Neural Population Geometry

ME 225NN, Winter 2025

Santiago Acosta[*1] and Jonathan Skaza[†1]

[1]Dynamical Neuroscience Graduate Program, University of California, Santa Barbara

1 Introduction

1.1 Problem Description & Motivation

Experimental neuroscience has seen rapid advances in recording techniques, enabling simultaneous recordings from thousands of neurons (and potentially up to a million [1]). These technological breakthroughs have driven theoretical neuroscientists to develop new methodologies for analyzing neural populations. Such large-scale analyses face distinct challenges: neurons typically respond to multiple variables simultaneously, complicating functional interpretation, while real-world tasks require robustness to complex variability that traditional tuning-based approaches struggle to capture.

In consideration of these factors, the focus of neural population analysis has transitioned from single-neuron tuning to *geometric* approaches that consider population-level representations [2, 3]. Traditional single-neuron tuning approaches characterize individual neurons by their response profiles to specific stimuli or task variables. For example, in the visual cortex, neurons might be classified by their preferred orientation, spatial frequency, or direction of motion, while in the motor cortex, neurons might be described by their correlation with movement direction or force. This approach, pioneered by Hubel and Wiesel's work on the visual cortex [4], has been instrumental in mapping functional properties across brain regions but struggles to capture the complex, multi-dimensional nature of neural responses during naturalistic behaviors.

The rationale behind the shift to population-level analysis is that neural computations emerge from structured, high-dimensional activity patterns rather than isolated responses of individual neurons. These patterns are conceptualized as *neural manifolds*—low-dimensional geometric structures embedded in high-dimensional neural state space. The geometric properties of these manifolds—dimensionality, curvature, separability, and capacity (the number of distinct manifolds a neural system can support)—reveal computational principles governing neuronal networks [5].

A compelling aspect of the neural manifold framework is its ability to link geometric structures to specific behavioral or cognitive variables. For example, ring-shaped manifolds have been observed in the mammalian hippocampus encoding head direction [6] and in the mouse visual cortex representing stimulus orientation [7]. These direct mappings between manifold geometry and task variables provide mechanistic insights into how neural populations encode and process information relevant to behavior.

This report examines the application of geometric techniques to analyze large neuronal populations—a field known as *neural population geometry*. This framework provides a unified approach for understanding both biological neural circuits and artificial neural networks (ANNs). By conceptualizing neural computation as geometric transformations, we can analyze how population-level representations enable robust, efficient, and scalable information processing across biological and artificial systems. This parallel analysis reveals fundamental computational principles that transcend the specific implementation details of different neural architectures.

---

[*]santiago_acosta@ucsb.edu

[†]skaza@ucsb.edu

1.2 Literature Review

1.3 Summary

In this paper, we demonstrate how neural population geometry provides a unified framework for understanding information encoding in both biological and artificial neural systems. We focus specifically on the representation of circular variables—a fundamental computational challenge that requires preserving topological structure. Through parallel analyses of orientation encoding in the mouse visual cortex and a trained convolutional neural network (CNN), we show how similar ring-shaped manifolds emerge in both systems despite their different implementations.

Our analysis reveals that these circular manifolds arise naturally from the computational demands of the task rather than from specific architectural constraints. In the mouse visual cortex, we characterize how populations of neurons collectively form a ring-shaped manifold that encodes orientation information, with the manifold's geometry reflecting the circular topology of orientation space. Similarly, in our CNN simulation, we demonstrate how the network spontaneously develops a circular representation in its latent space when trained to discriminate orientation angles.

By comparing the geometric properties of these representations—including dimensionality, curvature, and topological structure—we identify common principles that govern how neural systems, both biological and artificial, efficiently encode circular variables. This comparative approach highlights how neural population geometry can be applied to both neuroscience and machine learning, offering insights into fundamental computational strategies that transcend specific neural implementations.

2 Preliminaries

2.1 Mathematical Foundations of Neural Manifolds

In mathematics, a *manifold* is a topological space that locally resembles Euclidean space. Formally, an m-dimensional manifold $\mathcal{M}$ is a topological space where each point $p \in \mathcal{M}$ has a neighborhood U that is homeomorphic to an open subset of $\mathbb{R}^m$.

In neuroscience, the concept of a *neural manifold* extends this mathematical definition to describe geometric structures in neural population activity (see Fig. 1). Consider a population of n neurons, where the activity of each neuron can be represented as a coordinate in an n-dimensional neural state space $\mathbb{R}^n$. Let $\mathbf{r} = (r_1, r_2, \ldots, r_n) \in \mathbb{R}^n$ represent the firing rates of these neurons at a given time. The neural manifold hypothesis posits that, during specific task execution, these population activity patterns do not explore the entire n-dimensional space but instead are constrained to a lower-dimensional subspace or manifold $\mathcal{M} \subset \mathbb{R}^n$, where typically $\dim(\mathcal{M}) \ll n$.

Formally, we can define a neural manifold as follows: Let $\mathbf{r}(t) \in \mathbb{R}^n$ represent the population activity at time t. For a given task or stimulus condition s from a set of possible conditions $\mathcal{S}$, the neural manifold $\mathcal{M}_s$ is defined as:

$$\mathcal{M}_s = \{\mathbf{r}(t) \in \mathbb{R}^n : \text{condition } s \text{ is presented at time } t\} \tag{1}$$

It is important to note that real neural data exhibits characteristics that extend beyond the strict mathematical definition of a manifold in two key ways:

1. *Neural variability*: When the same stimulus is presented repeatedly, neural responses vary due to stochastic processes in neural firing. This variability is thought to be fundamental to neural coding and transforms what would be a sparse set of isolated points (one per stimulus) into richer *point clouds* (clusters of points for each stimulus). These point clouds collectively define the empirical neural manifold, with the distribution and statistical properties of responses containing important information about the neural code.

2. *Sparse sampling*: Experimental limitations often allow only a finite sampling of possible stimuli or conditions, providing discrete observations of what is theoretically a continuous underlying manifold structure.

2.2 Types of Neural Manifolds

Several types of neural manifolds have been identified in neuroscience research:

1. *Object manifolds* or *perceptual manifolds* arise in sensory neural populations as a result of identity-preserving variations in stimulus properties. For example, if $\mathbf{x} \in \mathbb{R}^d$ represents a stimulus (e.g., an image) and $T_\theta : \mathbb{R}^d \to \mathbb{R}^d$ represents a transformation with parameter θ (e.g., rotation), then the object manifold is defined as:

$$\mathcal{M}_{\text{obj}} = \{f(T_\theta(\mathbf{x})) : \theta \in \Theta\} \tag{2}$$

   where $f : \mathbb{R}^d \to \mathbb{R}^n$ represents the neural encoding function mapping stimuli to population responses, and Θ is the set of possible transformation parameters.

2. *Point-cloud manifolds* represent the empirical manifestation of neural manifolds in experimental data. Given a set of stimuli $\{s_1, s_2, \ldots, s_k\}$ and corresponding neural responses $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k\}$, the point-cloud manifold is:

$$\hat{\mathcal{M}} = \{\mathbf{r}_i : i = 1, 2, \ldots, k\} \tag{3}$$

   This discrete set of points approximates the underlying continuous manifold structure, with the approximation quality depending on sampling density and noise levels.

   To estimate the continuous manifold from these discrete point clouds, several approaches have been developed that focus on uncovering the intrinsic geometry of neural representations [5]:

   - *Linear dimensionality reduction*: Methods such as Principal Component Analysis (PCA) [8] identify the shape, location, and orientation of neural data within the neural state space by providing a Cartesian coordinate basis describing subspaces in which the data lie. While computationally efficient, these methods can only capture linear subspaces and may miss important nonlinear structure in the data.

   - *Nonlinear dimensionality reduction*: To characterize the intrinsic geometry of neural manifolds, which are often curved and nonlinear, methods such as Isomap [9], Locally Linear Embedding (LLE) [10], t-SNE [11], Multidimensional Scaling (MDS) [12], PHATE [13], and UMAP [14] can be employed. These techniques attempt to preserve local or global geometric relationships in the data while reducing dimensionality. However, they typically assume topologically simple manifolds and may fail to capture complex topological structures.

   - *Topologically-motivated methods*: For neural data with complex topological structure, specialized techniques have been developed. Spline Parameterization for Unsupervised Decoding (SPUD) [6] uses cubic splines to parameterize low-dimensional manifolds, enabling the discovery of ring-shaped manifolds in head direction cells. This approach leverages persistent homology [15, 16] to determine the intrinsic dimension and uncover nontrivial topological structures in neural data.

   - *Dynamics-based approaches*: Methods like Manifold Inference from Neural Dynamics (MIND) [17, 18] define distances between neural states based on transition probabilities, which gives rise to intrinsic dimensions relevant for topological maps underlying task implementation. This approach has been successfully applied to characterize neural activity in the hippocampus during spatial navigation tasks, revealing toroidal manifold structures in grid cells.

   - *Geometric approximation methods*: Traditional approaches such as convex hulls [19] and alpha shapes [20] provide geometric approximations of the manifold boundary. While convex hulls offer a simple outer boundary, alpha shapes can capture non-convex features by removing simplices whose circumscribing spheres exceed a threshold radius α. These methods are particularly useful for visualizing and analyzing the shape of neural manifolds.

- *Density-based approaches*: For noisy neural data, methods like kernel density estimation (KDE) [21] create continuous representations by estimating the probability density function of the point cloud. The manifold boundary can then be defined as a level set of this density estimate, providing robustness to outliers and noise.

The choice of estimation method depends on the hypothesized structure of the neural manifold and the characteristics of the data. Linear methods are appropriate when the manifold is expected to be flat, while nonlinear methods are necessary for curved manifolds. Topologically-motivated methods are essential when the manifold has complex topology, such as circular or toroidal structures often found in neural representations of periodic variables like orientation or spatial location [6, 18]. For a comprehensive review of manifold estimation techniques in neuroscience, see [8, 22].

2.3 Neural Population Geometry

Neural population geometry refers to the study of configurations, relationships, and properties of neural manifolds embedded in the high-dimensional neural state space. This framework provides quantitative measures to characterize neural representations:

1. *Dimensionality*: The intrinsic dimensionality d of a neural manifold $\mathcal{M}$ can be estimated using techniques such as principal component analysis (PCA) or more sophisticated nonlinear dimensionality reduction methods.

2. *Curvature*: For a smooth manifold, the curvature at point $p \in \mathcal{M}$ quantifies how the manifold deviates from a flat Euclidean space locally. Intuitively, curvature measures how much a surface "bends" at a given point. In neural manifolds, regions of high curvature often represent areas where small changes in stimuli produce large changes in neural responses. This can be characterized mathematically using concepts like the second fundamental form or Ricci curvature, which provide formal ways to measure this bending in different directions.

3. *Separability*: For multiple manifolds $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k$ corresponding to different stimulus conditions, separability measures how distinguishable these manifolds are. One metric is the minimum distance between manifolds:

$$d(\mathcal{M}_i, \mathcal{M}_j) = \min_{\mathbf{r}_i \in \mathcal{M}_i, \mathbf{r}_j \in \mathcal{M}_j} \|\mathbf{r}_i - \mathbf{r}_j\| \tag{4}$$

4. *Capacity*: The number of distinct manifolds a neural system can support, which relates to the information-carrying capacity of the population.

The geometric properties of neural manifolds provide insights into the computational principles governing neural information processing, including how information is encoded, transformed, and decoded across neural circuits.

3 Mouse Experiment

4 Simulation Study

To demonstrate the emergence of geometric structures in neural representations of ANNs, we implemented a simulation experiment investigating how ANNs encode "circular variables". This experiment provides an example of neural population geometry principles and illustrates how topological structures may naturally emerge in neural representations in response to the nature of the task.

We trained a CNN to predict the angle of orientation, θ, of visual grating stimuli, a task analogous to orientation selectivity in the visual cortex. The network was presented with sinusoidal gratings at various orientations ranging from 0 to 2π rad and tasked with predicting the orientation angle. Fig. 2 contains example stimuli.
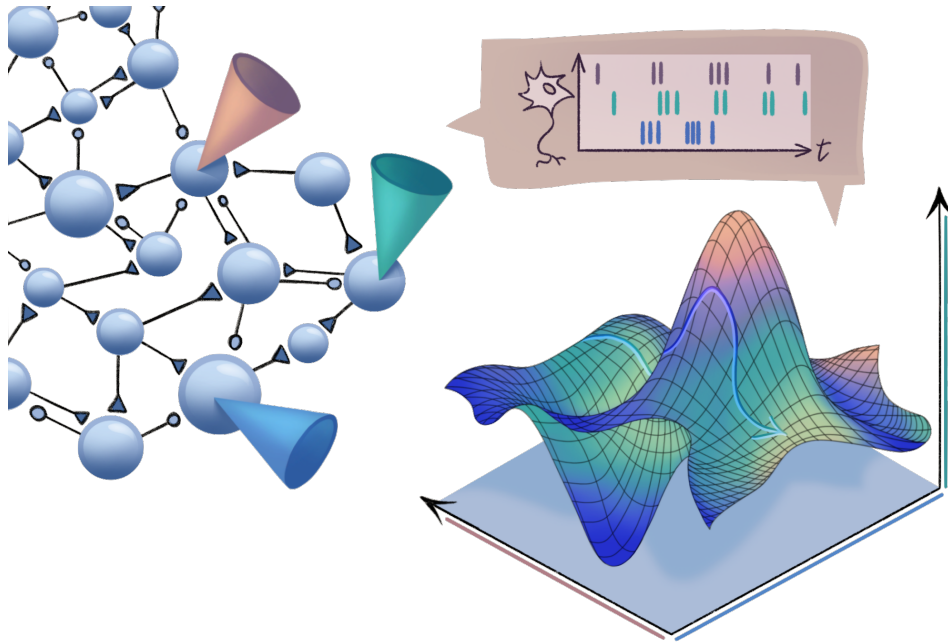
Figure 1: A neural manifold is formed by plotting neural activity (e.g., firing rate) in state space, where each axis is a single neuron. Note that the repeated presentation of the same stimulus does not lead to the occupancy of an identical point within the state space. Rather, neuronal variability induces fluctuations in the points derived from different trials. Consequently, each stimulus is associated not with a single point but with a point cloud, the dimensions and configuration of which are contingent upon the magnitude and nature of the neuronal variability. One may be interested in the manifold that emerges in response to a particular stimulus or whether there is separability among manifolds from multiple stimuli. Fig. from [23].
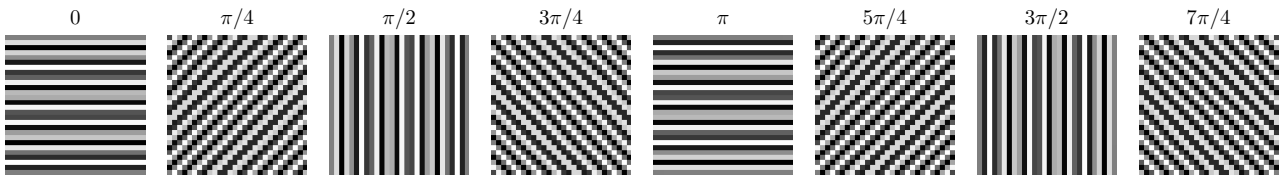

Figure 2: Sample grating stimuli at different orientations.

The architecture consisted of convolutional layers followed by fully connected layers, with a 32-dimensional latent space that we analyzed for geometric properties. Importantly, the network was trained to predict the sine and cosine components of the orientation angle rather than the raw angle, which naturally handles the circular topology of the orientation space[1].

After training, we extracted the 32-dimensional latent representations for test stimuli and applied dimensionality reduction techniques (PCA, Isomap, and MDE) to visualize the structure of these representations. The results revealed a striking circular manifold in the latent space, where stimuli with similar orientations were positioned close together, and the entire orientation space was represented as a continuous ring.

This circular topology emerged naturally from the training process, despite no explicit topological constraints being imposed on the network. Through backpropagation, the network discovered that a circular representation is the most efficient way to encode a periodic variable, demonstrating how the geometry of the task space (in this case, the circular nature of orientation angles) becomes reflected in the geometry of the neural representations.

The experiment highlights several key principles of neural population geometry. First, the network's

---

[1]For example, for small ϵ, the angles of ϵ and $2\pi - \epsilon$ are nearly identical orientations but numerically distant. Using sine-cosine encoding, these angles map to points $(\cos(\epsilon), \sin(\epsilon)) \approx (\cos(2\pi - \epsilon), \sin(2\pi - \epsilon)) \approx (1, 0)$, i.e., the orientations become close in Euclidean space, better reflecting their similarity.
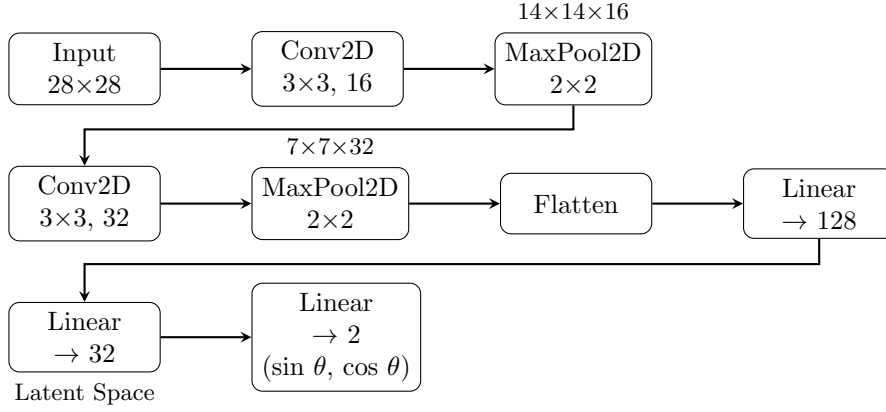
Figure 3: Neural network architecture used for the circular manifold experiment. The network processes grating stimuli through convolutional layers, followed by fully connected layers that produce a 32-dimensional latent space. The output layer predicts the sine and cosine of the orientation angle.
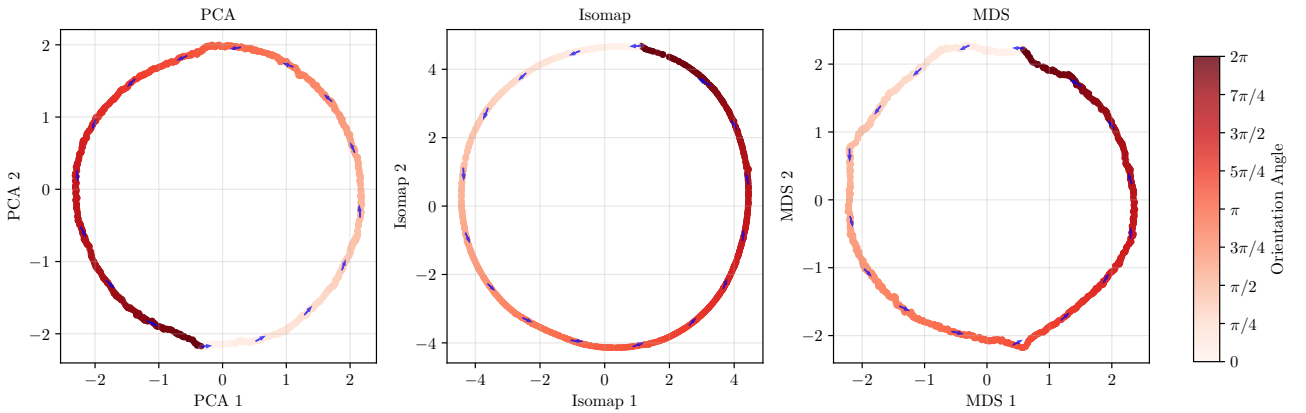

Figure 4: Dimensionality reduction of the CNN's 32-dimensional latent space representations reveals a circular manifold structure. Each point along the curve represents a grating stimulus, colored according to its true orientation angle. The three techniques (PCA, Isomap, and MDS) all recover the circular topology, demonstrating that the network has learned to encode the periodic orientation variable as a ring-shaped manifold. Blue arrows indicate the direction of increasing orientation angle along the manifold.

internal representations form a low-dimensional manifold (a circle) embedded in the high-dimensional latent space. Second, the topology of the manifold (circular) matches the topology of the task space (orientation angles), demonstrating topological correspondence. Third, similar stimuli are mapped to nearby points on the manifold, preserving the continuity of the input space through continuous representation. Finally, the network learns to compress the high-dimensional input (grating images) into a low-dimensional representation that captures the essential variable (orientation), effectively performing dimensionality reduction.

This simple experiment provides an intuitive foundation for understanding how more complex neural networks might represent more complicated manifolds in their latent spaces, and how the geometry of these representations relates to the computational tasks being performed.

5 Conclusion

References

[1] Jeffrey Demas, Jason Manley, Frank Tejera, Kevin Barber, Hyewon Kim, Francisca Martínez Traub, Brandon Chen, and Alipasha Vaziri. High-speed, cortex-wide volumetric recording of

neuroactivity at cellular resolution using light beads microscopy. *Nature Methods*, 18(9):1103–1111, 2021.

[2] Rafael Yuste. From the neuron doctrine to neural networks. *Nature reviews neuroscience*, 16(8):487–497, 2015.

[3] Shreya Saxena and John P Cunningham. Towards the neural population doctrine. *Current opinion in neurobiology*, 55:103–111, 2019.

[4] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, 1959.

[5] SueYeon Chung and Larry F Abbott. Neural population geometry: An approach for understanding biological and artificial neural networks. *Current opinion in neurobiology*, 70:137–144, 2021.

[6] Rishidev Chaudhuri, Berk Gerçek, Biraj Pandey, Adrien Peyrache, and Ila Fiete. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature neuroscience*, 22(9):1512–1520, 2019.

[7] Kosio Beshkov, Marianne Fyhn, Torkel Hafting, and Gaute T Einevoll. Topological structure of population activity in mouse visual cortex encodes densely sampled stimulus rotations. *Iscience*, 27(4), 2024.

[8] John P Cunningham and Byron M Yu. Dimensionality reduction for large-scale neural recordings. *Nature neuroscience*, 17(11):1500–1509, 2014.

[9] Joshua B Tenenbaum, Vin De Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[10] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[11] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[12] Joseph B Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

[13] Kevin R Moon, David van Dijk, Zheng Wang, Scott Gigante, Daniel B Burkhardt, William S Chen, Kristina Yim, Antonia van den Elzen, Matthew J Hirn, Ronald R Coifman, et al. Visualizing structure and transitions in high-dimensional biological data. *Nature biotechnology*, 37(12):1482–1492, 2019.

[14] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[15] Herbert Edelsbrunner and John Harer. Persistent homology—a survey. *Contemporary mathematics*, 453:257–282, 2008.

[16] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[17] Rishidev Chaudhuri and Ila Fiete. The intrinsic dimensionality of neural representations from population activity. *Neuron*, 97(5):1006–1007, 2018.

[18] Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. Toroidal topology of population activity in grid cells. *Nature*, 602(7895):123–128, 2022.

[19] C Bradford Barber, David P Dobkin, and Hannu Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4):469–483, 1996.

[20] Herbert Edelsbrunner and Ernst P Mücke. Three-dimensional alpha shapes. *ACM Transactions on Graphics (TOG)*, 13(1):43–72, 1994.

[21] Larry Wasserman. *All of nonparametric statistics*. Springer Science & Business Media, 2006.

[22] Juan A Gallego, Matthew G Perich, Lee E Miller, and Sara A Solla. Neural manifolds for the control of movement. *Neuron*, 94(5):978–984, 2017.

[23] Matthew Perich. Neural manifolds: Latest buzzword or pathway to understand the brain? *The Transmitter*, 2024.