fluctuations, reflecting irregular, unstable, or intermittently expanding latent dynamics.

Statistics. Before hypothesis testing, data were first tested for normality by the Shapiro-Wilk normality test and for homoscedasticity by the F test. For normally distributed data with homogeneous variances, parametric tests were used; otherwise, non-parametric tests were used. All of the ANOVA analyses

694 are corrected by the recommended options of Prism 8.0. No data in this work are removed. All related
695 data are included in the analysis.
696
697

Data availability

We provide all model weights and the validation data used in this paper at https://doi.org/10.6084/m9.figshare.30773648, https://doi.org/10.6084/m9.figshare.30781931, https://doi.org/10.6084/m9.figshare.30771338, https://doi.org/10.6084/m9.figshare.30772454, https://doi.org/10.6084/m9.figshare.30772685, https://doi.org/10.6084/m9.figshare.30772766, https://doi.org/10.6084/m9.figshare.30773072. They also contain the necessary intermediate data for the validation and visualization of this study. This paper can be replicated using the downloadable data and supplied code.

Code availability

All codes to reproduce this work is at https://github.com/YNCris/behavioral_timescales.

References

1 Mathis, M. W. & Mathis, A. Joint modelling of brain and behaviour dynamics with artificial intelligence. *Nature Reviews Neuroscience*, 1–14 (2025).

2 Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience* **17**, 1661–1663 (2014). https://doi.org/10.1038/nn.3862

3 Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019).

4 Testard, C. *et al.* Neural signatures of natural behaviour in socializing macaques. *Nature* **628**, 381–390 (2024). https://doi.org/10.1038/s41586-024-07178-6

5 Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012). https://doi.org/10.1038/nature11129

6 Gallego, J. A. *et al.* Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature communications* **9**, 4233 (2018).

7 Suresh, A. K. *et al.* Neural population dynamics in motor cortex are different for reach and grasp. *elife* **9**, e58848 (2020).

8 Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M. & Cunningham, J. P. Reorganization between preparatory and movement population responses in motor cortex. *Nature communications* **7**, 13239 (2016).

9 Recanatesi, S., Pereira-Obilinovic, U., Murakami, M., Mainen, Z. & Mazzucato, L. Metastable attractors explain the variable timing of stable behavioral action sequences. *Neuron* **110**, 139–153. e139 (2022).

10 Mazzucato, L. Neural mechanisms underlying the temporal organization of naturalistic animal behavior. *Elife* **11**, e76577 (2022). https://doi.org/10.7554/eLife.76577

11 Mimica, B. *et al.* Behavioral decomposition reveals rich encoding structure employed across neocortex in rats. *Nature Communications* **14**, 3947 (2023).

12 Han, Y. *et al.* MouseVenue3D: A Markerless Three-Dimension Behavioral Tracking System for Matching Two-Photon Brain Imaging in Free-Moving Mice. *Neurosci Bull* **38**, 303–317 (2022). https://doi.org/10.1007/s12264-021-00778-6

13 Han, Y. *et al.* Revealing the Body Language of Social Interaction in Free-Moving Mice Using BL-BERT. *Tsinghua Science and Technology* **31**, 1487–1500 (2026).

14 Huang, K. *et al.* A hierarchical 3D-motion learning framework for animal spontaneous behavior mapping. *Nat Commun* **12**, 2784 (2021). https://doi.org/10.1038/s41467-021-22970-y

15 Tang, G. *et al.* Anti-drift pose tracker (ADPT), a transformer-based network for robust animal pose estimation cross-species. *Elife* **13**, RP95709 (2025). https://doi.org/10.7554/eLife.95709

16 Mathis, A., Schneider, S., Lauer, J. & Mathis, M. W. A primer on motion capture with deep learning: principles, pitfalls, and perspectives. *Neuron* **108**, 44–65 (2020).

17 Pereira, T. D., Shaevitz, J. W. & Murthy, M. Quantifying behavior to understand the brain. *Nat Neurosci* **23**, 1537–1549 (2020). https://doi.org/10.1038/s41593-020-00734-z

18 Han, Y. *et al.* Multi-animal 3D social pose estimation, identification and behaviour embedding with a few-shot learning framework. *Nature Machine Intelligence* **6**, 48–61 (2024). https://doi.org/10.1038/s42256-023-00776-5

19 Pereira, T. D. *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat Methods* **19**, 486–495 (2022). https://doi.org/10.1038/s41592-022-01426-1

20 Lauer, J. *et al.* Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nature Methods* **19**, 496–504 (2022).

21 Zong, W. *et al.* Miniature two-photon microscopy for enlarged field-of-view, multi-plane and long-term brain

imaging. *Nat Methods* **18**, 46–49 (2021).

22 Zong, W. *et al.* Fast high-resolution miniature two-photon microscopy for brain imaging in freely behaving mice. *Nat Methods* **14**, 713–719 (2017). https://doi.org/10.1038/nmeth.4305

23 Zhou, F. & De la Torre, F. in *2012 IEEE conference on computer vision and pattern recognition.* 1282–1289 (IEEE).

24 Zhou, F., Torre, F. D. l. & Hodgins, J. K. Hierarchical Aligned Cluster Analysis for Temporal Clustering of Human Motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**, 582–596 (2013). https://doi.org/10.1109/TPAMI.2012.137

25 Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019). https://doi.org/10.1038/s41586-019-1787-x

26 Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nature neuroscience* **22**, 1677–1686 (2019).

27 Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013). https://doi.org/10.1038/nature12742

28 Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013). https://doi.org/10.1038/nature12160

29 Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* **548**, 92–96 (2017). https://doi.org/10.1038/nature23020

30 Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat Neurosci* **21**, 102–110 (2018). https://doi.org/10.1038/s41593-017-0028-6

31 Zeraati, R., Levina, A., Macke, J. H. & Gao, R. Neural timescales from a computational perspective. *Nat Neurosci* **29**, 1534–1547 (2026). https://doi.org/10.1038/s41593-026-02343-8

32 Remington, E. D., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible Sensorimotor Computations through Rapid Reconfiguration of Cortical Dynamics. *Neuron* **98**, 1005–1019.e1005 (2018). https://doi.org/10.1016/j.neuron.2018.05.020

33 Sohn, H., Narain, D., Meirhaeghe, N. & Jazayeri, M. Bayesian Computation through Cortical Latent Dynamics. *Neuron* **103**, 934–947.e935 (2019). https://doi.org/10.1016/j.neuron.2019.06.012

34 Rajan, K., Harvey, C. D. & Tank, D. W. Recurrent Network Models of Sequence Generation and Memory. *Neuron* **90**, 128–142 (2016). https://doi.org/10.1016/j.neuron.2016.02.009

35 Egger, S. W., Remington, E. D., Chang, C. J. & Jazayeri, M. Internal models of sensorimotor integration regulate cortical dynamics. *Nat Neurosci* **22**, 1871–1882 (2019). https://doi.org/10.1038/s41593-019-0500-6

36 Paton, J. J. & Buonomano, D. V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**, 687–705 (2018). https://doi.org/10.1016/j.neuron.2018.03.045

37 Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M. & Harris, K. D. High-dimensional geometry of population responses in visual cortex. *Nature* **571**, 361–365 (2019). https://doi.org/10.1038/s41586-019-1346-5

38 Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A. & Miller, L. E. Long-term stability of cortical population dynamics underlying consistent behavior. *Nat Neurosci* **23**, 260–270 (2020). https://doi.org/10.1038/s41593-019-0555-4

39 Sani, O. G., Abbaspourazad, H., Wong, Y. T., Pesaran, B. & Shanechi, M. M. Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification. *Nat Neurosci* **24**, 140–149 (2021). https://doi.org/10.1038/s41593-020-00733-0

40 Sani, O. G., Pesaran, B. & Shanechi, M. M. Dissociative and prioritized modeling of behaviorally relevant neural dynamics using recurrent neural networks. *Nat Neurosci* **27**, 2033–2045 (2024). https://doi.org/10.1038/s41593-024-01731-2

41 Rickgauer, J. P., Deisseroth, K. & Tank, D. W. Simultaneous cellular-resolution optical perturbation and imaging of place cell firing fields. *Nat Neurosci* **17**, 1816–1824 (2014). https://doi.org/10.1038/nn.3866

42 Adesnik, H. & Abdeladim, L. Probing neural codes with two-photon holographic optogenetics. *Nat Neurosci* **24**, 1356–1366 (2021). https://doi.org/10.1038/s41593-021-00902-9

43 Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M. B. & Moser, E. I. The neural bases for timing of durations. *Nat Rev Neurosci* **23**, 646–665 (2022). https://doi.org/10.1038/s41583-022-00623-3

44 Perich, M. G., Narain, D. & Gallego, J. A. A neural manifold view of the brain. *Nature Neuroscience* **28**, 1582–1597 (2025).

45 Hafner, D., Pasukonis, J., Ba, J. & Lillicrap, T. Mastering diverse control tasks through world models. *Nature* **640**, 647–653 (2025). https://doi.org/10.1038/s41586-025-08744-2

46 Tiezzi, M. *et al.* Back to recurrent processing at the crossroad of transformers and state-space models. *Nature Machine Intelligence* **7**, 678–688 (2025). https://doi.org/10.1038/s42256-025-01034-6

47 Singh, S. H., van Breugel, F., Rao, R. P. N. & Brunton, B. W. Emergent behaviour and neural dynamics in artificial agents tracking odour plumes. *Nature Machine Intelligence* **5**, 58–70 (2023). https://doi.org/10.1038/s42256-022-00599-w

Acknowledgements

This work was supported in part by STI2030-Major Projects grant 2021ZD0202205 (HC), STI2030-Major Projects grant 2021ZD0203900 (PW), National Natural Science Foundation of China grant T2394533 (PW), National Natural Science Foundation of China grant T2600297 (YH), Jiangsu Provincial Major Science and Technology Special Project grant BG2025040 (PW), Jiangsu Provincial Advanced Technology Research and Development Program grant BF2025060 (PW), and Start-up Research Fund of Southeast University grant RF1028625096 (PW).

Author contributions

Conceptualization: YH, QL, SL, HC, PW
Methodology: YH, PW
Investigation: YH
Visualization: YH, QL, SL, HC, PW
Funding acquisition: HC, PW
Project administration: QL, SL, PW
Supervision: YH, QL, SL, HC, PW
Writing – original draft: YH
Writing – review & editing: YH, QL, SL, PW

Competing interests

The authors declare no competing interests.

Figures and legends

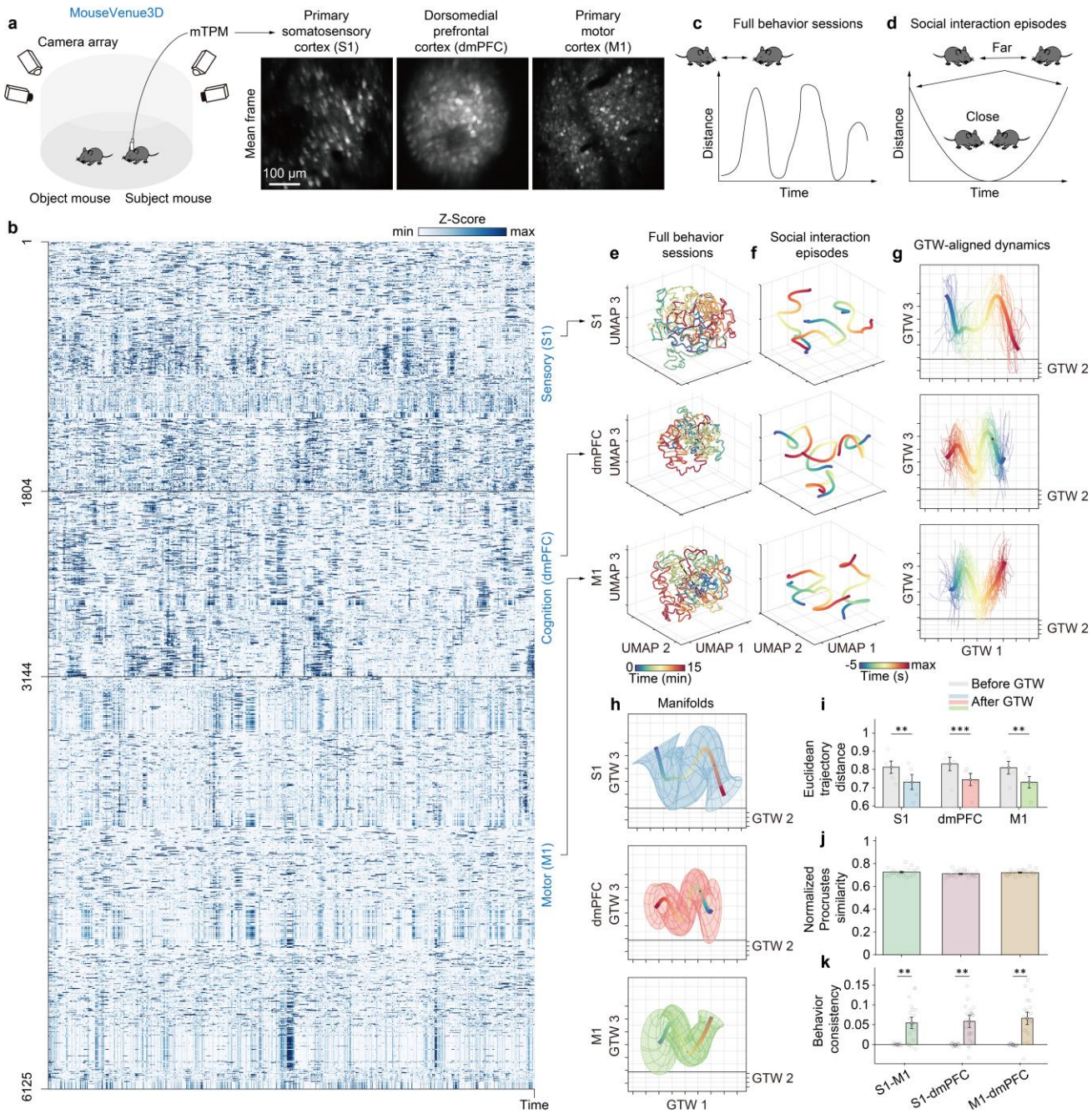


Fig. 1 | Temporally stable distributed cortical population dynamics during naturalistic social behavior.

a, Experimental setup for simultaneous recording of naturalistic social behavior and cortical population activity. Freely interacting mice were recorded using a multi-camera system (MouseVenue3D)[12] combined with miniaturized two-photon microscopy (mTPM) targeting distributed cortical regions, including the primary somatosensory cortex (S1), dorsomedial prefrontal cortex (dmPFC), and primary motor cortex (M1). The mouse implanted with the mTPM is referred to as the subject mouse, whereas the interacting partner without neural implants is referred to as the object mouse throughout the study.

b, Z-scored neural activity from S1, dmPFC, and M1. Neurons are grouped by cortical area: 1,804

from S1, 1,340 from dmPFC, and 2,981 from M1, for a total of 6,125 neurons.
c, Schematic illustration of a full behavioral session. The distance between two freely interacting mice continuously fluctuated over time.
d, Extraction of social interaction episodes from full behavioral sessions. Individual close-interaction events were identified based on inter-animal distance and used for subsequent alignment and population-level analysis.
e, Low-dimensional representations of cortical population activity during full behavioral sessions. Neural population activity from S1, dmPFC, and M1 was projected into three-dimensional UMAP space, with trajectories color-coded by elapsed time throughout the 15-min recording session.
f, Low-dimensional population trajectories during individual social interaction episodes extracted from the full behavioral sessions. Neural activity was projected into the same UMAP space and aligned to the onset of close social interaction, revealing recurrent trajectory structures across interaction episodes.
g, Generalized Time Warping (GTW)[23] alignment of social interaction episodes pooled across all animals. Neural population trajectories from S1, dmPFC, and M1 were temporally warped to a common reference axis to account for variability in episode duration and timing across animals, revealing reproducible population dynamics across subjects.
h, Representative low-dimensional neural manifolds reconstructed from GTW-aligned population trajectories in S1, dmPFC, and M1. Colored trajectories indicate temporally ordered population activity embedded within the corresponding manifold structure.
i, GTW reduced the mean pairwise Euclidean distance between social-episode neural trajectories in S1, dmPFC, and M1. Pairwise distances were first averaged across all social-episode pairs within each session, and the same sessions were then compared before and after GTW (one-sided paired *t*-test followed by Holm's multiple-comparisons correction, mean ± SEM, $N = 5$, 5, and 5 paired sessions; the adjusted *P* values from left to right are 0.0036, 0.0007, and 0.0017).
j, Normalized Procrustes similarity revealed comparable shapes of GTW-aligned neural manifolds across cortical-area pairs, where a value of 1 indicates identical manifold shape after rotation and scaling (one-way ANOVA followed by pairwise Welch's *t*-tests with Holm's multiple-comparisons correction, mean ± SEM, $N = 25$, 25, and 25 cross-area session pairs for S1-M1, S1-dmPFC, and M1-dmPFC, respectively, the adjusted pairwise *P* values for S1-M1 versus S1-dmPFC, S1-M1 versus M1-dmPFC, and S1-dmPFC versus M1-dmPFC are 0.19727, 0.60320, and 0.20055, respectively).
k, Cross-validated representational similarity analysis showed greater correspondence between neural manifolds carrying matched movement labels than between those carrying shuffled labels (1,000 within-session-pair label permutations followed by Holm's multiple-comparisons correction, mean ± SEM, $N = 23$, 25, and 23 cross-area session pairs for S1-M1, S1-dmPFC, and M1-dmPFC, respectively; the adjusted *P* values from left to right are 0.0030, 0.0030, and 0.0030). $^{**}P < 0.01$, $^{***}P < 0.001$.

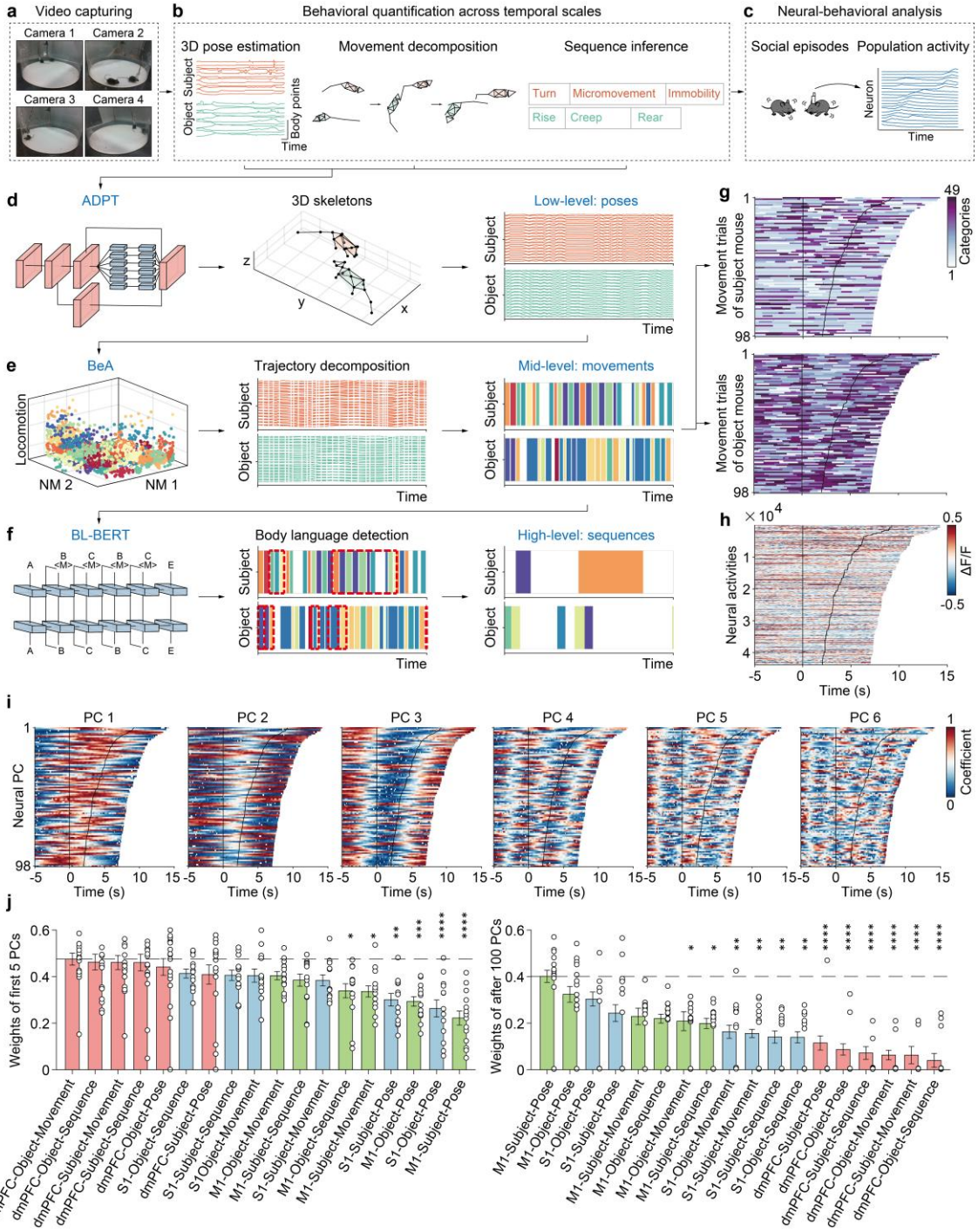


Fig. 2| Distributed cortical dynamics encode behavior across temporal scales.

a, Multi-view video acquisition of freely interacting mice using four synchronized cameras positioned around the behavioral arena.

b, Behavioral quantification across temporal scales. Three-dimensional body poses of the subject and object mice were reconstructed from multi-view recordings and progressively quantified as low-level poses, mid-level movements, and high-level behavioral sequences.

c, Neural-behavioral analysis framework. Social interaction episodes were aligned with simultaneously recorded cortical population activity to examine how neural dynamics encode behavior across temporal scales.

d, Quantification of low-level poses. Three-dimensional skeletal configurations of the subject and object mice were reconstructed using Anti-Drift Pose Tracker (ADPT)[15], providing continuous representations of instantaneous body posture over time.
e, Quantification of mid-level movements. Continuous skeletal trajectories were decomposed using Behavior Atlas (BeA)[14] into discrete movement categories, generating temporally resolved movement sequences for both the subject and object mice.
f, Quantification of high-level behavioral sequences. Body Language-Bidirectional Encoder Representation from Transformers (BL-BERT)[13] identified temporally extended body-language sequences from the movement streams of the interacting mice, capturing behavioral structure over longer timescales.
g, Trial-aligned behavioral phenotypes of the subject and object mice during social interaction episodes. Movement categories are shown over time for each episode, aligned to interaction onset at 0 s. Black curves indicate the temporal progression of interaction episodes across trials.
h, Trial-aligned single-neuron calcium activity during social interaction episodes. Neural activity from S1, dmPFC, and M1 was aligned to the same interaction onset used in **g**. Although behavior exhibited a stereotyped far-close-far structure, single-neuron activity showed heterogeneous and weakly aligned temporal patterns.
i, Trial-aligned cortical principal-component dynamics during social interaction episodes. The first six neural PCs are shown across the same interaction episodes as in **g** and **h**. Low-order PCs displayed reproducible temporal structure aligned to the interaction episode, whereas higher-order PCs showed progressively less stereotyped alignment.
j, Dimension-dependent mapping between cortical dynamics and hierarchical behavior. Linear models were used to predict behavioral variables from neural PCs across cortical areas and behavioral levels. Left, model weights for the first five PCs were enriched for higher-level behavioral features, including movement and sequence variables, with prominent contributions from dmPFC-related features. Right, model weights for PCs after the first 100 dimensions were enriched for low-level pose variables, with stronger contributions from S1- and M1-related features. (One-way ANOVA followed by Dunnett's multiple comparisons test, mean±SEM, N=18, 18, 18, 18, 18, 13, 18, 13, 13,14,14, 13,14,14, 13,14,13, and 14, the P values from left to right are >0.9999, >0.9999, >0.9999, 0.9972, 0.8142, 0.6305, 0.6814, 0.6618, 0.6248, 0.3000, 0.3105, 0.0190, 0.0152, 0.0010, 0.0004, <0.0001, and <0.0001). Right: weights from PCs beyond the 100th (One-way ANOVA followed by Dunnett's multiple comparisons test, mean±SEM, N=14, 14, 13, 13, 14, 14, 14, 14, 13, 13, 13, 13, 18, 18, 18, 18, 18, and 18, the P values from left to right are 0.9158, 0.7339, 0.1593, 0.0847, 0.0583, 0.0365, 0.0236, 0.0049, 0.0033, 0.0014, 0.0013, <0.0001, <0.0001, <0.0001, <0.0001, <0.0001, and <0.0001). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

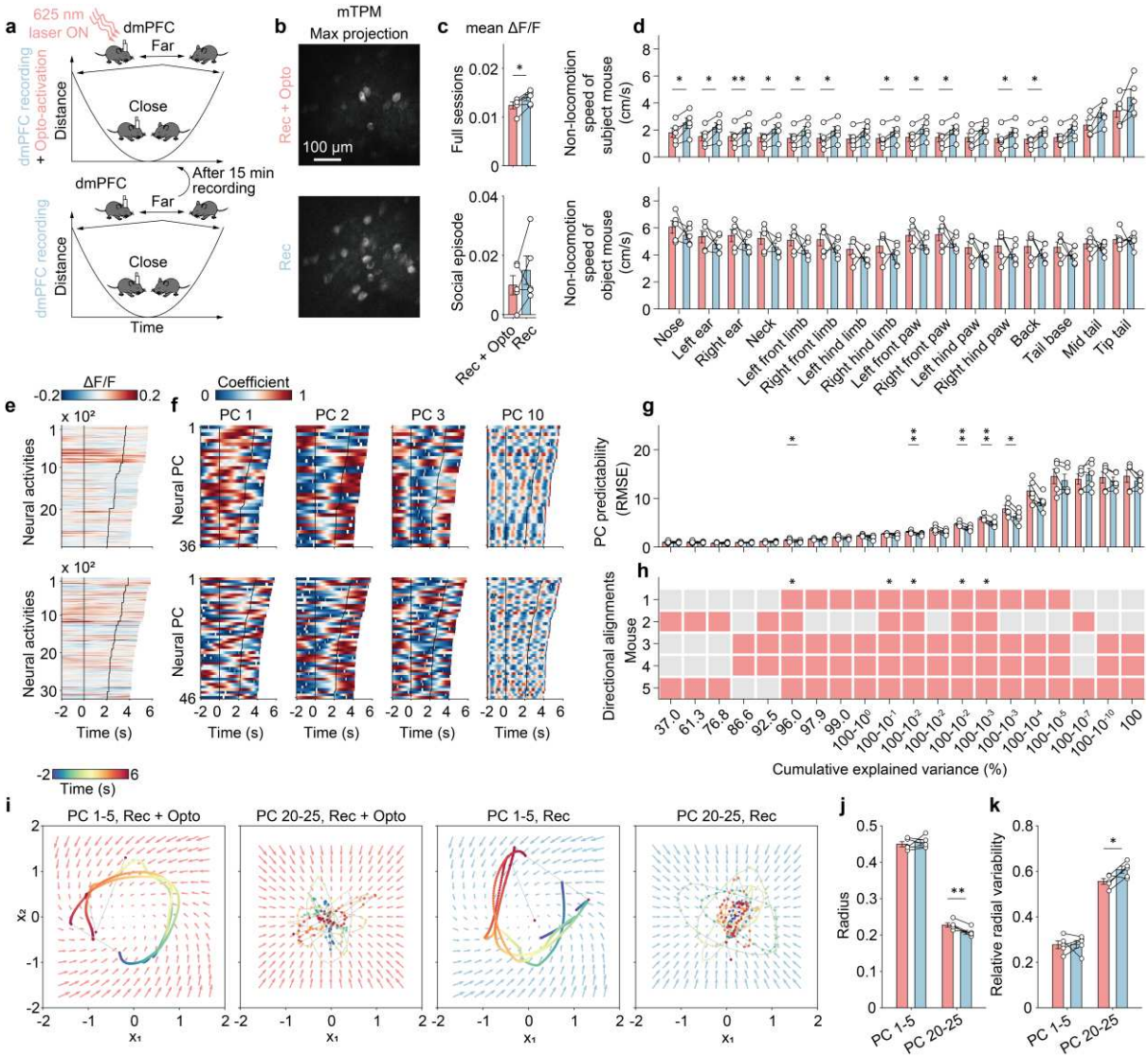


Fig. 3| Optogenetic perturbation of dmPFC disrupts high-dimensional neural dynamics and selectively alters behavioral details.

a, Social interaction episodes were extracted for comparison between dmPFC optogenetic activation (Rec + Opto, upper) and no-stimulation control (Rec, lower). Each mouse underwent 15 min of recording without stimulation, followed by 15 min of continuous 625-nm activation. For both sessions, social interaction episodes were segmented post hoc from the full recordings based on distance.

b, Max-projection mTPM images from dmPFC for the Rec + Opto and Rec conditions, taken from the full behavior sessions. Optogenetic perturbation was associated with an overall reduction in dmPFC neuronal activity.

c, Mean ΔF/F of dmPFC activity for the Rec + Opto and the Rec conditions. Top: mean ΔF/F across the full 15-min recording sessions (Paired t-test, mean±SEM, N=5, the P value is 0.0184). Bottom: mean ΔF/F computed only within the extracted social interaction episodes (Paired t-test, mean±SEM, N=5, the P value is 0.3015).

d, Non-locomotion speed across body parts during the extracted social interaction episodes for the Rec + Opto and Rec conditions. Optogenetic perturbation of dmPFC modulated non-locomotor speeds in specific body parts (paired t-test and sign-rank test, mean±SEM, N=5, the P values of subject non-locomotion speed from left to right are 0.0137, 0.0112, 0.0099, 0.0114, 0.0446, 0.0117, 0.0925, 0.0252,

0.0318, 0.0164, 0.0601, 0.0138, 0.0224, 0.0518, 0.0643, and 0.1334. The P values of object non-locomotion speed from left to right are 0.3914, 0.3857, 0.3797, 0.3715, 0.3005, 0.4303, 0.3987, 0.4523, 0.3100, 0.3568, 0.4278, 0.4401, 0.4813, 0.4009, 0.6369, and 0.8359).

e, Activity of each neuron during the social interaction episodes.

f, Neural PC 1, 2, 3, 10, and 15 extracted from **g**.

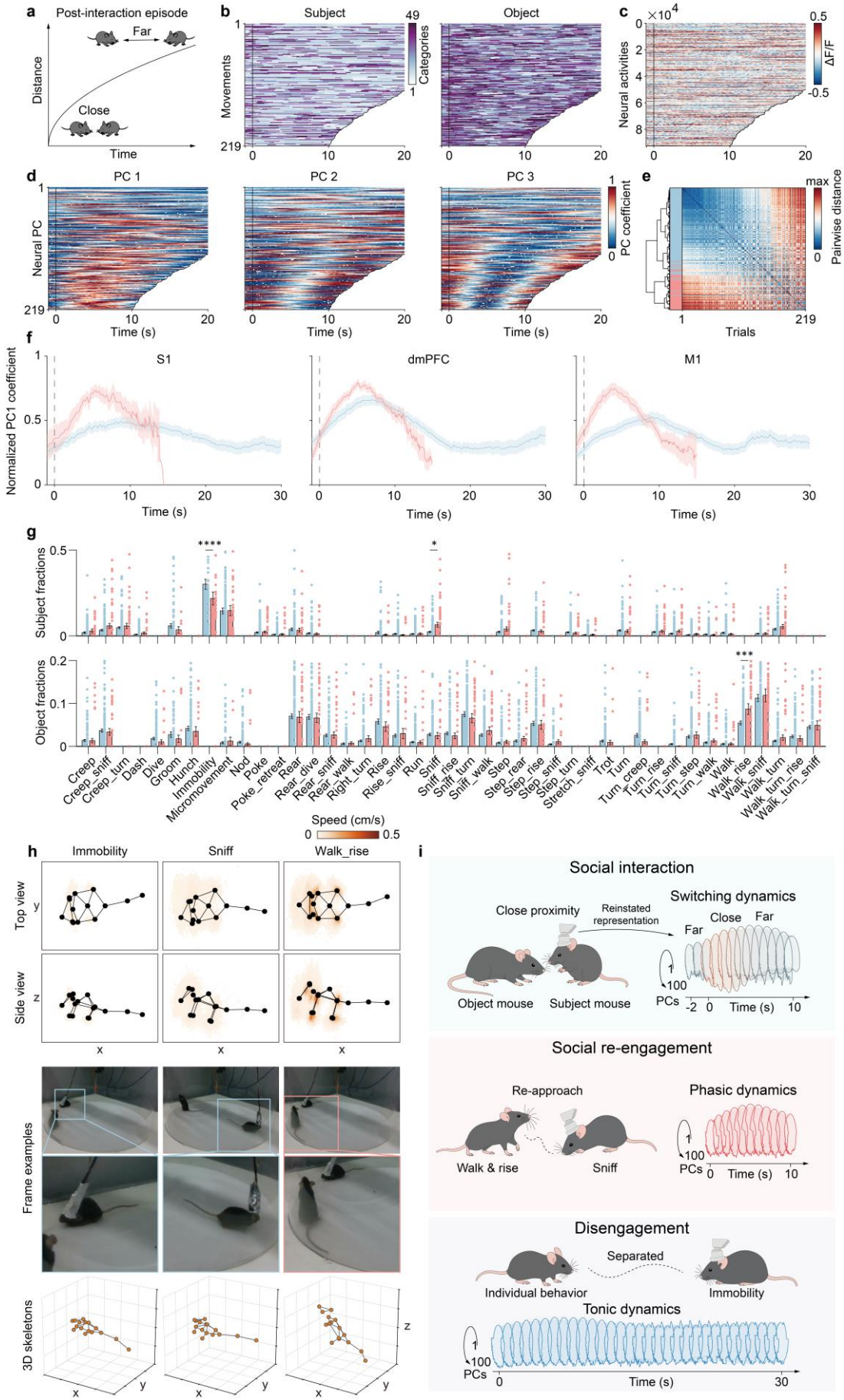
g, RMSE of PC predictability for the optogenetic activation (Rec + Opto) and control (Rec) conditions (paired t-test, mean±SEM, N=5, the P values from left to right are 0.8679, 0.9616, 0.4299, 0.9701, 0.3125, 0.0409, 0.1296, 0.1525, 0.1189, 0.0625, 0.0048, 0.0857, 0.0091, 0.0064, 0.0560, 0.4099, 0.2210, 0.3415, and 0.2567).

h, Directional alignments between non-locomotion speed and PC predictability in each mouse (one-sided exact binomial test, N=5, P values from left to right are 0.8125, 0.8125, 0.8125, 0.8125, 0.5000, 0.0312, 0.1875, 0.1875, 0.1875, 0.0312, 0.0312, 0.1875, 0.0312, 0.0312, 0.1875, 0.1875, 0.5000, 0.8125, 0.5000, and 0.5000).

i, Flow fields reconstructed from low-dimensional (PCs 1-5) and high-dimensional (PCs 20-25) neural PCs, modeled using rSLDS for the Rec + Opto and Rec conditions.

j, Radius of trajectories in **i**, reflecting the spatial extent of neural state trajectories (paired t-test, mean±SEM, N=5, P values are 0.6450 and 0.0064).

k, Relative radial variability of trajectories in **i**, quantifying the heterogeneity of radial components in the flow field (paired t-test, mean±SEM, N=5, P values are 0.9270 and 0.0379). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.


a
Post-interaction episode
Far
Close
Distance
Time
b
Subject
Object
Movements
Categories
49
1
219
c
Neural activities
ΔF/F
0.5
-0.5
d
PC 1
PC 2
PC 3
Neural PC
PC coefficient
Time (s)
e
Pairwise distance
max
Trials
f
S1
dmPFC
M1
Normalized PC1 coefficient
Time (s)
g
Subject fractions
Object fractions
Creep
Creep_sniff
Creep_turn
Dash
Dive
Groom
Hunch
Immobility
Micromovement
Nod
Poke
Poke_retreat
Rear
Rear_dive
Rear_sniff
Rear_walk
Right_turn
Rise
Rise_sniff
Run
Sniff
Sniff_rise
Sniff_turn
Sniff_walk
Step
Step_rear
Step_rise
Step_sniff
Step_turn
Stretch_sniff
Trot
Turn
Turn_creep
Turn_rise
Turn_sniff
Turn_step
Turn_walk
Walk
Walk_rise
Walk_sniff
Walk_turn
Walk_turn_rise
Walk_turn_sniff
Speed (cm/s)
h
Immobility
Sniff
Walk_rise
Top view
Side view
Frame examples
3D skeletons
i
Social interaction
Close proximity
Reinstated representation
Switching dynamics
Close
Far
Object mouse
Subject mouse
PCs
Time (s)
Social re-engagement
Re-approach
Phasic dynamics
Walk & rise
Sniff
Disengagement
Separated
Individual behavior
Immobility
Tonic dynamics

977

Fig. 4| Distinct cortical dynamics accompany transitions across behavioral contexts.
a, Post-interaction behavioral episodes. Each episode began with a close-interaction baseline and ended at the onset of the next close social interaction. This segmentation captured the interval after one social encounter and before the next social engagement.
b, Trial-aligned movement sequences of the subject and object mice during post-interaction episodes. Each row represents one episode, sorted by duration. The initial close-interaction period was retained as a baseline, and the subsequent interval was extended until the next close interaction.
c, Trial-aligned single-neuron calcium activity during the same post-interaction episodes. Neurons from S1, dmPFC, and M1 were pooled and aligned to the post-interaction baseline. Single-neuron activity showed heterogeneous temporal patterns across episodes.
d, Trial-aligned neural principal-component trajectories during post-interaction episodes. The first three neural PCs were extracted from the population activity shown in c. PC trajectories displayed structured temporal patterns that differed across post-interaction episodes.
e, Pairwise distance matrix of trial-level PC trajectories. Hierarchical clustering of trajectory distances separated post-interaction episodes into two neural trajectory groups, indicating that cortical population dynamics predicted distinct post-interaction states.
f, Normalized PC1 coefficients during post-interaction periods in S1, dmPFC, and M1. Trials were separated according to behavioral outcome, showing distinct temporal evolution of cortical dynamics across post-interaction trajectories.
g, Movement fractions of the subject and object mice in the two PC trajectory-defined groups. Blue and red bars with points indicate individual post-interaction episodes from the two clustered trajectory groups. The two trajectory groups differed in specific movement categories, distinguishing episodes associated with rapid social re-engagement from those associated with disengagement. (Two-way ANOVA followed by Sidak's multiple comparisons test, adjusted P values of subject fractions (mean±SEM, N=153) are >0.9999, 0.7408, >0.9999, >0.9999, >0.9999, 0.8489, >0.9999, <0.0001, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.0173, >0.9999, >0.9999, >0.9999, 0.9990, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.9998, 0.9999, >0.9999. Adjusted P values of object fractions (mean±SEM, N=66) are >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.9995, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.9796, >0.9999, >0.9999, >0.9999, >0.9999, 0.9999, 0.9990, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.7783, >0.9999, >0.9999, >0.9999, >0.9999, >0.9999, 0.0002, >0.9999, >0.9999, >0.9999, >0.9999).
h, Representative behavioral examples associated with distinct post-interaction outcomes. Top and side views of 3D body poses, example video frames, and reconstructed three-dimensional skeletons are shown for immobility, sniffing, and walk-rise behaviors.
i, Summary of distributed cortical population dynamics associated with distinct behavioral contexts. Close social interaction was characterized by switching dynamics accompanying transitions between far and close proximity. Social re-engagement was associated with phasic dynamics during re-approach, whereas disengagement was accompanied by sustained tonic dynamics as the animals separated and returned to individual behavior. *P < 0.05, ***P < 0.001, ****P < 0.0001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DistributedcorticaldynamicsorganizebehavioracrosstimescalesSupplementaryinformationnewV1.pdf