Extended Abstract Track

Topological Neural Data Analysis with Behavioral Constraint

Editors: List of editors' names

Abstract

Recently, Topological Data Analysis (TDA) has revealed insights into the topological structure of neural population activity. However, existing TDA methods for neural population activity are computationally demanding, noise-sensitive, and sometimes difficult to interpret. We develop a simple and more interpretable analysis approach to infer the topological structure of behaviorally relevant neural response variability. Our approach first maps the neural activity onto firing rate maps of behavioral variables, and then performs analysis based on these rate maps. Application of our method to grid cell recordings demonstrates its effectiveness without sophisticated preprocessing as required in prior methods. Further test of the methods based on synthetic data suggests that our method is more informative of the deviations from standard topological shapes. Our results also point to the importance of joint analysis of the geometry and topology of neural manifolds.

Keywords: Grid Cell, Neural Manifold, Topological Data Analysis, Neural Data Science

1. Introduction

While classic work in neuroscience emphasized individual neurons, recently there is surge of interest to understand the population-level structure of neural activity (Vyas et al., 2020) by conceptualizing it as a "neural manifold" (Seung and Lee, 2000; Kriegeskorte and Wei, 2021; Perich et al., 2025; Chung and Abbott, 2021). At each moment, the neural population activity represents a point on the manifold, which is determined by stimulus, internal states, and noise. The geometrical (Kriegeskorte and Wei, 2021) and topological (Giusti et al., 2015) structure of neural manifolds can be informative of the underlying neural computation.

Advances in modern recording techniques for simultaneously recording large populations of neurons (Grienberger and Konnerth, 2012; Jun et al., 2017) provide unprecedented opportunity to investigate the structures of neural manifolds. Recent studies applied Topological Data Analysis (TDA; Wasserman (2017)) to analyze neural data in various neural systems, e.g., V1 (Singh et al., 2008), the hippocampus (Giusti et al., 2015), head direction cells (Chaudhuri et al., 2019), and grid cells (Gardner et al., 2022). In particular, Gardner et al. (2022) provided evidence for toroidal structures in the grid cells by evaluating the persistent homology (Zomorodian and Carlsson, 2004). However, existing methods based on neural population activity are prone to noise and rely on the pre-processing steps, thus complicating the interpretation and the reliability of the result.

Here, we introduce a simple method that extracts the topological structures of neural population activity that are relevant for encoding certain behavioral variables. Our approach achieves comparable topology characterization with simpler data curation steps compared to previous methods, while enabling robust interpretation of topological features. We also demonstrate a challenge for Neural TDA: when neural manifolds deviate from standard shapes (e.g., circle/torus), it can be difficult to detect these deviations with TDA.

We show that our method better captures these deviations. These findings suggest that future research can benefit from the joint analysis of the geometry and topology of neural manifolds (Kriegeskorte and Wei, 2021; Ye and Wessel, 2025).

2. Methods

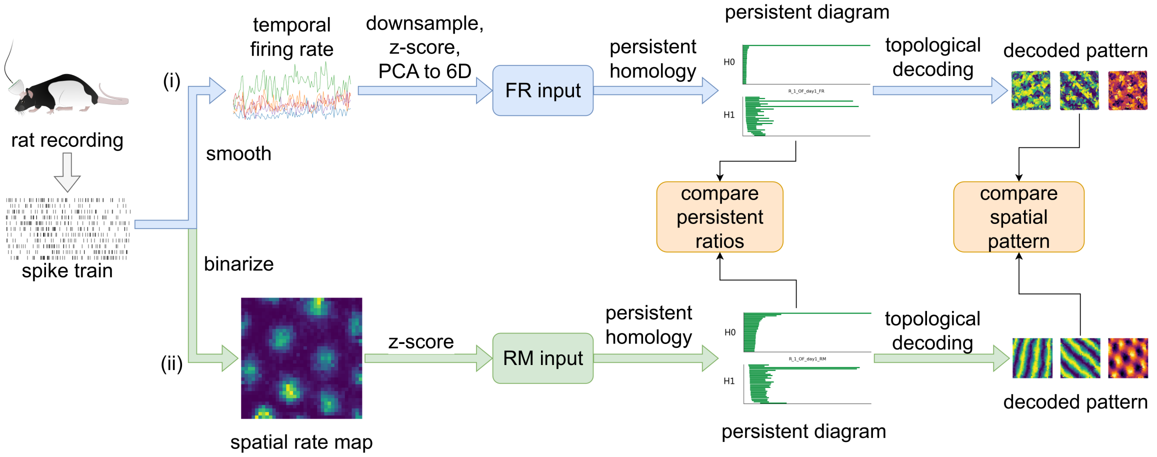


Figure 1: A comparison of the two TDA pipelines.

Analysis pipeline Fig. 1 illustrates the comparison of the previous analysis approach and ours. For the previous approach (see Fig. 1(i)) developed in (Gardner et al., 2022), first the firing rate of individual neurons were computed based on the spike trains. Substantially downsampling were needed to select the most active time bins. After further downsampling, z-scoring, and projecting the data onto the first 6 principal components, the *persistent diagrams* were computed. The *persistent barcodes* in a persistent diagram indicate topological features of a certain dimension (see Appendix B.4). When applying to grid cell data, two outstanding long barcodes in the H1 diagram indicated two independent circular features, jointly composing the torus manifold. Each circular feature was used to assign a circular parameter to all data points. The dual circular parameters were summed up across neurons to compute two arrays of spatial firing strengths over time bins, visualizing two circular firing patterns with a 60° angle, together yielding a hexagonal grid pattern.

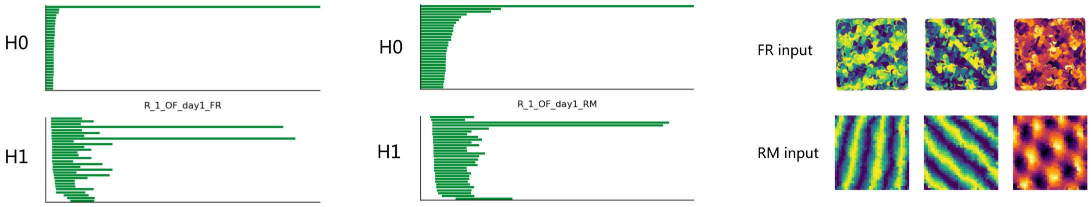
In our analysis pipeline (Fig. 1(ii)), we first project the neural activity onto the spatial locations to yield the firing rate maps of individual neurons. We then compute the persistent homology using the z-scored firing rate maps. No additional preprocessing steps are needed. By projecting the neural activity onto behavioral variables, our method enables analyzing topological features of behavior-relevant neural variability.

Persistence ratio Since the length of a persistent barcode can indicate the significance of a topological feature, we propose the following metric to evaluate the relative significance of a topological feature: we first rank the lengths of all the barcodes in the bottom H1 diagram (see Fig. 1), and then compute the ratio of the i-th and $(i+1)$-th longest barcodes. We will refer to it as the *i-th persistence ratio*, denoted as $\text{PR}(i)$. A larger $\text{PR}(i)$ indicates stronger significance of the first i dominant circular features relative to other features, thus indicating the whole manifold closer to a product of i circles in its shape. For instance, the ratio of lengths between the second and third longest barcodes in the H1 persistent diagram is the second persistence ratio $\text{PR}(2)$, and a manifold with a larger $\text{PR}(2)$ has a shape closer to a torus.

3. Results

Rat grid cell recordings We analyzed the Neuropixels (Jun et al., 2017) recordings of rat MEC provided in (Gardner et al., 2022). We performed TDA analyses on all the open-field (OF) foraging task recordings. See Appendix B.2 about the dataset.

We first replicated the results reported in Gardner et al. (2022) by following their exact analysis procedure. The firing rate (FR) inputs to TDA computation are arrays of $[1200, N]$ where N is the number of recorded neurons in one session. For fair comparisons, in our method (RM), we chose a spatial bin size so the dimensionality of the RM inputs are comparable to theirs. Fig. 2 reports the TDA results from the OF session of rat R, module 1, recording day 1. Fig. 2(a) shows the persistent diagrams computed with firing rate (FR) input. Two significantly longer barcodes in the bottom (H1) diagram indicates two dominant circular feature that are irrelevant to each other, hence a toroidal topological feature. Fig. 2(b) displays persistent diagrams computed with our rate-map (RM) method. Fig. 2(c) depicts decoded circular features for FR (top) and RM (bottom) inputs. Two circular patterns tend to have a 60 degree angle, and form a hexagonal grid together. The results with all OF recording sessions are shown in Fig. 4. We find the mean PR(2) for FR input is slightly higher than that for the RM input (3.43 v.s. 3.13), suggesting that method based on FR has more evidence for toroidal topological shapes. For decoded circular patterns, the results from our proposed RM method are significantly cleaner.

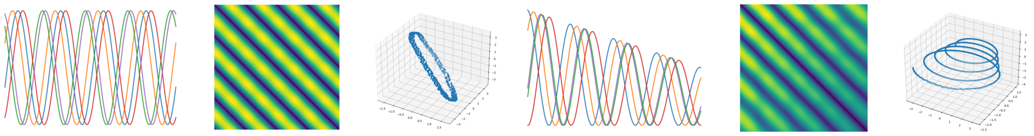


(a) FR input (b) RM input (c) Decoded patterns

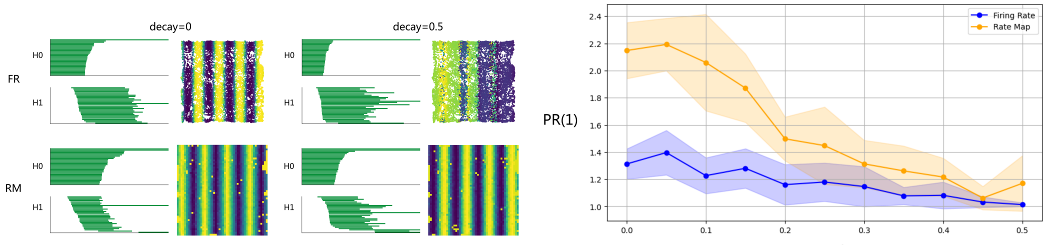
Figure 2: A comparison of two TDA methods with different types of input.

We further analyzed the robustness of two methods when removing certain data curation processes. Results in Fig. 5 suggest that our method, while simple, can robustly recover clear topological features. In contrast, firing-rate-based method may fail to recover clear features when missing certain preprocessing steps.

Synthetic heterogeneous 1D "grid cells" While applying TDA on grid cells reveals a toroidal feature, in reality the underlying topology often deviates from a perfect torus due to the heterogeneity within firing rate tuning. To examine how well the TDA methods can detect deviations from simple shapes like circle/torus, we conducted an analysis on simulated data, focusing on populations of 1-dimensional "circular cells" due to their simplicity and high interpretability.



(a) Tuning geometry with 0 decay (b) Tuning geometry with 0.5 decay



(c) Decoded patterns with heterogeneity (d) PR(1) of FR and RM inputs

Figure 3: Visualization of heterogeneous tuning amplitude.

In this simulation (for details see Appendix B.3), we introduce heterogeneity in the response gain of individual firing fields for each neuron. Specifically, we construct the tuning curves by multiplying a periodic tuning with a linear function over space. The slope of the linear function (linear decay factor) controls the heterogeneity of firing. Importantly, when the decay factor is larger than 0, the manifold is topologically equivalent to a line, but not a cicle. As the decay factor increases, the ground-truth manifold becomes like a spring being stretched by force. Fig. 3(b) shows the representational distance matrix and a 3D visualization using Multi-dimensional Scaling (Kruskal and Wish, 1978) when the decay factor is 0.5, from which it is clear that the geometry deviates from a circle.

Fig. 3(c) shows the representative persistent diagrams and decoding results with heterogeneity. The results suggest that the H1 (bottom) diagram becomes messy with an induced decay factor, and our RM-based method can better decode the circular pattern. We then systematically analyzed models with a range of decaying factors from 0 to 0.5 with a 0.05 increment, repeating the computation of persistent diagrams 10 times for each setting. Fig. 3(d) shows the inferred persistence ratios for different decaying factors. We find that the circular feature becomes less dominant with the decay, in accordance with tuning geometries shown above. While the persistence ratio inferred from both methods are affected by increasing heterogeneity, our method is substantially more informative of the change of the decaying factor and thus the deviations in tuning amplitudes.

Together, these results suggest that practically it may be challenging for the existing TDA method to detect mild deviations from standard shapes. Our RM-based method is more informative for revealing these deviations, at least for the settings we have tested. Integrating TDA with the analysis of geometry may lead to more powerful approaches.

References

Ulrich Bauer. Ripser: efficient computation of Vietoris–Rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, 6 2021. doi: 10.1007/s41468-021-00071-5. URL https://doi.org/10.1007/s41468-021-00071-5.

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, 8 2019. doi: 10.1038/s41593-019-0460-x. URL https://doi.org/10.1038/s41593-019-0460-x.

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.

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, 1 2022. doi: 10.1038/s41586-021-04268-7. URL https://doi.org/10.1038/s41586-021-04268-7.

Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44):13455–13460, 2015.

Christine Grienberger and Arthur Konnerth. Imaging calcium in neurons. *Neuron*, 73(5): 862–885, 2012.

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2. URL https://doi.org/10.1038/s41586-020-2649-2.

James J. Jun, Nicholas A. Steinmetz, Joshua H. Siegle, Daniel J. Denman, Marius Bauza, Brian Barbarits, Albert K. Lee, Costas A. Anastassiou, Alexandru Andrei, Çağatay Aydın, Mladen Barbic, Timothy J. Blanche, Vincent Bonin, João Couto, Barundeb Dutta, Sergey L. Gratiy, Diego A. Gutnisky, Michael Häusser, Bill Karsh, Peter Ledochowitsch, Carolina Mora Lopez, Catalin Mitelut, Silke Musa, Michael Okun, Marius Pachitariu, Jan Putzeys, P. Dylan Rich, Cyrille Rossant, Wei-Lung Sun, Karel Svoboda, Matteo Carandini, Kenneth D. Harris, Christof Koch, John O'Keefe, and Timothy D. Harris. Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236, 11 2017. doi: 10.1038/nature24636. URL https://doi.org/10.1038/nature24636.

Nikolaus Kriegeskorte and Xue-Xin Wei. Neural tuning and representational geometry. *Nature Reviews Neuroscience*, 22(11):703–718, 2021.

Joseph Kruskal and Myron Wish. *Multidimensional scaling*. 1 1978. doi: 10.4135/9781412985130. URL https://doi.org/10.4135/9781412985130.

Matthew G. Perich, Devika Narain, and Juan A. Gallego. A neural manifold view of the brain. *Nature Neuroscience*, 28(8):1582–1597, August 2025. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-025-02031-z. URL https://www.nature.com/articles/s41593-025-02031-z.

H Sebastian Seung and Daniel D Lee. The manifold ways of perception. *science*, 290(5500):2268–2269, 2000.

Gurjeet Singh, Facundo Memoli, Tigran Ishkhanov, Guillermo Sapiro, Gunnar Carlsson, and Dario L Ringach. Topological analysis of population activity in visual cortex. *Journal of vision*, 8(8):11–11, 2008.

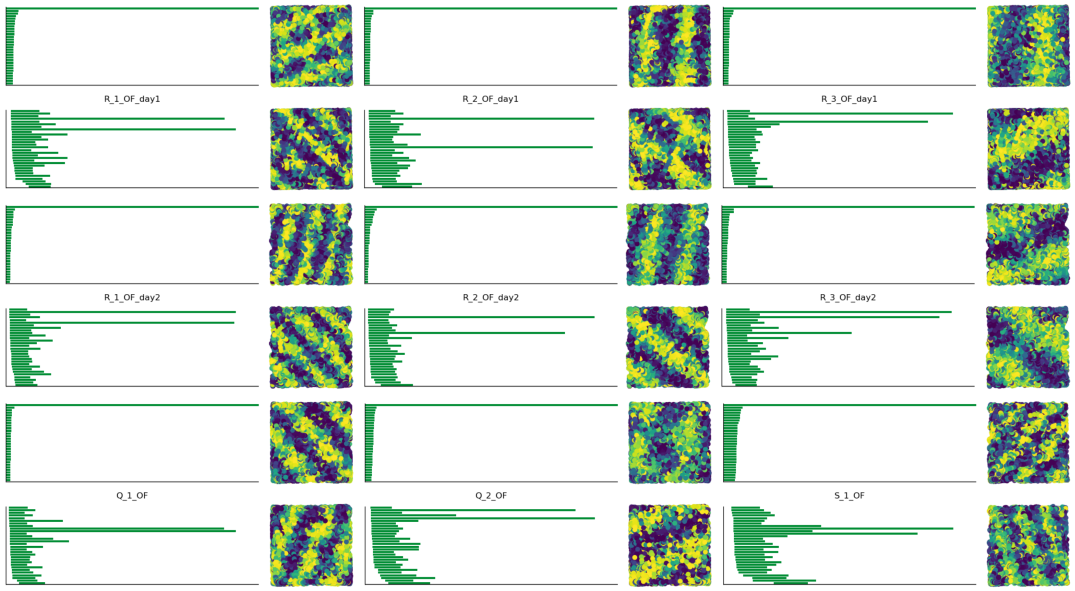
Saurabh Vyas, Matthew D. Golub, David Sussillo, and Krishna V. Shenoy. Computation through neural population dynamics. *Annual Review of Neuroscience*, 43(Volume 43, 2020):249–275, 2020. ISSN 1545-4126. doi: https://doi.org/10.1146/annurev-neuro-092619-094115. URL https://www.annualreviews.org/content/journals/10.1146/annurev-neuro-092619-094115.

Larry Wasserman. Topological data analysis. *Annual Review of Statistics and Its Application*, 5(1):501–532, 12 2017. doi: 10.1146/annurev-statistics-031017-100045. URL https://doi.org/10.1146/annurev-statistics-031017-100045.

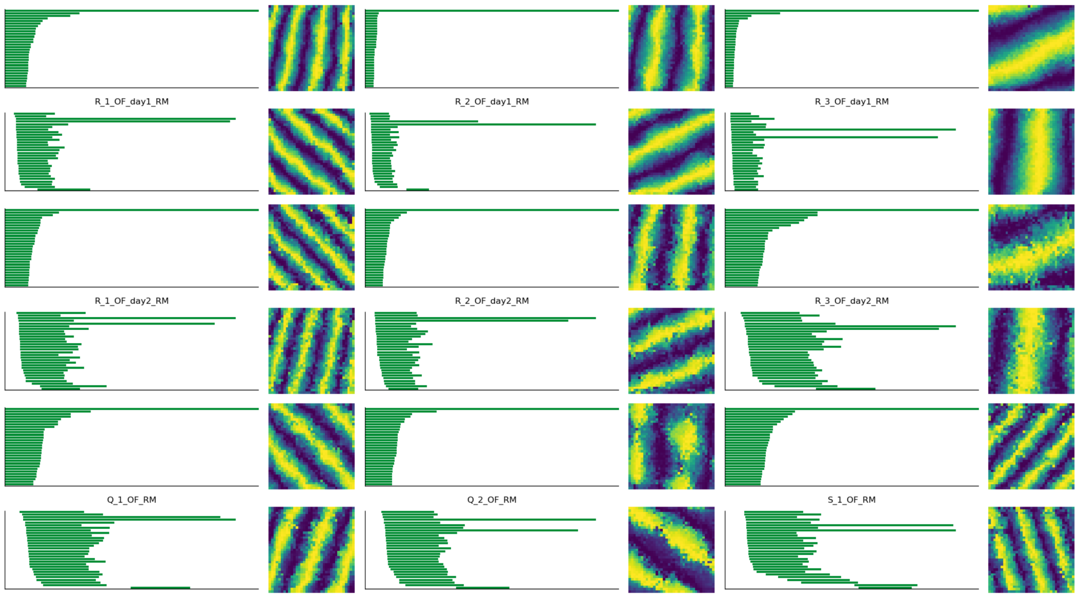
Zeyuan Ye and Ralf Wessel. Speed modulations in grid cell information geometry. *Nature Communications*, 16(1):7723, 2025.

Afra Zomorodian and Gunnar Carlsson. Computing Persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 11 2004. doi: 10.1007/s00454-004-1146-y. URL https://doi.org/10.1007/s00454-004-1146-y.

Appendix A. Extended figures and tables



(a) TDA results with FR input



(b) TDA results with RM input

Figure 4: Computed persistent diagrams of all the open-field (OF) sessions, together with two circular patterns decoded from the two longest barcodes in the bottom (H1) diagram.

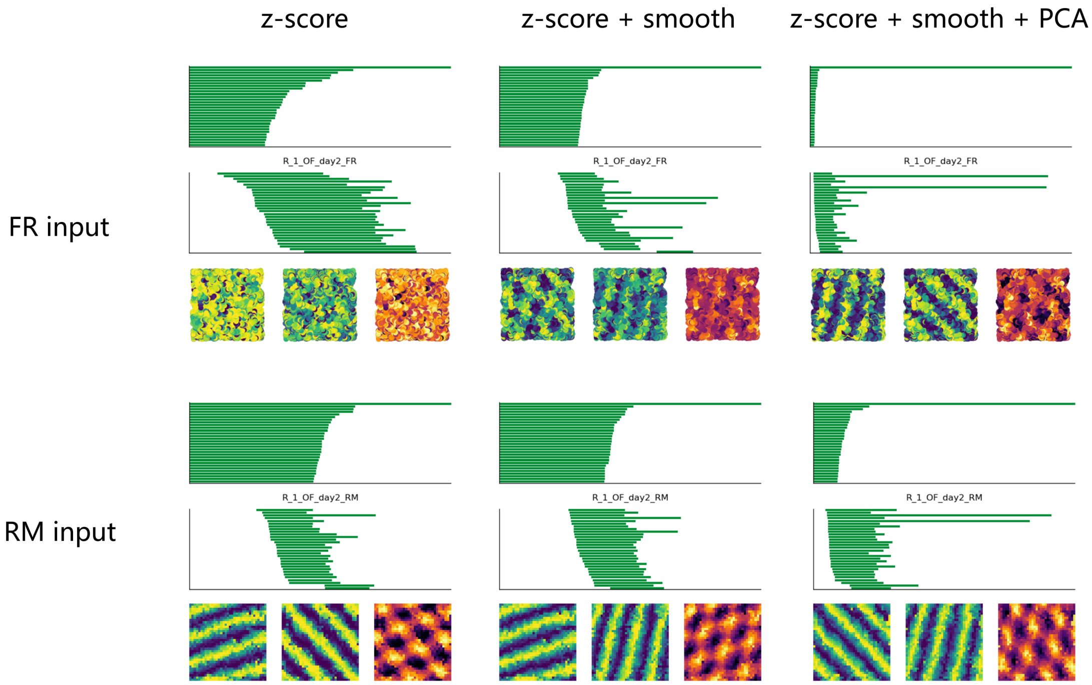


Figure 5: Computed persistent diagrams of rat R, grid cell module 1, open field session in recording day 2; together with two circular patterns decoded from the two longest barcodes in the bottom (H1) diagram, and their combined grid structure. FR inputs are downsampled to 1,200 time points in advance, while RM inputs are set as 1,225 (35×35) spatial bins. Each row: results with firing rate and rate map inputs. Each column: results with different data curation steps.

Appendix B. Experimental details

B.1. Code base and computing devices

We use NumPy (Harris et al., 2020) to store, formulate, and compute our data as arrays, and Ripser (Bauer, 2021) to perform all the topological data analysis and topological decoding steps. All the code runs on a CPU with 32GB RAM.

B.2. Grid cell recording dataset

The grid cell recording dataset is provided in (Gardner et al., 2022), and can be available at `https://figshare.com/articles/dataset/Toroidal_topology_of_population_activity_in_grid_cells/16764508`.

The dataset consists of recordings of rat Medial Entorhinal Cortex (MEC) with Neuropixels silicon probes (Jun et al., 2017). In the dataset, there are recordings of 3 animals named “rat Q”, “rat R”, and “rat S”. The grid cells in rat Q are classified as 2 modules, and the grid cells in rat R are classified as 3 modules. The recording sessions cross 2 days, and are classified as open-field foraging sessions (OF), wagon-wheel foraging sessions (WW), rapid-eye-movement sleeping sessions (REM), and slow-wave sleeping sessions (SWS). For a detailed introduction of the experimental techniques and procedures, please refer to “Methods” of (Gardner et al., 2022).

In (Gardner et al., 2022), the authors retracted toroidal topological features from all of the recording sessions with firing rate based inputs. While our rate map based method requires an input of animal movement positions, so we only performed our method with OF sessions in comparison with their OF results.

B.3. Simulation settings

We performed simulations of circularly tuned “1-dimensional grid cells”, or “circular cells”.

We first constructed circular tuning curves for 100 artificial “circular cells”, while each cell has a random shift in its tuning phase, sampled from a uniform distribution of $[0, 2\pi]$. After that, we apply a linear decay on the tuning curves, with a heterogeneity factor set as the lowest peak amplitude.

We then apply the spatial tuning curves on the real movement trajectory of rat R, OF session, day 1 to obtain a ground-truth firing rate. After that, we apply a Poisson firing model to get spike trains of 100 circular cells. We then smooth the spike trains into a $[15000, 100]$ empirical firing rate tensor.

The firing rate tensor was z-scored, projected to 6 principal components, and downsampled to shape $[2500, 6]$ as the FR input; and the firing rate tensor was binned against 50×50 positional bins to form a firing rate map, z-scored, projected to 6 principal components, and sent in as a shape $[2500, 6]$ RM tensor.

B.4. Topological analysis

Here we introduce the topological analysis process implemented in our work.

Topology configuration After applying topological data analysis (TDA) to the neural activity, we obtain two persistent diagrams denoted as H0 and H1 diagrams, each indicating the 0-th and first dimensional topological features. Intuitively, H0 features indicate disconnected clusters within the data, while H1 topological features correspond to circles in the data manifold.

Since our focus is on neural manifolds with hypothesized toroidal (grid cells) or circular ("circular cells") topology, we analyze the longest bars in the H1 diagram to determine whether the data supports the presence of one or two prominent circles. To quantify this, we rank the lengths of the persistent barcodes in the H1 diagram, then compute the PR(1) and PR(2) values to evaluate the dominance of a "single circle" or "two disentangled circles (forming a torus)" in the topological shape.

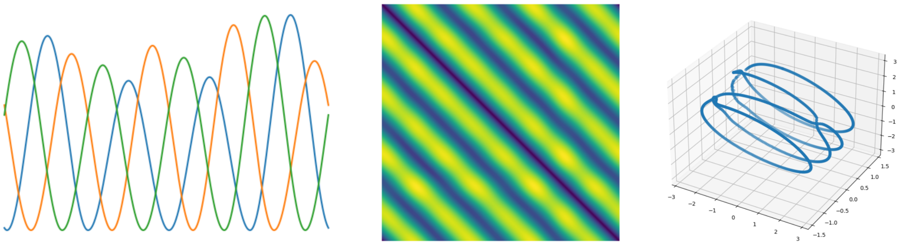
Topological decoding Once circular features are identified, the next step is to interpret what they represent. The decoding procedure follows Gardner et al. (2022).

Each circular feature assigns a *cocycle* value to the data points, with which we solve a least-square problem ($Ax = b$) to obtain an angular parameter for each data point, so that the angular parameters can span the edges to obtain the distance matrix among cocycles.

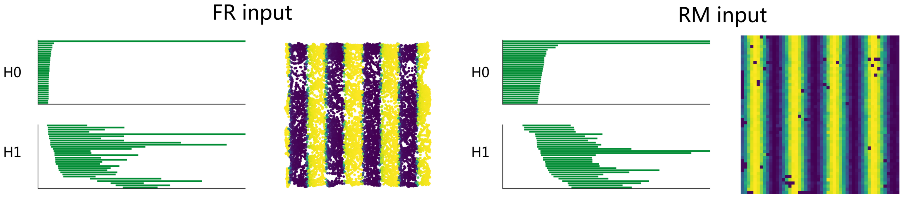
To connect these circular parameters with spatial behavior, we compute weighted activity maps. Specifically, we compute a scaled activity value for each spatial bin across all neurons by multiplying the RM value with the angular parameter. Finally, by summing up across all neurons, we can obtain spatial coordinates associated with each spatial bin, and visualize them as a circular spatial firing pattern.

In (Gardner et al., 2022), however, the FR input was first downsampled to 1,200 temporal bins. To extend decoding to the full dataset, they introduced an interpolation step: each neuron was assigned a weighted factor by summing up the scaled values across downsampled temporal bins, then the original firing rates were scaled with the weighted factors, where each factor is shared for all original activity across one neuron. Then temporal coordinates were computed by summing up across all neurons, and scattered on a 2D open-field map in correspondence to their x-y positions to visualize a circular spatial firing pattern.

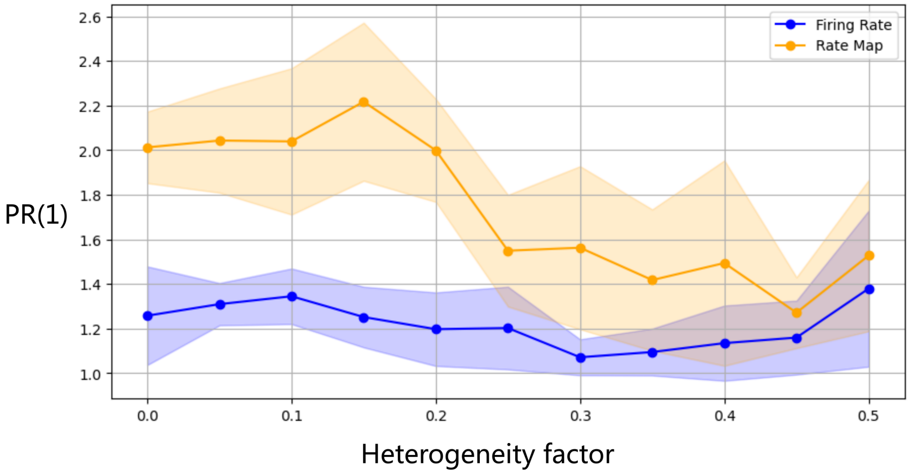
Appendix C. Alternative simulation experiments



(a) Tuning geometry with decay factor 0.5



(b) Decoded patterns with fluctuation



(c) PR(1) of FR and RM inputs

Figure 6: A demonstration of “fluctuating” tuning curve simulation

In addition to the “linear decay” simulation scheme described in the main text, we also implemented an alternative simulation setting. Each artificial “circular cell” retains an intrinsic phase shift sampled from a uniform distribution over $[0, 2\pi]$. To introduce heterogeneity in tuning amplitude, we assign a “fluctuation factor” within $[0, 1]$ and sample a scaler for each waveform (between two local minima) of a circular cell, hereby allowing each waveform to exhibit a random scaling of tuning amplitude. We argue that this mechanism of tuning heterogeneity better approximates realistic neural variability compared to the linear decay scheme, although its effects on the topology of tuning curves, firing rates, and open-field rate maps are less straightforward and need further investigation.

Following the simulation, we derived empirical firing rate tensors and converted them into FR and RM inputs using the same preprocessing steps described in Appendix B.3. We then applied the same TDA analysis and topological decoding procedure in the main text.

As shown in Fig. 6(*a*), with a fluctuation factor of 0.5, the tuning geometry deviates noticeably from a perfect ring topology. The tuning curves of three representative circular cells illustrate variability in both phases and amplitudes across neurons. Correspondingly, the distance matrix and MDS visualization of the tuning curves also deviate from the idealized setting. In Fig. 6(*b*), we present the persistent diagrams computed with FR and RM inputs. Notably, the RM-based method recovers a one-dimensional circular spatial pattern with discernible color gradient, whereas the FR-based decoding result presents a step-like switching pattern. Finally, Fig. 6(*c*) plots the average PR(1) values across 11 fluctuation factors with 10 repeated runs each. The results indicate that our RM-based method captures the influence of tuning heterogeneity more effectively, demonstrating greater sensitivity to perturbations under behavioral constraints in neural data.