A neural manifold view of the brain

Matthew G. Perich [1,2] ✉, Devika Narain [3] & Juan A. Gallego [4] ✉

Animal behavior arises from the coordinated activity of neural populations that span the entire brain. The activity of large neural populations from an increasing number of brain regions, behaviors and species shows low-dimensional structure. We posit that this structure arises as a result of neural manifolds. Neural manifolds are mathematical descriptions of a meaningful biological entity: the possible collective states of a population of neurons given the constraints, both intrinsic (for example, connectivity) and extrinsic (for example, behavior), to the neural circuit. Here, we explore the link between neural manifolds and behavior, and discuss the insights that the neural manifold framework can provide into brain function. To conclude, we explore existing conceptual gaps in this framework and discuss their implications when building an integrative view of brain function. We thus position neural manifolds as a crucial framework with which to describe how the brain generates behavior.

Neuroscientists seek to understand how the many interactive scales of brain architecture—from molecules to synapses and neurons—give rise to animal cognition and behavior. Since the days of Ramón y Cajal[1], the neuron has been posited as the fundamental computational unit of the brain. This 'neuron doctrine' became the basis for subsequent approaches in which the activity of individual neurons was typically correlated to task variables that the experimentalist deemed relevant[2-5], enabling new insights into the key functions and specializations of different brain regions[6]. Behavior, however, ultimately emerges from the complex coordination of neural populations[7-9]. Perhaps this is the reason why fundamental relationships between the brain and behavior have remained elusive from the perspective of individual neurons[3,10-12], and why it has been challenging to uncover neuroscientific 'first principles' that describe the neural basis of behavior spanning individuals and species.

With the advent of modern experimental techniques that allow the activity of many neurons to be recorded simultaneously[13-16], the focus in neuroscience has shifted to studying the coordinated activity of these neural populations[7,9,17]. Studying the brain at this population level allows us to incorporate the interactions among neurons that ultimately drive behavior[7,9]. Empirical results across many species have consistently demonstrated that neural population activity can be captured by a relatively small number of possible population-wide activity patterns[7,9,18]—a seemingly surprising finding

given the striking heterogeneity of single-neuron activity[10-12,19,20]. Here, we propose that these empirical observations reflect the existence of lower-dimensional latent dynamics that capture the activity of the brain region(s) under investigation. These latent dynamics arise from constraints imposed by a neural manifold underlying the activity of that population of neurons. Accordingly, we view these neural manifolds not as mere approximations of neural data, but as entities with meaningful and unexplored functions in the generation of behavior.

In this article, we review the central concepts and key advances of neural manifolds and discuss potential underlying causes for their emergence. We propose that, despite the myriad open questions in the field, neural manifolds will allow us to build more cohesive theories across different disciplines and toward a more unified theory of brain function.

What is a neural manifold?

Formally, a manifold is defined as a locally Euclidean topological space[21] (Box 1). Many types of continuous geometric object, such as a plane, a toroid or a sphere, fall within this definition. For example, the surface of a ball or a coffee mug can be thought of as a two-dimensional manifold embedded in the three-dimensional space of the physical world of your office. When analyzing data such as neural population recordings, neural manifolds are often conceptualized as lower-dimensional objects within a higher-dimensional ambient space representing the activity

[1]Département de neurosciences, Faculté de médecine, Université de Montréal, Montreal, Quebec, Canada. [2]Mila - Quebec Artificial Intelligence Institute, Montreal, Quebec, Canada. [3]Department of Neuroscience, Erasmus Medical Center, Rotterdam, the Netherlands. [4]Department of Bioengineering, Imperial College London, London, UK. ✉e-mail: matthew.perich@umontreal.ca; jgallego@imperial.ac.uk

BOX 1

Definition of manifolds

In the existing neuroscience literature, the term 'manifold' has taken on several meanings, including:

(i) the output of dimensionality reduction or manifold learning methods with the intent of visualizing or exploring data, or for engineering applications such as BCIs[53,152]

(ii) the geometric organization of neural population activity, presumably along a surface in a neural state space

(iii) the low-dimensional structure that captures the time-varying dynamics of neural population activity during a specific behavior

Mathematically, a manifold is defined as a locally Euclidean topological space. For instance, while we know the Earth to be spherical at a global scale, it can be defined as a manifold because, at local scales (such as our view when standing on its surface), our perception of it is approximately flat. Manifolds can thus be described by their geodesics, or the shortest line between any two points along its surface, even if curved. The equator could then be considered a ring-like structure spanning the geodesic of this spherical surface. Following such logic, manifolds encompass many types of object, such as rings, squares or tori. Here, we limit ourselves to the more general assumptions that manifolds[153]: (i) locally follow approximately Euclidean topology; (ii) follow homeomorphism properties[154]; and (iii) have a finite positive integer dimensionality. There are several categories of manifolds with interesting additional assumptions, such as Riemannian, symplectic or smooth manifolds. Smooth manifolds, notably, can be described by a continuous set of points, and these structures are differentiable. Theoretical work often pertains to idealized smooth manifolds, whereas empirically determined manifolds tend to be discrete approximations of such smooth manifolds. To the best of our knowledge, no study has specifically addressed whether the manifolds estimated from neural population recordings can rigorously satisfy the theoretical, mathematical definition of a

manifold. However, we would venture that many of the intuitions from mathematics—in particular, Euclidean topology, finite integer dimensionality and smoothness—should guide how we think about neural manifolds.

In this Review, we adopt the following specific definition of a neural manifold: a low-dimensional surface that captures the possible collective states that can be traversed by the dynamics of a neural population considering its biophysical constraints. Thus, we argue that the neural manifold provides meaningful constraints on the neural activity in the brain. Although neural manifolds are often estimated from neural population recordings using dimensionality reduction or manifold learning methods, as specified in definition (i) above, it is important to note that merely applying these methods does not necessarily imply the existence of a neural manifold as a meaningful constraint on the neural population activity. Indeed, dimensionality reduction is often used to great effect for data visualization and exploration. In neuroscience, it has also been very powerful for translational applications, such as building 'decoders' that extract command signals from neural population recordings in the context of BCIs[53,152]. Here, we propose that neural manifolds encompass both the geometric and dynamic features of definitions (ii) and (iii) because both aspects are tightly linked (see 'Neural manifolds and neural dynamics: the chicken or the egg?'). Although we believe there are relatively static geometric constraints on neural activity, behavior is ultimately caused by the dynamics of neural activity[9]. We can thus think of these two approaches as complementary in the sense that both geometry and dynamics can be informative about neural function.

Topology versus geometry. Neural manifolds can be characterized through numerous properties related to their topology and geometry—including curvature or nonlinearities of the surface, intrinsic dimensionality—as well as dynamical characteristics of the trajectories along their surface—including stability, flow
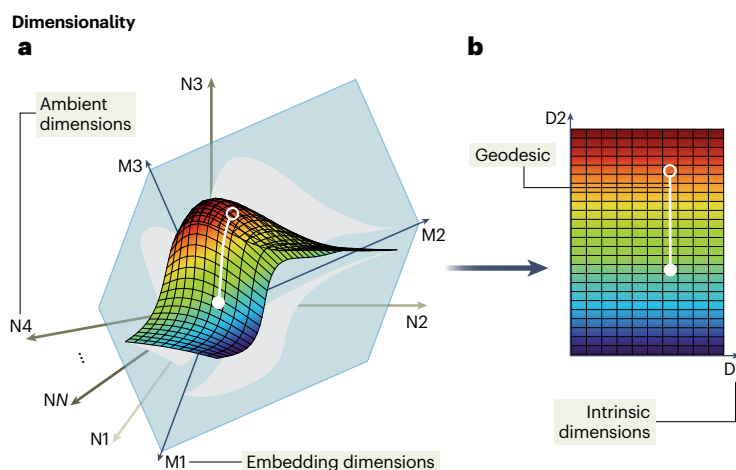
Box Fig. 1 | a,b, Manifold dimensionality. **a,** Example manifold object embedded in an *N*-dimensional ambient space. This manifold (colored to enable mapping between the two coordinate frames) has an embedding dimensionality (*M*) of three, and an intrinsic dimensionality (*D*) of two. Each manifold has its own geodesic, or a set of the shortest lines between any two points along its surface, as shown on this representation unfolded along the intrinsic dimensions (white line). **b,** The same manifold in its intrinsic coordinates (dimensionality of two).

(continued from previous page)

fields and attractors—among other features. Topology and geometry are closely related branches of mathematics. In the context of neural manifolds, topology encapsulates the intrinsic properties of the manifold that are preserved under continuous transformations, whereas geometry studies their properties using metrics such as distance, size and shape. Intuitively, topology studies global properties and geometry studies local properties. A clear distinction is that two geometrically different objects are topologically equivalent if they can be transformed into one another through continuous deformations, without tearing or glueing. Consequently, a torus—think of a hollow donut—and a coffee mug are topologically equivalent (homeomorphic) despite their vastly different geometric properties.

Dimensionality. This concept can take several meanings (Box 1 Figure):

(i) The intrinsic dimensionality is the number of independent variables required to fully parameterize the structure of the manifold[155].

(ii) The minimum number of orthogonal state space directions needed to fully subsume the manifold, that is, the space

in which the manifold lives. This concept has been referred to as the embedding dimensionality when studying low-rank systems or manifolds with low intrinsic dimensionality[113,155].

(iii) The ambient dimensionality is the number of axes available to estimate the neural manifold. In neuroscience experiments this will typically be the number of recorded neurons (Box 2).

Note that this taxonomy makes no explicit assumption about the definitions of 'low' and 'high' dimensional. These notions, although common in the neuroscience literature, including in this article, must be defined relative to features such as the number of recorded neurons or the dimensionality of behavior. In practice, we expect the intrinsic dimensionality of the underlying neural manifold to be much less than the number of recorded neurons owing to numerous constraints such as common inputs, recurrent connectivity and the animal's behavior. Note that if this intrinsically low-dimensional manifold is highly nonlinear, you may still need a very large number of independent linear embedding directions (for example, dimensions in principal component analysis (PCA)), even close to the limit of recorded neurons, to fully capture all of the intricacies of the manifold[101,156,157].
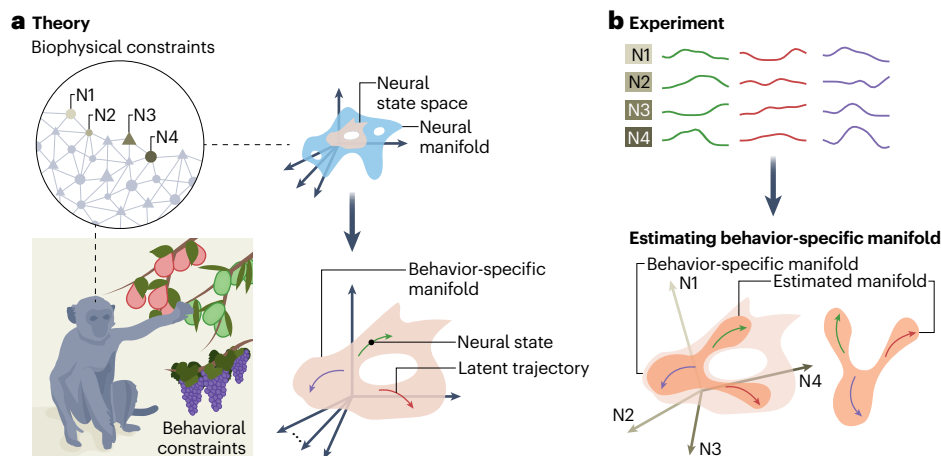

Fig. 1 | Neural manifolds underlying the generation of behavior by the brain.
a, Theory: a broad variety of biophysical properties and processes (for example, connectivity) constrain the activity of a neural population, resulting in the emergence of a neural manifold that captures its possible collective states (top right, shown in blue). When an animal engages in a particular behavior (for example, reaching to three different fruits in different locations), neural population activity lies on a behavior-specific portion of this manifold (bottom, shown in pink). Color-coded arrows on the neural manifold indicate the latent trajectories produced when the animal reaches to each of the three fruits. **b**, Experiment: recording the activity of a subset of neurons in a population (four in this example, N1–N4) allows estimation of the portion of the full neural manifold (pink) explored by the neural data (shown in orange).

of the neural population. Much work has been done in mathematics to identify and characterize the topology and geometry of manifolds, with applications in many fields[22].

In neuroscience, empirical evidence from across many brain regions, species and behaviors (see review sections below) shows that the latent dynamics of neural populations consistently explore only a subset of possible states. This led to the hypothesis that these activity patterns may reflect the constraints of an underlying neural manifold[7,23]. Empirically, one can estimate these neural manifolds within a neural state space in which each axis captures the firing rate of a recorded neuron. The activity at each point in time of the neural population becomes a point in this state space, a neural state. Over time, such as when an animal produces a behavior, this neural state evolves with a specific trajectory representing the latent dynamics of the neural population (color-coded trajectories in Fig. 1a). The neural manifold framework thus posits that

these latent trajectories during a behavior are confined to the surface of an intrinsically low-dimensional object—the neural manifold. That is, it is assumed that the neural manifold has relatively low intrinsic dimensionality (with respect to the total number of neurons in the population of interest), meaning that it can be defined with few independent variables (Box 1 explores interpretations of dimensionality). We can estimate the neural manifold using a variety of mathematical techniques[24] (Fig. 1b and Box 2) from even a subset of neurons in this population[7,17].

Neural manifolds are likely to emerge from the coordinated activity of interacting neurons within (and probably across) brain regions (Fig. 1a). As such, the neural manifold framework makes explicit the notion that the activity of a given neuron is a function of the other neurons that provide its input. These neuron-to-neuron interactions are by nature highly complex, as they are influenced by a wide variety of biophysical processes in the brain, including direct synapses and

BOX 2

Estimating neural manifolds

Neural state spaces. Neural manifolds are estimated within a neural state space in which each orthogonal axis represents the firing rate of a single recorded neuron (or independent multi-unit[158]) in the dataset (Box 2 Fig. part a). At each time step in the dataset, the neural population activity can be represented as a point in this state space. Over some span of time, the collection of these points traces a latent trajectory through the state space. In this framework, the activity of each neuron in a given population can be thought of as a (random) projection of this latent trajectory onto one of the axes that defines the neural state space[53,113,159]. The neural manifold framework proposes that the collection of latent trajectories over all time points in the dataset will explore a constrained subset of possible neural states, resulting in intrinsically lower-dimensional structure, which we posit is confined to the surface of a neural manifold.

Dimensionality reduction methods. The coordinated activity of neural populations can be well captured by relatively few population activity patterns owing to correlations and redundancy in the firing rates of the constituent neurons. Tools that enable the identification of such low-dimensional activity patterns are known as dimensionality reduction methods. A subset of these methods describes the original data as a linear combination of latent signals. These latent signals, which capture patterns of covariation across neurons (that is, directions in neural state space), thus define a hyperplane—a flat manifold that underlies the neural activity. Such linear dimensionality reduction methods include PCA or factor analysis, as well as targeted dimensionality reduction approaches that use supervised labeling to identify specific linear directions, which covary with experimentally observable phenomena[160,161]. In cases where the neural manifold is nonlinear, these linear methods will at best provide planar approximations to the true underlying manifold, resulting in higher estimated dimensionality than the true number of intrinsic dimensions[85]. In this case, nonlinear dimensionality reduction methods, including LFADS[162], TNDM[163] and MARBLE[164], can estimate the dynamics underlying the population activity. Further, nonlinear manifold learning methods (discussed below) can provide even better estimations of the underlying population structure by directly incorporating assumptions about the manifold (for example, continuity or geometry).
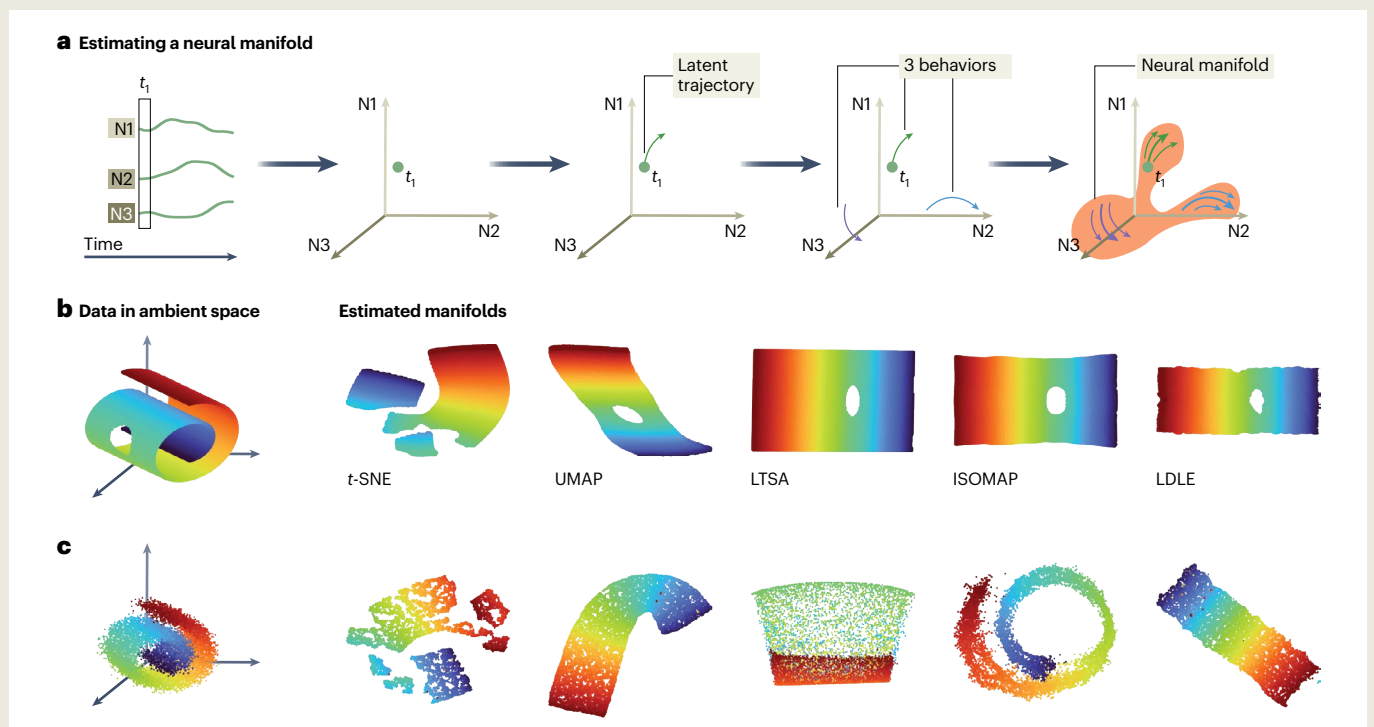
Box Fig. 2 | a–c, Simulations to demonstrate manifold learning methods. **a,** Illustration of a procedure to estimate a neural manifold by computing latent states from neural population firing rates. The firing rates of a population of recorded neurons (first diagram beginning from the left) are represented in a neural state space in which each axis represents the activity of a recorded neuron (second diagram). In this space, the simultaneous activity of all recorded neurons at a certain time point, for example, t_1, is represented as a point. The activity of a recorded population during a certain behavior describes a latent trajectory through this neural state space (third and fourth diagrams show the average trajectories for one and three conditions, respectively). Thus, all recorded behaviors—three in this example—describe a series of latent trajectories (fifth diagram shows average (thick lines) and single trial (thin lines) trajectories). We hypothesize that these latent trajectories arise from an underlying neural manifold (fifth diagram). **b,** Simulations of a manifold, a Swiss roll with a circular hole embedded along its intrinsic dimension (two) using several techniques: t-SNE, LTSA, UMAP, Isomap and LDLE. **c,** Simulations of a similar manifold, a Swiss roll with increased noise structure, embedded by the same techniques highlighted in **b**. The increased noise disrupts the identified embeddings of several methods.

(continued from previous page)

Nonlinear dimensionality reduction using manifold learning methods. As the complexity of behavioral and neural data increases, the underlying neural manifolds we uncover may become similarly complex[113] and probably nonlinear[85]. Manifold learning methods can allow us to uncover nonlinear structure in the data that cannot be effectively captured by linear dimensionality reduction techniques.

For instance, Isomap (isometric feature mapping)[165] constructs a low-dimensional embedding (that is, manifold estimate) that maintains the overall intrinsic geometry of the original ambient space by preserving the geodesic distances among neighboring points in the embedding. This is a widely used and versatile method but, like many manifold learning methods, relies on proper hyperparameter selection and can be vulnerable to noise[166]. Several other manifold learning methods also aim to preserve local information as they embed data from higher-dimensional ambient spaces into lower dimensions, including local linear embedding[167], diffusion maps[168] and Laplacian eigenmaps[169]. Some methods adopt a top-down approach by first assembling a globally optimized embedding that is then refined (for example, *t*-stochastic neighborhood embedding (*t*-SNE)[170], CEBRA[69], and uniform manifold approximation and projection (UMAP)[171]). Other methods adopt a bottom-up approach by finding global embeddings that optimize and preserve local structure by optimizing local views to construct a global embedding piecewise, for example, local tangent space alignment (LTSA)[172], low-distortion local eigenmaps (LDLE)[173] and Riemannian alignment of tangent spaces

(RATS)[174]. Notably, RATS minimizes recently developed distortion metrics—which quantify how much the distances between the newly embedded points deviate from the distances in the original ambient space—at both local and global views to generate an overall low-distortion embedding.

As each manifold estimation method has its own assumptions and advantages, the choice of techniques can greatly influence the outcomes of the embedding (see examples in Box 2 Fig. parts b,c). For example, the flat manifolds estimated with linear methods such as PCA may miss important topological characteristics if the researcher does not choose enough dimensions, because behaviorally relevant features may be captured in the high-variance components. Alternatively, nonlinear methods such as Isomap may give misleading embeddings if the observed neural states do not explore a sufficiently large portion of the manifold (see example in Fig. 1) or in the presence of noise (Box 2 Fig. part c). This is an issue because it is very hard to predict a priori the topology or geometry of the manifold (although see refs. 50,55,84 for a few exceptions in which normative models were used to predict the specific topology of a manifold). In other words, we can observe manifolds only through the activity of neurons that we can record. Each of these neurons provides only one partial projection of the true underlying manifold[113]. Therefore, sampling sufficient neurons and neural states (for example, through longer recordings or more behavioral conditions) and choosing an appropriate manifold learning method is critical to enable the full reconstruction of the neural manifold.

multisynaptic connections, glial cells, widely distributed neuromodulators and underlying synaptic and membrane changes. In addition to these biophysical constraints, the behavior an animal is engaged in determines which portion of the manifold is explored by the neural population activity. Ultimately, we conceptualize neural manifolds as hypothetical constructs capturing a meaningful biological entity: the possible collective states of a population of neurons given its biophysical constraints, which can be estimated by applying dimensionality reduction or manifold learning methods (Box 2) to neural recordings.

Behaviorally relevant organization of neural population latent dynamics

Low-dimensional latent dynamics underlying neural population activity have been described in many species—from invertebrates[25–28] to fish[29–31] and mammals[10,12,32–43], and even humans[44–47]—in all corners of cortex—from sensory[48,49] to cognitive and associative[10,19,39,40,46,50–52] to motor[32,36,53,54] regions—and in a variety of subcortical structures[37,55–67]. The analysis of these latent dynamics has enabled substantial conceptual advances toward describing the neural implementation of processes that give rise to behavior.

Many studies have identified behaviorally relevant organization within the full neural state space. This organizational structure persists over long time periods in monkeys performing the same task when the recorded neurons change[53] and even across different individuals[38,59,68,69]. Moreover, neural population activity is organized such that specific directions within the neural state space capture latent dynamics related to particular cognitive or motor functions, for example, integrating sensory evidence for decision-making[10], learning new behavioral output[70,71], acting impulsively[72] and engaging with a task[42]. Some of these directions may be private to individual brain regions—that is, they are 'output-null' with respect to activity in other regions[73]—an observation that has shed light onto unclear phenomena such as accumulating evidence before a decision[39], learning[70,71,74] and preparing a movement without causing

it[75,76]. Likewise, the organization of latent dynamics also seems to capture interactions across brain regions, accounting for information flow across various visual[77,78] and sensorimotor[48,71,75,79] regions. Crucially, these directions enabling interregional communication are orthogonal to directions that putatively capture local processes, which could allow a single brain region to flexibly perform distinct functions simultaneously[35,39,48,80].

Evidence for neural manifolds throughout the brain

The existence of low-dimensional latent dynamics with behaviorally relevant organization is consistent with the notion that neural population activity is subject to constraints imposed by an underlying neural manifold. As such, any possible behavioral state is associated with a specific state along the manifold[81]. This leads to the prediction that changes in behavioral output that require neural states within the existing manifold should be easier and more intuitive to achieve than those that require modification of the manifold itself. In a foundational body of work, researchers tested this in a learning paradigm based on a brain–computer interface (BCI) that mapped latent trajectories—hypothesized to emerge from the neural manifold—in monkey motor cortex to the velocity of a computer cursor[23]. They demonstrated that 'within-manifold' perturbations that preserved the coordination across neurons were easier to learn than 'outside-manifold' perturbations that required the monkeys to generate new patterns of activity, a process requiring extensive training[82]. Although this establishes potential constraints on the behaviorally relevant latent dynamics, it does not explicitly characterize the properties (for example, geometry or topology; Box 1) of the underlying neural manifold.

Recent work directly identified neural manifolds underlying neural population activity: (i) manifold learning techniques applied to recordings from the nervous system of *Caenorhabditis elegans* uncovered a clear geometric structure underlying the states of all neurons in the worm nervous system during locomotion[83] (Fig. 2a). Crucially, the
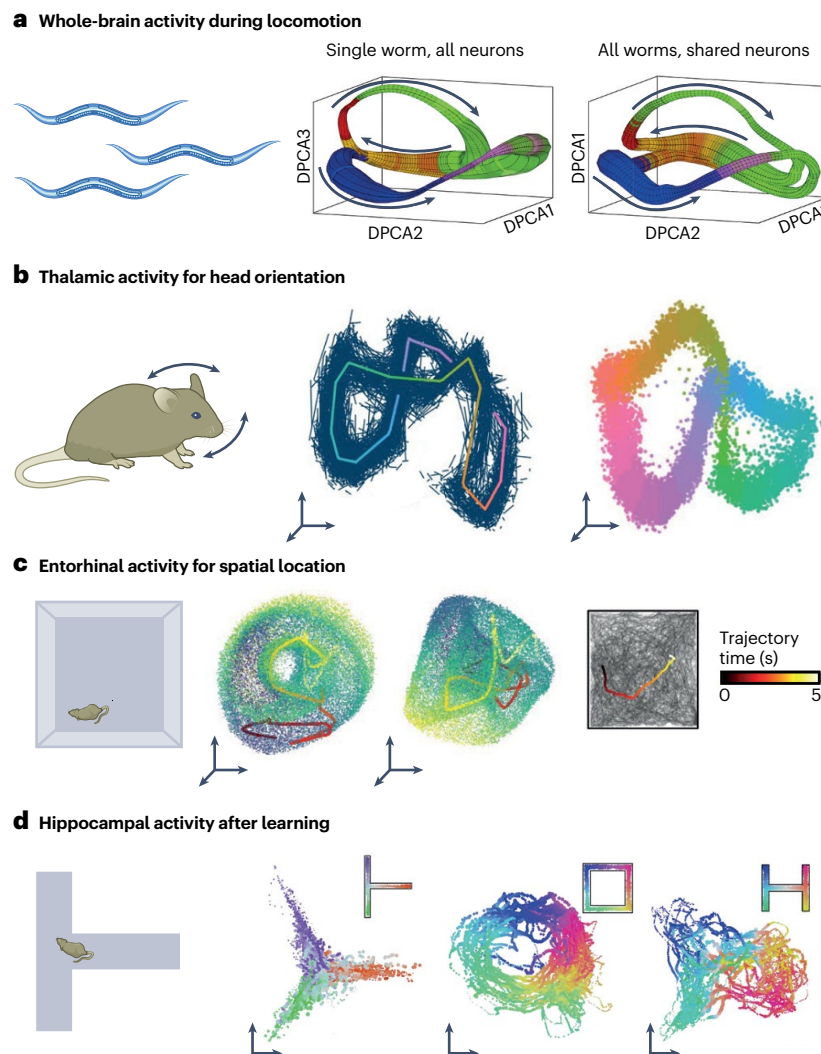
a Whole-brain activity during locomotion

Single worm, all neurons

All worms, shared neurons

b Thalamic activity for head orientation

c Entorhinal activity for spatial location

Trajectory time (s)
0 — 5

d Hippocampal activity after learning

Fig. 2 | Examples of neural manifolds estimated from behaving animals.
a, Consistent manifolds underlying all neurons recorded by Ca²⁺ fluorescence in different *C. elegans* individuals during locomotion. Center: a manifold estimated from 107 neurons of a single worm. DPCA1–DPCA3 are the first three principal components spanning the neural activity averaged with respect to key behavioral parameters. Right: a similar manifold estimated from 15 genetically specified neurons common to all recorded worms. **b**, A ring attractor manifold underlies head-direction cells in the thalamus. Center: visualization of neural states in a single animal during 30 min of wakeful movement in a three-dimensional embedding from Isomap. Each blue line represents two adjacent neural states. Color indicates the head-direction angle associated with states on the ring attractor. The overlaid line on the manifold represents a spline fit to the states colored according to head direction. Right: the same manifold as in the center with all neural states colored by head-direction angle. **c**, A toroidal manifold underlies grid cell activity in medial entorhinal cortex. Center: Two views of a

three-dimensional UMAP embedding of the activity of 149 grid cells recorded from a single mouse during free exploration of a square environment. Each point is a neural state colored by its projection on the first principal component identified in the population activity. Right: A top-down view of the environment explored by the mouse. Gray lines show the path taken by the animal. The colored lines in both panels show a 5-s trajectory during exploration. **d**, Manifold structure matches the geometry of learned environments in CA1 neurons. Each subpanel shows a two-dimensional manifold identified by Isomap from Ca²⁺ recordings of hippocampal CA1 during exploration of the associated environment shape (top right). Neural states are colored to illustrate the correspondence between neural states on the manifold and position in the environment. Part **a** reproduced from ref. 83, CC BY 4.0; part **b** reproduced from ref. 55, Springer Nature Limited; part **c** reproduced from ref. 50, CC BY 4.0; part **d** reproduced with permission from ref. 86, Elsevier.

geometry of these neural manifolds remained highly preserved across different individuals, and so was the mapping between location within the manifold and locomotor behavior; (ii) consistent neural manifold topology has been observed in head-direction cells in the thalamus of mice[55,84]: neural population dynamics could be described by an irregular ring attractor manifold (meaning that neural states are constrained to a one-dimensional ring), and this geometric organization persisted between waking and sleep states (Fig. 2b); (iii) populations of grid cells in the medial entorhinal cortex showed a preserved toroidal topology during navigation across different environments[50,84] that also persisted during sleep (Fig. 2c); (iv) evidence accumulation was represented along a continuous nonlinear manifold in the mouse hippocampus

that also represented space[68]; (v) aggressive states were linked to a line attractor manifold in the hypothalamus[62,65]; and (vi) nonlinear manifolds were shown to best capture movement-related activity across various cortical and subcortical motor regions from mice, monkeys and humans[85]. Together, these studies provide evidence that neural manifolds underlie population activity during behavior.

The properties of neural manifolds can be intimately related to features of behavioral function and the task being performed. At least in some brain regions, neural manifolds are shaped by the behavioral function of the recorded neural population. Indeed, the ring attractor manifolds in mouse thalamus[55] and zebrafish anterior hindbrain[31] allow accurate tracking of head direction, and the toroidal manifold structure

found in the medial entorhinal cortex of mice enables the construction of an accurate representation of space[50]. Hippocampal 'maps' of space also exploit manifold properties for navigation: as mice familiarize themselves with a novel enclosure, activity in their hippocampal CA1 manifolds increasingly reflects the enclosure's shape[86] (Fig. 2d).

Just as function can serve to shape manifold properties, manifold properties can play an important role in shaping behavioral function. In the medial frontal cortex, the organization of trajectories along a presumed manifold matches the internal estimates of timing in monkeys[87]. Furthermore, the task-related latent dynamics exhibit a curved geometry—presumably resulting from a nonlinear manifold—such that a linear readout of this curvature explains the response patterns of the monkeys when estimating temporal intervals[41]. Similarly, curved representational geometry in the intraparietal cortex could facilitate optimal decision-making in different tasks[88]. However, nonlinearity can also be a disadvantage, leading to suboptimal performance during value estimation[52]. These studies show links between manifold geometry, the function of a neural population and, ultimately, animal behavior.

Factors shaping neural manifolds

In our view, neural manifolds emerge from constraints on the possible activity patterns of the neural population that arise from sources both intrinsic (for example, connectivity, neuromodulators) and extrinsic (for example, behavior, task) to the physiological circuit. Here, we discuss how these factors shape the emergence of neural manifolds in the brain (Fig. 3).

Low-level biological properties

Evidence from several studies demonstrates that low-dimensional structure in neural population activity relates to intrinsic constraints in the network[89]. First, dominant population structure is conserved between spontaneous and stimulus-evoked activity in visual cortex in anesthetized cats[90] and awake mice[91], and the topology of neural manifolds underlying head-direction cells in the mouse thalamus and medial entorhinal cortex grid cells persists between behavior and sleep[50,55,84]. Second, holographic optogenetic activation of a sufficiently large number of stimulus-discriminating neurons in the mouse visual cortex recapitulates the latent trajectories described by a different population of neurons from the same region during stimulus-evoked activity[92]. Third, manifold topology—but not the individual activation patterns of shared neurons—is conserved between genetically identical *C. elegans* individuals[83]. Last, these constraints on population activity seem to similarly constrain the animal's behavior, such as their ability to learn in a BCI paradigm[23,82]. The theoretical links between neural manifolds and circuit constraints have also been explored through neural network simulations (Box 3).

Neural manifolds should be shaped by aspects of circuit organization (for example, synaptic connections, cell types, effective synaptic strength)[93]. Genetically specified cell types can have distinct firing properties, inputs and projection targets[94]. Further, they may have different roles in the production of behavior, as shown during motor control[95,96] and decision-making[97]—although genetically specified cell types do not always map onto functionally separated classes[67,98]. At a higher level, connectivity in the circuit (for example, connectivity motifs, laminar organization) should similarly shape and constrain population activity. Indeed, different brain regions—which are known to have distinct connectivity profiles, cellular composition and inputs—express different degrees of nonlinearity in the estimated manifold[85], distinct structure in latent trajectories[54,99,100] and varying dimensionalities in population activity[71,93,101–103].

Many other phenomena may provide other interesting avenues to explore the emergence of neural manifolds. These could include subcellular (for example, receptor expression or vesicular transport)[104,105] or dendritic[106] processes, precise spike timing events[107,108], neuromodulators released widely across the brain, non-neuronal activity such as glia[109,110] and even physiological signals such as heart
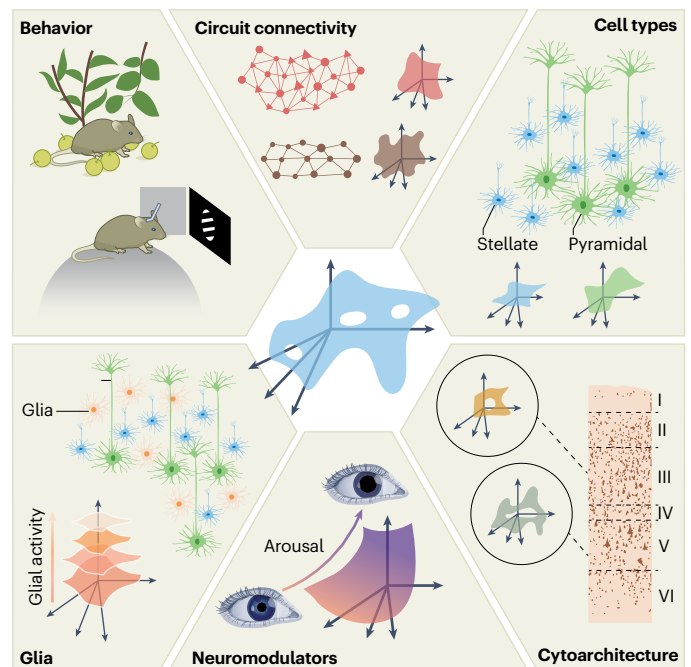

Fig. 3 | Factors shaping the emergence of neural manifolds. Neural manifolds are shaped by both extrinsic factors, such as the behavior the animal is engaged in, and numerous intrinsic factors, including circuit connectivity, constituent cell types, circuit cytoarchitecture, neuromodulator release associated with behavioral state changes such as arousal, and non-neural processes such as glial cell activity, which could change neural manifold properties as their activity modulates.

rate[111]. Each of these factors could influence not only the current neural states (that is, the position of neural activity along the surface of the neural manifold) but also the neural manifold properties. In particular, neuromodulators and glial cells could shape neural manifolds at much slower timescales[112] than those typically considered in behavioral studies. The influence of all of these subcellular and non-neuronal factors points toward a possible role of neural manifolds in bridging across different spatial and temporal scales of neural inquiry: from cellular processes up to brain-wide imaging, and from rapid behavioral adaptations to slow changes in mood, motivation and other internal states.

Behavior

Behavior shapes the properties of the observed neural manifold in two key ways. First, the intrinsic dimensionality of neural manifolds will be constrained by the behavioral conditions explored[113]—for the prefrontal cortex, this could be the range of choices explored during foraging, whereas for the motor cortex it could be the different movements an animal performs. Neural recordings are typically made during the performance of tasks in a laboratory setting. Thus, the relatively low dimensionality of sensory stimuli[114] and/or behavioral output[53,113] in many laboratory tasks constrains the range of states visited by the neural population (Fig. 3), and this is likely to contribute to the low intrinsic dimensionality of the observed neural manifolds. Second, if there is a geometry or structure in the task or behavioral output, we should expect to see this structure reflected in some aspects of the neural manifold. For example, the spatial organization of reaching targets in a motor task[113] is likely to ensure that there is a similar organization in the estimated neural manifolds[53].

The circularity here is unavoidable: behavior results from the organization of neural activity, and this organization is adapted through evolution and development to enable the behavioral repertoire of the animal. To illustrate, consider that different neural

BOX 3

Using recurrent neural networks to study neural manifolds

The properties of neural manifolds (geometry, topology, intrinsic dimensionality) and their relation to behavior cannot always be directly and reliably inferred from neural recordings. Computational modeling, especially recurrent neural networks (RNNs), can provide a platform to explore and generate hypotheses on the relation between neural manifolds and behavior.

What is an RNN? RNNs are function approximators built by recurrently connecting units, which can be regarded as analogous to biological neurons. These connections are then trained to achieve a certain objective using mathematical techniques. The dynamics of a typical RNN is governed by:

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij}\phi(x_j(t)) + \sum_{k=1}^{M} B_{ik}u_k(t)$$

Here, the dynamics of each ith unit, $\dot{x}_i(t)$, is influenced by three components: (i) its own state, $x_i(t)$; (ii) inputs from other units ($x_i(t)$), transformed nonlinearly and weighted by J_{ij}; and (iii) external inputs $u(t)$, where the kth input influences the ith unit weighted by B_{ik}. Many RNNs impose nonlinearities on the recurrent weights through a nonlinear activation function ϕ, and may also include stochastic terms.

Connectivity matrix rank and structure. Structure in the connectivity matrix J influences the function, dynamics and dimensionality of the RNN population activity[155]. A key component of this structure with relevance for neural manifolds is the 'rank', which refers to the minimum number of orthogonal vectors l needed to reconstruct the matrix. The $N \times N$ connectivity matrix is considered 'full rank' if $l \sim N$, and is 'low rank' if $l << N$. Structured connectivity has allowed researchers to study the emergence of low-dimensional dynamics in the brain. For example, untrained full-rank RNNs exhibit chaotic internally generated dynamics[175] that resemble spontaneous activity patterns recorded in vivo, but cannot perform any complex functions. By contrast, both the spontaneous and stimulus-evoked activity of low-rank networks are low-dimensional and can be predicted from the relationship between a small number of vectors that represent the connectivity structure and feedforward inputs[176].

Using RNNs to study neural manifold properties and dynamics. Here, we highlight some particularly compelling use cases for RNNs to study neural population activity. Recent work has started to explore RNNs that are directly constrained to replicate neural data[177–180] to infer principles of neural manifold properties or dynamics. These models are a potentially powerful approach to 'peek under the hood' and access properties of neural circuits that are difficult to obtain experimentally[8]. More commonly, RNNs—both full-rank[10,19,41,87,117,177,181–188] and low-rank[176,185,186,189]—have been trained to perform tasks such as those used in neuroscience experiments. Consistent with the observation of low-dimensional manifolds in neural recordings from animals performing laboratory tasks, the dynamics of many task-trained networks are low dimensional and, in several cases, exhibit manifold-like topologies[19,85,87,185]. These networks have been used in two primary ways.

First, RNNs have been used to identify, propose or confirm properties of the neural population latent dynamics underlying laboratory tasks[10,117,118,177,183,184,190–195]. For example, theoretical studies using RNNs have explored how neural networks can flexibly perform many tasks[117,192] and demonstrated that sequence propagation implemented in recurrent networks can be a possible implementation for working memory[177]. Additionally, RNN modeling of a BCI learning task[23,82] comparing within-manifold and outside-manifold perturbations identified possible solutions that could be achieved by biological brains[118] and proposed a form of learning that can extend to other contexts[194].

Second, RNNs have also been used to study basic principles of neural function by building models with specific assumptions and characterizing their properties and behavior[85,151,185,190,196–199]. For example, RNN models led to the proposal that the activity of grid cells in the entorhinal cortex would live on a toroidal manifold[190,200], a prediction that has been recently verified with experimental recordings[50]. This type of modeling is also especially useful to study the link between low-level circuit details and high-level neural population properties (including neural manifolds)[85,197]. Recent work has further leveraged RNN models to propose links between recurrent connectivity statistics and the properties of neural manifolds during the performance of the same task[85,196,198].

populations across the brain (for example, different regions) should have manifolds whose geometries are suited to the behavioral processes these neural populations are involved in. For example, hippocampal manifolds may reflect the structure of the environment[86], and manifolds in the hypothalamus may reflect properties of behavioral states such as aggression[62]. However, not all geometric properties of manifolds arise solely from behavioral output: the persistence of some manifold topologies between awake exploration and sleep—when overt behavior is absent—provide evidence against this notion[50,55]. Ultimately, we should be able to build links between manifolds and behavior in both directions: tying the properties of neural manifolds (for example, geometry, dynamics) to specific aspects of behavior, and building normative or computational models (Box 3) that anticipate the properties of manifolds.

Learning, development and evolution

As discussed in the previous section, at the short behavioral timescale, neural manifolds reflect aspects of the ongoing sensations, actions, intentions and experience. However, neural manifolds and their

properties are concurrently shaped on multiple longer timescales: learning, development and evolution (Fig. 4).

The consistent execution of the same behavior seems to be associated with equally stable manifolds[53,59,83], assuming all other internal states (for example, impulsivity[72], engagement[42]) are equivalent. Yet, such apparent stability does not imply that neural manifolds are immutable. At the learning timescale, neural manifold properties likely change to help improve the ability to perform certain behaviors through practice[115], or to acquire behaviors that are not innate[116]. This begs the question: when learning a new behavior, is the resultant neural manifold best viewed as a repurposing of the original manifold, an expansion of that manifold or an entirely new manifold? We posit that all three cases could be true depending on the needs of the specific behavior. Small adaptations to improve a learned behavior may only require exploring a different portion of the existing neural manifold[23,71,85] or compositionally combining aspects of previously acquired manifolds[117]. By contrast, a completely new skill (such as playing the violin or learning a tonal language as a native Portuguese speaker) is likely to require
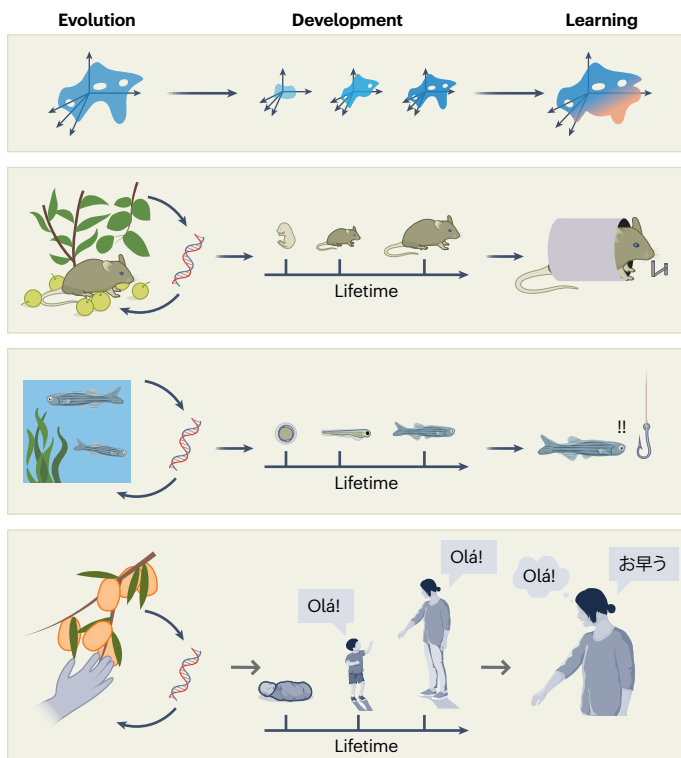
Fig. 4 | Neural manifolds are shaped over various timescales. Examples of how neural manifolds are shaped over evolutionary timescales (left), during development (middle) and during learning (right). Top: mice can learn to perform arbitrary laboratory tasks such as reaching, grasping and pulling on a lever that are far outside their evolved repertoire as herbivore foragers in their natural environment. Middle: zebrafish evolved to be omnivorous seawater fish but can learn during their lifetimes to recognize potentially dangerous fishing spots in streams. Bottom: humans evolved as social hunter gatherers who acquired language, and now many of them even learn many different languages as adults.

substantial modification of the existing manifold. This prediction offers a possible explanation for the difficulty in acquiring some new skills[82,118].

Neural manifolds that emerge on the timescales of behavior and learning are subject to constraints imposed on longer timescales of development and evolution. At the longest evolutionary timescale, animals acquire the adaptations that then become part of the species' innate behavioral repertoire, as well as their ability to learn[119]. These adaptations are expressed at the developmental timescale by running the 'genetic program'[120,121] that ultimately gives rise to an individual with the fundamental neural circuit properties of its species. These circuit properties help constrain the possible activity patterns generated by the brain as it develops innate behaviors[122,123] and acquires new, learned ones[124], in a process that is likely to be necessary to form the required neural manifolds. In this way, evolution and development can have a profound effect on the emergent neural manifolds. Indeed, genetically specified circuit constraints may explain the observation of preserved latent dynamics across individuals in the motor system[59], or the preserved ring attractor structure guiding head-direction cells across species[31,55,125], suggesting common (or at least convergent) evolutionary constraints.

Next steps toward a manifold view of how the brain generates behavior

Here, we discuss the next steps to build a more cohesive view of how various intrinsic and extrinsic constraints shape neural manifolds, and how these in turn shape behavior.

Causally manipulating neural states to characterize neural manifolds

Developing a pluralistic explanation of brain function will require causal manipulations that allow us to directly validate, characterize and manipulate neural manifolds, going beyond observation through neural recordings. We propose three overarching goals for causal experiments (Fig. 5). During validation, experimenters causally test whether and how the neural manifold constrains both neural and behavioral states. During exploration, experimenters interrogate the possible neural states within the existing manifold, and how this manifold constrains future or unobserved activity. During alteration, experimenters modify the existing manifold to study how it shapes an animal's behavior (and vice versa).

Causal manipulations to characterize neural manifolds have two key properties. First, there is the timescale of the perturbation; short, transient manipulations will typically have a different effect on the neural circuit and behavior than long, extended manipulations. Second, whether the effects of the perturbation bring the neural state on-manifold or off-manifold[81]. On-manifold perturbations are ones that move neural activity to states that can be achieved during animal behavior. If we assume that the relationship between neural manifold states and behavioral states is preserved, an expected outcome of an on-manifold perturbation is that the animal will ultimately produce or experience the expected behavioral state. Conversely, off-manifold perturbations lead to neural states that cannot be plausibly reached during natural behavior. Note that determining whether a perturbation is off-manifold can be difficult when there are not clearly observable 'biomarkers' for behavioral outcomes. Each of the three goals defined above have specific needs in terms of which type of manipulation is needed.

Validation is likely to be best done through transient, short-timescale perturbations that amount to 'pinging' the current neural state. The subsequent response in the latent dynamics can be used to validate predicted constraints on neural population activity or map the local properties of the manifold akin to system identification[126], allowing experimenters to characterize both the dynamical and geometric properties of the manifold and its presumed contribution to behavior[65]. The pinging approach can also allow exploration of what neural states are possible, although more meaningful exploration is likely to result from longer-timescale perturbations (for example, manipulating neural activity via open-loop stimulation, BCIs, and behavioral perturbation and adaptation paradigms). Alteration of neural manifolds will almost certainly require long-timescale perturbations that engage some form of plasticity or learning (for example, closed-loop stimulation, entrainment through long-term open-loop stimulation, learning paradigms, pharmacology, chemogenetics, knock-out models and lesions) can also be used to probe neural manifold properties, but also provide a means to alter the existing neural manifold. These manipulations are likely to be off-manifold with respect to the neural states observed in naturally behaving animals, given the difficulty in predicting long-timescale dependencies in recurrent (and likely nonlinear) neural networks. However, the development of next-generation neurostimulation technologies and computational models[127] could ultimately enable the steering of neural states toward specific on-manifold targets over long timescales. In the following sections we discuss several specific uses of these perturbations to define how neural manifolds emerge and their relationship with behavior.

Defining how neural manifold properties are shaped by low-level biological properties

Within a neural circuit, 'low-level' intrinsic factors including cell-type distributions, cortical layers, region architecture and the connectome are likely to shape properties of neural manifolds and their relation with behavior. In this section, we outline a few key directions to explore the link between low-level circuit properties and neural manifolds.
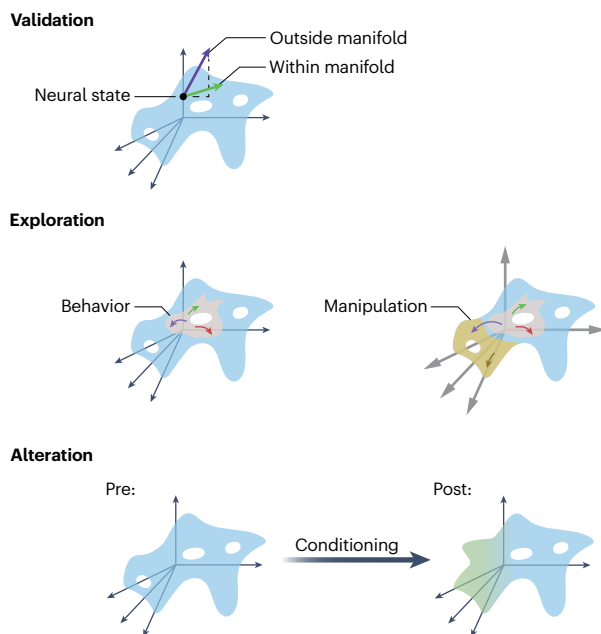
Fig. 5 | Using causal manipulations to characterize neural manifolds.
Validation: Brief, short-timescale 'pinging' perturbations can be used to causally test predictions about the neural and behavioral constraints imposed by neural manifolds. These perturbations (illustrated by arrows) are likely to push activity outside the natural repertoire of neural states (outside manifold), but with appropriate design could test possible states within manifolds. Exploration: The properties of neural manifolds can be explored using neural activity recorded during behavior (left) or during manipulations that require animals to volitionally control neural states (right). Alteration: A variety of manipulation techniques and behavioral learning paradigms ('conditioning') may be used to alter a pre-existing neural manifold.

First, new experiments can unpack the specialized contributions (if they exist) of different cell types to the neural manifold by simultaneously recording the activity of different cell types from a given region. This can be achieved through genetic labeling of each cell type using two or more indicators[128] or sophisticated post-experiment reconstruction and labeling through transcriptomics[129]. With these data, one can calculate both the various cell-type-specific manifolds[130,131] and the manifold including all cell types to study the relationship between the two.

Second, we can explore the influence of spatial and/or laminar organization on neural manifolds using targeted genetic expression[132,133], high-density electrophysiology probes[14] and multi-site recordings. For example, input and output projections are often organized on the basis of the cortical layer, and modern high-density probes allow continuous laminar sampling of spiking activity. As specific directions in the neural manifold seem to capture inter-regional communication[48,71,77,78,80], we would expect to see variation in the properties of manifolds specific to different cortical layers based on their input and output organization.

Third, we can try to bridge anatomical knowledge of the underlying circuitry and the properties of neural manifolds. Interactions between neural populations can occur directly through synaptic connections, and also indirectly through widespread neuromodulator release. The former is particularly interesting given modern efforts to map complete connectomes of experimentally tractable species[134]. Connectomes could help frame how interregional interactions shape neural manifolds. Although the link between the connectome and circuit function may be tenuous in some species[135,136], the organization of circuit connectivity likely shapes higher-order interactions between neural populations[137] and should thus shape manifold properties. New, neuroscience-focused manifold learning methods that include a priori

information about the intrinsic biological properties when estimating the manifold topology could allow for the empirical investigation of the link, if any, between low-level biophysical features and emergent population-wide structure in activity. Further, the influence of cell-type or circuit-level organization on neural manifold properties should be causally tested using the techniques outlined above. For example, we can test the effect of each cell type or projection pathway by removing their influence with transient manipulations that activate or inactivate them with short-timescale optogenetics[77], or long-timescale manipulations with designer receptors exclusively activated by designer drugs (DREADDs)[131,138].

Inferring the relationship between behavior and neural manifolds

Our goal is to establish clear links between the properties of neural manifolds and different aspects of behavior. Such links may only be apparent if we study animals performing a wide range of ethologically relevant behaviors and tasks of varying complexity[139–141]. This is important because simple, low-dimensional tasks may enforce that neural manifolds appear to be similar owing to behavioral constraints[113], disguising subtle functional differences in the regions that would be expressed during more naturalistic behaviors[48]. Further, looking at neural manifolds at a single timescale may also occlude their relationship with behavior. For example, brain regions that contain similar information about behavior during the execution of a well-practiced skill may have subserved different functions during their acquisition[141,142]. Thus, to understand the individual contributions of these various regions, we may need to characterize region-specific manifolds and their interactions across a broad range of increasingly complex behaviors and by looking at different timescales.

Ideally, these behaviorally rich experiments should be paired with causal, closed-loop manipulations to appropriately interpret the contributions of each brain region to the generation of behavior[8,139,141]. We propose this could take two forms. First, we should manipulate behavioral output (for example, through behavioral perturbations[70,71] or BCI mappings[23,82,143,144]) to test hypotheses of how activity within the manifold, or the manifold properties, should change based on theoretical predictions. For example, one could apply a perturbation that changes the relationship between a specific task feature and behavioral output and make predictions for differential effects on manifold properties across regions of the brain. Second, we should directly manipulate neural activity along the manifold in a closed loop with behavior in order to map the causal effect that specific manifold properties bear on the animal's behavior. For example, one could use holographic optical stimulation to 'teleport' neural activity[144] in the region being studied to a specific manifold state explored during natural behavior and verify that the animal produces the expected behavioral response.

Ultimately, we would like to anticipate changes in behavior simply by observing neural states and applying interpretable manifold-based models linking the two. Similarly, we would like to anticipate changes in manifold geometry using only behavioral information. These goals will be supported by normative models as well as principled experimental manipulations, which can help to dissociate and explore the complex relationships between neural manifolds and behavior at different timescales.

Big open questions
On emergence
Neural manifolds provide an intuitive way to conceptualize the coordinated activity from neural populations, enabling the integration of findings across studies, comparison across brain regions, and making testable experimental predictions grounded in geometry and dynamics that abstract away the details of implementation by neural circuits. An important question is whether the neural manifold level of description

holds more explanatory power[145] than its lower-level constituents, such as circuits, neurons and synapses. In this case, neural manifolds would be considered emergent[146,147]. To illustrate this, we draw an analogy: if your friend asks you about a movie you have just watched, you would not explain its plot by carefully describing a hundred frames from the movie (Krakauer, J. W. Personal correspondence). Instead, you would summarize key events in the plot. Similarly, while the neural manifold is rooted in biophysical properties, it captures structure in the neural population activity that directly informs the behavior of the animal. Thus, it may afford more compact and understandable descriptions of brain function than those drawn from its constituent neurons, synapses, and so on.

An important related question is whether one can freely move between the low and high levels of description, that is, predict the neural manifold properties from the low-level network properties and single unit activity (and vice versa)[148]. Moving between these levels may be impossible as there can be, in principle, a near infinite number of low-level states (for example, synaptic connections, membrane properties and vesicular release probabilities) that give rise to the same population activity, and subsequently neural manifold properties. This has been shown by theoretical, experimental and modeling work[135,149]. Classic work demonstrated that many combinations of biophysical parameters could lead to the same circuit output in the lobster stomatogastric ganglion[135]. Similarly, the same cells recorded from genetically identical worms exhibit very different activity patterns during the same locomotor behavior[83], yet the emergent neural manifolds estimated from the population recordings have remarkably consistent structure. We thus argue that the neural manifold is an apt level at which to formulate and test hypotheses about brain function.

Neural manifolds and neural dynamics: the chicken or the egg?
In this paper, we propose that neural dynamics are constrained by neural manifolds that arise, in part, from biophysical constraints. Yet, neural manifolds are continually shaped by behavior resulting from these neural dynamics. This is a process that may even start in utero and continue during subsequent development, when activity (dynamics) shapes circuit connectivity through activity-dependent plasticity[123,150]. These developmental constraints could shape and refine the manifold, which in turn enables some dynamics but not others in order to acquire an individual's behavioral repertoire[116,151]. This begs the question: what came first, the manifold or the dynamics? We feel that there is an inextricable link between the two that complicates the interpretation of experimental data, in particular when seeking to describe neural manifold properties. At this time, we can only estimate neural manifolds post hoc by finding an embedding that spans the trajectories described by the latent dynamics. This makes it challenging to ask principled questions about the relationship between these latent dynamics and the properties of neural manifolds. However, the experiments and conceptual innovations outlined in the previous sections may reduce our dependence on correlative methods to estimate manifolds from dynamics post hoc, thereby allowing us to better investigate how latent dynamics shape the manifold and vice versa.

How are neural manifolds organized brain wide?
Neural manifolds and their corresponding latent dynamics have historically been identified within single regions (or perhaps a small set of two or three regions). However, the neural activity that generates behavior is the product of continuous interactions between numerous, highly specialized regions[8]. Thus, there is a critical open question when relating neural manifold properties to behavior: what is the spatial scale at which we should consider manifolds in the brain, as we move between small (for example, cortical columns), medium (for example, regions) and large (for example, brain-wide) scales?

One possibility is that each of the sub-modules in the brain (cell types, cortical columns, regions) has its own independent manifold and dynamics dictated by its local circuit properties and behavior-specific inputs from other regions—perhaps a torus in one region and a ring attractor in another. Thus, a putative 'brain-wide manifold' can be fully described by assembling these region-specific manifolds. However, we argue that this is unlikely given the high degree of interconnectivity in the organization of these sub-modules. To illustrate the problem, we consider a population consisting of all neurons in a given cortical region. These neurons could be further subdivided into cortical columns. Now consider a neural manifold underlying the activity of all neurons in just one of these columns. This manifold reflects the coordinated activity—presumably attributable to some extent to constraints on the possible states—in the constituent neural population. This covariance (or constraints) can arise from local (recurrent) connections within that population, or from inputs driving the neurons, for example from other nearby columns. Every column-specific manifold is thus a product of inputs from other columns, each with their own column-specific manifold. We may not be able to characterize the overarching manifold underlying the activity of all neurons in the region without considering the interactions between all columns. Said differently, column-specific manifolds are not isolatable sub-manifolds within the region-wide manifold. Consequently, we could not use individual column-specific manifolds to predict the larger manifold, nor could we characterize individual columns knowing only the region-wide manifold. Similar logic applies when considering brain-wide activity made up of numerous interacting region-specific (or cell-type-specific[94]) manifolds. Perhaps, then, we should consider the brain-wide manifold as emergent from numerous sub-manifolds throughout the brain, in much the same way that each of those sub-manifolds may be emergent from the activity of individual neurons.

Summary

Neural manifolds provide a compact and often more interpretable means to describe the function of a neural population than the independent activity of its constituent neurons. This framework can lead to the generation of clear testable hypotheses about brain function. Furthermore, the neural manifold framework could also enhance our ability to use those insights to treat diseases of the nervous system and enable new neurotechnologies for interfacing with the brain. Ultimately, the neural manifold framework could help link all aspects of the neural basis of behavior—cognition, action, sensation and emotion—to arrive at a unified description of how the brain works.

References

1. Ramón y Cajal, S. *Textura del sistema nervioso del hombre y de los vertebrados: estudios sobre el plan estructural y composición histológica de los centros nerviosos adicionados de consideraciones fisiológicas fundadas en los nuevos descubrimientos* (N. Moya, 1899).

2. Georgopoulos, A. P., Kalaska, J. F., Caminiti, R. & Massey, J. T. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci.* **2**, 1527–1537 (1982).

3. Humphrey, D. R., Schmidt, E. M. & Thompson, W. D. Predicting measures of motor performance from multiple cortical spike trains. *Science* **170**, 758–762 (1970).

4. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).

5. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34**, 171–175 (1971).

6. Brodmann, K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, 1909).

7. Gallego, J. A., Perich, M. G., Miller, L. E. & Solla, S. A. Neural manifolds for the control of movement. *Neuron* **94**, 978–984 (2017).

8. Perich, M. G. & Rajan, K. Rethinking brain-wide interactions through multi-region 'network of networks' models. *Curr. Opin. Neurobiol.* **65**, 146–151 (2020).

9. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation through neural population dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).

10. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).

11. Churchland, M. M. & Shenoy, K. V. Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. *J. Neurophysiol.* **97**, 4235–4257 (2007).

12. Machens, C. K., Romo, R. & Brody, C. D. Functional, but not anatomical, separation of 'what' and 'when' in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).

13. Stevenson, I. H. & Kording, K. P. How advances in neural recording affect data analysis. *Nat. Neurosci.* **14**, 139–142 (2011).

14. Jun, J. J. et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).

15. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* **10**, 413–420 (2013).

16. Chen, T. -W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

17. Cunningham, J. P. & Yu, B. M. Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* **17**, 1500–1509 (2014).

18. Saxena, S. & Cunningham, J. P. Towards the neural population doctrine. *Curr. Opin. Neurobiol.* **55**, 103–111 (2019).

19. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).

20. Wang, Y. et al. Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* **9**, 534–542 (2006).

21. Riemann, B. *On the Hypotheses Which Lie at the Bases of Geometry* (Birkhäuser, 2016).

22. Dubrovin, B. A., Fomenko, A. T. & Novikov, S. P. *Modern Geometry—Methods and Applications: Part II: The Geometry and Topology of Manifolds* (Springer Science & Business Media, 2012).

23. Sadtler, P. T. et al. Neural constraints on learning. *Nature* **512**, 423–426 (2014).

24. Mitchell-Heggs, R., Prado, S., Gava, G. P., Go, M. A. & Schultz, S. R. Neural manifold analysis of brain circuit dynamics in health and disease. *J. Comput. Neurosci.* **51**, 1–21 (2023).

25. Stopfer, M., Jayaraman, V. & Laurent, G. Intensity versus identity coding in an olfactory system. *Neuron* **39**, 991–1004 (2003).

26. Briggman, K. L., Abarbanel, H. D. I. & Kristan, W. B. Jr. Optical imaging of neuronal populations during decision-making. *Science* **307**, 896–901 (2005).

27. Kato, S. et al. Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell* **163**, 656–669 (2015).

28. Bruno, A. M., Frost, W. N. & Humphries, M. D. A spiral attractor network drives rhythmic locomotion. *Elife* **6**, e27342 (2017).

29. Ahrens, M. B. et al. Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* **485**, 471–477 (2012).

30. Andalman, A. S. et al. Neuronal dynamics regulating brain and behavioral state transitions. *Cell* **177**, 970–985 (2019).

31. Petrucco, L. et al. Neural dynamics and architecture of the heading direction circuit in zebrafish. *Nat. Neurosci.* **26**, 765–773 (2023).

32. Churchland, M. M. et al. Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).

33. Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci.* **17**, 1784–1792 (2014).

34. Sauerbrei, B. A. et al. Cortical pattern generation during dexterous movement is input-driven. *Nature* **577**, 386–391 (2020).

35. Stringer, C. et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, 255 (2019).

36. Miri, A. et al. Behaviorally selective engagement of short-latency effector pathways by motor cortex. *Neuron* **95**, 683–696 (2017).

37. Rubin, A. et al. Revealing neural correlates of behavior without behavioral measurements. *Nat. Commun.* **10**, 4745 (2019).

38. Chen, H. -T., Manning, J. R. & van der Meer, M. A. A. Between-subject prediction reveals a shared representational geometry in the rodent hippocampus. *Curr. Biol.* **31**, 4293–4304 (2021).

39. Yoo, S. B. M. & Hayden, B. Y. The transition from evaluation to selection involves neural subspace reorganization in core reward regions. *Neuron* **105**, 712–724 (2020).

40. Cueva, C. J. et al. Low-dimensional dynamics for working memory and time encoding. *Proc. Natl Acad. Sci. USA* **117**, 23021–23032 (2020).

41. Sohn, H., Narain, D., Meirhaeghe, N. & Jazayeri, M. Bayesian computation through cortical latent dynamics. *Neuron* **103**, 934–947 (2019).

42. Hennig, J. A. et al. Learning is shaped by abrupt changes in neural engagement. *Nat. Neurosci.* **24**, 727–736 (2021).

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

44. Stavisky, S. D. et al. Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis. *Elife* **8**, e46015 (2019).

45. Pandarinath, C. et al. Neural population dynamics in human motor cortex during movements in people with ALS. *Elife* **4**, e07436 (2015).

46. Minxha, J., Adolphs, R., Fusi, S., Mamelak, A. N. & Rutishauser, U. Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* **368**, eaba3313 (2020).

47. Dekleva, B. M. et al. Motor cortex retains and reorients neural dynamics during motor imagery. *Nat. Hum. Behav.* **8**, 729–742 (2024).

48. Perich, M. G. et al. Motor cortical dynamics are shaped by multiple distinct subspaces during naturalistic behavior. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.30.228767 (2020).

49. Kobak, D., Pardo-Vazquez, J. L., Valente, M., Machens, C. K. & Renart, A. State-dependent geometry of population activity in rat auditory cortex. *Elife* **8**, e44526 (2019).

50. Gardner, R. J. et al. Toroidal topology of population activity in grid cells. *Nature* **602**, 123–128 (2022).

51. Ebitz, R. B. & Hayden, B. Y. The population doctrine in cognitive neuroscience. *Neuron* **109**, 3055–3068 (2021).

52. Jurewicz, K., Sleezer, B. J., Mehta, P. S., Hayden, B. Y. & Ebitz, R. B. Irrational choices via a curvilinear representational geometry for value. *Nat. Commun.* **15**, 6424 (2024).

53. 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).

54. Russo, A. A. et al. Motor cortex embeds muscle-like commands in an untangled population response. *Neuron* **97**, 953–966 (2018).

55. Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A. & Fiete, I. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nat. Neurosci.* **22**, 1512–1520 (2019).

56. Lindén, H., Petersen, P. C., Vestergaard, M. & Berg, R. W. Movement is governed by rotational neural dynamics in spinal motor networks. *Nature* **610**, 526–531 (2022).

57. Gmaz, J. M. & van der Meer, M. A. A. Context coding in the mouse nucleus accumbens modulates motivationally relevant information. *PLoS Biol.* **20**, e3001338 (2022).

58. Lanore, F., Cayco-Gajic, N. A., Gurnani, H., Coyle, D. & Silver, R. A. Cerebellar granule cell axons support high-dimensional representations. *Nat. Neurosci.* **24**, 1142–1150 (2021).

59. Safaie, M. et al. Preserved neural dynamics across animals performing similar behaviour. *Nature* https://doi.org/10.1038/s41586-023-06714-0 (2023).

60. Ayar, E. C., Heusser, M. R., Bourrelly, C. & Gandhi, N. J. Distinct context- and content-dependent population codes in superior colliculus during sensation and action. *Proc. Natl Acad. Sci. USA* **120**, e2303523120 (2023).

61. Allsop, S. A. et al. Corticoamygdala transfer of socially derived information gates observational learning. *Cell* **173**, 1329–1342 (2018).

62. Nair, A. et al. An approximate line attractor in the hypothalamus encodes an aggressive state. *Cell* **186**, 178–193 (2023).

63. Jahn, C. I. et al. Learning attentional templates for value-based decision-making. *Cell* **187**, 1476–1489 (2024).

64. Markanday, A., Hong, S., Inoue, J., De Schutter, E. & Thier, P. Multidimensional cerebellar computations for flexible kinematic control of movements. *Nat. Commun.* **14**, 2548 (2023).

65. Vinograd, A., Nair, A., Kim, J. H., Linderman, S. W. & Anderson, D. J. Causal evidence of a line attractor encoding an affective state. *Nature* **634**, 910–918 (2024).

66. Talpir, I. & Livneh, Y. Stereotyped goal-directed manifold dynamics in the insular cortex. *Cell Rep.* **43**, 114027 (2024).

67. Bush, N. E. & Ramirez, J. -M. Latent neural population dynamics underlying breathing, opioid-induced respiratory depression and gasping. *Nat. Neurosci.* **27**, 259–271 (2024).

68. Nieh, E. H. et al. Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80–84 (2021).

69. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360–368 (2023).

70. Sun, X. et al. Cortical preparatory activity indexes learned motor memories. *Nature* **602**, 274–279 (2022).

71. Perich, M. G., Gallego, J. A. & Miller, L. E. A neural population mechanism for rapid learning. *Neuron* **100**, 964–976 (2018).

72. Cowley, B. R. et al. Slow drift of neural activity as a signature of impulsivity in macaque visual and prefrontal cortex. *Neuron* **108**, 551–567 (2020).

73. Churchland, M. M. & Shenoy, K. V. Preparatory activity and the expansive null-space. *Nat. Rev. Neurosci.* **25**, 213–236 (2024).

74. Vyas, S. et al. Neural population dynamics underlying motor learning transfer. *Neuron* **97**, 1177–1186 (2018).

75. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).

76. 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. *Nat. Commun.* **7**, 13239 (2016).

77. Javadzadeh, M. & Hofer, S. B. Dynamic causal communication channels between neocortical areas. *Neuron* **110**, 2470–2483 (2022).

78. Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M. & Kohn, A. Cortical areas interact through a communication subspace. *Neuron* **102**, 249–259 (2019).

79. Veuthey, T. L., Derosier, K., Kondapavulur, S. & Ganguly, K. Single-trial cross-area neural population dynamics during long-term skill learning. *Nat. Commun.* **11**, 4057 (2020).

80. Semedo, J. D. et al. Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nat. Commun.* **13**, 1099 (2022).

81. Jazayeri, M. & Afraz, A. Navigating the neural space in search of the neural code. *Neuron* **93**, 1003–1014 (2017).

82. Oby, E. R. et al. New neural activity patterns emerge with long-term learning. *Proc. Natl Acad. Sci. USA* **116**, 15210–15215 (2019).

83. Brennan, C. & Proekt, A. A quantitative model of conserved macroscopic dynamics predicts future motor commands. *Elife* **8**, e46814 (2019).

84. Hermansen, E., Klindt, D. A. & Dunn, B. A. Uncovering 2-D toroidal representations in grid cell ensemble activity during 1-D behavior. *Nat. Commun.* **15**, 5429 (2024).

85. Fortunato, C. et al. Nonlinear manifolds underlie neural population activity during behaviour. Preprint at *bioRxiv* https://doi.org/10.1101/2023.07.18.549575 (2023).

86. Guo, W., Zhang, J. J., Newman, J. P. & Wilson, M. A. Latent learning drives sleep-dependent plasticity in distinct CA1 subpopulations. *Cell Rep.* **43**, 115028 (2024).

87. Remington, E. D., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron* **98**, 1005–1019 (2018).

88. Okazawa, G., Hatch, C. E., Mancoo, A., Machens, C. K. & Kiani, R. Representational geometry of perceptual decisions in the monkey parietal cortex. *Cell* **184**, 3748–3768 (2021).

89. Langdon, C., Genkin, M. & Engel, T. A. A unifying perspective on neural manifolds and circuits for cognition. *Nat. Rev. Neurosci.* **24**, 363–377 (2023).

90. Tsodyks, M., Kenet, T., Grinvald, A. & Arieli, A. Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science* **286**, 1943–1946 (1999).

91. Okun, M. et al. Diverse coupling of neurons to populations in sensory cortex. *Nature* **521**, 511–515 (2015).

92. Marshel, J. H. et al. Cortical layer-specific critical dynamics triggering perception. *Science* **365**, eaaw5202 (2019).

93. Dahmen, D. et al. Strong and localized recurrence controls dimensionality of neural activity across brain areas. Preprint at *bioRxiv* https://doi.org/10.1101/2020.11.02.365072 (2023).

94. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).

95. Park, J. et al. Motor cortical output for skilled forelimb movement is selectively distributed across projection neuron classes. *Sci. Adv.* **8**, eabj5167 (2022).

96. Azcorra, M. et al. Unique functional responses differentially map onto genetic subtypes of dopamine neurons. *Nat. Neurosci.* **26**, 1762–1774 (2023).

97. Musall, S. et al. Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. *Nat. Neurosci.* **26**, 495–505 (2023).

98. Bugeon, S. et al. A transcriptomic axis predicts state modulation of cortical interneurons. *Nature* **607**, 330–338 (2022).

99. Nakai, S., Kitanishi, T. & Mizuseki, K. Distinct manifold encoding of navigational information in the subiculum and hippocampus. *Sci. Adv.* **10**, eadi4471 (2024).

100. Lara, A. H., Cunningham, J. P. & Churchland, M. M. Different population dynamics in the supplementary motor area and motor cortex during reaching. *Nat. Commun.* **9**, 2754 (2018).

101. 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).

102. Recanatesi, S., Bradde, S., Balasubramanian, V., Steinmetz, N. A. & Shea-Brown, E. A scale-dependent measure of system dimensionality. *Patterns* **3**, 100555 (2022).

103. Park, J. et al. Conjoint specification of action by neocortex and striatum. *Neuron* **113**, 620–636 (2025).

104. Kutsuwada, T. et al. Molecular diversity of the NMDA receptor channel. *Nature* **358**, 36–41 (1992).

105. Nakanishi, S. Molecular diversity of glutamate receptors and implications for brain function. *Science* **258**, 597–603 (1992).

106. Fişek, M. et al. Cortico-cortical feedback engages active dendrites in visual cortex. *Nature* https://doi.org/10.1038/s41586-023-06007-6 (2023).

107. Srivastava, K. H. et al. Motor control by precisely timed spike patterns. *Proc. Natl Acad. Sci. USA* **114**, 1171–1176 (2017).

108. Riehle, A., Grün, S., Diesmann, M. & Aertsen, A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278**, 1950–1953 (1997).

109. Perea, G. et al. Activity-dependent switch of GABAergic inhibition into glutamatergic excitation in astrocyte-neuron networks. *Elife* **5**, e20362 (2016).

110. Mu, Y. et al. Glia accumulate evidence that actions are futile and suppress unsuccessful behavior. *Cell* **178**, 27–43 (2019).

111. Hsueh, B. et al. Cardiogenic control of affective behavioural state. *Nature* **615**, 292–299 (2023).

112. Cahill, M. K. et al. Network-level encoding of local neurotransmitters in cortical astrocytes. *Nature* **629**, 146–153 (2024).

113. Gao, P. et al. A theory of multineuronal dimensionality, dynamics and measurement. Preprint at *bioRxiv* https://doi.org/10.1101/214262 (2017).

114. Chung, S. & Abbott, L. F. Neural population geometry: an approach for understanding biological and artificial neural networks. *Curr. Opin. Neurobiol.* **70**, 137–144 (2021).

115. Zhou, J. et al. Evolving schema representations in orbitofrontal ensembles during learning. *Nature* **590**, 606–611 (2021).

116. Shuvaev, S., Lachi, D., Koulakov, A. & Zador, A. Encoding innate ability through a genomic bottleneck. *Proc. Natl Acad. Sci. USA* **121**, e2409160121 (2024).

117. Driscoll, L. N., Shenoy, K. & Sussillo, D. Flexible multitask computation in recurrent networks utilizes shared dynamical motifs. *Nat. Neurosci.* **27**, 1349–1363 (2024).

118. Feulner, B. & Clopath, C. Neural manifold under plasticity in a goal driven learning behaviour. *PLoS Comput. Biol.* **17**, e1008621 (2021).

119. Simpson, G. G. The Baldwin effect. *Evolution* **7**, 110–117 (1953).

120. Hiesinger, P. R. *The Self-Assembling Brain: How Neural Networks Grow Smarter* (Princeton University Press, 2022).

121. Mitchell, K. J. *Innate: How the Wiring of Our Brains Shapes Who We Are* (Princeton University Press, 2020).

122. Martini, F. J., Guillamón-Vivancos, T., Moreno-Juan, V., Valdeolmillos, M. & López-Bendito, G. Spontaneous activity in developing thalamic and cortical sensory networks. *Neuron* **109**, 2519–2534 (2021).

123. Glanz, R. M., Sokoloff, G. & Blumberg, M. S. Neural decoding reveals specialized kinematic tuning after an abrupt cortical transition. *Cell Rep.* **42**, 113119 (2023).

124. Maimon-Mor, R. O., Schone, H. R., Henderson Slater, D., Faisal, A. A. & Makin, T. R. Early life experience sets hard limits on motor learning as evidenced from artificial arm use. *Elife* **10**, e66320 (2021).

125. Seelig, J. D. & Jayaraman, V. Neural dynamics for landmark orientation and angular path integration. *Nature* **521**, 186–191 (2015).

126. O'Shea, D. J. et al. Direct neural perturbations reveal a dynamical mechanism for robust computation. Preprint at *bioRxiv* https://doi.org/10.1101/2022.12.16.520768 (2022).

127. Dowling, M., Zhao, Y. & Park, I. M. eXponential FAmily Dynamical Systems (XFADS): large-scale nonlinear Gaussian state-space modeling. In *Advances in Neural Information Processing Systems 38* (2024).

128. Zhang, F. et al. Multimodal fast optical interrogation of neural circuitry. *Nature* **446**, 633–639 (2007).

129. Lovett-Barron, M. et al. Ancestral circuits for the coordinated modulation of brain state. *Cell* **171**, 1411–1423 (2017).

130. Sylwestrak, E. L. et al. Cell-type-specific population dynamics of diverse reward computations. *Cell* **185**, 3568–3587 (2022).

131. Esparza, J. et al. Cell-type-specific manifold analysis discloses independent geometric transformations in the hippocampal spatial code. *Neuron* https://doi.org/10.1016/j.neuron.2025.01.022 (2025).

132. Chan, K. Y. et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).

133. Cardin, J. A. et al. Targeted optogenetic stimulation and recording of neurons in vivo using cell-type-specific expression of Channelrhodopsin-2. *Nat. Protoc.* **5**, 247–254 (2010).

134. Economo, M. N. et al. A platform for brain-wide imaging and reconstruction of individual neurons. *Elife* **5**, e10566 (2016).

135. Prinz, A. A., Bucher, D. & Marder, E. Similar network activity from disparate circuit parameters. *Nat. Neurosci.* **7**, 1345–1352 (2004).

136. Randi, F., Sharma, A. K., Dvali, S. & Leifer, A. M. Neural signal propagation atlas of *Caenorhabditis elegans*. *Nature* **623**, 406–414 (2023).

137. Pospisil, D. A. et al. The fly connectome reveals a path to the effectome. *Nature* **634**, 201–209 (2024).

138. Roth, B. L. DREADDs for neuroscientists. *Neuron* **89**, 683–694 (2016).

139. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A. & Poeppel, D. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93**, 480–490 (2017).

140. Gomez-Marin, A. & Ghazanfar, A. A. The life of behavior. *Neuron* **104**, 25–36 (2019).

141. Gmaz, J. M., Keller, J. A., Dudman, J. T. & Gallego, J. A. Integrating across behaviors and timescales to understand the neural control of movement. *Curr. Opin. Neurobiol.* **85**, 102843 (2024).

142. Mizes, K. G. C., Lindsey, J., Escola, G. S. & Ölveczky, B. P. Dissociating the contributions of sensorimotor striatum to automatic and visually guided motor sequences. *Nat. Neurosci.* **26**, 1791–1804 (2023).

143. Ganguly, K. & Carmena, J. M. Emergence of a stable cortical map for neuroprosthetic control. *PLoS Biol.* **7**, e1000153 (2009).

144. Lai, C., Tanaka, S., Harris, T. D. & Lee, A. K. Volitional activation of remote place representations with a hippocampal brain-machine interface. *Science* **382**, 566–573 (2023).

145. Barack, D. L. & Krakauer, J. W. Two views on the cognitive brain. *Nat. Rev. Neurosci.* **22**, 359–371 (2021).

146. Humphreys, P. *Emergence: a Philosophical Account* (Oxford University Press, 2016).

147. Kim, J. Emergence: core ideas and issues. *Synthese* **151**, 547–559 (2006).

148. Anderson, P. W. More is different. *Science* https://doi.org/10.1126/science.177.4047.393 (1972).

149. Marder, E. & Goaillard, J. -M. Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.* **7**, 563–574 (2006).

150. Kilgard, M. P. et al. Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J. Neurophysiol.* **86**, 326–338 (2001).

151. Chang, J. C., Perich, M. G., Miller, L. E., Gallego, J. A. & Clopath, C. De novo motor learning creates structure in neural activity that shapes adaptation. *Nat. Commun.* **15**, 4084 (2024).

152. Degenhart, A. D. et al. Stabilization of a brain-computer interface via the alignment of low-dimensional spaces of neural activity. *Nat. Biomed. Eng.* **4**, 672–685 (2020).

153. Lee, J. *Introduction to Topological Manifolds* (Springer Science & Business Media, 2010).

154. Dijkstra, J. J. & van Mill, J. *Erdos Space and Homeomorphism Groups of Manifolds* (American Mathematical Society, 2010).

155. Jazayeri, M. & Ostojic, S. Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Curr. Opin. Neurobiol.* **70**, 113–120 (2021).

156. Manley, J. et al. Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number. *Neuron* **112**, 1694–1709 (2024).

157. Marshall, N. J. et al. Flexible neural control of motor units. *Nat. Neurosci.* **25**, 1492–1504 (2022).

158. Trautmann, E. M. et al. Accurate estimation of neural population dynamics without spike sorting. *Neuron* **103**, 292–308 (2019).

159. DePasquale, B., Sussillo, D., Abbott, L. F. & Churchland, M. M. The centrality of population-level factors to network computation is demonstrated by a versatile approach for training spiking networks. *Neuron* **111**, 631–649 (2023).

160. Kobak, D. et al. Demixed principal component analysis of neural population data. *Elife* **5**, e10989 (2016).

161. Aoi, M. C., Mante, V. & Pillow, J. W. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).

162. Pandarinath, C. et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat. Methods* **15**, 805–815 (2018).

163. Hurwitz, C. et al. Targeted neural dynamical modeling. *Adv. Neural Inf. Process. Syst.* **34**, 29379–29392 (2021).

164. Gosztolai, A., Peach, R. L., Arnaudon, A., Barahona, M. & Vandergheynst, P. MARBLE: interpretable representations of neural population dynamics using geometric deep learning. *Nat. Methods* **22**, 612–620 (2025).

165. Tenenbaum, J. B., de Silva, V. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323 (2000).

166. Balasubramanian, M. & Schwartz, E. L. The isomap algorithm and topological stability. *Science* **295**, 7 (2002).

167. Roweis, S. T. & Saul, L. K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**, 2323–2326 (2000).

168. Lafon, S., Keller, Y. & Coifman, R. R. Data fusion and multicue data matching by diffusion maps. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 1784–1797 (2006).

169. Belkin, M. & Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396 (2003).

170. van der Maaten, L. & Hinton, G. Visualizing data using *t*-SNE. *J. Machine Learn. Res.* **9**, 2579–2605 (2008).

171. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at https://arxiv.org/abs/1802.03426 (2018).

172. Zhang, Z. & Zha, H. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM J. Sci. Comput.* **26**, 313–338 (2004).

173. Kohli, D., Cloninger, A. & Mishne, G. LDLE: Low distortion local eigenmaps. *J. Mach. Learn. Res.* **22**, 1–64 (2021).

174. Kohli, D., Nieuwenhuis, J., S., Cloninger, A., Mishne, G. & Narain, D. RATS: Unsupervised manifold learning using low-distortion alignment of tangent spaces. Preprint at *bioRxiv* https://doi.org/10.1101/2024.10.31.621292 (2024).

175. Rajan, K., Abbott, L. F. & Sompolinsky, H. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **82**, 011903 (2010).

176. Mastrogiuseppe, F. & Ostojic, S. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron* **99**, 609–623 (2018).

177. Rajan, K., Harvey, C. D. & Tank, D. W. Recurrent network models of sequence generation and memory. *Neuron* **90**, 128–142 (2016).

178. Perich, M. G. et al. Inferring brain-wide interactions using data-constrained recurrent neural network models. Preprint at *bioRxiv* https://doi.org/10.1101/2020.12.18.423348 (2021).

179. Fisher, D., Olasagasti, I., Tank, D. W., Aksay, E. R. F. & Goldman, M. S. A modeling framework for deriving the structural and functional architecture of a short-term memory microcircuit. *Neuron* **79**, 987–1000 (2013).

180. Dinc, F., Shai, A., Schnitzer, M. & Tanaka, H. CORNN: convex optimization of recurrent neural networks for rapid inference of neural dynamics. In *Advances in Neural Information Processing Systems 36* https://proceedings.neurips.cc/paper_files/paper/2023/file/a1035297387069793317783773f2d5864-Paper-Conference.pdf (2023).

181. Laje, R. & Buonomano, D. V. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat. Neurosci.* **16**, 925–933 (2013).

182. Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).

183. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–1033 (2015).

184. Sussillo, D. & Barak, O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput.* **25**, 626–649 (2013).

185. Beiran, M., Meirhaeghe, N., Sohn, H., Jazayeri, M. & Ostojic, S. Parametric control of flexible timing through low-dimensional neural manifolds. *Neuron* **111**, 739–753 (2023).

186. Valente, A., Ostojic, S. & Pillow, J. W. Probing the relationship between latent linear dynamical systems and low-rank recurrent neural network models. *Neural Comput.* **34**, 1871–1892 (2022).

187. Pollock, E. & Jazayeri, M. Engineering recurrent neural networks from task-relevant manifolds and dynamics. *PLoS Comput. Biol.* **16**, e1008128 (2020).

188. Enel, P., Procyk, E., Quilodran, R. & Dominey, P. F. Reservoir computing properties of neural dynamics in prefrontal cortex. *PLoS Comput. Biol.* **12**, e1004967 (2016).

189. Schuessler, F., Dubreuil, A., Mastrogiuseppe, F., Ostojic, S. & Barak, O. Dynamics of random recurrent networks with correlated low-rank structure. *Phys. Rev. Res.* **2**, 013111 (2020).

190. Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput. Biol.* **5**, e1000291 (2009).

191. Michaels, J. A., Schaffelhofer, S., Agudelo-Toro, A. & Scherberger, H. A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. *Proc. Natl Acad. Sci. USA* **117**, 32124–32135 (2020).

192. Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T. & Wang, X. -J. Task representations in neural networks trained to perform many cognitive tasks. *Nat. Neurosci.* **22**, 297–306 (2019).

193. Litwin-Kumar, A. & Doiron, B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.* **15**, 1498–1505 (2012).

194. Feulner, B. et al. Small, correlated changes in synaptic connectivity may facilitate rapid motor learning. *Nat. Commun.* **13**, 5163 (2022).

195. Feulner, B., Perich, M. G., Miller, L. E., Clopath, C. & Gallego, J. A. A neural implementation model of feedback-based motor learning. *Nat. Commun.* **16**, 1805 (2025).

196. Darshan, R. & Rivkind, A. Learning to represent continuous variables in heterogeneous neural networks. *Cell Rep.* **39**, 110612 (2022).

197. Pezon, L., Schmutz, V. & Gerstner, W. Linking neural manifolds to circuit structure in recurrent networks. Preprint at *bioRxiv* https://doi.org/10.1101/2024.02.28.582565 (2024).

198. Dubreuil, A., Valente, A., Beiran, M., Mastrogiuseppe, F. & Ostojic, S. The role of population structure in computations through neural dynamics. *Nat. Neurosci.* **25**, 783–794 (2022).

199. Hennequin, G., Vogels, T. P. & Gerstner, W. Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron* **82**, 1394–1406 (2014).

200. Fuhs, M. C. & Touretzky, D. S. A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.* **26**, 4266–4276 (2006).

Acknowledgements

Author contributions

M.G.P. and J.A.G. conceived of the overall concept for the manuscript. M.G.P., D.N. and J.A.G. wrote and edited the manuscript.

Competing interests

J.A.G. receives funding from Meta Platform Technologies and InBrain Neuroelectronics. M.G.P. and D.N. declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Matthew G. Perich or Juan A. Gallego.

Peer review information *Nature Neuroscience* thanks Joao Barbosa and Srdjan Ostojic for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.