

ZIGZAG PERSISTENCE OF NEURAL RESPONSES TO TIME-VARYING STIMULI

Yuri Gardinazzi

Area Science Park, Padriciano 99, Trieste; University of Trieste, Department of Mathematics, Informatics and Geosciences

Alessio Ansuini

Area Science Park, Padriciano 99, Trieste

Eugenio Piasini

SISSA, Via Bonomea 265, Trieste

Fabio Anselmi

University of Trieste, Department of Mathematics, Informatics and Geosciences, Via Economo 12/3, Trieste

Matteo Biagetti

MATTEO.BIAGETTI@AREASCIENCEPARK.IT

Area Science Park, Padriciano 99, Trieste

Editors: Michael Bleher, Freya Jensen, Levin Maier, Diaaeldin Taha, and Anna Wienhard

ABSTRACT

We use topological data analysis to study neural population activity in the Sensorium 2023 dataset, which records responses from thousands of mouse visual cortex neurons to diverse video stimuli. For each video, we build frame-by-frame cubical complexes from neuronal activity and apply zigzag persistent homology to capture how topological structure evolves over time. These dynamics are summarized with persistence landscapes, providing a compact vectorized representation of temporal features. We focus on one-dimensional topological features—loops in the data—that reflect coordinated, cyclical patterns of neural co-activation. To test their informativeness, we compare repeated trials of different videos by clustering their resulting topological neural representations. Our results show that these topological descriptors reliably distinguish neural responses to distinct stimuli. This work highlights a connection between evolving neuronal activity and interpretable topological signatures, advancing the use of topological data analysis for uncovering neural coding in complex dynamical systems.

1. INTRODUCTION

A fundamental challenge in neuroscience is understanding how information about the world and an animal’s internal state and goals is encoded in the brain. Modern electrophysiological and imaging recording techniques allow for simultaneous recording of very large populations of neurons with single-cell resolution, and have driven the design of methods for the analysis of collective patterns of neuronal activity and their representational properties (Saxena and Cunningham, 2019).

A central concept in this research is that of the *neural manifold* – a low-dimensional geometric embedding of population activity within a high-dimensional neural state space (Gallego

et al., 2017). Experimental and theoretical studies suggest that neuronal activity typically lies on such manifolds, whose structure reflects the underlying task variables, e.g. object identity in the case of object recognition in visual cortex (DiCarlo et al., 2012; Chung and Abbott, 2021; Perich et al., 2025). Neural manifold learning methods—including both linear (e.g., PCA) and non-linear (e.g., Isomap, UMAP) techniques—have been applied to provide insight into systems as diverse as the head direction circuits in mouse (Chaudhuri et al., 2019) and zebrafish (Petrucco et al., 2023), grid cells (Gardner et al., 2022), motor regions across multiple species (Gao et al., 2017; Fortunato et al., 2024) and head ganglia during locomotion in *C. Elegans* (Brennan and Proekt, 2019).

Beyond a geometric characterization of the neural manifold, *topological data analysis* provides another layer of understanding by focusing on shape invariants rather than coordinates. Persistent homology—a key technique in topological data analysis—captures multi-scale topological features (such as loops and voids) in a robust, noise-tolerant manner (Edelsbrunner et al., 2000; Carlsson, 2009). Several works have used these tools in the context of computational neuroscience (e.g. (Giusti et al., 2015; Reimann et al., 2017)). However, classical persistent homology assumes a *static* point cloud filtration. Neuronal activity, by contrast, evolves dynamically, especially under time-varying stimuli or behavior. *Zigzag persistent homology* (Carlsson et al., 2009) extends persistence to sequences that allow additions and deletions of points, in our context the neuronal activity in time, enabling the direct tracking of topological structure as it changes.

In this work, we leverage zigzag persistence to analyze population responses from the Sensorium 2023 dataset (Turishcheva et al., 2024), which contains records of visual cortex neuronal activity of mouse exposed to grayscale videos. We develop a pipeline that takes the neuronal activity for each video, frame-by-frame, and transforms it into a sequence of evolving topological shapes underlying the neuronal activity. We then use zigzag persistence to track how these shapes change over time, focusing on cyclical patterns of neural co-activation. We summarize these evolving features using persistence landscapes (Bubenik, 2015), providing a compact, vectorized representation of the topological changes in time.

Our experiments address three questions: (A) do the zigzag descriptors separate repeated presentations of different videos to the same mouse? (B) do they carry signal about coarse video families (e.g., Naturalistic, Gaussian, Waves)? and (C) do they carry mouse identity? Our results show that:

- (A1) Repeated presentations of distinct videos reliably cluster by individual movie, with high Adjusted Rand Index (Hubert and Arabie, 1985) across video types.
- (A2) Shuffling frame order within each video collapses clustering performance, showing that temporal order is critical. Scrambling the spatial grid (per plane, per video) degrades performance toward chance, showing that spatial contiguity is also critical.
- (B) Linear classifiers trained on our descriptors to predict video types achieve accuracy modestly above chance.
- (C) Linear classifiers trained on our descriptors for mouse identification perform only slightly above chance.

Overall, these results indicate that applying zigzag persistence to frame-wise activity fields produces compact, interpretable topological signatures that are informative about specific dynamic stimuli. This approach opens a promising direction for future research, that of extending topological data analysis to capture the temporal structure of neural representation, potentially across behavioral contexts, brain regions, or in the presence of learning and adaptation in time.

2. DATASET

The Sensorium 2023 movie dataset (Turishcheva et al., 2024) contains two-photon calcium imaging data recorded from the visual cortex of behaving mice during repeated presentations of video clips. For each mouse, neuronal activity was recorded simultaneously from multiple imaging planes (z -slices) in V1. The dataset provides, for each video frame, single-trial fluorescence responses for all recorded neurons in each plane, along with 2D spatial coordinates of cell bodies and behavioural data (e.g. pupil diameter, treadmill velocity).

We consider a subset of 5 mice in the training set release. We select videos with the following classes: NATURALISTIC, GAUSSIAN, WAVES and MOVING DOT. Each class features different videos with several repeats.¹ In this subset, each scan comprises 10 z -planes, and each plane contains on average approximately 800 neurons, yielding several thousand units per mouse. Stimuli are grayscale videos presented at a fixed frame rate; responses are provided preprocessed by the authors of the dataset, aligned to stimulus frame times.²

We organize the data as follows: for each mouse a , plane p , video v , and frame t , we extract the vector $r_{a,p,v}(t)$ of single-trial neural responses across neurons in that plane. We keep only frames belonging to the active video (excluding inter-trial blanks). For downstream analysis, we transform r into $\delta_{a,p,v}(t) = (r_{a,p,v}(t) - \bar{r}_{a,p,v})/\bar{r}_{a,p,v}$, where the average is across frames, such that δ is a normalized response, which ranges from -1 (no response) to any positive value, and it is zero at average response per mouse. Each δ is interpolated on a 2D grid using Piecewise Cubic Spline interpolation. These per-frame grids are the inputs to the cubical complexes used by our zigzag pipeline, explained below.

3. METHODOLOGY

We build a sequence of complexes from per-frame 2D activity grids (one per z -plane) and extract their evolving topological features via zigzag persistence.

For each mouse a , plane p , video v , and frame t , we have a 2D activity $\delta_{a,p,v}(t)$ on a fixed 2D grid (Sec.2). We form a superlevel-set cubical complex $\mathcal{K}_{a,p,v}(t)$ by selecting all grid points (and their incident edges/squares) whose value exceeds the threshold $\delta_{\text{threshold}} = 0$ (activity above the per-neuron mean across frames). This yields one 2D cubical complex per frame.³

1. For more details, see Appendix A.

2. We do not include behavioural data in our analysis, which would allow for a more fine-grained examination of neural responses. We leave this for future work.

3. Cubical complexes have also been used in time-varying functional Magnetic Resonance Imaging (fMRI), see (Rieck et al., 2020).

For practical use of publicly available code computing zigzag persistence⁴, we map each $\mathcal{K}_{a,p,v}(t)$ to a simplicial complex by closing the corner set of every active square into an abstract simplex and then taking the 2-skeleton:

$$\mathcal{S}_{a,p,v}(t) = \text{Sk}_2\left(\bigcup\{C \text{ active (square/edge/vertex) at } t\} \langle V(C) \rangle\right). \quad (1)$$

Here, C ranges over active cubical cells (vertices, edges, squares), $V(C)$ is the set of its grid-corner vertices, and $\langle V(C) \rangle$ denotes the abstract simplex on $V(C)$ (in a simplex tree, inserting $\langle V(C) \rangle$ automatically inserts all its faces). Concretely, we: (i) insert all vertices (0-simplices); (ii) insert all active grid edges (1-simplices); (iii) for each active square with corners $v_{00}, v_{10}, v_{01}, v_{11}$, insert the 3-simplex $[v_{00}, v_{10}, v_{01}, v_{11}]$ and retain its 2-skeleton. This is not a triangulation of the square; it is an abstract “closure” that adds both diagonals and all triangular faces among the four corners.⁵

Having set up the simplicial complexes for each frame, we now construct an interleaved sequence over consecutive frames using intersections at intermediate steps:

$$\mathcal{S}_{a,p,v}(t_1) \leftarrow \mathcal{S}_{a,p,v}(t_1) \cap \mathcal{S}_{a,p,v}(t_2) \hookrightarrow \mathcal{S}_{a,p,v}(t_2) \leftarrow \dots \quad (2)$$

This “intersection–zigzag” matches our implementation, where even layers contain the common simplices of adjacent frames. We compute zigzag persistent homology in dimension 1 (H_1) on this sequence using code based on the implementation of [Gardinazzi et al. \(2025\)](#) for neural network layers. Each H_1 zigzag barcode is mapped to persistence landscapes Λ_k , where k is the landscape layer, over the time index. We sample the landscapes at a resolution of 50 and retain the first 5 landscape layers ($k = 1, \dots, 5$), yielding a 250-dimensional vector per plane.

4. EXPERIMENTS

We evaluate whether zigzag descriptors capture stimulus information across three settings: (A) clustering of repeated presentations of distinct videos of the same video type for a given mouse; (B) classification across video type for a given mouse; and (C) across-mouse classification. Each trial is represented by a 2500-dimensional vector obtained by concatenating the 10 z -planes.

Clustering protocol (A). For a given mouse and a given video type, we cluster repeated presentations of different videos of that type. Before clustering, we normalise all data vectors in the range $[0, 1]$ and reduce the dimensionality to 10 dimensions via PCA fit on the pooled set under evaluation. We then apply agglomerative clustering (Ward linkage, Euclidean distance) with the number of clusters set to the ground truth. Performance is reported as Adjusted Rand Index, for which chance corresponds to ≈ 0 , mean \pm std over 20 resamplings. In each resampling, we draw balanced subsets (typically 10 samples per class when available). As a consistency check, we run two different tests: (i) we permute frames within each video

4. We use a combination of DIONYSUS2 ([Morozov, 2012](#)) and FASTZIGZAG ([Dey and Hou, 2022](#)), see details in Appendix B.

5. For details about the algorithm and a discussion about this choice, we refer to Appendix B.

identically across planes (disrupt temporal order, preserve per-frame spatial structure) and (ii) we permute pixel values on each plane once per video (disrupt spatial contiguity, largely preserve marginal activity statistics). Results for these experiments are reported in Table 1.⁶

Condition	Baseline ARI	Frame-shuffled ARI	Grid-scrambled ARI
Naturalistic	0.945 ± 0.107	0.194 ± 0.118	0.180 ± 0.153
Gaussian	0.702 ± 0.156	0.018 ± 0.024	0.081 ± 0.072
Waves	0.619 ± 0.197	-0.005 ± 0.011	0.028 ± 0.004

Table 1: Clustering ARI (mean \pm std over 20 runs) for the mouse labelled 2–10. Baseline uses intact data. Frame-shuffled permutes frames within each video. Grid-scrambled permutes pixels per plane once per video.

Classification protocol (B) video type. For a given mouse, we train a linear logistic regression classifier to predict video type (e.g., Naturalistic, Gaussian, Waves) from the 2500-dimensional descriptors. We create stratified 80/20 train/test splits and repeat this procedure over 5 random splits; we find a cross-validation accuracy of 0.69 ± 0.06 . We also report the f1-score per class and confusion matrix aggregated on test-set predictions across splits (see Figure 7 in Appendix C).

Classification protocol (C) mouse identity. We similarly train a linear logistic regression classifier to predict mouse identity from the 2500-dimensional descriptors, pooling all available videos of the selected types across 5 mice. Because the same video is never shared across mice (due to dataset construction), this task probes whether descriptors predominantly reflect individual mouse features. In this case, the cross-validation accuracy is 0.36 ± 0.03 . The aggregated confusion matrix is shown in Figure 7 in Appendix C).

5. RESULTS AND CONCLUSIONS

We presented a pipeline that applies zigzag persistent homology to frame-by-frame neural population activity. On Sensorium 2023, H_1 zigzag descriptors based on persistence landscapes robustly separate repeated presentations of individual videos within mouse (high ARI; ≈ 0.94 Naturalistic, ≈ 0.70 Gaussian, ≈ 0.62 Waves) and only modestly separate stimulus type ($\approx 70\%$ accuracy). Two controls indicate that temporal order and spatial contiguity are essential: shuffling frames collapses clustering, and scrambling spatial grids drives performance toward chance. Mouse identification is just above chance when classifying mice that viewed disjoint movies ($\approx 35\%$ accuracy), suggesting limited sensitivity to mouse-specific characteristics.

Methodologically, we converted cubical complexes into simplicial complexes by closing active squares with an abstract simplex and retaining its 2-skeleton; this may over-connect cells and accelerate cycle filling. Future work will compare this choice with triangulated and

6. In Appendix C we quote an extended range of results for multiple mice and including other metrics (accuracy of cluster identification and adjusted mutual information).

cubical alternatives. The analysis can also be improved by incorporating behavioral data to generally reduce unexplained variance in neural responses. Overall, our results indicate that time-resolved cyclic coordination patterns in neural populations carry stimulus-specific information that zigzag persistence can extract and quantify.

ACKNOWLEDGEMENTS

M.B. thanks Mathieu Carrière and Karthik Viswanathan for useful feedback on the zigzag implementation. M.B., Y.G. and F.A. thank Nikos Karantzas for discussion about the dataset in the early stages of the project. Y.G. is supported by the Programma Nazionale della Ricerca (PNR) grant J95F21002830001 with title “FAIR-by-design”. E.P. was partly supported by the PRIN Project no. 2022XE8X9E (CUP:G53D23004590001) and by the SISSA 5x1000 IRPEF 2023 funds.

We thank Area Science Park supercomputing platform ORFEO made available for conducting the research reported in this paper and the technical support of the Laboratory of Data Engineering staff.

REFERENCES

- Connor Brennan and Alexander Proekt. A quantitative model of conserved macroscopic dynamics predicts future motor commands. *eLife*, 8:e46814, July 2019. ISSN 2050-084X. doi: 10.7554/eLife.46814.
- Peter Bubenik. Statistical topological data analysis using persistence landscapes. *Journal of Machine Learning Research*, 16(1):77–102, 2015.
- Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.
- Gunnar Carlsson, Vin de Silva, and Dmitriy Morozov. Zigzag persistent homology and real-valued functions. *Proceedings of the Annual Symposium on Computational Geometry*, pages 247–256, 2009.
- Rishidev Chaudhuri, Ari S Gershon, and Haim Sompolinsky. Intrinsic dimensionality of behavioral and neural representations. *Annual review of neuroscience*, 42:381–410, 2019.
- SueYeon Chung and L. F. Abbott. Neural population geometry: An approach for understanding biological and artificial neural networks. *Current Opinion in Neurobiology*, 70: 137–144, October 2021. ISSN 0959-4388. doi: 10.1016/j.conb.2021.10.010.
- Tamal K. Dey and Tao Hou. Fast Computation of Zigzag Persistence. In Shiri Chechik, Gonzalo Navarro, Eva Rotenberg, and Grzegorz Herman, editors, *30th Annual European Symposium on Algorithms (ESA 2022)*, volume 244 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 43:1–43:15, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-247-1. doi: 10.4230/LIPIcs.ESA.2022.43. URL <https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ESA.2022.43>.

- James J. DiCarlo, Davide Zoccolan, and Nicole C. Rust. How does the brain solve visual object recognition? *Neuron*, 73(3):415–434, February 2012. ISSN 1097-4199. doi: 10.1016/j.neuron.2012.01.010.
- Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 454–463, 2000.
- Cátia Fortunato, Jorge Bannasar-Vázquez, Junchol Park, Joanna C. Chang, Lee E. Miller, Joshua T. Dudman, Matthew G. Perich, and Juan A. Gallego. Nonlinear manifolds underlie neural population activity during behaviour, April 2024.
- 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.
- Peiran Gao, Eric Trautmann, Byron M Yu, Gopal Santhanam, Stephen Ryu, Krishna V Shenoy, and Surya Ganguli. A theory of multineuronal dimensionality, dynamics and measurement. *bioRxiv*, page 214262, 2017.
- Yuri Gardinazzi, Karthik Viswanathan, Giada Panerai, Alberto Cazzaniga, Matteo Biagetti, et al. Persistent topological features in large language models. In *Forty-second International Conference on Machine Learning*, 2025.
- 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, February 2022. ISSN 1476-4687. doi: 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.
- Lawrence J. Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2: 193–218, 1985. URL <https://api.semanticscholar.org/CorpusID:189915041>.
- Dmitriy Morozov. Dionysus2, 2012. URL <https://www.mrzv.org/software/dionysus2/>.
- 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 1546-1726. doi: 10.1038/s41593-025-02031-z.
- Luigi Petrucco, Hagar Lavian, You Kure Wu, Fabian Svara, Vilim Štíh, and Ruben Portugues. Neural dynamics and architecture of the heading direction circuit in zebrafish. *Nature Neuroscience*, 26(5):765–773, May 2023. ISSN 1546-1726. doi: 10.1038/s41593-023-01308-5.
- Michael W. Reimann, Max Nolte, Martina Scolamiero, Katharine Turner, Rodrigo Perin, Giuseppe Chindemi, Paweł Dłotko, Ran Levi, Kathryn Hess, and Henry Markram. Cliques of neurons bound into cavities provide a missing link between structure and

function. *Frontiers in Computational Neuroscience*, Volume 11 - 2017, 2017. ISSN 1662-5188. doi: 10.3389/fncom.2017.00048. URL <https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2017.00048>.

Bastian Rieck, Tristan Yates, Christian Bock, Karsten Borgwardt, Guy Wolf, Nicholas Turk-Browne, and Smita Krishnaswamy. Uncovering the topology of time-varying fmri data using cubical persistence. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Shreya Saxena and John P Cunningham. Towards the neural population doctrine. *Current Opinion in Neurobiology*, 55:103–111, April 2019. ISSN 0959-4388. doi: 10.1016/j.conb.2019.02.002.

Polina Turishcheva, Paul G. Fahey, Laura Hansel, Rachel Froebe, Kayla Ponder, Michaela Vystrčilová, Konstantin F. Willeke, Mohammad Bashiri, Eric Wang, Zhiwei Ding, Andreas S. Tolias, Fabian H. Sinz, and Alexander S. Ecker. The dynamic sensorium competition for predicting large-scale mouse visual cortex activity from videos, 2024. URL <https://arxiv.org/abs/2305.19654>.

APPENDIX A. MORE INFO ON SENSORIUM (2023)

In this section, we provide further details on the Sensorium dataset ([Turishcheva et al., 2024](#)) used in this work.

Video classes. For each mouse, ≈ 700 videos are provided with corresponding neural responses and behavioural metadata. These videos are characterized by subgroups of repeated entries. Video types change across mice, we select 4 video types and label them as NATURALISTIC, GAUSSIAN, WAVES and MOVING DOT. We show a sample frame for each type in Figure 1. Groups of repeated videos have typically 10 samples each. Each video contains 324 frames, with the last 24 frames being blank. The format of each video is (36, 64) in grayscale.

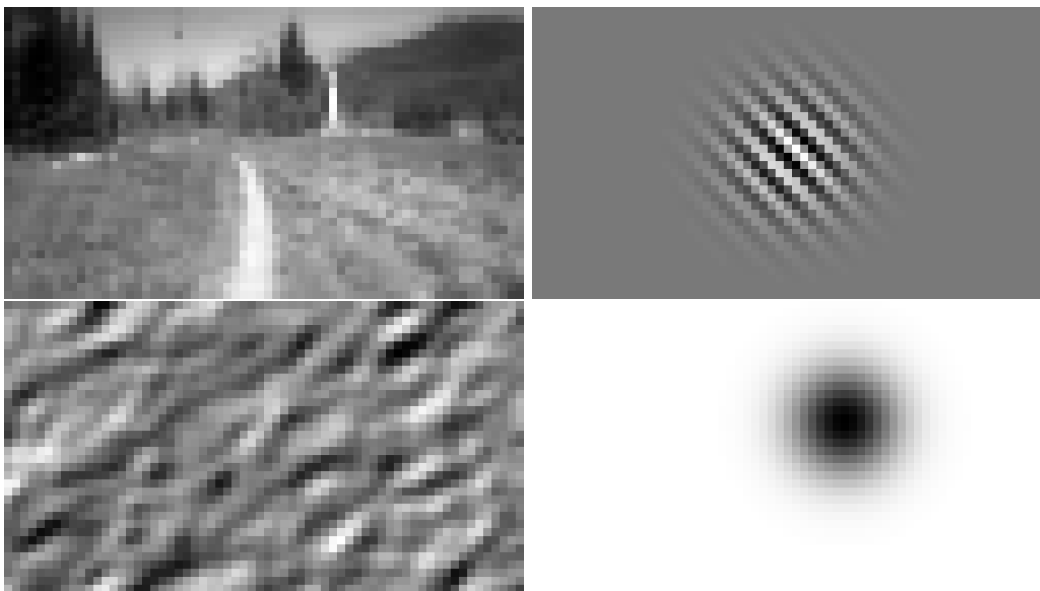


Figure 1: Middle frame for each video type considered in this work: NATURALISTIC (upper left), GAUSSIAN (upper right), WAVES (bottom left), MOVING DOT (bottom right).

Spatial sampling and geometry of neuron positions. For a representative mouse–video pair (mouse 2-10, video 34), we show cell positions in the three projections (X–Y, X–Z, Y–Z) in Figure 2. Neurons populate a stack of regularly spaced z-planes; after coordinate normalization, the spacing between adjacent planes is ≈ 0.11 (marked in the distance plots as a reference scale). Within-plane sampling density appears comparable across slices: pairwise X–Y distances have similar distributions for each z-plane, indicating a fairly uniform coverage in each image plane. Cell counts per plane are also broadly balanced, with only minor variability across z. We show these characteristics for the same sample case in Figure 3. We also checked for potential duplicates across adjacent planes: only two (x, y) locations recur at different z, and their fluorescence time series are numerically indistinguishable, indicating rare double detections that are negligible at the dataset scale.

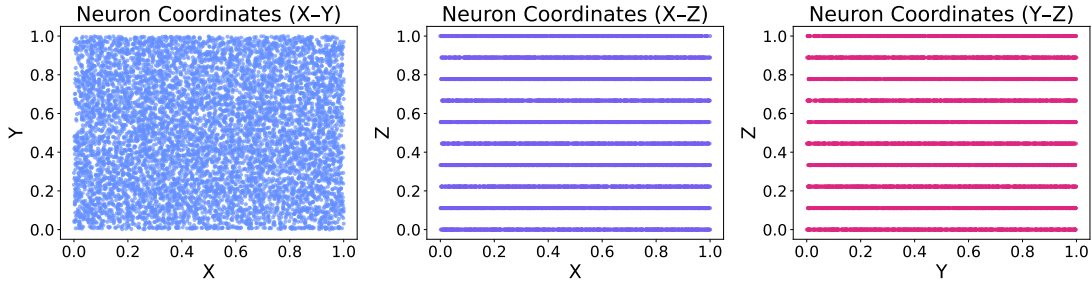


Figure 2: **Spatial sampling geometry.** Neuron positions for a representative mouse–video pair, shown in three projections: X–Y (left), X–Z (center), and Y–Z (right). The laminar stack of regularly spaced z-planes is visible, and within-plane coverage appears uniform. Coordinates are normalized to the imaging volume.

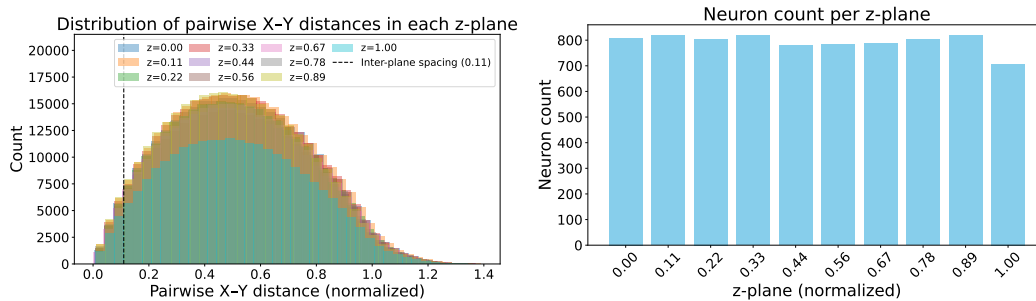


Figure 3: Left: **Within-plane sampling density.** Distribution of pairwise X–Y distances between neurons for each z-plane (overlaid histograms). The vertical dashed line marks the normalized inter-plane spacing (≈ 0.111), serving as a reference scale. The similar distributions across planes indicate uniform sampling density within each image plane. Right: **Depth sampling balance.** Bar plot of neuron counts per z-plane for the same mouse–video pair. Cell counts are broadly balanced across planes, with only minor variability, confirming uniform sampling in depth.

Neural responses. For each neuron we use single-trial fluorescence traces time-locked to the video frames. The duplicate check above confirms that when a cell is inadvertently captured in neighboring planes, its activity traces are identical, as expected. In the main analyses we work with per-frame response interpolated on 2D grids; for descriptive purposes, it is useful to visualize raw traces and simple aggregates per plane to illustrate trial structure and variability across neurons and depths, see Figure 4.

APPENDIX B. ZIGZAG ALGORITHM

Here we describe the full analysis pipeline. A schematic representation of each step is shown in Figure 5. The algorithm is presented below: For each video and each z-plane, we load the interpolated activity grid, normalize it frame-wise, and optionally apply either frame shuffling or spatial scrambling as controls. At each frame we build the active cubical cells from the thresholded mask ($\delta > 0$), then map them to a simplicial complex via the closure

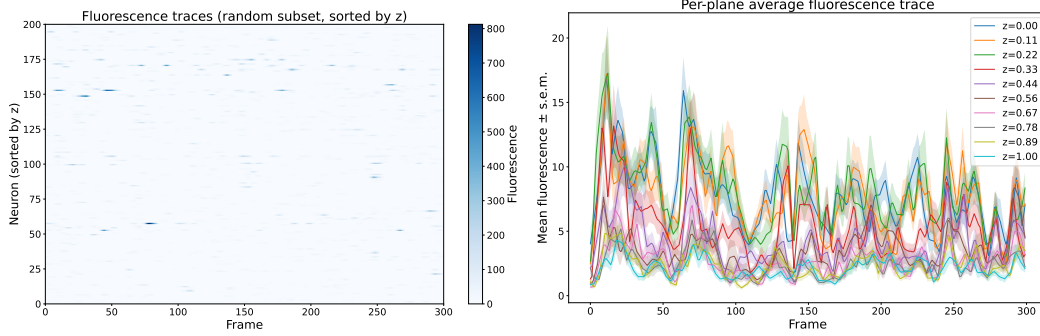


Figure 4: Left: **Fluorescence response overview.** Heatmap of single-trial fluorescence traces for a random subset of 200 neurons, sorted by z-plane. Each row is a neuron, each column a video frame. This illustrates the diversity and temporal structure of neural responses across the imaging volume. Right: **Per-plane average fluorescence dynamics.** Mean fluorescence trace (solid line) and standard error (shaded area) for each z-plane, showing depth-dependent response dynamics and variability across neurons and planes.

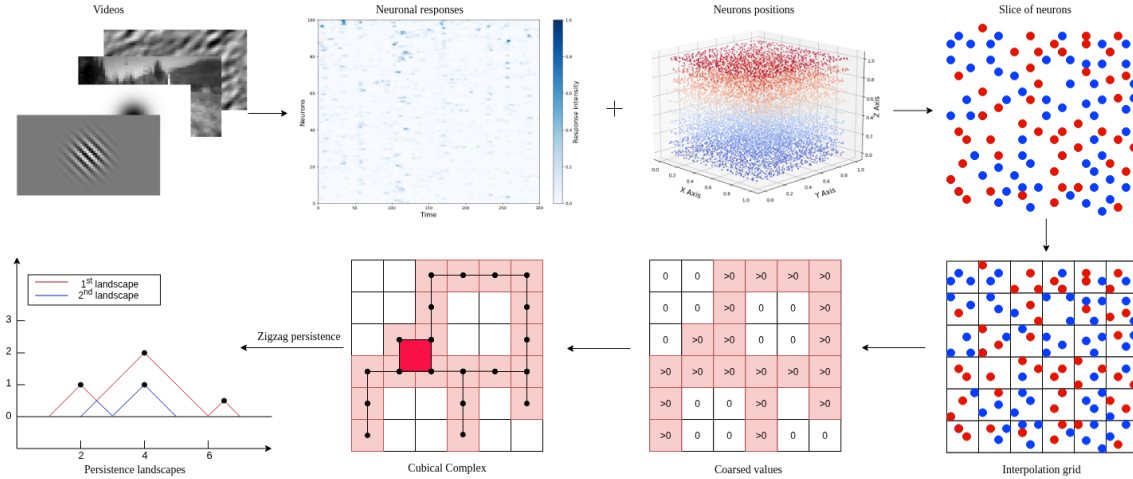


Figure 5: Schematic representation of the full zigzag pipeline.

adapter described below. We interleave adjacent-frame intersections to form the zigzag sequence, compute zigzag persistence, select the H_1 bars, and vectorize them with persistence landscapes (sampling R points and retaining k layers) to obtain one descriptor per plane.

For each mouse a , video v , plane p and frame t , we load the interpolated activity grid $r_{a,p,v}(t)$ and convert it to a per-pixel activity $\delta_{a,p,v}(t)$ which centers each pixel at zero and makes the threshold $\delta > 0$ correspond to above-average activity within that plane and video.

To probe what the descriptor uses, we optionally apply:

- a frame-order shuffle $t \mapsto \pi(t)$ (identical across planes) to disrupt temporal order while preserving per-frame spatial structure;

- a fixed spatial permutation π per (i, z) to scramble the grid indices across all frames, disrupting spatial contiguity while largely preserving marginal activity statistics.

Per-frame complexes and simplicial adapter. For each frame t , we form a binary mask of active pixels ($\delta > 0$) and construct the corresponding 2D cubical cells (active vertices, grid-edges, and unit-squares). To interface with zigzag software operating on simplicial complexes, we map this cubical complex to a simplicial complex S_t using an abstract closure: (i) insert all active vertices as 0-simplices and all active grid-edges as 1-simplices; (ii) for each active square with four grid-corner vertices, insert the abstract 3-simplex on those four vertices and retain only its 2-skeleton (thus adding both diagonals and all triangular faces). This is not a triangulation; it over-connects to ensure consistent 2D faces across frames and stable computation. Vertex identifiers are the grid coordinates, so identical simplices are recognized across time.

Interleaved intersection zigzag. We assemble the time sequence with intersections at intermediate steps:

$$S_1 \leftarrow S_1 \cap S_2 \hookrightarrow S_2 \leftarrow S_2 \cap S_3 \hookrightarrow \dots \leftarrow S_{T-1} \cap S_T \hookrightarrow S_T. \quad (3)$$

Intersections are set-theoretic intersections of simplices using global vertex labels. This produces a valid zigzag of inclusions that captures gains and losses of simplices between consecutive frames while maintaining well-defined maps.

Zigzag persistence and vectorization. We compute H_1 zigzag persistent homology on the sequence X using two public libraries:

- DIONYSUS2 (Morozov, 2012), which provides data structures and routines to encode the zigzag filtration and its time indices (birth/death layers);
- FASTZIGZAG (Dey and Hou, 2022), which converts the zigzag filtration to an equivalent non-zigzag filtration of an auxiliary complex, computes standard persistence efficiently, and maps intervals back to the zigzag setting to obtain the H_1 barcode.

This combination of codes and algorithm allows for a computational time of about 150 ms on a single CPU per single data vector, i.e. roughly 100 minutes to run the full pipeline on a single mouse (≈ 700 video stimuli).

Topological descriptors. Each H_1 barcode is mapped to persistence landscapes Λ_k sampled on a uniform grid over the time index; in our experiments we use $R = 50$ sample points and $k = 5$ landscape layers, yielding a 250-dimensional vector per plane. Trial-level descriptors concatenate all planes. In Figure 6 we show a representative data vector for mouse 2-10 on a single z -plane both for the full 250-dimensional vector for a naturalistic video and a comparison of the first landscapes between a naturalistic video and a Gaussian video.

APPENDIX C. ADDITIONAL RESULTS

In this section, we quote additional results to complement the results quoted in the main text.

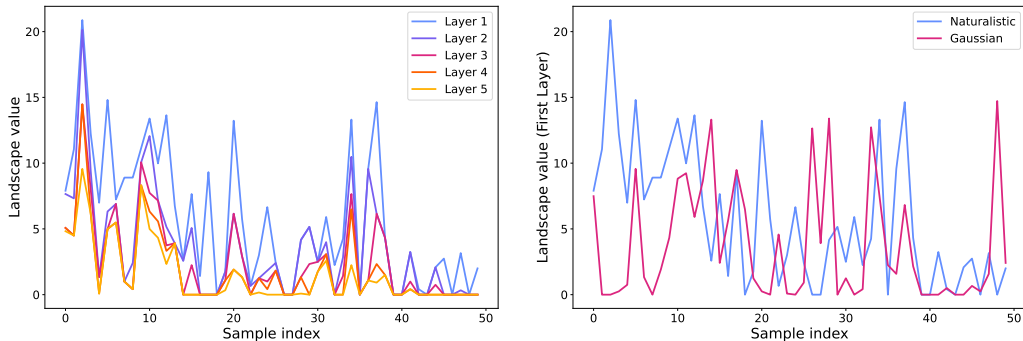


Figure 6: **Representative persistence landscape vectors.**

Left: Each curve shows one of the first five landscape layers (Λ_k) for a single trial and a single z -plane, summarizing the time-varying prominence of topological cycles (H_1) in the neuronal activity. Right: Each curve shows the first landscape layer (Λ_1) of a single z -plane of trials corresponding to a naturalistic video and a Gaussian video. All trials refer to mouse 2-10.

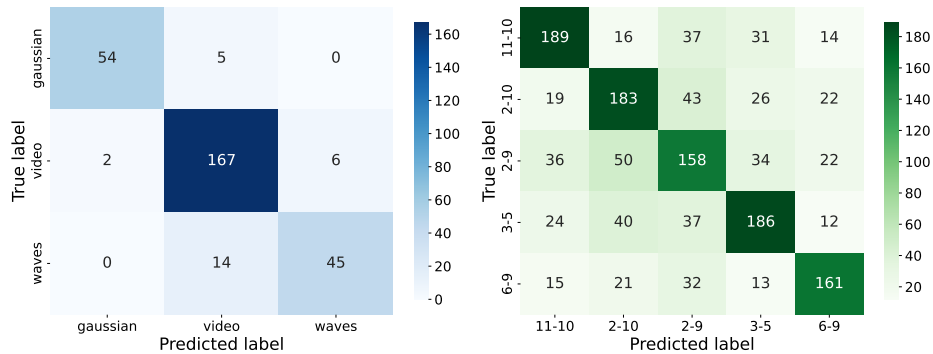
Clustering protocol (A). For the clustering experiments, we run the pipeline on all 5 mice, shown in Table 2, where we also include the accuracy of classification by clustering as a metric. For these, we did not run control-type checks. Results are broadly coherent with what shown in Table 1 for mouse 2-10.

Classification protocol (B) video family. Here we provide for all mice the F1-score for each class and the cross-validation accuracies on the task of predicting the right video family, Table 3. Additionally, we show the aggregated confusion matrix for 5 train/test splits for the mouse 2-10, see left panel of Figure 7.⁷

Classification protocol (C) mouse identity. Here we provide for the aggregated confusion matrix for 5 train/test splits on the mouse identity classification task, see Right panel of Figure 7.

7. Note: discrepancy of accuracy rates between the aggregated confusion matrix and the cross-validation score is because the latter uses mean per-fold accuracy, which is more reliable for unbalanced classes and small datasets, as in this case. The aggregated confusion matrix is quoted for interpretability purposes.

Mouse 2-10		
Condition	Baseline ARI	Baseline Accuracy
Naturalistic	0.945 ± 0.118	0.955 ± 0.076
Gaussian	0.702 ± 0.156	0.837 ± 0.104
Waves	0.603 ± 0.218	0.799 ± 0.126
Mouse 2-9		
Condition	Baseline ARI	Baseline Accuracy
Naturalistic	0.734 ± 0.161	0.863 ± 0.103
Gaussian	0.820 ± 0.141	0.921 ± 0.071
Moving Dot	0.463 ± 0.239	0.720 ± 0.148
Mouse 11-10		
Condition	Baseline ARI	Baseline Accuracy
Naturalistic	0.791 ± 0.185	0.884 ± 0.117
Waves	0.423 ± 0.225	0.683 ± 0.128
Moving Dot	0.982 ± 0.069	0.993 ± 0.026
Mouse 3-5		
Condition	Baseline ARI	Baseline Accuracy
Naturalistic	0.911 ± 0.098	0.963 ± 0.044
Waves	0.446 ± 0.196	0.706 ± 0.120
Moving Dot	0.791 ± 0.078	0.921 ± 0.031
Mouse 6-9		
Condition	Baseline ARI	Baseline Accuracy
Naturalistic	0.912 ± 0.138	0.955 ± 0.081
Gaussian	0.953 ± 0.067	0.983 ± 0.025

 Table 2: Clustering ARI and accuracy (mean \pm std over 20 runs) for the all mice.

 Figure 7: Aggregated confusion matrix for 5 train/test splits. *Left*: video type. *Right*: Mouse identity.

Mouse 2-10		
Class	F1	Support
Naturalistic	0.88	35
Gaussian	0.82	12
Waves	0.78	12
CV accuracy: 0.686 ± 0.055		

Mouse 2-9		
Class	F1	Support
Naturalistic	0.95	37
Gaussian	0.96	12
Moving Dot	0.83	11
CV accuracy: 0.743 ± 0.064		

Mouse 11-10		
Class	F1	Support
Naturalistic	0.93	35
Gaussian	0.86	11
Moving Dot	0.91	12
CV accuracy: 0.721 ± 0.115		

Mouse 3-5		
Class	F1	Support
Naturalistic	0.89	36
Waves	0.73	12
Moving Dot	0.82	12
CV accuracy: 0.672 ± 0.056		

Mouse 6-9		
Class	F1	Support
Naturalistic	1.0	37
Gaussian	1.0	12
CV accuracy: 0.888 ± 0.064		

Table 3: Per-class F1 scores and supports per mouse for a single train/test split. Cross-validated accuracy is run on 5 train/test splits.